NGU Rapport 94.080

Geochemical processes, Weathering and Ground Water Recharge in Catchments - Short Course Notes

NORGES GEOLOGISKE UNDERSØKELSE

Postboks 3006 - Lade
7002 TRONDHEIM
Tlf. 73 90 40 11
Telefax 73 92 16 20

RAPPORT

Rapport nr. 94.080	ISSN 0800-3416	Gradering: Åpen

Tittel: Geochemical Processes, Weathering and Ground Water Recharge in Catchments - Short Course Notes.

Forfatter: Ola M. Sæther, Patrice de Caritat & Knut Sand (Editors)	Oppdragsgiver: NGU NorFa
Fylke: Sør-Trøndelag	Kommune: Selbu
Kartbladnavn (M=1:250.000)	Kartbladnr. og -navn (M=1:50.000)
Forekomstens navn og koordinater:	Sidetall: 516 Pris: 250,- + Field excursion guides Kartbilag:

Feltarbeid utført:	Rapportdato: 13/10-1994	Prosjektnr.: 63.2628.00	Ansvarlig:

Sammendrag:

These are the Course Notes prepared for the NorFA-sponsored Short Course on "Geochemical processes, weathering and ground water recharge in catchments" held at Selbusjøen Hotell & Gjestegård October 16-23, 1994.

There were 26 students from the Nordic Countries including two from Latvia, and 15 lecturers from Denmark (1), Norway (8), Sweden (1), UK (3) and the USA (2).

The one-week course was aimed at doctoral students who plan to do research within the interdisciplinary field of catchment research. The disciplines involved are, among others, biology, soil chemistry, geology and hydrogeology. This field of research is of increasing international interest in connection with climatic changes, which may affect the ecology of catchments (e.g. Americal Geophysical Union EOS, v. 75, no. 38, Sept. 1994, p. 433-439).

The Scandinavian countries have been active contributors to research related to the acidification of drainage basins for over twenty years.

The purpose of this Short Course is to give an overview of aspects that need to be integrated when one is to carry out catchment research. It is hoped that these Course Notes will be of assistance in bringing the candidates closer to the research front, and to establish contacts and cooperation between the participating scientists.

Emneord: Geochemistry	Catchment	Weathering
Ground Water	Processes	Ground Water Recharge
NorFa short course	Scientific Report	

GEOCHEMICAL PROCESSES, WEATHERING AND GROUND WATER RECHARGE IN CATCHMENTS

SHORT COURSE
16 to 23 October 1994, Selbusjøen (Trondheim area), Norway

PROGRAMME

PART I. REGIONAL CONTEXT AND BACKGROUND RELEVANT TO NORDIC COUNTRIES

1) Geological context of the Nordic countries. Bedrock geology	*Brian Sturt*	*16/10*	*1400-1445*
2) Geological context of the Nordic countries. Quaternary geology	*Eiliv Larsen*	*16/10*	*1445-1530*
3) Climatic context of the Nordic countries	*Arvid Skartveit*	*16/10*	*1545-1700*
4) Hydrologic context of the Nordic countries	*Åmund Killingtveit*	*16/10*	*1715-1830*

PART II. GEOCHEMICAL PROCESSES AND WEATHERING

1) Introduction to weathering processes	*James Drever*	*17/10*	*0800-1215*
2) Introduction to soil development processes	*Ole Borggaard*	*17/10*	*1400-1815*

PART III. GROUND WATER RECHARGE

1) Introduction to catchment hydrology	*Allan Rodhe*	*18/10*	*0800-1215*
2) Introduction to ground water recharge	*David Lerner*	*18/10*	*1400-1815*

PART IV. INTEGRATING WEATHERING AND GROUND WATER RECHARGE

1) Inorganic aspects of catchment mass balance	*James Drever*	*21/10*	*0800-1215*
2) Organic aspects of catchment mass balance	*Don Macalady*	*21/10*	*1400-1815*
3) Relationship between rock, soil and ground water compositions	*John Mather*	*22/10*	*0800-1215*
4) Towards coupling hydrological, soil and weathering processes - a modelling perspective	*Colin Neal*	*22/10*	*1400-1815*

PART V. TECHNICAL ASPECTS OF CATCHMENT STUDIES

1) Chemical analysis of rocks and soils	*Magne Ødegård*	*19/10*	*0800-1000*
2) Chemical analysis of water	*John Mather*	*19/10*	*1400-1600*
3) Tracer studies	*Allan Rodhe*	*19/10*	*1015-1215*
4) Manual measurements and automatic instrumentation	*Åmund Killingtveit & Knut Sand*	*19/10*	*1615-1815*

PART VI. FIELD TRIPS

1) Sagelva and Kaldvella field stations	*Åmund Killingtveit, Knut Sand & Gaute Storrø*	*20/10*	*0830-1800*
2) Selbu-Tydal-Røros-Ålen-Selbu	*Ola Sæther & Patrice de Caritat*	*23/10*	*0830-1630*

Dear participant!

It is with great pleasure that we welcome you to this NorFa-sponsored Short-Course on *«Geochemical processes, weathering and ground water recharge in catchments»*. We have worked towards this over the past year, and are very pleased to have received the wholehearted support from the lecturers and a positive response from the students within the Nordic countries and the Baltic States.

Our intention has been to tailor a one-week course aimed at doctoral students who plan to do research within the interdisciplinary field of catchment research. The disciplines involved are among others, biology, soil chemistry, geology and hydrogeology. This field of research is of increasing international interest in connection with climatic changes, which may affect the ecology of catchments (e.g. American Geophysical Union EOS, v.75, no. 38, Sept. 1994, p. 433-439). The Scandinavian countries have been active contributors of research related to the acidification of drainage basins for over twenty years.

The purpose of this Short-Course is to give an overview of aspects which need to be integrated when one is to carry out catchment research. It is hoped that these course notes will be of assistance in bringing the candidates closer to the research front in a shorter time than would otherwise be the case, and to establish contacts and cooperation between the participating scientists.

We at the Geological Survey of Norway (NGU), the Norwegian Technical University (NTH) and the Norwegian Hydrotechnical Laboratories (NHL), feel that it is very important to join forces within the fields of geology and hydrology to provide scientists within the disciplines of agriculture, biology, forestry and climatology with the best possible foundation for their work.

Please enjoy this week in Selbu!

Brian Sturt, NGU
Chairman org. comm.

Ola M. Sæther, NGU
Patrice de Caritat, NGU
Gaute Storrø, NGU
Aanund Killingtveit, NTH
Knut Sand, NHL

GEOCHEMICAL PROCESSES, WEATHERING AND GROUND WATER RECHARGE IN CATCHMENTS

SHORT COURSE
16 to 23 October 1994, Selbusjøen (Trondheim area), Norway

SCHEDULE

Sunday 16 October 1994

1000-1300	Arrival and registration	
1300-1345	Lunch	
1345-1400	Introductory remarks	*Brian Sturt*
1400-1445	***Part I.1-Regional context and background relevant to Nordic countries:*** ***Geological context of the Nordic countries. Bedrock geology***	*Brian Sturt*
1445-1530	***Part I.1-Regional context and background relevant to Nordic countries:*** ***Geological context of the Nordic countries. Quaternary geology***	*Eiliv Larsen*
1530-1545	Coffee break	
1545-1700	***Part I.2-Regional context and background relevant to Nordic countries:*** ***Climatic context of the Nordic countries***	*Arvid Skartveit*
1700-1715	Break	
1715-1830	***Part I.3-Regional context and background relevant to Nordic countries:*** ***Hydrologic context in the Nordic countries***	*Ånund Killingtveit*
1900-2030	Dinner	
2030-2200	Icebreaker	

Monday 17 October 1994

0700-0800	Breakfast	
0800-1215	***Part II.1-Geochemical Processes and Weathering:*** ***Introduction to weathering processes*** (including a short coffee break at ca. 1000)	*James Drever*
1215-1300	Buffer zone (discussions, free time)	
1300-1400	Lunch	
1400-1815	***Part II.2-Geochemical Processes and Weathering:*** ***Introduction to soil development processes*** (including a short coffee break at ca. 1600)	*Ole Borggaard*
1815-1900	Buffer zone (discussions, free time)	
1900-2030	Dinner	
2030-2200	Presentation of local history, culture and handicrafts	

Tuesday 18 October 1994

0700-0800	Breakfast	
0800-1215	***Part III.1-Ground water recharge: Introduction to catchment hydrology*** (including a short coffee break at ca. 1000)	*Allan Rodhe*
1215-1300	Buffer zone (discussions, free time)	
1300-1400	Lunch	
1400-1815	***Part III.2-Ground water recharge: Introduction to ground water recharge*** (including a short coffee break at ca. 1600)	*David Lerner*
1815-1900	Buffer zone (discussions, free time)	
1900-2000	Dinner	
2000-2200	***Tutorial on Soil Classification***	*Ole Borggaard*

Wednesday 19 October 1994

0700-0800	Breakfast	
0800-1000	***Part V.1-Technical aspects of catchment studies: Chemical analysis of rocks and soils***	*Magne Ødegård*
1000-1015	Coffee break	
1015-1215	***Part V.3-Technical aspects of catchment studies: Tracer studies***	*Allan Rodhe*
1215-1300	Buffer zone (discussions, free time)	
1300-1400	Lunch	
1400-1600	***Part V.2-Technical aspects of catchment studies: Chemical analysis of water***	*John Mather*
1600-1615	Coffee break	
1615-1815	***Part V.4-Technical aspects of catchment studies: Manual measurements and automatic instrumentation***	*Ånund Killingtveit & Knut Sand*
1815-1900	Buffer zone (discussions, free time)	
1900-2030	Dinner	
2030-2200	Editorial meeting for lecturers	

Thursday 20 October 1994

0700-0800	Breakfast	
0830-1800	***Part VI.1-Field trip 1: Sagelva and Kaldvella field stations***	*Ånund Killingtveit, Knut Sand & Gaute Storrø*
1800-1900	Free time	
1900-2030	Dinner	
2030-2200	Informal discussions	

Friday 21 October 1994

0700-0800	Breakfast	
0800-1215	***Part IV.1-Integrating weathering and ground water recharge: Inorganic aspects of catchment mass balance*** (including a short coffee break at ca. 1000)	*James Drever*
1215-1300	Press conference	
1300-1400	Lunch	
1400-1815	***Part IV.2-Integrating weathering and ground water recharge: Organic aspects of catchment mass balance*** (including a short coffee break at ca. 1600)	*Don Macalady*
1815-1900	Buffer zone (discussions, free time)	
1900-2030	Dinner	
2030-2200	Informal discussions	

Saturday 22 October 1994

0700-0800	Breakfast	
0800-1215	***Part IV.3-Integrating weathering and ground water recharge: Relationship between rock, soil and ground water compositions*** (including a short coffee break at ca. 1000)	*John Mather*
1215-1300	Buffer zone (discussions, free time)	
1300-1400	Lunch	
1400-1815	***Part IV.4-Integrating weathering and ground water recharge: Towards coupling hydrological, soil and weathering processes - a modelling perspective*** (including a short coffee break at ca. 1600)	*Colin Neal*
1815-1830	Closing remarks	*Brian Sturt*
1830-1900	Buffer zone (discussions, free time)	
1900-2030	Dinner (Banquet)	
2030-2300	Folk music & dance	

Sunday 23 October 1994

0700-0800	Breakfast	
0830-1630	***Part VI.2-Field trip 2: Selbu-Tydal-Røros-Ålen-Selbu***	*Ola Sæther &* *Patrice de Caritat*
1700-1900	Drop off at airport and Hell & Storlien train station	

...THE END...

NB:

Monday 24 October 1994 The national seminar on "Environmental Hydrogeology and Geochemistry" will take place this day in Trondheim for those interested to participate on a personal basis (not part of the Short Course)

GEOCHEMICAL PROCESSES, WEATHERING AND GROUND WATER RECHARGE IN CATCHMENTS

SHORT COURSE
16 to 23 October 1994, Selbusjøen (Trondheim area), Norway

KEY WORDS

PART I. REGIONAL CONTEXT AND BACKGROUND RELEVANT TO NORDIC COUNTRIES

1) Geological context of the Nordic countries. Bedrock geology *Brian Sturt* *16/10 1400-1445*
Key words: *Pangea, Baltica, Precambrian complexes, Archean complexes, Greenstone Belts, Svecofennian Orogeny, Iapetus Ocean, Caledonides, Oslo Graben, North Sea, Fault reactivation, Neotectonic movements*

1) Geological context of the Nordic countries. Quaternary geology *Eiliv Larsen* *16/10 1445-1530*
Key words: *Climate change, Quaternary geology, Scandinavian ice sheet, Glaciation, Deglaciation, Sediments*

2) Climatic context of the Nordic countries *Arvid Skartveit* *16/10 1545-1700*
Key words: *Climatic controls, Temporal and spatial climatic variability, Air pollution, Acid precipitation, Snow, Freezing/thawing*

3) Hydrologic context of the Nordic countries *Ånund Killingtveit* *16/10 1715-1830*
Key words: *Precipitation, Evaporation and runoff distribution, Groundwater recharge, Hydrological regions, River flow regimes*

PART II. GEOCHEMICAL PROCESSES AND WEATHERING

1) Introduction to weathering processes *James Drever* *17/10 0800-1215*
Key words: *Weathering, Silicate minerals, Alkalinity, Dissolution kinetics, Stoichiometry, Flow paths, Mass balance, Loch Vale (Colorado, USA)*

2) Introduction to soil development processes *Ole Borggaard* *17/10 1400-1815*
Key words: *Soils, Genesis, Constituents, Soil organic matter, Clay, Silicates, Oxides, Interactions, Reactions, Transport, Leaching, Classification, Alfisol, Ultisol, Spodosol, Histosol*

PART III. GROUND WATER RECHARGE

1) Introduction to catchment hydrology *Allan Rodhe* *18/10 0800-1215*
Key words: *The water cycle; The catchment: Water divide, Water budget, Temporary storage in snow, Surface water, Soil water, Ground water; Soil water storage and flow: Basic physical processes, Darcy's Law for unsaturated and saturated flow, Particle velocity versus flow propagation velocity, Effect of vertical variations in soil texture and structure, Flow in macropores and soil matrix; Stream flow generation: Historical perspective, Recharge and discharge areas for ground water, Flow paths from infiltration to stream, Ground water response mechanisms, Role of topography; Stream runoff: Seasonal variation in the Nordic countries, Relation between physiographical factors and stream runoff (mean values and short time variations), The HBV-model for runoff modelling*

2) Introduction to ground water recharge *David Lerner* *18/10 1400-1815*
Key words: *Definitions, Hydrological cycle, Processes, Hydrogeological environments, Direct measurements, Water balances, Darcy's Law, Tracer techniques, Empirical methods, Accuracy, Lysimeters, Soil moisture budgets, Point tracer techniques, Aquifer-wide tracers, River recharge, Surface water modelling, base flow estimation, Inter-aquifer flows, Man-made recharge, Irrigation, Urbanisation, Regional recharge, Water table rise, Groundwater discharge, Inverse techniques, Time and spatial variability, Recharge estimation procedure*

PART IV. INTEGRATING WEATHERING AND GROUND WATER RECHARGE

1) Inorganic aspects of catchment mass balance *James Drever* *21/10 0800-1215*
Key words: *Weathering, Silicate minerals, Alkalinity, Dissolution kinetics, Stoichiometry, Flow paths, Mass balance, Loch Vale (Colorado, USA)*

2) Organic aspects of catchment mass balance *Don Macalady* *21/10 1400-1815*
Key words: *Natural organic matter (NOM), Humic substances, Sorption/partitioning, Complexation, Colloidal materials, Chemical weathering, Redox transformations, Pedogenic/aquagenic, Hydrolysis, Soils, Decomposition reactions*

3) Relationship between rock, soil and ground water compositions *John Mather* *22/10 0800-1215*
Key words: *Composition of the precipitation source-term, Water/soil interactions, Interactions in the unsaturated and saturated zones, Significance of residence time and hydraulic pathways, Environmental isotopes of carbon, hydrogen and oxygen, Origin of saline waters, ground water contamination from diffuse and point sources, Natural attenuation processes, Immiscible pollutants*

4) Towards coupling hydrological, soil and weathering processes - a modelling perspective *Colin Neal* *22/10 1400-1815*
Key words: *Hydrochemistry, Catchments, Acidification, Weathering, Hydrograph separator, Modelling, Environmental impact, Thermodynamics, Hydrology*

PART V. TECHNICAL ASPECTS OF CATCHMENT STUDIES

1) Chemical analysis of rocks and soils *Magne Ødegård* *19/10 0800-1000*
Key words: *Spectrochemical analysis, Geological materials, Total analysis, Partial analysis, Quality control*

2) Chemical analysis of water *John Mather* *19/10 1015-1215*
Key words: *Range of parameters, Precision and detection limits, Analysis of ground waters for major cations, anions, trace elements and organics, Determination of isotopic components, Field versus laboratory measurements, Representation and interpretation of hydrochemical data, Water quality standards*

3) Tracer studies *Allan Rodhe* *19/10 1400-1600*
Key words: *Types of tracers: Environmental, Artificial, Chemical, Stable and radiogenic isotopes; General principles: Different information with different tracers; The catchment and its constituents as water reservoirs: Age, Transit time, Turnover time, Definitions and examples; Stable isotopes: Oxygen-18, Deuterium, Fractionation processes, Geographical and temporal variation, Applications to catchment hydrology (ground water recharge, stream flow generation, mean transit time of ground water); Tritium: Geographical and temporal distribution, Applications to catchment hydrology (transit time distributions, age)*

4) Manual measurements and automatic instrumentation *Ånund Killingtveit & Knut Sand* *19/10 1615-1815*
Key words: *Observation of meteorological variables, Observation of hydrological variables, Observation of hydrochemical variables, Sensors, Automatic measurements, Automatic data transmission*

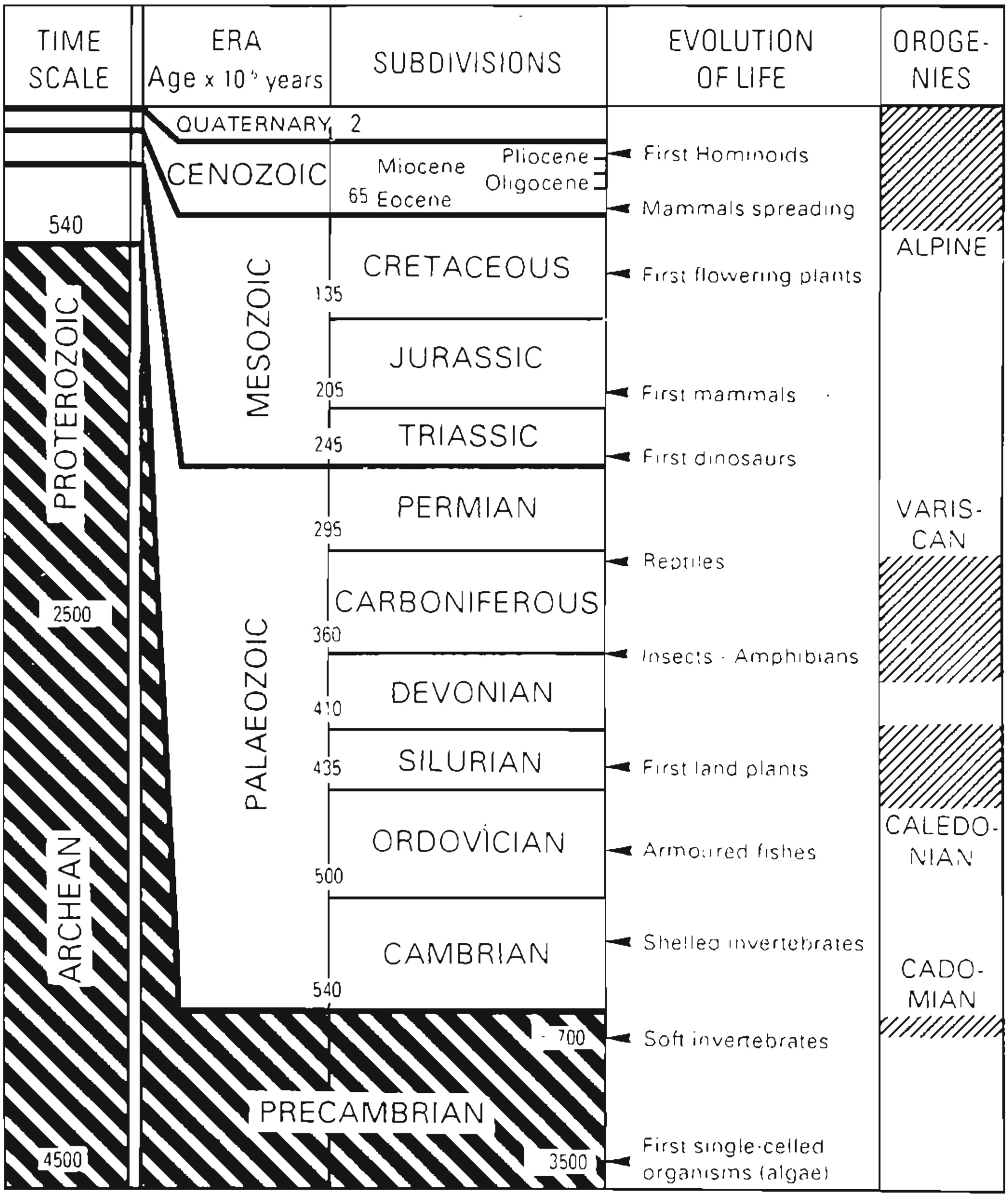
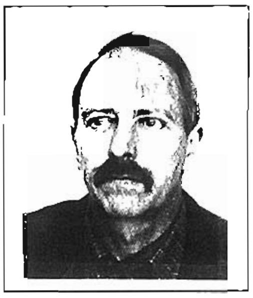


Fig. 2.1 Time-scale and main stratigraphical divisions (isotopic ages from IUGS (Commission of Stratigraphy) and G. S. Odin (1990) *Geochronique 35*).

LECTURER PROFILES
(sorted alphabetically)



Prof. Ole K. BORGGAARD

Chemistry Department
Royal Veterinary and Agricultural University Denmark

Research interests:
Soil chemistry, pedology, environmental soil chemistry.

Three most significant publications:
1) Borggaard, 1990. Dissolution and Adsorption Properties of Soil Iron Oxides. Thesis, Copenhagen.
2) Borggaard, 1992. Dissolution of poorly crystalline oxides in soils by EDTA and oxalate. Z. Phlanzenernährung Bodine, 155: 431-436.
3) Hansen, Borggaard & Sørensen, 1994. Evaluation of the free energy of formation of iron (II) iron (III)-hydroxide-sulphate (Green Rust) and its reduction of nitrite. Geochim. Cosmoschim. Acta, in press.

Dr. Patrice de CARITAT

Section for Geochemistry and Hydrogeology
Geological Survey of Norway, Norway

Research interests:
Sediment diagenesis, weathering, (clay) mineralogy, environmental geochemistry, hydrogeology, ground water acidification and pollution, quantitative geology.

Three most significant publications:
1) Abercrombie, Hutcheon, Bloch & Caritat, 1994. Silica activity and the smectite-illite reaction. Geology, 22: 539-542.
2) Caritat, Bloch & Hutcheon, 1994. LPNORM: a linear programming normative analysis code. Computers & Geosciences, 20: 313-347.
3) Caritat, 1989. Note on the maximum upward migration of pore water in response to sediment compaction. Sedimentary Geology, 65: 371-377.

Prof. James I. DREVER

Dept. of Geology and Geophysics
University of Wyoming, USA (presently on sabbatical in France)

Research interests:
Aqueous geochemistry, mineral weathering, including kinetics of silicate dissolution, effect of acid deposition on surface water chemistry, organic-inorganic interactions, mechanisms controlling contaminant migration.

Three most significant publications:
1) Drever, 1982, 1988. The Geochemistry of Natural Waters. Prentice-Hall.
2) Swoboda-Colberg & Drever, 1993. Mineral dissolution rates in plot-scale field and laboratory experiments. Chem. Geol., 105: 51-69.
3) Drever & Hurcomb, 1986. Neutralization of atmospheric acidity by chemical weathering... Geology, 14: 221-224.

Prof. Ånund KILLINGTVEIT

Dept. of Hydraulic and Environmental Engineering
Norwegian Institute of Technology, Norway

Research interests:
Hydrological models, water resources planning, arctic hydrology, water balance computations.

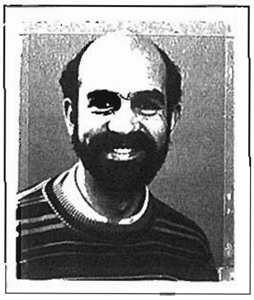
Three most significant publications:
1)
2)
3)

Prof. Eiliv LARSEN

Dept. of Geology
University of Tromsø, Norway

Research interests:
Paleoclimate, paleoenvironment, glacial history, sedimentology.

Three most significant publications:
1) Larsen & Mangerud, 1981. Erosion rate of a Younger Dryas ... Annals of Glacial., 2: 153-158.
2) Larsen et al., 1987. Cave stratigraphy ... Boreas, 16: 267-292.
3) Larsen & Sejrup, 1990. Weichselian land-sea interactions. Quaternary Science Reviews, 9: 85-97.

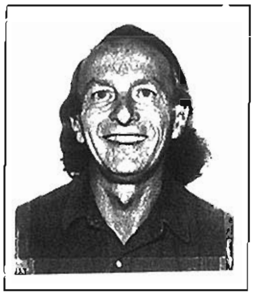


Prof. David N. LERNER

Dept. of Civil and Structural Engineering
University of Bradford, UK

Research interests:
Ground water resources, particularly recharge, ground water in urban areas, organic pollution of ground water.

Three most significant publications:
1) Lerner, Issar & Simmers, 1990. Groundwater recharge: a guide to estimating and understanding natural recharge. Heise
2) Lerner, 1992. Well catchments and time-of-travel zones in aquifers with recharge. Water Resources Research, 28: 261-268.
3) Lerner et al., 1993. Coventry groundwater investigation: sources and movement of chlorinated hydrocarbon solvents. Special Section of Journal of Hydrology, 149: 111-272.



Prof. Donald L. MACALADY

Dept of Geochemistry
Colorado School of Mines, USA

Research interests:
Chemistry and geochemistry of natural and anthropogenic organic chemicals in aqueous systems, particle size effects in aqueous systems, metal-organic interactions.

Three most significant publications:
1) Wolfe & Macalady, 1992. New Perspectives in aquatic redox chemistry: abiotic transformations of pollutants in groundwater and sediments. J. Contam. Hydrol., 9: 17-34.
2) Macalady, Tratnyek & Wolfe, 1989. Influences of natural organic matter on abiotic hydrolysis of organic contaminants in aqueous systems. In: «Aquatic Humic Substances Influence on Fate and Treatment of Pollutions», Suffet & Macarty, Eds. ACS Advances in Chemistry, 219: 323-332.
3) Tratnyek & Macalady, 1989. Abiotic reduction of nitroaromatic pesticides in anaerobic laboratory systems. J. Agric. Food Chem, 37: 248-254.

Prof. John D. MATHER

Dept. of Geology
Royal Holloway University of London, UK

Research interests:
Ground water pollution and remediation, waste disposal by landfill, contaminated land, environmental geochemistry.

Three most significant publications:
1) Mather, 1993. The impact of contaminated land on groundwater: a UK appraisal. Land Contamination and Reclamation, 1: 187-195.
2) Mather, 1989. Groundwater pollution and the disposal hazardous and radioactive wastes. J. Int. Wat. Environ. Management, 3: 31-38.
3) Mather, 1989. The attenuation of the organic component of landfill leachate in the unsaturated zone: a review. Q. J. Eng. Geol., 22: 241-246.

Dr. Colin NEAL

Hydrochemistry
Institute of Hydrology, UK

Research interests:
Hydrochemistry, environmental impacts at the catchment and regional scale, low temperature geochemistry.

Three most significant publications:
1) Neal et al., 1994. Chemical variations in near surface drainage waters for an acidic spruce forested UK upland catchment. Sci. Tot. Environ., 154: 47-61.
2) Neal. Describing anthropogenic impacts on stream water quality: the problem of integrating chemistry variability. Sci. Tot. Environ., 115: 207-218.
3) Neal, Robson & Sauth, 1990. Acid neutralization capacity variations for the Hofney forest stream, mid Wales, inferences for hydrological processes. J. Hydrol., 121: 85-101.

Dr. Allan M. RODHE

Institute of Earth Science
Uppsala University, Sweden

Research interests:
Stream flow generation, environmental isotope techniques in hydrology, shallow ground water-soil water flow.

Three most significant publications:
1) Rodhe, 1987. The origin of streamwater traced by oxygen-18. Uppsala Univ., Dept. of Phys. Geog., Div. Hydrol., Report Series, A41.
2) Grip & Rodhe, 1988. Vattnets väg från regn till bäck. Hallgren & Fallgren.
3) Halldin, Rodhe & Bjurman, 1990. Urban storm water transport and wash-off of caesium-137 after the Chernobyl accident. Water, Air and Soil Pollution, 65: 83-100.

Dr. Knut SAND

Norwegian Hydrotechnical Laboratory
SINTEF, Norway

Research interests:
Hydrological processes in permafrost areas, hydrological modelling, remote sensing applications in hydrology.

Three most significant publications:
1) Sand, 1990. Modeling Snowmelt Runoff Processes in Temperate and Arctic Environments. Ph.D. thesis. Norw. Inst. of Technology, Trondheim.
2) Sand, Hagen, Repp & Berntsen, 1990. Climate-related research in Svalbard. Int. Conf. on the Role of the Polar Regions in Global Change. June 11-15, 1990, Fairbanks, Alaska, USA.
3) Sand, Kjensli, Sundøen & Aamodt, 1992. Data acquisition systems for meteorology and hydrology. Guidebook for users. Trondheim.

Prof. Arvid SKARTVEIT

Geophysical Institute
University of Bergen, Norway

Research interests:
Micro- and local meteorology, precipitation chemistry, hydrology, solar and atmospheric radiation.

Three most significant publications:
1) Skartveit, Rydén & Kärenlampi, 1975. Climate and Hydrology of Some Fennoscandian Tundra Ecosystems. Ecological Studies, Springer Verlag, 16: 41-53.
2) Skartveit, 1981. Relationships between precipitation chemistry, hydrology and runoff acidity. Nordic Hydrology: 65-80.
3) Skartveit & Olseth, 1992. The probability density and autocorrelation of short-term global and beam irradiance. Solar Energy, 49: 477-487.

Dr. Gaute STORRØ

Section for Geochemistry and Hydrogeology
Geological Survey of Norway, Norway

Research interests:
Hydrogeology in general, especially ground water chemistry, water-rock interaction and the changes in water chemistry due to pollution.

Three most significant publications:
1) Storrø, 1990. Hydrochemistry, Hydrology and Geology of the Kaldvella Catchment. Ph.D. thesis. Norw. Inst. of Technology, Trondheim.
2)
3)

Prof. Brian A. STURT

Research Director
Geological Survey of Norway, Norway

Research interests:
Tectonics, tectonostratigraphy, significance of unconformities, timing of orogenic events, ancient ophiolites and island arcs, sequence and how all of these features contribute to the understanding of developmental histories of orogenic belts

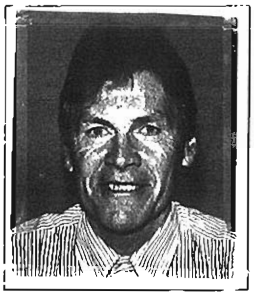
Three most significant publications:
1) Gee & Sturt, 1985. The Caledonide Orogen - Scandinavia and Related Areas. New York, Wiley, 1266 pp.
2) Sturt, Pringle & Ramsay, 1978. The Finmarkian phase of the Caledonian Orogen. J. Geol. Soc. London, 135: 597-610.
3) Melezhik & Sturt, 1994. General geology and history of the Polmak-Pasvik-Pechenga-Imandra/Varsuga-Ust'Ponoy Greenstone Belt in the northeastern Baltic Shield. Earth Science Reviews, 36: 205-241.

Dr. Ola M. SÆTHER

Section for Geochemistry and Hydrogeology
Geological Survey of Norway, Norway

Research interests:
Contaminant hydrogeology/aqueous geochemistry, mapping international background levels, interaction humic substances/rocks, fluoride in the environment.

Three most significant publications:
1) Sæther, Runnells, Ristinen & Smythe, 1981. Fluorine: its mineralogical residence in the oil shale of the Mahogany Zone of the Green River Formation, Piceance Creek Basin, Colorado, USA. Chem. Geol., 31: 169-184.
2) Sæther, Bølviken, Låg & Steinnes, 1988. Concentration and chemical form of lead during natural transportation in groundwater. Chem. Geol., 69: 309-319.
3) Sæther, Misund, Ødegård, Andreassen & Voss, 1992. Groundwater contamination at Trandum Landfill, southeastern Norway. NGU Bull., 422: 83-95.



Mr. Magne ØDEGÅRD

Section for Laboratories
Geological Survey of Norway, Norway

Research interests:
Analytical inorganic chemistry based on spectrochemical techniques for solving geochemical and environmental problems.

Three most significant publications:

1) Ødegård, 1981. The use of inductively coupled argon plasma (ICAP) atomic emission spectroscopy in the analysis of stream sediments. J. Geochem. Explor., 14: 119-130.
2) Ødegård, 1979. Determination of major elements in geological materials by ICAP spectroscopy. Jarrell- Ash Plasma Newsletter, 2(1).
3) Ødegård, 1984. A selenium-rich sulphide assemblage in the Caledonides of northern Norway. Norsk Geologisk Tidsskrift, 64: 287-292.

STUDENT PROFILES
(sorted by land, then alphabetically)

SWEDEN

T. Marie-Louise BENGTSSON

Department of Geology,
Chalmers University of Technology, Sweden

Research topic:
Thesis: «Vulnerability classification as a tool in municipal planning». Risk assessment of municipal activities with respect to surface water and ground water, geological and hydrogeological model.

Johan R. GRANHAGEN

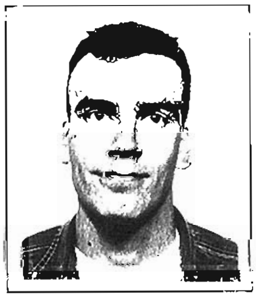
Dept. of Geology and Geochemistry
Stockholm University, Sweden

Research topic:
Acid mine drainage from a reclamation site. leaching experiments. pyrite-carboxylic acid equilibrium.

Anna-Maria HULTÉN

Department of Geology
Chalmers University of Technology, Sweden

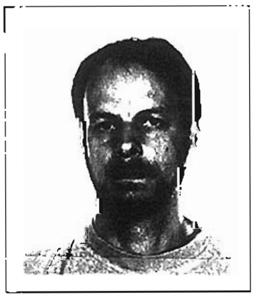
Research topic:
Thesis: «The hydrogeology of filling materials in an urban area». How water composition, filling material and ground water level affects constructions in the ground.



Magnus LAND

Dept. Environmental Planning
Luleå University of Technology, Sweden

Research topic:
Thesis: «Weathering and water quality in northern Sweden». Weathering profiles in podzolised tills, water sampling, heavy metals. soil sensitivity to acid deposition.



Ulf A. NYSTRÖM

Department of Geology
University of Gothenburg, Sweden

Research topic:
The effects of soil treatment. esp. soil liming, on ground water and runoff-water. Applied acidification research (NITREX and roofed catchment experiments).

Malin H. E. OLIN

Dept. of Geology
Chalmers University of Technology, Sweden

Research topic:
Thesis: «Modelling of groundwater recharge and aquifer characteristics from level recordings on aquifers». Ground water level variations to evaluate recharge parameters and geological conditions.

Anna SANDER

Div. of Land and Water Resources
Royal Institute of Technology, Sweden

Research topic:
Effects of gravel extraction on ground water and soil processes.

LATVIA

Agrita BRIEDE

Lab. of Hydrobiology
Institute of Biology, Latvia

Research topic:
Assessment of load and impact of phosphorus and related elements on rivers and lakes.

Inguna FOLBERGA

Hydrogeology
Baltec Assoc., Latvia

Research topic:
Assessment of hydrocarbon contamination in soil and ground water and development of remedial measures, morphodynamic processes in Baltic Sea coastal zone, modelling.

Mats E. ÅSTRÖM

Dept. of Geology
Åbo Akademi University, Finland

Research topic:
Thesis: «Impact of acid sulfate soils and regional till geochemistry on stream-water quality in W. Finland». Weathering and stream-water geochemistry, leaching of H^+, sulfate and metals.

ICELAND

Arni HJARTARSON

National Energy Authority, Iceland

Research topic:
Thesis: «Hydrogeological maps and mapping methods in the field». Geochemical methods in ground water mapping, water works and wells, hydrogeological map of the Reykjavik area 1:25 000.

NORWAY

Leif BASBERG

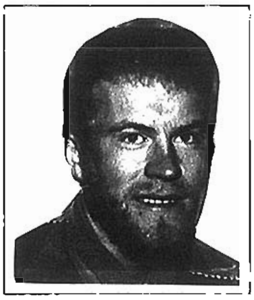
Inst. for Geology and Mineral Resources Engineering
Norwegian Institute of Technology, Norway

Research topic:
Modelling of pollutants in groundwater.

Oddveig BREDESEN

Section for Geochemistry and Hydrogeology
Geological Survey of Norway, Norway

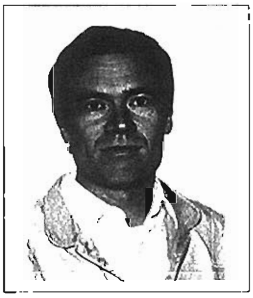
Research topic:
National ground water program, use of ground water in Norway, field investigations, ground water supply, drinking water supply.



Jo H. HALLERAKER

Section for Geochemistry and Hydrogeology
Geological Survey of Norway, Norway

Research topic:
Thesis: «River mosses as bio-indicator for some heavy metals in the border water systems of Norway and Russia». Ecogeochemical mapping of the Barents region



Jens KVÆRNER

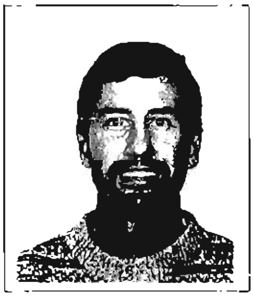
Center for Soil and Environmental Research
Jordforsk, Norway

Research topic:
Environmental geology, strategies for liming against acidification in catchments, unsaturated zone in agricultural land, preferential flow.

Thomas POULSEN

Dept. of Chemistry
University of Trondheim, Norway

Research topic:
Thesis: «A hydrogeochemical study of heavy metal pollution in selected catchment areas in eastern Finnmark and on the Kola Peninsula». Hydrological and hydrochemical dispersion of heavy metal pollutants (e.g., Cu, Ni, Co) in the Kola region.



Kim RUDOLPH-LUND

Inst. of Geology
University of Oslo, Norway

Research topic:
Biological degradation of petroleum products in the unsaturated zone, Trandum.

DENMARK

Marina B. JENSEN

Soil Faculty and Plant Nutrition
The Royal Veterinary and Agricultural University, Denmark

Research topic:
Thesis: «Mobilisation and transport of phosphorus in structured soil». Dynamics of phosphorus as influenced by soil redox potential, soil microbial activity and wetting/drying events.

Lars O. LARSEN

Chemistry Department
The Royal Veterinary and Agricultural University, Denmark

Research topic:
Ph.D. project on soil/environmental science. Previous work on granitic soil formation in Thailand.

Gitte H. RUBÆK

Department of Soil Science
Danish Institute of Plant and Soil Science, Denmark

Research topic:
Ph.D. project on lability and vertical distribution of phosphorus in soils with different types of land-use and cultivation.

Lars VESTERDAL

Unit of Forestry
The Royal Veterinary and Agricultural University, Denmark

Research topic:
Nutrient mineralisation in forest ecosystems, mineralisation patterns.

FINLAND

Kirsti I. KORKKA-NIEMI

Dept. of Geology
University of Turku, Finland

Research topic:
Ph.D. project on national well water quality data, and geological, environmental and technical influences. Acidification in Finnish wells.

Jussi LEVEINEN

Dept. of Geology
University of Helsinki, Finland

Research topic:
Ph.D. project on fracture zones and ground water.

T. Harriet LONKA

Dept. of Geology and Mineralogy
University of Helsinki, Finland

Research topic:
Fracture zones and hard-rock groundwater. Ph.D. project on «Developing methods of groundwater protection in fracture-zone aquifers».

Maija E. PAASONEN-KIVEKÄS

Laboratory of Water Resources Engineering
Helsinki University of Technology, Finland

Research topic:
Ph.D. project on «Water table management for reduction of nutrient load from agricultural fields». Nitrogen cycle in the soil, effects of soils moisture on mineral nitrogen forms.

Seija S. ROMAN

Dept. of Geology
University of Oulu, Finland

Research topic:
Thesis «The quality and suitability for water supply of the glacio-fluvial aquifers in the Pudasjärvi region, Northern Finland, with special references to the Viinivaara interlobate complexes». Hydrogeology, ground water recharge in catchments.

Geological context of the Nordic countries. Bedrock geology

Brian A. Sturt
Geological Survey of Norway

INTRODUCTION

The bedrock geology of the Nordic Countries is dominated by the Precambrian basement complex of the Fennoscandian Shield. In Precambrian times those rocks all formed part of the great Pangean supercontinent which began its break-up during late Proterozoic times (Fig. 1). Thus the Archaen and Palaeoproterozoic (pre -2500 Ma) complexes can be related to similar complexes in now widely separated regions e.g. Greenland, Siberia etc. The western rim of the Fennoscandia Shield is the regime of the great thrust sheets of the Caledonian orogen which form a mountainous spine, though the present-day morphology of this mountain tract is essentially the results of Tertiary and Quaternary uplift and erosional sculpting. The southern part of the shield area was the site of major rifting during Carboniferous-Permian times resulting in the formation of the Oslo Graben with development of the well-known Oslo magma-series of plutonic and volcanic rocks of essentially alkaline to sub-alkaline character. There is a very clear distinction between the on-shore and off-shore geology, where the offshore areas contain large sedimentary basins of Late Palaeozoic, Mesozoic and Tertiary strata. These strata form the major basis for the modern Hydrocarbons Industry of particularly Norway and Denmark.

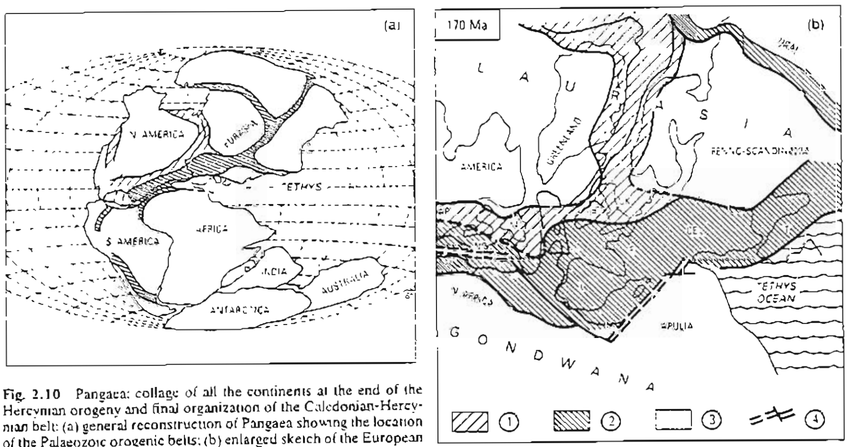


Fig. 2.10 Pangaea: collage of all the continents at the end of the Hercynian orogeny and final organization of the Caledonian-Hercynian belt: (a) general reconstruction of Pangaea showing the location of the Palaeozoic orogenic belts; (b) enlarged sketch of the European part of Pangaea showing the Caledonian (1) and Hercynian (2) belts between Precambrian shields (3) and the position of the new spreading ridges (4) that initiated the beginning of the Alpine cycle.

Ap: Appalachians. N-S. Nova Scotia. N: Newfoundland; M: Morocco; I: Iberia. F: France; IR: Ireland. UK: United Kingdom. CE: Central Europe; T: Turkey (from Olivet *et al.* (In press). *Synthèse géologique des Pyrénées*. BRGM-ITGE).

Figure 1. From Lumsden (1992)

The Precambrian complexes of the Fennoscandian Shield lie to the east and south of the great NorthSea - North Germany-Russia sedimentary basins. The Precambrian rocks continue under the younger sedimentary cover of the Russian Platform, to reappear in the Ukraine. A succession of orogenic cycles between 3500 and 900 Ma ago constructed this resistant shield which remained a single block from the Atlantic to the Urals (Fig. 2). The shield was built up, on the margins of the Archaen nucleus, by a series of orogenic cycles which essentially began with a rifting phase leading to the development of the «Greenstone Belts», which latter have great importance for their deposits of Ni-Cu-Cr-Pt-Au-Zn-Co.

The Archaen Complexes consists essentially of high-grade gneisses (amphibolite-granulite facies) and themselves record a number of orogenic cycles interspersed with a number of stages of sedimentary basin formation. A number of major plutons were emplaced during the Archaen and are mainly the result of the rise of mantle-derived magmas. A change of in the dynamics of earth evolution during the Palaeoproterozoic is well illustrated in the Fennoscandian Shield and also the progressive formation of continental crust by the extraction of magmas from the mantle to form igneous rocks, as in Archaen times, and the recycling of older continental crust by deformation, metamorphism, and partial melting to form new magmas during succeeding orogenic cycles.

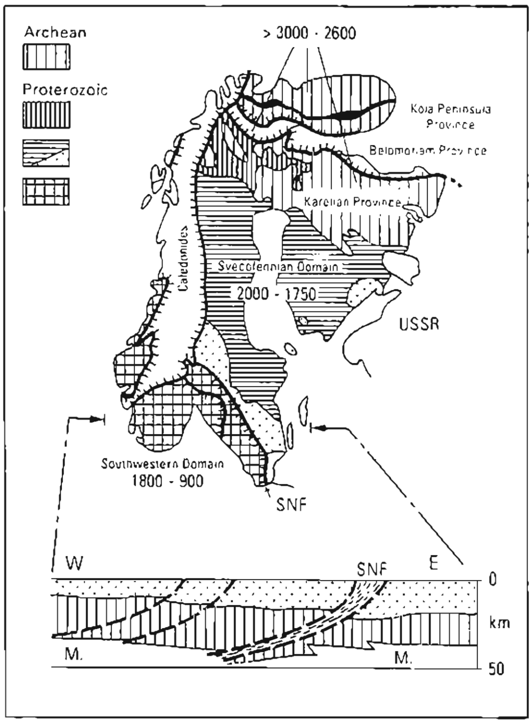


Fig. 2.6 The polyorogenic organization of the Fennoscandian Shield with its Archaean northern part and the surrounding Early and Middle Proterozoic orogenies. The seismic section of the crust across the southern domain shows the crustal strips thrust eastwards during the last Sveconorwegian orogeny. M: Mantle; vertical lines: lower dense crust; dots: upper crust; SNF: Sveconorwegian front (from EUGENO-S project: *Tectonophysics* 1989 vol. 162. n°1).

Figure 2. From Lumsden (1992)

The Late Archaen and Palaeoproterozoic periods saw the development of the «Greenstone Belts» which in their initial stages involved volcanism including effusion of both ultramafic (Komatiitic/ferro-picritic) and mafic (basaltic) types. The Palaeoproterozoic saw the first real development of oceanic crust, with effusion of MORB-type basalts, representing spreading, and of high-Al_2O_3 andesites, representing island arc volcanism related to subduction systems (Melezhik and Sturt, 1994). The largest part of the Fennoscandian Shield (almost 2000 km across) was affected by the Svecofennian orogenic cycle (2000 - 1750 Ma) and the identification of relics of oceanic crust, island arcs and sedimentary basins has made it possible to reconstruct a geodynamic setting where the various stages in a plate-tectonic scenario can be recognised. These commence with the opening of oceanic basins followed by development of subduction cycles, resulting in arc volcanism and eventual collision with continent-directed thrusting resulting in great far-travelling slabs over-riding the continental margins. The lower metamorphic grade of the greenstone belts is often in marked contrast to that of the Archaen nucleus, especially where they are separated by a primary stratigraphic unconformity. In deeper parts of the Svecofennian orogen, however, the metamorphic grade may rise into amphibolite or even granulite facies. The greenstone belts and their borderlands have been the site of mining activity for many decades e.g. Ni-Cu ores of the Pechenga region, Cu-Zn-Co deposits of Outokumpu, Skellefte, Orijävri and Bergslagen districts.

The Svecofennian Orogeny was followed in the coeval Transscandinavian Belt by a period (1750-1550 Ma) of intense intrusive activity where great plutons invaded the now cratonized continential crust. These plutons form a sequence with an east-west chemical zonation according to Gorbatschev (1985). Orogenic calc-alkaline tonalities and granodiorites in the west are followed, eastwards, by more alkaline granites, syenites and monzonites (Fig. 3). Farther east, isolated rapakivi granite and bimodal rapakivi granite-gabbro complexes form a coeval continuation of the magmatism. Generally Sr-isotopic ratios of the granites increase eastwards. Gorbatschev (1985) took these variations in whole-rock, trace-element and isotopic signatures to indicate possible subduction occurring beneath a westward extension of the present west coast of Norway. There are, however, no simple answers to this problem as there has been shown, in Western Norway to be a progressive evolution in time from orogenic tonalities to rapakivi-like intrusions. Furthermore augen gneisses derived from rapakivi granite and rapakivi granite-gabbro complexes (1700 - 1600 Ma old) have been detected over wide areas in the basement plinths of eastward directed Caledonian nappes in the Central Caledonides.

IAPETUS OCEAN/CALEDONIDES

Major intercontinental crustal rupture started in Middle to Late Riphean times (1000 - 900 Ma ago). During the Late Riphean and Vendian, aulacogens, rift basins, and continental depressions formed within Baltica from the Ukraine in the southeast to the Scandinavian part of the craton in the northest (Fig. 4). The general senario for sedimentation and associated minor volcanism is shown in Fig. 5, and indicates rifting on the continental craton with the formation of grabens and half-grabens initially receiving a supply of siliciclastic deposits of predominantly alluvial nature. In the westward basin (Tossåsfjellet Basin) these pass upwards to marine sequences containing considerable carbonate deposits. This westward basin represented the emerging continental margin of Baltica during the initial development of the Iapetus Ocean System, and its deposits are now seen in lower and middle Allochthons of the

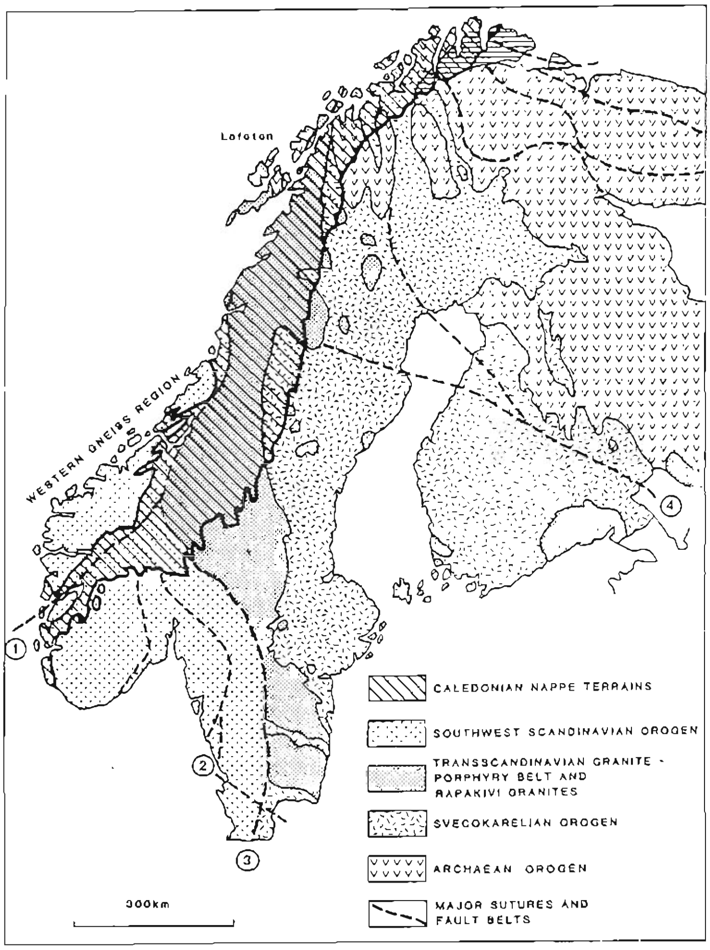


Fig. 1. Orogens of the Baltic Shield. The horizontally ruled area in northernmost Scandinavia is the Timanide Aulacogen resting on Archaean crust covered by Svecokarelian sediments. The sutures and fault belts indicated by digits are: 1 'Faltungsgraben' delimiting the Western Gneiss Region in Norway, 2 'Mylonite Zone' suture, 3. 'Protogine Zone' suture, 4 Ladoga–Lofoten belt of faulting

Figure 3. From Gorbatschev (1985)

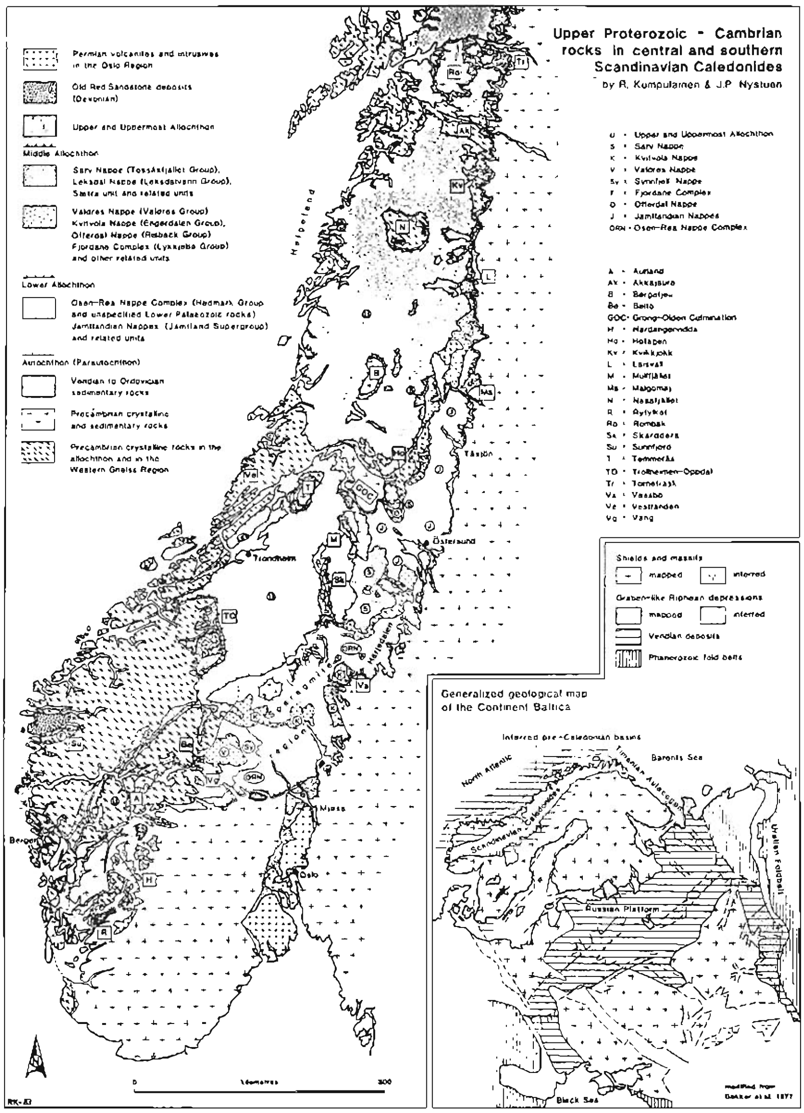
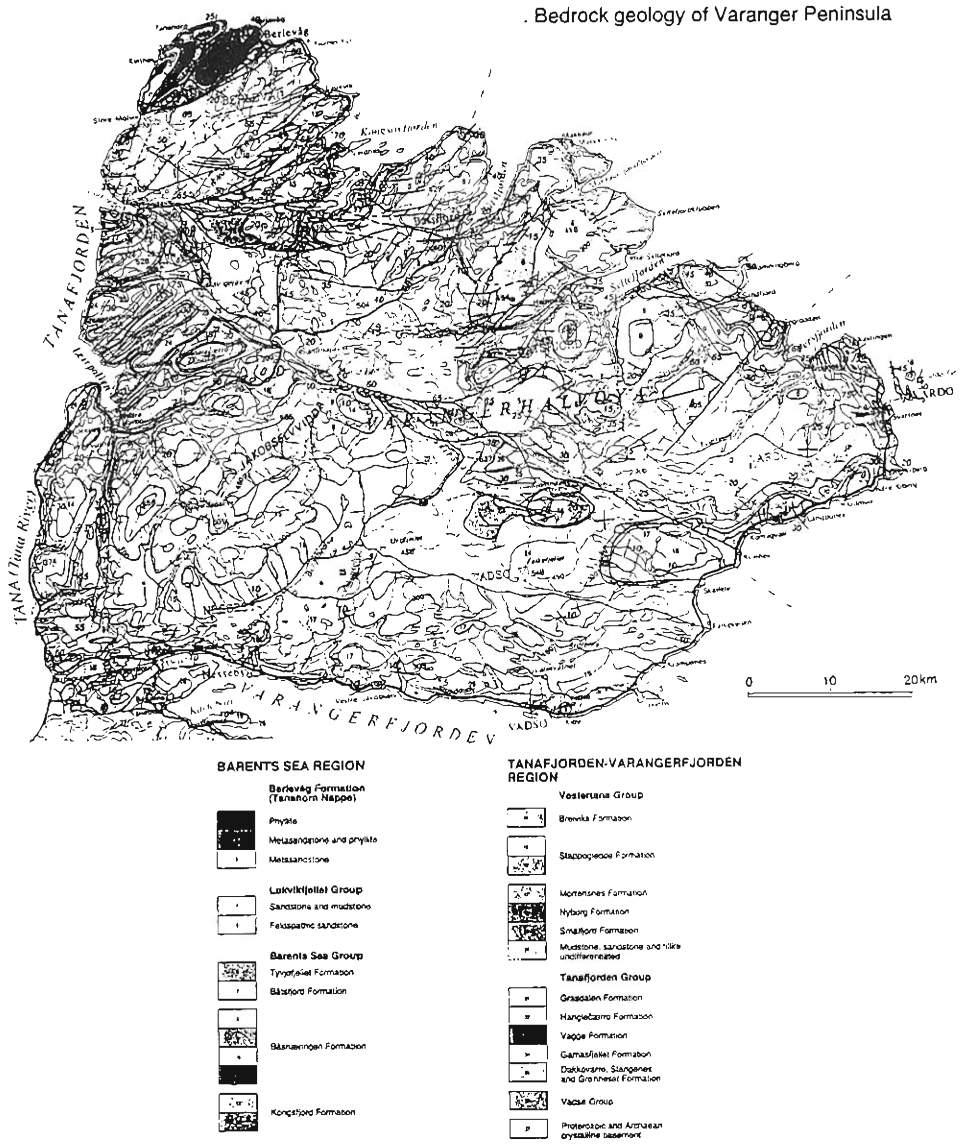


Figure 4. From Kumpulainen and Nystrom (1985)

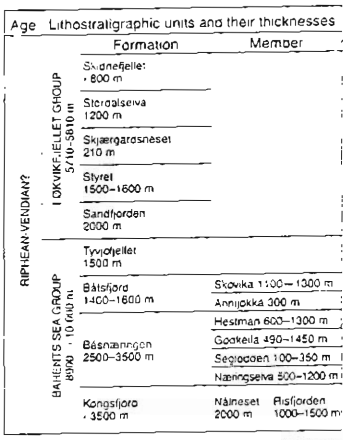
Scandinavian Caledonides (Fig. 4), where they pass upwards via Vendian tillites into the Lower Palaeozoic rocks on the Caledonian Miogeocline. In the far north of Norway (Fig. 6) a similar situation pertains to the south of the Trollfjord-Komagelv Fault Zone (TKFZ) which seperates a southern assemblage (attached to the craton) of mainly fluviatile to shallow marine sedimentary rocks (ca. 4 km thick), including two tillite horizons (Vadsø, Tanafjorden, Vestertana Groups). This assemblage rests directly on crystalline rocks of the craton. To the north of the TKFZ) the northern assemblage of the Barents Sea Region occurs (Fig. 7). These latter sediments, also of Riphean-Vendian age, consist of turbidites submarine fan deposits passing upwards into deltaic sequences and then into shallow marine sediments (>9 km thick).



. Bedrock geology of Varanger Peninsula

Figure 6. From Siedlecka and Roberts (1992)

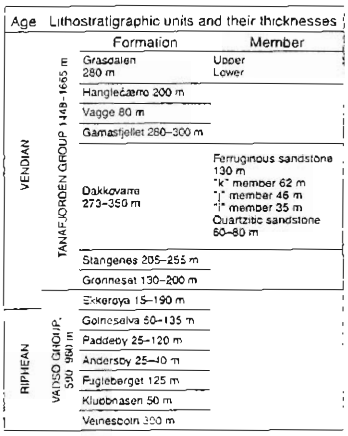
North



Age	Lithostratigraphic units and their thicknesses		
		Formation	Member
RIPHEAN-VENDIAN?	LOKVIKFJELLET GROUP 5710–5810 m	Skidnefjellet >800 m	
		Stordalselva 1200 m	
		Skjærgårdsneset 210 m	
		Styret 1500–1600 m	
		Sandfjorden 2000 m	
	BARENTS SEA GROUP 8900–10 600 m	Tyvjofjellet 1500 m	
		Båtsfjord 1400–1600 m	Skovika 1100–1300 m
			Annijokka 300 m
		Båsnæringen 2500–3500 m	Hestman 600–1300 m
			Godkeila 490–1450 m
			Seglodden 100–350 m
			Næringselva 500–1200 m
		Kongsfjord >3500 m	Nålneset 2000 m Risfjorden 1000–1500 m

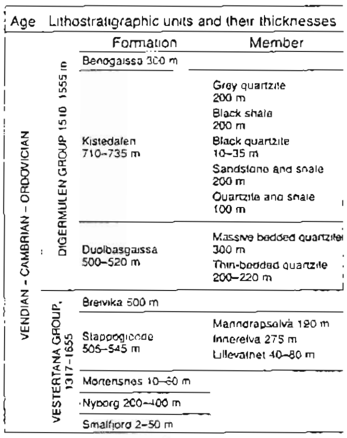
Fig. 4. Lithostratigraphy of the Lokvikfjellet Group and Barents Sea Group in the Barents Sea Region (after Siedlecka 1978, 1989, Siedlecka & Edwards 1980, Siedlecka & Siedlecki 1967, Siedlecki & Levell 1978).

-----------------Trollfjord-Komagelv Fault Zone (TKFZ)-----------------



Age	Lithostratigraphic units and their thicknesses		
		Formation	Member
VENDIAN	TANAFJORDEN GROUP 1448–1665 m	Grasdalen 280 m	Upper Lower
		Hanglečærro 200 m	
		Vagge 80 m	
		Gamasfjellet 280–300 m	
		Dakkovarre 273–350 m	Ferruginous sandstone 130 m "k" member 62 m "j" member 46 m "i" member 35 m Quartzitic sandstone 60–80 m
		Stangenes 205–255 m	
		Grønneset 130–200 m	
	VADSØ GROUP, 590–960 m	Ekkerøya 15–190 m	
RIPHEAN		Golneselva 50–135 m	
		Paddeby 25–120 m	
		Andersby 25–40 m	
		Fugleberget 125 m	
		Klubbnasen 50 m	
		Veinesbotn 300 m	

Fig. 2. Lithostratigraphy of the Tanafjorden Group and the Vadsø Group in the Tanafjorden - Varangerfjorden Region (mainly after Siedlecka & Siedlecki 1971 and Banks et al. 1974).



Age	Lithostratigraphic units and their thicknesses		
		Formation	Member
VENDIAN - CAMBRIAN - ORDOVICIAN	DIGERMULEN GROUP 1510–1555 m	Berlogaissa 300 m	
		Kistedalen 710–735 m	Grey quartzite 200 m Black shale 200 m Black quartzite 10–35 m Sandstone and shale 200 m Quartzite and shale 100 m
		Duolbasgaissa 500–520 m	Massive bedded quartzite 300 m Thin-bedded quartzite 200–220 m
	VESTERTANA GROUP, 1317–1655	Breivika 500 m	
		Stappogiedde 505–545 m	Manndrapselva 120 m Innerelva 275 m Lillevatnet 40–80 m
		Mortensnes 10–60 m	
		Nyborg 200–400 m	
		Smalfjord 2–50 m	

Fig 3. Lithostratigraphy of the Digermulen Group and the Vestertana Group (after Reading 1965).

South

Figure 7. From From Siedlecka and Roberts (1992)

This Barents Sea Group is succeeded by a nearly 6 km-thick package of mainly shallow marine sedimentary rocks (Løkvikfjellet Group). The seperating TKFZ is a long-lived fault zone that was initiated at some stage during the protracted Caledonian history with a dextral strike-slip motion. Younger components of movement, extensional and/or strike-slip, extend up into the Cenozoic and also post-glacial times (Siedlecka & Roberts 1992).

The Caledonian era was heralded by this scenario of an emerging Vendian-Cambrian rifted continental margin with a major crustal extension resulting in formation of oceanic crust, flooring the Iapetus Ocean. This latter separated Baltica from another continental landmass, conventionally regarded as Laurentia though, as the results of recent palaeomagnetic work, is considered to have been the Siberian Craton. During Cambrian times there was a rapid expansion of Iapetus and a clear distinction between miogeoclinal and eugeoclinal regimes was established. The first signs of orogeny was evidenced by Tremandocian Island Arc development with essentially coeval subduction of continental crust in which eclogite facies metamorphism (P>18kb) occurred in part of the miogeoclinal assemblage, now part of the Seve Nappe (Fig. 8). The rapid gravitational rebound of the continental slab was probably responsible for cratonwards propogating thrusts, and obduction of ophiolitic and island arc assemblages onto a westward extension of the continental miogeocline (Finnmarkian event). These thrust slabs which had been accreted onto the miogeocline were rapidly eroded and drowned by basinal assemblage of sedimentary and volcanic rocks during early Ordovician times (Fig. 9). The Ordovician history of the active Caledonian Baltic Margin was a scenario, not unlike that of the SE Pacific today, with the development of a plethora of island arcs, back-arc basins and periodic subduction (Sturt & Roberts, 1991). The polarity of Ordovician subduction is not known, but there was likely to have been various subduction zones with different orientations and polarities relative to the continental margins (Fig. 8). It was also during the Ordovician that the majority of the stratabound sulphide deposits so important historically to the Norwegian and Swedish mining industries were formed. The great granite batholiths of western Norway were emplaced mainly during the Upper Ordovician.

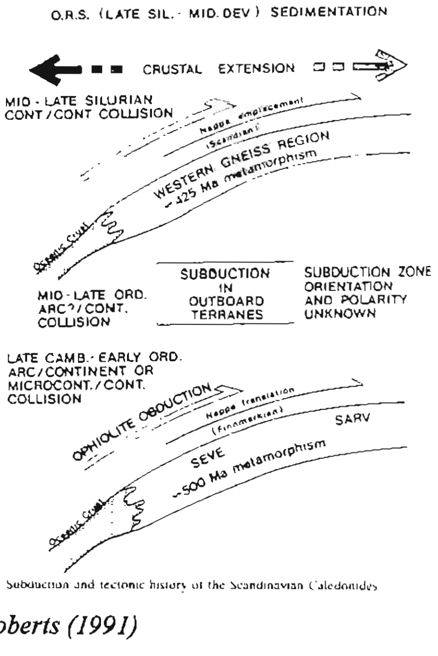


Figure [illegible]: Subduction and tectonic history of the Scandinavian Caledonides

Figure 8. From Sturt and Roberts (1991)

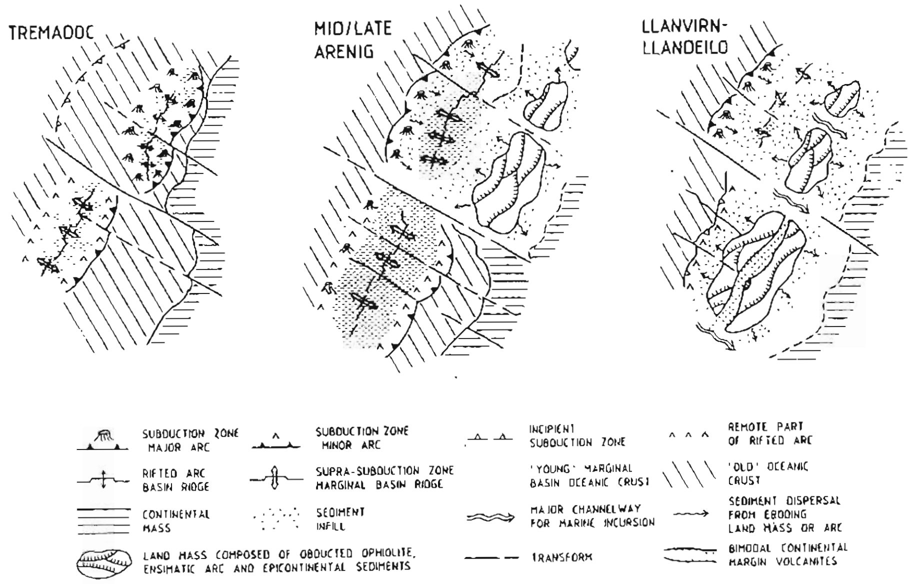


Figure 3. Diagrammatic plate tectonic/palaeogeographic setting of the Southwest and Central Scandinavian Caledonide segments during the period Tremadoc to Llandeilo. The upper half of each diagram (NE of the main NW-SE transform) is meant to portray the Central segment; and the lower half the Southwestern segment. For full explanation, compare the legend and the main text. Tremadoc - the pre-obduction situation. Mid/Late Arenig -obduction has already accurred in the Central segment, represented by the eroding landmass composed of thrust shees. Llanvirn-Llandeilo – in the Central Segment, a mature arc is present in the northwest and a marginal basin occurs between this arc and the eroding, dimishing landmass. In the Southwest segment, obduction has occurred and the uplifted, eroding landmass is affected by rifting with the production of bimodal volcanic rocks.

Figure 9. From Sturt and Roberts (1991)

According to modern palaeomagnetic reconstructions (Torsvik and Trench 1992) Baltica separated from the Siberian block, during Middle Ordovician to Early Silurian times. It rotated anticlockwise into essentially its present position opposite to Laurentia (Fig. 10). A result, of this rotation and southwards movement of Baltica, was a collision with Avalonia (European Craton) in the Silurian initiating the major movement along the Tornqvist-Tessier Fault Zone (TTFZ). This latter is also a long-lived zone which records periodic movements through Cenozoic and even into post-glacial times. Major subduction of continental crust was initiated, along the western margin of Baltica during the Middle Silurian resulting in continent-continent collision (Scandian event) between Baltica and either Laurentia or an interveening continental block. It was during this stage that the great eastward propogating thrust nappes were emplaced to produce the tectenostratigraphic collage that we see today (Fig. 11). Caledonian metamorphism was polycyclic, though the major regional distributions of metamorphism we observe are probably mainly of Scandian imprint. The Scandian continental subduction is well documented in the Western Gneiss Region (WGR) of Norway (Fig. 8) where eclogites (max p>21 kb) were developed. The local occurrences of coesite and microdiamonds confirm that this continental slab had undergone ultra-high pressure metamorphism.

During Caledonian times the Baltic Craton was covered intermittently by an epi-continental sea with local depocentres e.g. Oslo region, site of the present Baltic Sea etc. The overthrusting and folding of the Lower Palaeozoic assemblages of the Oslo area occurred in Early Devonian times.

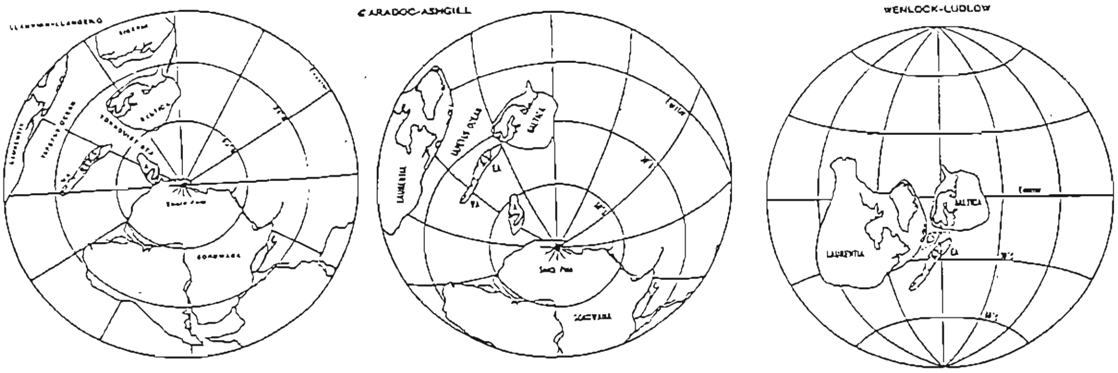


Figure 10. Relative positions of Baltica during the Ordovician and Silurian. From Trench and Torsvik (1992)

The Devonian represented a period, in Western and Central Norway, of deposition of great thicknesses of Old Red Sandstone facies of fluviatile conglomerates and sandstones (Fig. 12) in tectonically confined local basins. The basins are considered to be controlled by extensional faults in an extensional detachment mode. In such a scenario the metamorphic rocks of the WGR belong to the lower plate rising and decompressing rapidly with extensional collapse in Late Silurian - Middle Devonian times (Fig. 13). That processes were diachronous, during Caledonian times, is evidenced in that Baltica-directed thrusting continued into at least the Early Devonian in the northernmost parts of the Scandinavian Caledonies (Roberts & Sundvoll, 1991).

POST-CALEDONIAN TIMES

Virtually no evidence is found on the Fennoscandian Shield of sedimentary or volcanic rocks of Late Palaeozoic age. Analogies with the sequences in Eastern Greenland, the Russian Platform and in the substrate of the North, Norwegian and Baltic seas, however, suggest that much of the craton may have been blanketed by such malterials, though now removed by erosion (Fig. 14). A major change in geodynamical regime occurred in Late Carboniferous to Permian times when continental rift tectonics began to dominate. The earliest stage of this rifting was seen in the Oslo (Fig. 15) and Horn Grabens where the rift-related volcanics date back to the Late Carboniferous (300 Ma) and the youngest volcanic or plutonic rocks are ca. 240 Ma old. The source region for the igneous rocks appears to be a mildly depleted source in the lower sub-continental lithosphere (Neuman 1988). The rift propogation continued in the North Sea with the development of the Viking Graben (Fig. 16) which appears to have been initiated in Late Permian to Early Triassic times (Glennie 1986). The Viking Graben had two main phases of subsidence, the first through to early Bathonian times (Mid Jurassic) and was followed by a second rifting phase which continued through the Upper Jurassic until it abruptly ceased in the Ryazanian (basal Cretaceous). In the Ryazanian there is an important Lower Cretaceous unconformity. Subsidence continued through the Cretaceous into the Tertiary and there was a progressive migration of active faulting towards the basin margins.

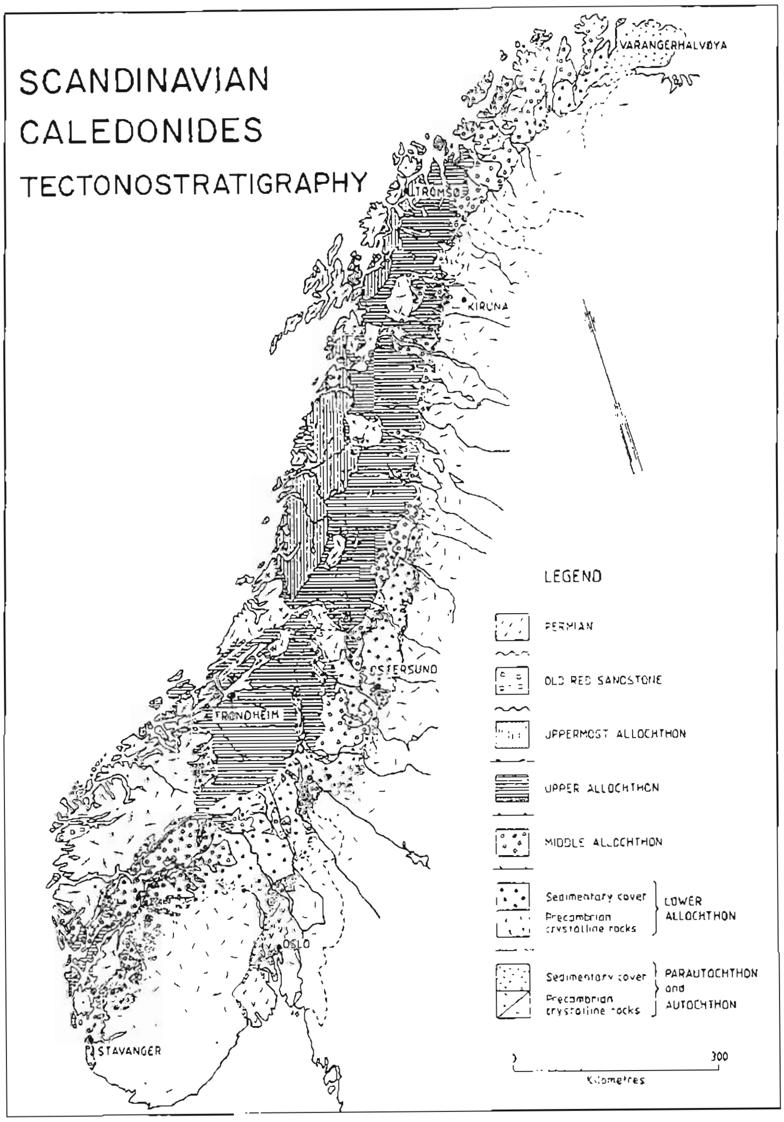


Figure 11. From Roberts and Gee(1985)

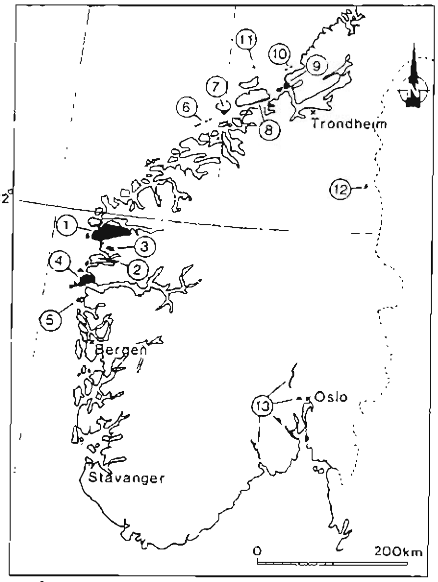


Fig. 1 Map showing location of the 13 districts/areas with Old Red Sandstone strata: 1—Hornelen, 2—Kvamshesten, 3—Haasteinen, 4—Solund/Bulandet, 5—Byrknesøyene/Holmengrå, 6—Føllingen/Inngripen, 7—Smøla, 8—Hitra, 9—Fosen, 10—Asenøya, 11—Froøyane, 12—Røragen, 13—Oslo

Figure 12. From Steel et al. (1985)

Extensional collapse of the southwest Norwegian Caledonides

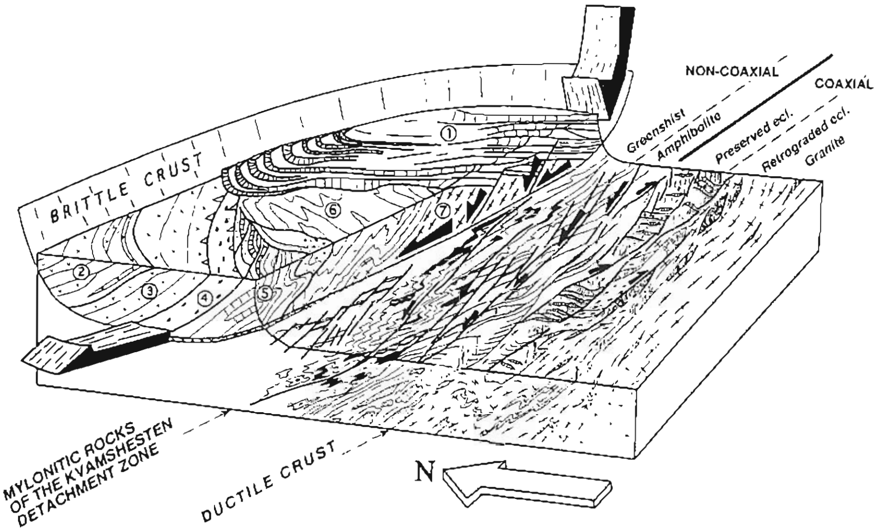


Fig. 6. Generalized block diagram (not to scale) for the geology and the structural relationships between the Upper and Lower Plate in the Sunnfjord region of the southwest Norwegian Caledonides. Note W-vergent folds and extensional reactivation of contractional faults in the Upper Plate and that sedimentation in the Devonian Kvamshesten Basin was partly controlled by normal faults in the Upper Plate that are rooted in the Kvamshesten Detachment Zone. The detachment is characterized by a several kilometres thick zone of mylonites formed by non-coaxial, top-to-the-west shear at amphibolite to greenschist facies metamorphism. The extensional mylonites are rooted in the lower crust that underwent coaxial vertical flattening and E–W extension. Legend: (1) Devonian sediments, (2) Staveneset Group, cover to the Solund–Stavfjord Ophiolite Complex; (3) Solund–Stavfjord Ophiolite Complex, (4) Sunnfjord Melange; (5) the Herland Group, Silurian continental margin deposits with W-vergent folds related to the extensional collapse; (6) the Hoyvik Group, Late-Precambrian (?) continental margin deposits; and (7) the Dalsfjord Suite, Precambrian (?) allochthonous basement. See the text for a detailed discussion of the fabrics in the Lower Plate.

Figure 13. From Anderson et al. (1994)

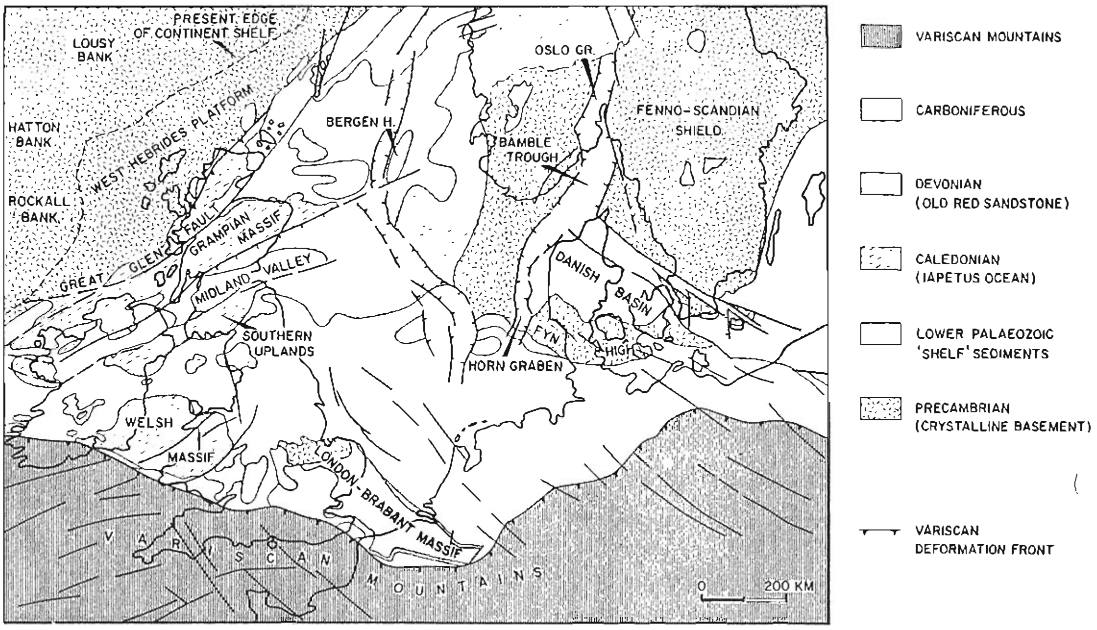


Fig. 2.4. Pre-Permian geological map. (Modified from Ziegler, 1982a, by addition of North Sea Graben system.)

Figure 14. From Glennie (1986)

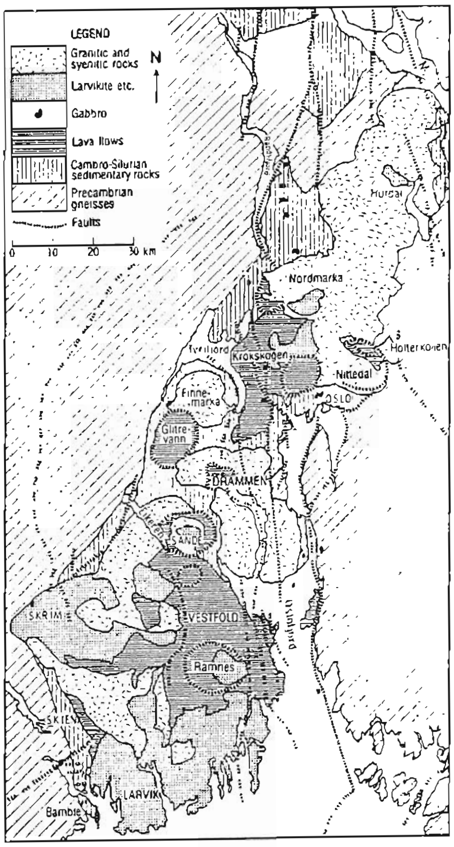


Fig. 1. Geological map of the Oslo rift, based on maps by Oftedahl (1969) and Ramberg & Larsen (1978). H: the Holmestand-Horten area.

Figure 15. From Neuman (1988)

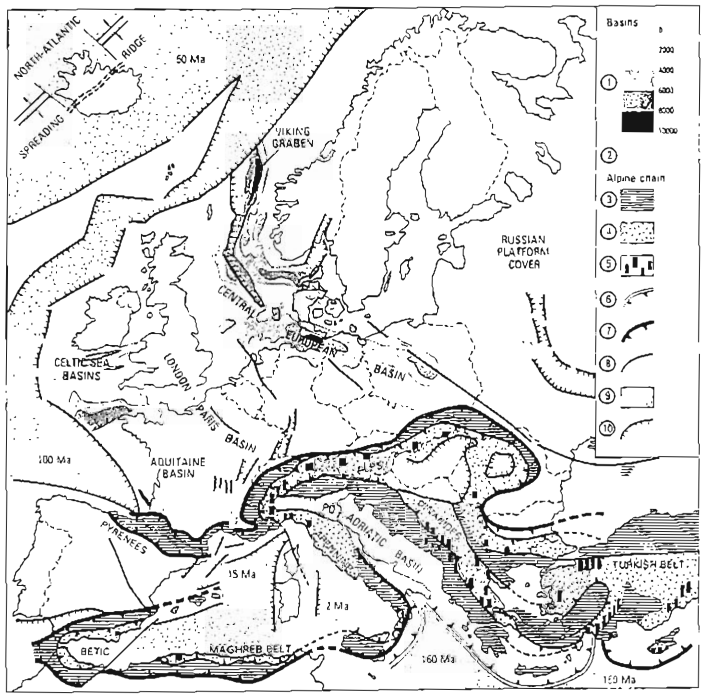


Fig. 2.11 Alpine chain and sedimentary basins in Europe.
Sedimentary basins 1. These are indicated by the thickness of the post-Permian deposits that filled them. The major central European basin is shown, with its extensions on the Atlantic passive continental margin, in the North Sea basin and in the Celtic Sea. 2. Basement to the sedimentary cover, outcropping around the basins
Alpine chain. 3. External zone. 4. Internal zone. 5. Remnants of oceanic crust (ophiolites). 6 Subduction zone. 7 Opposed thrust directions in the Alpine margins. 8 Faults. 9 Areas of oceanic crust with their age of formation. 10. Continental rise (see Fig. 2.4), a zone of transition from continental to oceanic crust.

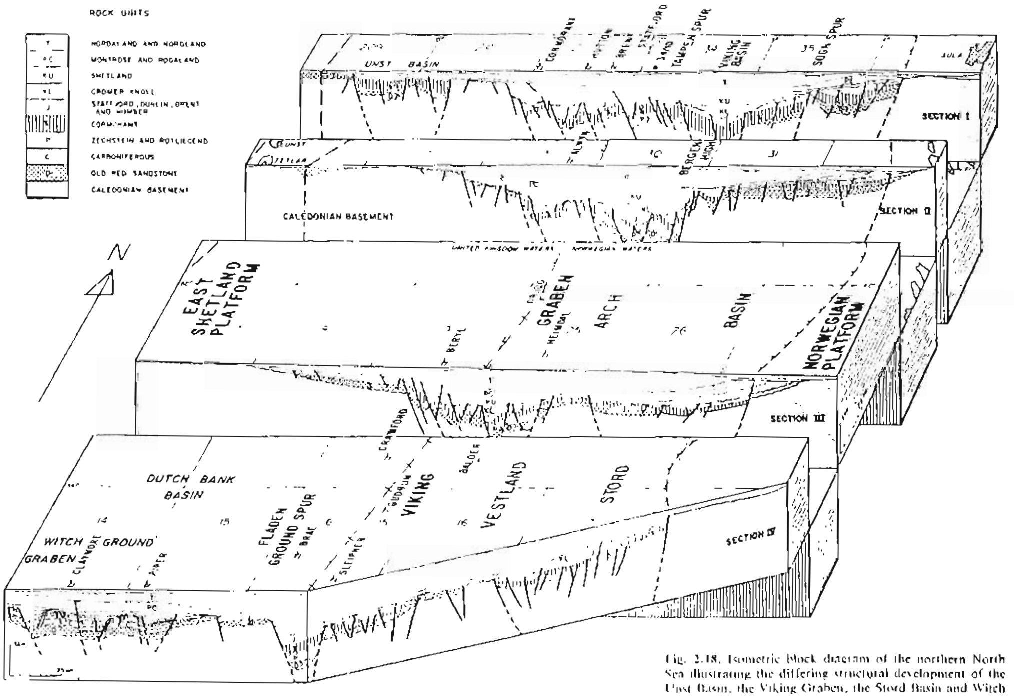


Fig. 2.18. Isometric block diagram of the northern North Sea illustrating the differing structural development of the Unst Basin, the Viking Graben, the Stord Basin and Witch Ground Graben.

Figure 16. Upper from Lumsden (1992), lower from Glennie (1986)

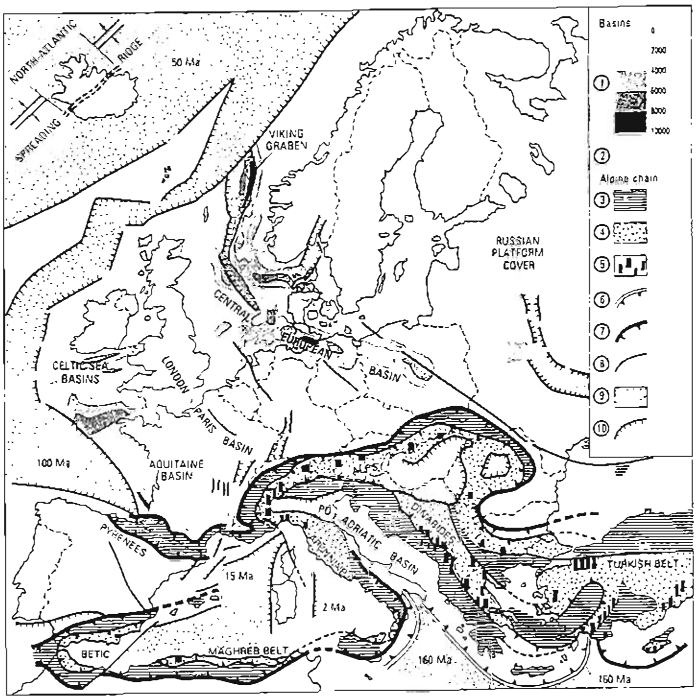


Fig. 2.11 Alpine chain and sedimentary basins in Europe.

Sedimentary basins. 1. These are indicated by the thickness of the post-Permian deposits that filled them. The major central European basin is shown, with its extensions on the Atlantic passive continental margin, in the North Sea basin and in the Celtic Sea. 2. Basement to the sedimentary cover, outcropping around the basins.

Alpine chain: 3. External zone. 4. Internal zone. 5. Remnants of oceanic crust (ophiolites). 6. Subduction zone. 7. Opposed thrust directions in the Alpine margins. 8. Faults. 9 Areas of oceanic crust with their age of formation. 10. Continental rise (see Fig. 2.4), a zone of transition from continental to oceanic crust.

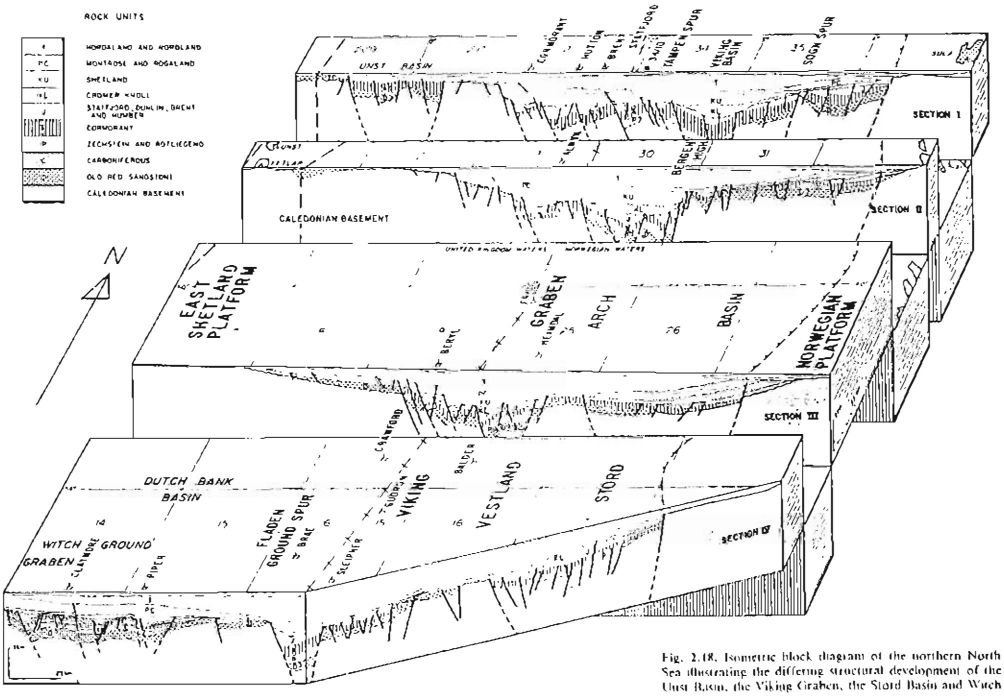


Fig. 2.18. Isometric block diagram of the northern North Sea illustrating the differing structural development of the Unst Basin, the Viking Graben, the Stord Basin and Witch Ground Graben.

Figure 16. Upper from Lumsden (1992), lower from Glennie (1986)

Thus in the off-shore areas there is a virtually complete section of Late Palaeozoic and Mesozoic sequences as opposed to on the craton. Here tectonics in Tertiary and Cenozoic times have resulted in major uplift of the craton which is recorded in the plateau areas by Tertiary marine land surfaces. Mesozoic sequences are little preserved on the craton, though near its western margin down-faulted sequences are found on Andøya (Lofoten islands), in Beistadfjord and Eddøyfjord (Trondheim region) and these may well represent down-faulted remnants of a once more extensive Mesozoic cover. In Denmark and more locally in southern Sweden Mesozoic and Tertiary deposits are more extensive due to the influence of the TTFZ. Late Palaeozoic and Mesozoic dyke intrusions are fairly common in southeastern Sweden, again reflecting the influence of the TTFZ and are also found at various intervals along the W. Norwegian coast reflecting major extensional phases in the North Sea.

Attention has been paid to the existence of long-lived fault zones on the marginal parts of the Fennoscandian Shield, and many evidences have been presented for fault-reactivation during extensional phases in the North Sea (Torsvik et al. 1992). The Møre-Trøndelag Fault Zone (MTFZ) is a furthur example of a deep-seated long-lived fault zone which has had a very varied history of strike-slip/extensional faulting, which can be traced from its landwared expression in the North Sea at least as far as the Viking Graben (Grønlie et al. 1994). Faulting in a number of cases can be shown to affect the glacial deposits and in certain areas the base of the Quaternary has been displaced by over 100 m (Cloetingh et al. 1991, Olesen et al. 1991). An interesting example of how multiple data sets including satellite imagery, geology and geophysics can be used to integrate information on faulting, joint-patterns, lineaments etc. is given in Karputz et al. (1992).

An increasing numbers of neotectonic movements on faults are being detected and the present-day stress field is characterised by a significant component of compression perpendicultar to the margin (Spann et al. 1991). The effects of such late fault movements have considerable implications concerning the nature of fractures in the crystalline rocks of the Fennoscandian Shield and their relationships to the present-day stress pattern are not without significance for hydrogeologists.

REFERENCES

Andersen, T.B., Osmundsen, P.T. and Jolivet, L. 1994: Deep crustal fabrics and a model for the extensional collapse of the southwest Norwegian Caledonides. Journal. Structural Geology, 16,1191-1203

Cloetingh, S., Reemst, D., Kooi, H., and Fanavoll, S., 1991: Intraplate stresses and the post-Cretaceous uplift and subsidence in northern Atlantic basins. Norsk geol. Tidsskr. 72, 229-235

Glennie, K.W., 1986: Introduction to the Petroleum Geology of the North Sea. Blackie, Oxford, 278 p.

Gorbatschev, R., 1985: Precambrian basement of the Scandinavian Caledonides. In Gee. D.G. and Sturt, B.A. (Eds). The Caledonide Orogen - Scandinavia and related areas. p. 197-212. Wiley, New York.

Karputz, M.R., Roberts, D., Olesen, O., Gabrielsen, R.H. and Herrevold T. 1992: Application of multiple data sets to structural studies on the Varanger Peninsula, Northern Norway, Int. Journal Remote Sensing, 14, 979-1003.

Kumpalainen, R., and J.P. Nystuen, 1985: Late Proterozoic basin evolution and sedimentation in the western part of Baltoscandia. In Gee, D.G., and Sturt, B.A. (Eds). The Caledonide Orogen-Scandinavia and related areas. p 213-232. Wiley, New York.

Lumsden, G.I. 1992: Geology and the Environment in Western Europe. Oxford University, Press, Oxford, 325 p.

Melezhik, V. and Sturt, B.A. 1994: General geology and evolutionary history of the early Proterozoic Polmak-Pasvik-Pechenga-Imandra/Varzuga Ust'Ponoy Greenstone Belt in the northeastern Baltic Shield. Earth Science Reviews 36, 205-241.

Neumann, R. 1988: Isotopic and petrological relations of the crust and upper mantle under the Oslo Graben, Southeast Norway. Norges geol. unders. Special Publication, 3, 7-13.

Olesen, O., Henkel, H., Lille, O.B., Mauring, E., Rønning, J.S. and Torsvik, T.H., 1991: Neotectonics in the Precambrian of Finnmark, northern Norway. Norsk geol. Tidsskr. 72, 301-306.

Roberts, D. and Sundvoll, B., 1990: Rb-Sr whole-rock and thin-section dating of mylonites from the Kalak Thrust Zone, near Børselv, Finnmark. Norsk geol. Tidsskr. 70, 259-266.

Roberts, D., and Gee, D.G. 1985: An introduction to the structure of the Scandinavia Caledonides. In Gee, D.G. and Sturt, B.A. (Eds). The Caledonide Orogen - Scandinavia and related areas., p. 56-68, Wiley, New York.

Siedlecka , A. and Roberts, D. 1992: The bedrock geology of the Varanger Peninsula, Finnmark, North Norway: An Excursion Guide. Norges geol. unders. Special Publication, 5, 1-45.

Spann, H., Brudy, A., and Fuchs, K., 1991: Stress evaluation in offshore regions of Norway. Terra Nova 3, 148-152.

Steel, R.J., Siedlecka A., and Roberts, D. 19085: The Devonian Basins of Norway and their deformation: a review in Gee D.G. and Sturt B.A. (Eds) . The Caledonide Orogen - Scandinavia and related areas p. 293-315, Wiley, New York.

Sturt, B.A. and Roberts, D. 1991: Tectonostratigraphic Relationships and Obduction Histories of Scandinavian Ophiolitic Terranes. In, Peters, T.et al (Eds) Ophiolite Genesis and Evolution of the Oceanic Lithosphere. Kleuer, Amsterdam.

Trench, A. and Torsvik, T.H.,1992: The closure of the Iapetus Ocean and Tornquist Sea: new palaeomagnetic constraints. Special Paper: Journal Geological Society of London, 148, 423-425.

Torsvik, T.H., Sturt B.A., Swensen E., Andersen, T.P. and Dewey, J.F., 1992: Palaeomagnetic dating of fault rocks: evidence for Permian and Mesozoic movements and brittle deformation along the extensional Dalsfjord Fault. Geophysics Journal. International 109, 565-580.

Geological context of the Nordic countries. Quaternary geology

Eiliv Larsen
University of Tromsø

CLIMATIC CHANGE AND QUATERNARY GEOLOGY

Most Quaternary deposits in the Nordic countries were deposited as the last ice sheet (the Scandinavian ice sheet) wasted back some 10,000-15,000 years ago. The great variety of sediment types formed during this time span, are utilized in many ways in the modern society. The type of utilization is largely dependent upon sediment properties and thus genesis of the sediments. In a human perspective these natural resources are non-renewable.

The recession of the last ice sheet was a result of climatic change from something 'worse-than-present' and into the present climatic regime. The changing climate over this time slice is only one example from a record of an ever-changing climate. Many of these changes are cyclic with different periods and amplitudes.

In the following, a brief discussion of climate variability will be followed by a description of climatic changes over the last interglacial-glacial cycle (i.e. the last 130,000 years), and the resulting imprints (mainly deposits) left behind on the landscape. There were three «ice-ages» before this during the last 1 million years.

TRIGGERS OF CLIMATIC CHANGE

Climate variations giving rise to shifts between glacial and interglacial periods are caused by three quasi periodic variations in the Earth's orbit and axial tilt (eccentricity of the orbit, tilt of the axis and precession of the axis, see Fig. 1). These variations cause large regional changes in solar insolation through time. The global mean insolation, however, remains almost constant through time, whereas global mean temperature varies considerably. Thus, there are internal (land - ocean - atmosphere) mechanisms for amplification that seem to be triggered by the external variations (Fig. 2). In this way small external variations causes large climatic changes at our latitudes.

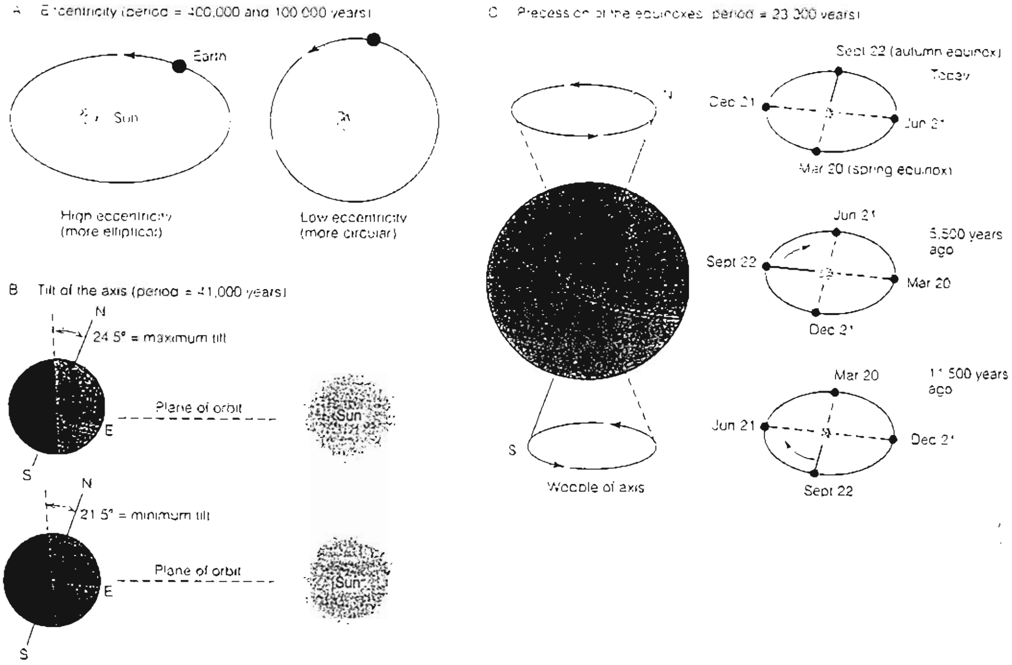


Fig. 1
Geometry of the Earth's orbit and axial tilt. A. Eccentricity. The shape of the orbit changes from almost circular (low eccentricity) to more elliptical (high eccentricity) over periods of 100,000 and 400,000 years. B. Tilt. The tilt of the axis ranges from 21.5 to 24.5^0, with each cycle lasting 41,000 years. C. Precession. The Earth wobbles on its axis like a spinning top, making one revolution every 26,000 years. All these factors cause a redistribution of insolation through time across the face of the Earth.

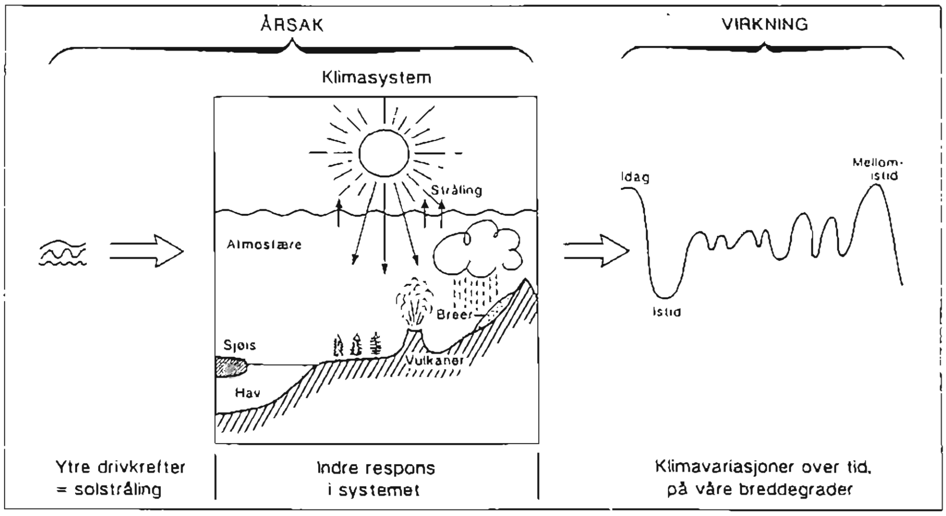
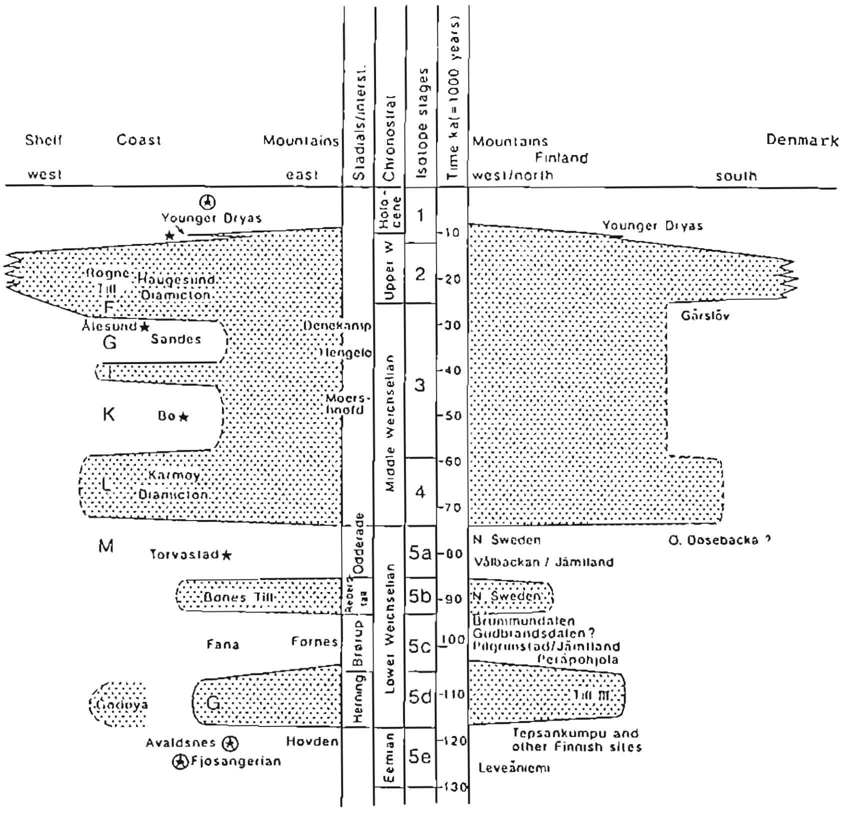


Fig. 2
Schematic illustration of cause and effect in the climate system (Mangerud 1992). Small external forcing cause feed-backs in the Earth climate system. This results in large natural climate variations.

CLIMATE AND GLACIER VARIATIONS FROM THE LAST INTERGLACIAL TO THE LAST GLACIAL MAXIMUM

The most dramatic expression of climatic changes caused by the small perturbations in the orbit and axial tilt (Fig. 1) are the mid-latitude (40-60°) ice sheets. During the last glacial stage (the Weichselian, 117,000-10,000 years B.P.), the Scandinavian ice sheet expanded and receeded several times (Fig. 3). Each time the landscape was affected strongly both due to erosion and deposition. The knowledge about the older part of the statigraphy is rather sparse because each successive glaciations removed most of previously deposited sediments. One area of exception may be part of northern Sweden where the landscape may have remained practially unaltered during several glacial episodes.

After the last interglacial (the Eemian, 129,000-117,000 years B.P.), the Scandinavian ice sheet started to grow somewhat later than ice sheets in North America/Antarctica. A major Scandinavian ice sheet developed some 110,000 years ago (Fig. 3). It's lateral extent is quite uncertain: It seems to have extended to the coastal areas of Norway, but the extent into Sweden and Finland is even more speculative.



Ⓐ Fauna suggesting climate as today, or warmer.
★ Fauna with Atlantic Current "guide fossils."

Fig. 3
Schematic glaciation curves for the last interglacial/glacial cycle in fennoscandia (Mangerud 1991). The left curve is for the west side of the mountains, the right curve for the east side.

During two Early Weichselian ice free periods (Brørup and Odderade), most, if not all, ice wasted away. The climate, however, was not as favourable as today. Organogenic and marine sediments from these periods are found both within and outside the Weichselian maximum ice limit (Fig. 4). The two ice free periods were separated by relatively restricted ice advances on either side of the ice divide.

In Middle to Late Weichselian time ice volumes generally were larger and duration of periods with ice cover were longer compared to the Early Weichselian. One or more periods with restricted ice volumes separate the periods with largest volumes. The maximum ice extent was reached some 20,000 years ago. At this time the ice sheet covered most of Scandinavia, except for the western part of Jutland, and possibly also some coastal mountains in western Norway (Fig. 4).

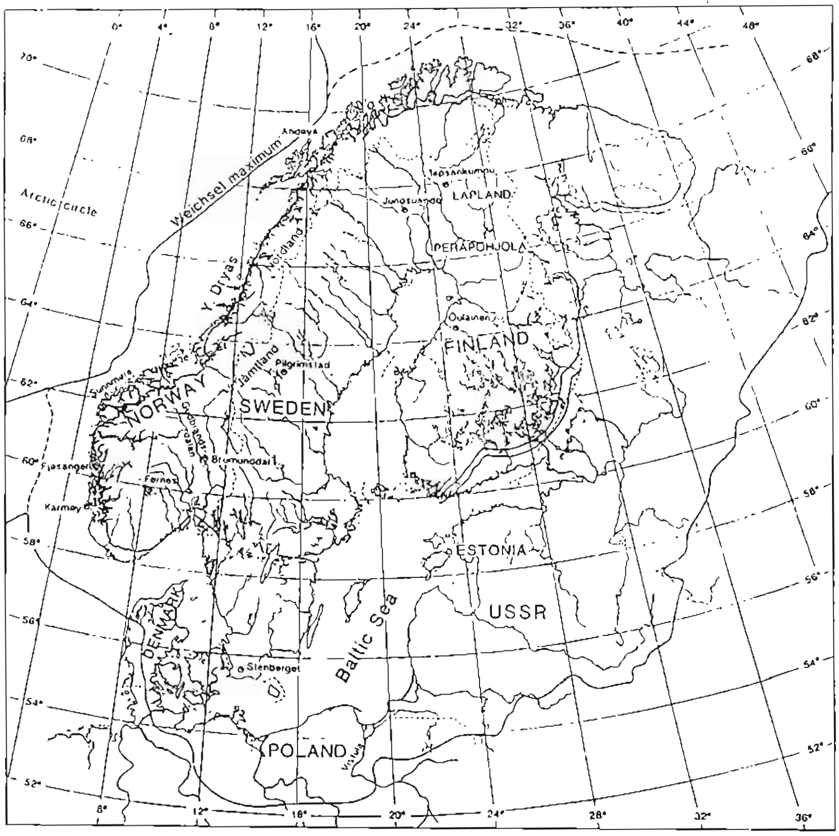


Fig. 4
The Weichselian maximum and Younger Dryas (11,000-10,000) ice margins of the Scandinavian ice sheet.

DEGLACIATION FROM THE LAST GLACIAL MAXIMUM

A considerably warming of the climate took place from some 15,000-16,000 years ago, and the large ice sheets started receeding rapidly. From that period and onward most sediments presently covering Scandinavia were formed: End moraines, basal tills, a variety of glaciofluvial sediments, marine/ glaciomarine sediments, glaciolacustrine sediments and many others. Most of these sediments were deposited in the proximity to the retreating ice front. Accordingly, the same type of sediments gets progressively younger as one moves towards the glaciation center(s). A major cooling took place between 11,000 and 10,000 years before present resulting in halts and readvances of the ice sheet. This resulted in the formation of the most prominent of all ice front deposits both in Sweden, Finland and Norway. Around 10,000 years before present climate rapidly improved again. This marks the end of the last glaciation, and the ice sheet was soon gone.

In mid Holocene time most, if not all, ice melted away, even from the mountaneous areas of Norway (Fig. 5) due to a climate generally somewhat warmer than today. Climate then gradually became worse from some 3,000 years ago, and glaciers started to form in the mountains. The glaciers attained a Holocene maximum position only some 250-300 years ago during the so-called 'Little Ice Age.'

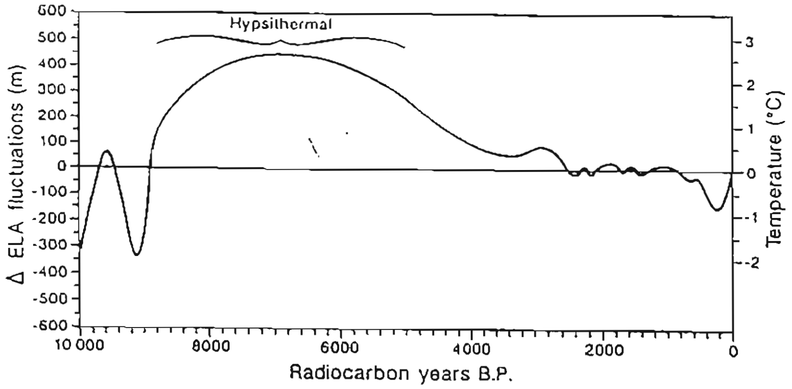


Fig. 5
Holocene equilibrium-line altitude (ELA) fluctuations and temperature variations relative to present in the Jostedalsbreen region (Norway), according to Nesje & Kvamme (1991). During the hypsithermal (temperature optimum), mountain glaciers were completely melted away.

SEDIMENT TYPES AND PROPERTIES

Tills, end moraines and glaciofluvial sediments are by far the most widespread of the Quaternary sediments in our region. The main glacial geological features in central Norway are shown in Fig. 6. Tills are principally formed in two ways:
-By lodging; Material frozen to the sole of the glacier is mechanically detached from the glacier as the ice moves over the bedrock. This is called 'Lodgement till.'
-By melting; Material is melting from the sole of the glacier and deposited right beneath it. This may take place both in dynamically active and inactive ice. This is called 'Melt-out till.'
Most basal tills probably are transitional types between the two end members. Tills also grade to a number of rapid and slow mass movement deposits dependent resedimentation upon deposition.

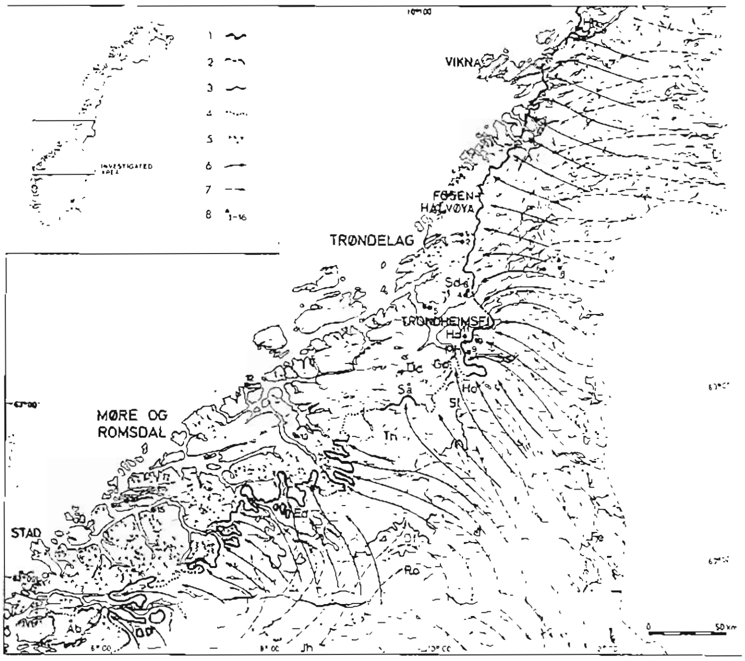


Figure 43 The main glacial geological features of Central Norway, after Sollid & Sorbel (1979, 1981). Abbreviations of geographical names given are as follows: Hh = Heilhornet. Li = Lierne. Sd = Skaudalen. Hd = Heimdal. Mh = Melhus. Od = Orkdalen. Gd = Gauldalen. Ss = Selbusjoen. Så = Storås. St = Storen. Ho = Hovin. Th = Trollheimen. Ed = Eikesdalen. Df = Dovrefjell. Ro = Rondane. Fe = Femund. Åb = Ålfoten. Jh = Jotunheimen. 1 = Maximum ice limit of the Younger Dryas advance, the main advance. 2 = Terminal moraines older than the Younger Dryas. 3 = Ice margin during the late Younger Dryas or during the transition to the Holocene. 4 = Assumed ice marginal positions. 5 = Younger Dryas cirque moraines. 6 = Younger Dryas ice-flow directions. 7 = Early Holocene ice-flow directions. 8 = Sites of radiocarbon dates (site number and name, radiocarbon date, reference, submitted by)

1	Åfjord	11480 ± 160 BP	T-3655	H. Selnes
2	Åfjord	10520 ± 230 BP	T-3660A	H. Selnes
3	Leksvik	11190 ± 200 BP	T-3858A	H. Selnes
4	Leksvik	10750 ± 130 BP	T-3860A	H. Selnes
	Leksvik	10420 ± 130 BP	T-4101A	H. Selnes
5	Rissa	11780 ± 90 BP	T-3034	M. Lolaldli
6	Levanger	9880 ± 40 BP	T-3997	H. Sveian
7	Verdal	9990 ± 130 BP	T-3999	H. Sveian
8	Tromsdalen	9930 ± 130 BP	T-3998	H. Sveian
9	Kræbu	9810 ± 120 BP	T-3113	A. J. Reite
10	Tiller	10990 ± 190 BP	T- 787	A. J. Reite
	Ekle	10230 ± 130 BP	T- 786	A. J. Reite
	Ekle	10150 ± 100 BP	T- 854	A. J. Reite
11	Steinan	11020 ± 90 BP	T-3296	A. J. Reite
12	Kristiansund	12090 ± 100 BP	T-1605	R. Kræmer
13	Skard	11520 ± 140 BP	T-3783	B. A. Follestad
14	Ålesund	12060 ± 160 BP	T-2305	J. Mangerud
15	Sykkylven	11620 ± 120 BP	T- 413	A. J. Reite
16	Kaldvatn	± 80 BP	T-3357	J. L. Sollid

Note the regional pattern of the sites in relation to the Younger Dryas maximum ice limit

Fig. 6
The main glacial geological features of central Norway (Sollid & Sørbel 1979).

Most tills are rather compact and poorly sorted reflecting the ability of glaciers both to produce, transport and deposit material of all size classes. As a result both porosity and permeability is rather low. As the ice moves over the bedrock, it picks up local bedrock material incorporating it to its load. The transport distance, however, normally is very short meaning that the composition of tills mainly will reflect the nature of the underlying bedrock. End moraines are produced at the ice margin in a number of ways:

-By the bulldozing effect of advancing glaciers in which case the material becomes rather unsorted and 'till-like.'

-By several frontal advances and retreats that may produce a stacking of debris-laden ice. Upon melting, this leaves a rather chaotic type of sediment along the former ice margin with highly varied texture and structure.

-By deposition directly from the glacier into a lake or the sea. This may produce glaciomarine or glaciolacustrine ice contact deltas or fans. These gets progressively finer in grain size and better sorted away from the ice margin. Due to land uplift since deglaciation, such deposits very often are found well above present sea level. Sediment composition may be highly variable, but normally the finest graines (silt and clay) are transported away from the source area leaving this zone rather sorted. Accordingly porosity and permeability may be rather high. Such ice contact deposits are very valuable resources for several purposes, a.o. as ground water reservoirs. Formation and composition is schematically illustarted in Fig. 7.

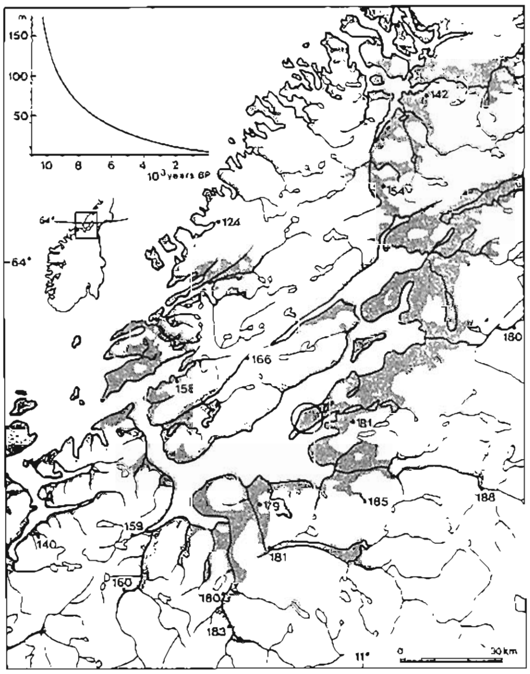
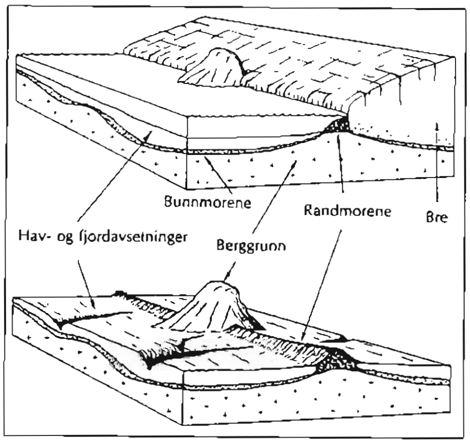


Fig. 7

Schetch showing how a marginal moraine may be formed in a fjord environment (top), and cross-section of such a moraine (bottom). Relatively sorted layers (D, E, F, G, H) may alternate with more till-like material (C). On top (B) is more fine grained marine material deposited as the ice wasted away. The uppermost bed is more gravelly due to resedimentation as the ridge emerged from the sea during the post-glacial land uplift.

Ice-contact deposit also is transitional to outwash sediments (or sandur sediments) deposited above contemporary sea level. This type of glaciofluvial sediments are deposited by melt-water rivers in front of the glaciers. Typically, the discharge of melt-water streams vary considerably as does the course of the stream. Accordingly, a section through such a deposit may reveal a very complex pattern both in terms of texture and structure. The general pattern, however, is a progressiv fining away from the source of sediments (the glacier).
Another common type of glaciofluvial sediments are the so-called eskers. These are sediments deposited by melt-water streams in tunnels under or in the ice. Upon melting of the ice, very distinct, long ridges, mainly composed of gravelly material appear. These are valuable resources in many areas.

All these sediment types are very common in Scandinavia, in addition to raised marine/glaciomarine and glaciolacusterine sediments.
During the Holocene, sediment accretion has mainly taken place due to resedimentation, partly due to slope processes and partly due to sea level changes since the last glaciation. This has given rise to slope derived sediments especially along the west coast of Norway. In many regions, clay deposited shortly after deglaciation, have given rise to clay slides, the most well-known being the Verdalen clay slide that took place in 1893. A sketch showing areas of marine clays in Trøndelag are shown in Fig. 8.



A schematic sketch showing how a marginal moraine is formed in a fjord (top). A cross-section of such a moraine ridge (bottom) often shows sorted layers (D, E, F, G, H) alternating with moraine material (C). Layers B consists of glaciomarine deposits, deposited after the glacier retreated, whereas A-A' is sand and gravel, caused by wave activity.

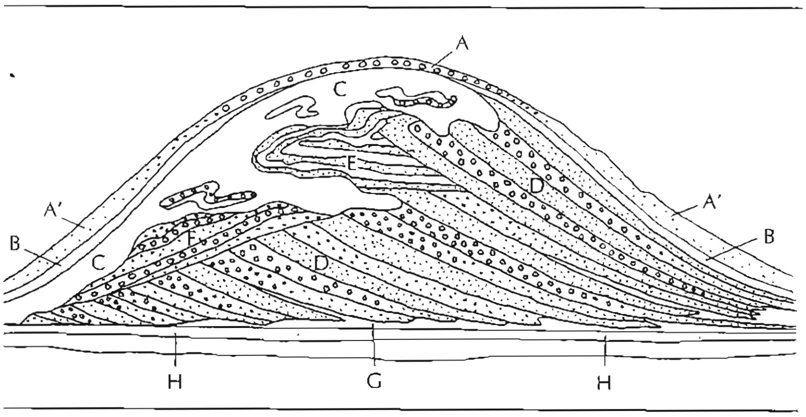


Fig. 8
Areas of marine clays in Trønelag (Sollid & Sørbel 1981)

REFERENCES

Mangerud, J. 1992: Naturlige klimavariasjoner (III). Dynamisk forståelse av klimaendringene. Naturen 4, 157-164.

Mangerud, J. 1991: The Scandinavian ice sheet through the last interglacial/glacial cycle. In Frenzel, B. (ed.): Klimageschichtliche Probleme der letzen 130,000 Jahre, 307-330. Gustav Fischer, Stuttgart.

Nesje, A. & Kvamme, M. 1991: Holocene glacier and climate variations in western Norway: evidence for Early Holocene glacier demise and multiple neoglacial events. Geology 19, 610-612.

Sollid, J.L. & Sørbel, L. 1979: Deglaciation of western central Norway. Boreas 8, 233-239.

Climatic context of the Nordic countries

Arvid Skartveit
Geophysical Institute, University of Bergen, Allégaten 70, N-5007 Bergen

1. CLIMATIC CONTROLS

On the one hand, the temperatures of the earth-atmosphere system are consistent with a global equilibrium between solar heating and infrared radiation loss to space. On the other hand, the temperatures are not governed by this global equilibrium requirement alone. Gradients in radiative heating/cooling produce temperature gradients which keep the atmosphere in motion. Through the exchange of matter, heat and momentum the atmosphere interacts strongly with the hydrosphere, the cryosphere, the biosphere, and even the litosphere. Taken together these interacting subsystems form the climatic system.

External climatic controls affect climate, but are not themselves affected by climate:

1) The flux of solar radiation.
2) The rotation and the orbital parameters of the earth.
3) Topography of the solid earth.
4) Amount of volcanic dust in the air.
5) Anthropogenic air pollution.

Internal climatic controls affect climate and are themselves affected by climate (feedback):

1) Glaciers, sea ice, snow cover.
2) Ocean currents.
3) Atmospheric water vapour and clouds.
4) Other atmospheric greenhouse gases.

Note that CO_2 is at present added to the atmosphere as an anthropogenic pollutant. CO_2 may thus be claimed to be an external climatic control. On the other hand, the huge natural cycle of CO_2 through the atmosphere is affected by climate. CO_2 may thus be regarded a combined external/internal climatic control.

Fig.1 (Wallace & Hobbs 1977) shows solar radiation incident on a horizontal surface at the top of the atmosphere as a function of latitude and month. In the summer hemisphere, the lengthening of the daylight hours with increasing latitude compensates for the increasing solar zenith angle, so that insolation is nearly constant over a wide range of latitude, and there is actually a slight maximum at the pole. In the winter hemisphere, the decreasing length of the day and the increasing solar zenith angle both contribute to the rapid decrease of insolation with increasing latitude until it reaches zero at the edge of the polar night region. The ratio of insolation at the summer solstice to that at the winter solstice increases substantially with increasing latitude.

Of the solar energy intercepted by the entire globe (342 Wm^{-2} = 100%), 30% are reflected back to space, 25% are absorbed in the atmosphere, and 45% are absorbed in the ground/ocean (Fig.2, Neiburger et al. 1982). The absorbtion of 70% solar energy is balanced by the earth-atmosphere system emitting 70% of longwave (infrared) radiation to space, of which 15% derive directly from the earth's surface and 55% from the atmosphere. The earth's surface thus accounts for two thirds of the global absorption of solar energy, while it accounts for only one fifth of the infrared emission to space. This atmospheric greenhouse effect is due to some 1/3% of the atmospheric mass, and it raises the surface temperature by some 30°C.

While the earth-atmosphere system as a whole is in radiative equilibrium, the entire atmosphere has a net radiative loss of 29% and the entire earth's surface has a net gain of 29%. "Steady-state" is maintained by non-radiative net transfer (conduction and convection) of 29% of latent and "sensible" heat from the earth's surface into the atmosphere.

Moreover, the earth-atmosphere system is not in radiative equilibrium at any given latitude. The solar energy absorption exceeds the infrared emission equatorward of some 35° latitude, and vice versa at higher latitude (Fig.3, Wallace & Hobbs 1977). "Steady-state" is maintained by atmospheric and oceanic heat transfer from lower to higher latitudes (up to the order of 200 MWm^{-1}). For example, evaporation exceeds precipitation between latitudes 10 and 40°, and "steady-state" is maintained by a net atmospheric transport of water vapour into higher and lower latitudes, allowing precipitation to exceed evaporation there (Fig.4, Wallace & Hobbs 1977). Another example is shown in Fig.5 (Wallace & Hobbs 1977) where the annual net energy flux through the Atlantic sea surface is plotted. The sea surface is heated by some 50 Wm^{-2} in equatorial regions, while it releases up to 250 Wm^{-2} off the American east coast and up to 100 Wm^{-2} north of Norway. This exemplifies the huge capacity of the ocean to store and transport energy, and to exchange heat between itself and the atmosphere. The ocean therefore has a substantial influence on the distribution of surface temperature. For example, the annual range in temperature at 60°N is less than 8°C in the Atlantic while it exceeds 50°C in the interior Eurasian continent (Fig.6, Wallace & Hobbs 1977).

2. CLIMATE DURING THE LAST 2000 YEARS

For the past one million years, our globe has experienced a sequence of glacial and interglacial periods, with global temperatures varying within a range of the order 5°C (Fig.7, Callaghan 1993). Compared to these variations, the last 8000 years have been very stable. Enhanced greenhouse warming may, however, produce a future very rapid global temperature rise comparable to the glacial-interglacial oscillations.

The last 2000 years was preceded by a several thousand years long considerably warmer period during which the largest glacier on the Scandinavian peninsula (Jostedalsbreen) most probably disappeared (Fig.8 (Nesje & Kvamme 1991 and Fig.9 (Wishman)). New glaciers formed, however, some 3100-3700 years ago. In the Viking era these glaciers were significantly smaller than today. But in the thirteenth and fourteenth century the climate deteriorated again. A general advance of the glaciers took place in this period, but the advance was small from about 1300 to 1550 when the climate of Scandinavia seems to have been relatively warm, with few severe winters. The most significant climatic event during the last 2000 years is the "little ice age" 1500-1850 A.D. during which period the mean temperature in the Nordic region was about 1°C colder than today and advancing glaciers caused much damage to human activity in general and to agriculture in particular (Johannessen 1970).

While temperatures more than 200 years back are inferred from indirect evidence, the most recent climatic history can be described by instrumental records, which allow identification of more short-lived climatic fluctuations than before. The causes of such short-term fluctuations should, however, be inferred with care. Coupled ocean-atmosphere models thus indicate that global temperatures may increase for several decades solely due to the interaction between internal climatic controls under fixed external climatic controls (Fig.10, IPCC 1992). It remains a major scientific problem to distinguish between such "natural" climatic variability and changes in external climatic controls.

During the instrumental period a global temperature rise is seen during the first four decades of the present century, and then a new temperature rise during the latest two decades (Fig.11, IPCC 1992). These features are not evenly distributed over the globe. In particular, the most recent of these two warming periods are not very obvious within the Nordic countries (Fig.12, Hanssen-Bauer 1990).

3. TODAYS CLIMATE

3.1 Climatic types

Latitude, elevation, and topographical barriers against intrusion of maritime air are main factors determining the climate in the Nordic countries, and five of Köppens' climatic types occur within the region (Fig.13, Klemsdal). In Denmark, southern Sweden and along the Norwegian coast up to the polar circle there is Cfb climate, which means a temperate (C: 22°C > warmest month > 10°C and coldest month > -3°C) rainy climate without dry seasons (f) and with long summers (b: 4 months or more > 10°C). North of this zone there is a narrow zone extending from the southern coast of Finland, across Sweden to the west coast of Norway having a snow-forest climate (D: 22°C > warmest month > 10°C and coldest month < -3°C) with no dry season and with long summers (Dfb). The northernmost Norwegian west coast has a Cfc climate, a temperate rainy climate with short summers (c: less than 4 months > 10°C). The Cfc climate is found also on the southern coast of Iceland. A tundra climate ET (0°C < warmest month < 10°C)) occurs in the high mountains of Norway and Sweden, in arctic Norway and in Iceland. The remainder of Scandinavia and parts of northern Iceland have the Dfc climate, which is a snow-forest climate with short summers.

3.2 Temperature, evapotranspiration

The normal length (1931-1960) of the vegetation period (daily mean air temperature > 6°C) ranges from above 210 days in southern Denmark to below 110 days in mountainous areas and coastal areas in the far north (Fig.14, Johannessen 1970). Similarly, the evapotranspiration ranges from above 500 mm yr^{-1} in parts of Denmark Sweden to below 200 mm yr^{-1} at high elevations and in the far north (Fig.15, Forsman 1976).

3.3 Precipitation patterns and chemistry

Several moisture-bearing air currents invade the Nordic countries and produce rain or snow at different times of the year. But none of these air currents are as moisture-bearing, strong or persistent as the westerlies and southwesterlies. The precipitation regime is, therefore, to a great extent determined by the amounts of water released within these air currents under the influence of the topography. Particularly, the topography plays an important role in Norway, Sweden and Iceland.

As outlined by Johannessen (1970) and by Forsman (1976, Fig.16) only small geographical differences occur in the mean annual precipitation in Denmark ($\approx$ 500-900 $mmyr^{-1}$) and Finland ($\approx$ 500-700 $mmyr^{-1}$) where the influence of topography is negligible. In Sweden there are three areas of high annual precipitation caused by orographic upglide: A narrow zone along the west coast (850-900 $mmyr^{-1}$), the western slopes of the Småland uplands in southern Sweden (1000 $mmyr^{-1)}$, and the western slopes of the high mountains of Norrland near the Norwegian border where the Sarek Massif receives about 2000 mm annually. Large local differences in mean annual precipitation also occur in the zone of maximum precipitation on the western slopes of the coastal mountains in western and northern Norway, with annual precipitation over 4000 mm in some areas. In some districts the screening effects of mountains causes very small amounts of precipitation over larger areas. This applies to some stations at the inland end of the Sognefjord (350-450 $mmyr^{-1}$), to the Otta Valley (250-300 $mmyr^{-1}$) which is a tributary valley to the

Gudbrand Valley, to the interior of Finnmark and Finnish Lapland where the annual amounts of precipitation are less than 500 mm in some areas. The annual precipitation in Iceland is highest in southeast and lowest in northeast (Lemmelä and Hyyverinen, Fig.17).

The typical maritime regions of Norway, west of the Scandinavian mountain chain, experience most precipitation in autumn and winter. The remainder of Norway, practically all parts of Sweden, Finland and Denmark experience most precipitation in summer. The months of least precipitation are February - April, except in western and northern Norway where March - May have the least precipitation. This maritime/inland contrast is exemplified by the seasonal variation at Skjåk in the northern Gudbrand Valley and at Brekke in Sogn close to the Norwegian west coast (Fig.18, Aune 1993).

The average annual frequency of precipitation days (≥ 0.1 mm) is 210-240 days on the western coast of Norway, 95-180 days in interior Norway, and 120-200 days in Denmark, Sweden and Finland. In all Scandinavian countries the greatest frequency of days with precipitation occurs from July through December.

Due to emissions from human activities, the chemical climate in Europe has changed significantly during the last century. The historical sulphur dioxide emissions in Norway and Poland (Fig.19, Mylona 1993) exemplifies the effects of e.g. industrialization, World War II, and even recent steps to reduce the emissions. The calculated sulphur concentrations in the air and the calculated deposition of oxidized sulphur at one grid square on the Kola peninsula and one grid square in southern Norway, reflects the impact of these emissions on atmospheric chemistry (Fig.20-21, Mylona 1993). It is seen that the 5-percentile critical load is substantially exceeded in both regions (5-percentile critical load = the maximum deposition that the ecosystem in question can tolerate in order that 95% of it is protected). The total deposition of oxidized sulphur accumulated over the period 1880-1991 ranges from above 500 $g(S)m^{-2}$ in parts of central Europe and UK, to 10-20 $g(S)m^{-2}$ in the northernmost parts of the Scandinavian peninsula and 5-10 $g(S)m^{-2}$ in northern Iceland (Mylona 1993). Except for significant sulphur emissions in the Kola peninsula, the major European emissions of sulphur and nitrogen compounds today occur outside the Nordic countries.

Precipitation removes a number of air pollutants with extreme efficiency. The precipitation distribution therefore strongly affects the deposition pattern of these pollutants. Modelled average annual precipitation for seven years (1985, 87-92, Fig.22, Berge 1994) displays a smoothed picture of the precipitation distribution outlined above.

The calculated average air concentration of sulphur dioxide ranges from above 20 $\mu g(S)m^{-3}$ in central parts of Europe to less than 1 $\mu g(S)m^{-3}$ in most of the Nordic countries (Fig.23, Berge 1994). The annual dry deposition of oxidized sulphur ranges from above 5000 $mg(S)m^{-2}$ in central parts of Europe to less than 100 $mg(S)m^{-2}$ in Iceland and parts of Scandinavia (Fig.24, Berge 1994). The annual wet deposition of oxidized sulphur ranges from above 2000 $mg(S)m^{-2}$ in central parts of Europe to 100-200 $mg(S)m^{-2}$ in parts of Scandinavia and even lower in Iceland (Fig.25, Berge 1994). The total deposition of sulphur is thus substantially less in the Nordic countries than in central parts of Europe, while the fraction wet deposition : dry deposition is considerably higher in the Nordic countries than it is closer to the main emissions. The dry and wet deposition of oxidized and reduced nitrogen (Figs 26-29, Berge 1994) show nitrogen depositions that are quite significant as fertilization contributions.

3.4 Runoff

The annual runoff (= precipitation - evapotranspiration) ranges from above 4000 mm yr^{-1} in some mountainous regions of western Norway to below 200 mm yr^{-1} in some coastal regions around the Gulf of Bothnia (Fig.30, Gottschalk 1979). Runoff regions may be based on the determination of boundaries between areas with the following runoff characteristics (Tollan 1975):

H_1: *Dominant snowmelt high water. The three months with the highest average runoff belong to spring or early summer (typically May-July).*
H_2: *Transition to secondary rain high water. The second or third highest monthly runoff takes place in autumn (typically October-November).*
H_3: *Dominant rain high water. The highest monthly runoff takes place in autumn or early winter (typically November-December).*

L_1: *Dominant low water flow in winter, caused by snow accumulation. The two months with the lowest runoff both belong to winter or early spring (typically February-March).*
L_2: *Transition zone. The two months with the lowest runoff do not belong to the same time of the year (typically: February and July).*
L_3: *Dominant summer low water caused by high evapotranspiration and/or low precipitation. The two months with the lowest runoff belong to summer and early autumn (typically June-August).*

Gottschalk et al (1979) distinguish five main combinations of regimes in the Nordic countries (Fig.31, Gottschalk 1979):

H_1L_1 *Mountain regime: Inner and northeast parts of Norway, Sweden and Finland.*
H_2L_1 *Inland regime: Fjord area along Norway's western coast, forest and coastal area to the west of the Gulf of Bothnia, Finland except south-western parts and northern Lapland.*
H_2L_2 (H_3L_2) *Transition regime: Inlands in Østlandet (Norway), Värmland, Dalecarlia and Gestrikland (Sweden) and the western and southern coast of Finland.*
H_2L_3 *Baltic regime: Parts of Møre-Trøndelag (Norway), southeastern Sweden, southwestern coast and the archipelago of Finland.*
H_3L_3 *Atlantic regime: Frontier coastal regions of Norway, southwestern Sweden and Denmark.*

3.5 Seasonal freezing and thawing

Normally the temperature of the earth's surface falls below 0°C just before the air temperature does it. As long as the ground is bare in autumn and in early winter, glazed ice is a common phenomenon when warm air masses invade the Nordic countries and frost have started to penetrate into the soil. The depth of ice in the soil ranges greatly from place to place depending on snow cover (Fig.32, Williams & Smith 1989) and even on soil type and soil water content (Fig.33, Williams & Smith 1989). But on average the maximum depth of ice in the soil is about 2.5 m in northern Scandinavia, 1.5 m in central and 1 m or less in southernmost Norway and Sweden (Johannessen 1970). The conditions may vary considerably from year to year depending upon how early the continuous snow cover begins, as well as upon the air temperature..

Snow lies on the ground for more than half the year in most parts of Scandinavia. Snow cover begins to form in the high mountains of Norway and in northern Sweden by the end of September and persists until the end of May or early June. In arctic Scandinavia snow cover generally lasts from the end of October until the middle of May, while in southern Norrland, central Finland and southeastern Norway, from the beginning of November until the middle or end of April. The duration of the snow cover varies considerably ,however, generally being shorter along the coasts and at lower elevations. Because of the thawing effect of the mild Atlantic air, snow cover is an irregular phenomenon in Denmark, southwestern Finland, and along the western and southernmost coast of Sweden and Norway. Except at coastal points where frequent thawing usually occurs, the snow cover attains its greatest depth in March. Snow depth reaches an average of 1 m and more in northern Norway/Sweden, and less than 15 cm in the snowiest month in southern Sweden, extreme southern Norway and at some coastal points in western Norway. In Finland the average snow depths range from 20 cm at the Åland Island to about 80 cm in the northern and eastern districts, and as much as about 3 m snow accumulates on the ground at a number of stations in the Scandinavian mountain chain. Above the so-called "snow-line" the snow cover lasts throughout the year. On the average the snow-line in southern Norway occurs at about 1450 and 2000 m in the western and eastern districts, respectively, and in norther Norway/Sweden at about 800 m (Johannessen 1970).

Permafrost distribution is reasonably well predicted from mean winter snow cover and mean daily air temperatures. The positive mean daily air temperature for the year are added together to provide the total annual thawing index, while the sum of the negative temperatures for the year yield the total annual freezing index (Harris 1986). For stations with a mean winter snow cover of under 50 cm, there is a clear relationship between permafrost zonation and freezing and thawing indices (Fig.34, Harris 1986). It is stressed, however, that the occurrence of permafrost is heavily affected by the kind of surface material present (vegetation, peat, soil) and by snow cover(Fig.35, Williams & Smith 1989).

In the Nordic countries continuous permafrost is found in Svalbard, where ground water recharge takes place mainly from below glaciers and from the bottom of some lakes and rivers. Zones of discontinuous permafrost are found in Iceland and in the northern parts of the Scandinavian peninsula (Fig.36, Péwé 1977). Fig.37 (Skartveit et al. 1975) shows, as an example, the seasonal freezing/thawing cycle of peat in a mire at Stordalen, Abisko (345 m a.s.l.). The cycle can be regarded as starting during the earliest days of May, when the frost disappears from the surface. The very slow diffusion of heat through the peat layers results in an average thawing depth of 60 cm during the latest part of the growing season. The maximum depth of thaw occurs in November, and seems to be not more than some few centimetres greater. But during the middle of October the first freezing starts at the surface, and normally penetrates the whole unfrozen layer of elevated areas in 3-4 weeks.

4. FUTURE CLIMATE

Under the IS92a scenario (annual economic growth 2.3%, world population 11.3B by 2100 A.D.) the global mean temperature is predicted to increase by some 2-4°C over the next century (Fig.38, IPCC 1992). Predictions of changes relating to the hydrological cycle, are more insecure than predictions of temperatures. Moreover, the relatively low confidence attached to regional projections of any sort should be emphasized (IPCC 1992).

However, most models agree that warming will be greatest in the high latitudes of the Northern Hemisphere. A larger warming over land areas compared to that over the oceans is common to all models in both winter and summer. This ocean-land asymmetry contributes to a more rapid warming in the Northern Hemisphere compared to that in the Southern Hemisphere. In particular over the Arctic Ocean winter warming is expected to exceed summer warming (Fig.39, IPCC 1992). Precipitation is expected to increase in the Northern Hemisphere in high latitudes throughout the year, and in much of mid-latitudes in winter (Fig.40, IPCC 1992). In the Northern Hemisphere, winter soil moisture is expected to increase over the mid-latitude continents. In the summer many areas of drying are expected, in particular outside the Nordic region (Fig.41, IPCC 1992).

Global warming may result in a net release of carbon from land to the atmosphere because of (a) the greater responsiveness of respiration to initial temperature changes, compared with photosynthesis, (b) increased decomposition of surface litter and soil organic matter, particularly in boreal forests and the arctic tundra, and (c) replacement of water-stressed southern temperate forests of the Northern Hemisphere by grasslands. Increased concentrations of atmospheric CO_2 also has the potential to increase plant production in several ways and increased nutrient availability from faster decomposition may also lead to higher productivity, at least initially (Callaghan 1993).

5. REFERENCES

B. Aune (1993): Nasjonalatlas for Norge. Hovedtema 3 Luft og vann. Klima. Statens Kartverk.

E. Berge (1994): Personal communication (EMEP/MSC-W).

T.V. Callaghan (1993): An application of the Intergovernmental Panel on Climate Change's scientific assessment to northern ecosystems. Pp. 12-23 in Holten, Paulsen and Oechel (eds.): Impacts of climatic change on natural ecosystems with emphasis on boreal and arctic/alpine areas. Norwegian Institute for Nature Research and The Directorate for Nature Management.

A. Forsman (1976) Water balance maps of the Nordic countries. Vannet i Norden, 4 - 1976.

L. Gottschalk, J. Lundager Jensen, D. Lundquist, R. Solantie and A. Tollan (1979): Hydrologic Regions in the Nordic Countries. Nordic Hydrology, 10, 273-286.

I. Hanssen-Bauer (1990): Klimaendringer de siste 100 år. Kap. 6 i NILU OR 21/90: Drivhuseffekten og klimautviklingen. Bidrag til den interdepartementale klimautredningen.

S.A. Harris (1986): The Permafrost Environment. Croom Helm.

Intergovernmental Panel on Climate Change (1992): Climate Change 1992. The supplementary Report to The IPCC Scientific Assessment. World Meteorological Organization/United Nations Environment Program.

T.W. Johannessen (1970): The Climate of Scandinavia. Pp. 23-79 in Landsberg: World Survey of Climatology Volume 5. Elsevier Publishing Company.

T. Klemsdal: Köppens klimasystem og Nordens klima. Norsk geografisk Tidsskr. 34, 107-114.

Lemmelä and Hyyverinen (incomplete reference).

S. Mylona (1993): Trends of sulphur dioxide emissions, air concentrations and depositions of sulphur in Europe since 1880. EMEP/MSC-W Report 2/93.

A. Nesje and M. Kvamme (1991): Holocene glacier and climate variations in western Norway: Evidence for early Holocene glacier demise and multiple Neoglacial events. Geology, 19, 610-612.

M. Neiburger, J.G. Edinger and W.D. Bonner (1982): Understanding our Atmospheric Environment. W.H. Freeman and Company.

T.L. Péwé (1977): Permafrost research. A workshop survey of some recent activities. Frost i Jord Nr. 19, 3-10.

A. Skartveit, B.E. Rydén and L. Kärenlampi (1975): Climate and Hydrology of Some Fennoscandian Tundra Ecosystems. In Wielgolaski (Ed.): Ecological Studies. Analysis and Synthesis, Vol. 16, Fennoscandian Tundra Ecosystems, Part 1. Springer-Verlag.

A. Tollan (1975): Hydrologiske regioner i Norden. Vannet i Norden, No.1.

J.M. Wallace and P.V. Hobbs (1977): Atmospheric Science. An Introductory Survey. Academic Press.

P.J. Williams and M.W. Smith (1989): The Frozen Earth. Fundamentals of geocryology. Cambridge University Press.

E. Wishman: Klimahistorie på Jæren gjennom 8000 år. In: "Viva på Hå. AmS-Småtrykk 16, Arkeologisk museum i Stavanger.

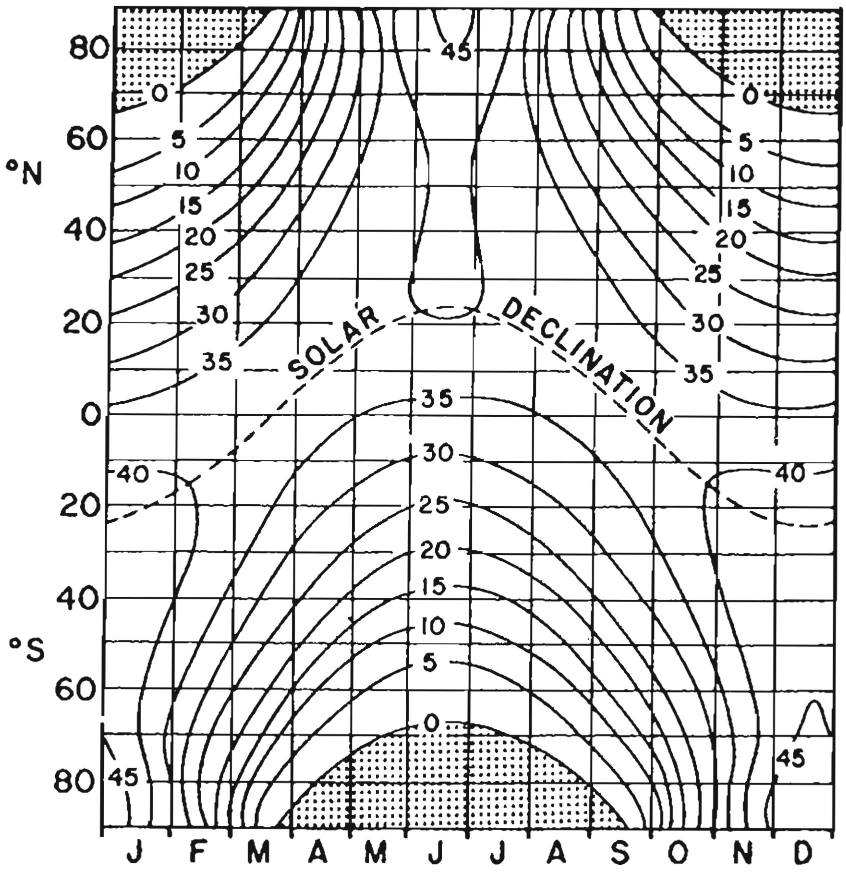


Fig.1 Solar radiation incident on a horizontal surface (given in $MJm^{-2}day^{-1}$*) at the top of the atmosphere as function of latitude and date (from Wallace & Hobbs 1977).*

63

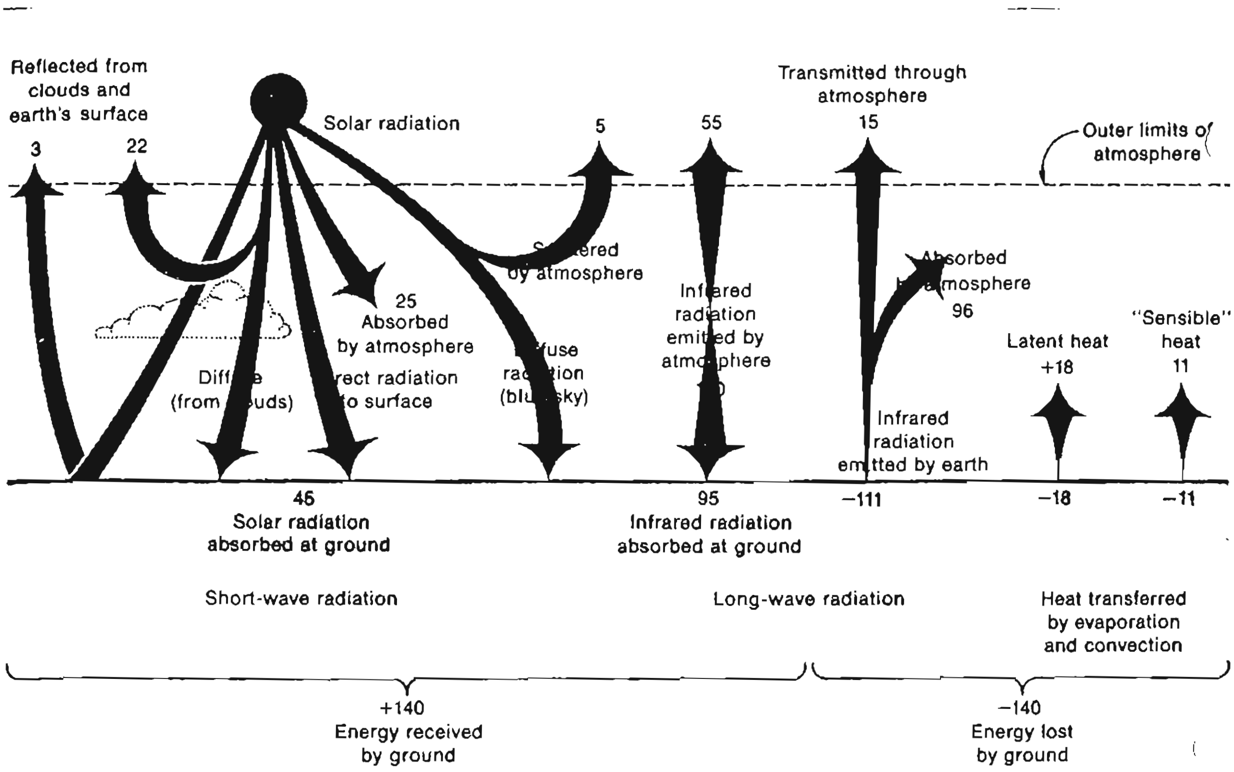


Fig. 2 The annual mean global energy balance for the earth-atmosphere system (from Neiburger et al. 1982). Numbers are given as percentages of the globally averaged solar irradiance incident upon the top of the atmosphere (342 Wm^{-2}).

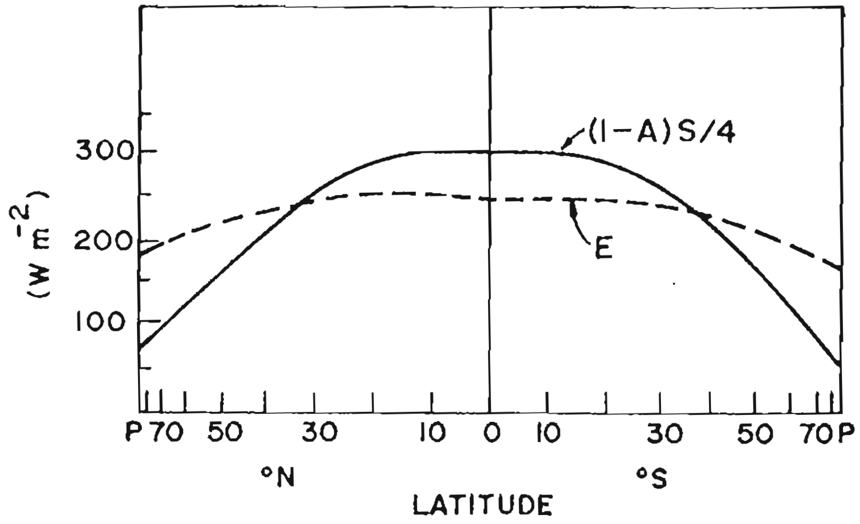


Fig.3 Annual average absorption of solar radiation (full line) and outgoing terrestrial radiation (broken line) as a function of latitude (from Wallace & Hobbs 1977).

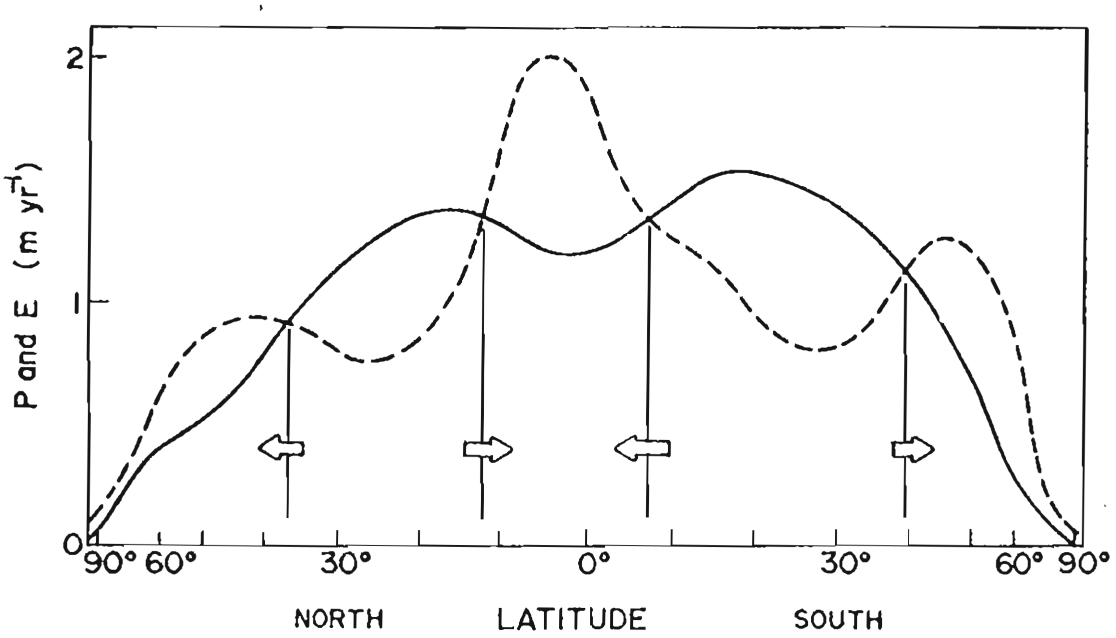


Fig. 4 Annual average evaporation (full line) and precipitation (broken line) given in meters per year. Arrows represent the sense of the water vapour flux in the atmosphere required to maintain "steady state" (from Wallace & Hobbs 1977).

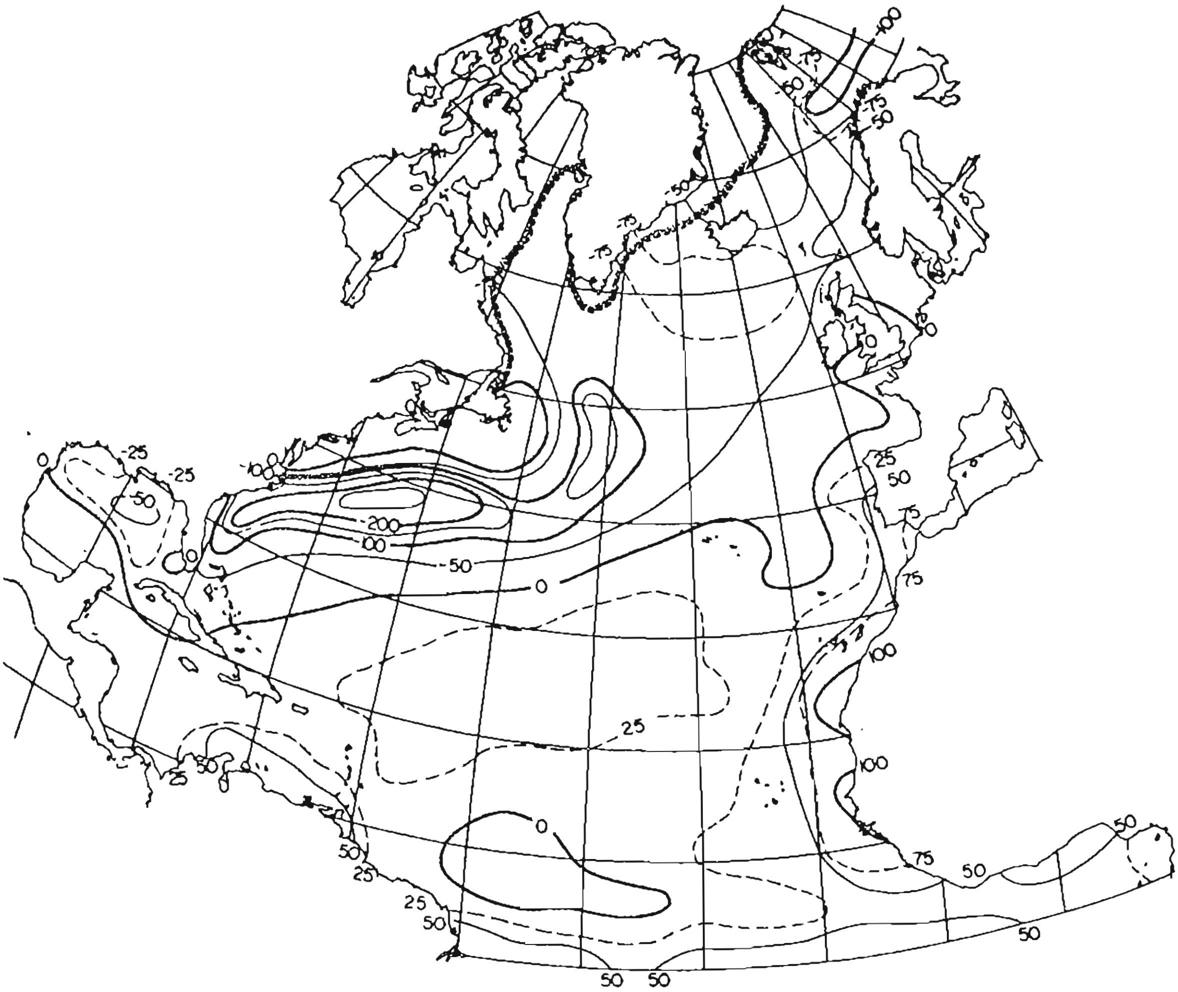


Fig.5 Average annual energy flux (Wm^{-2}) through the sea surface. Positive values indicate a downward flux (from Wallace & Hobbs 1977).

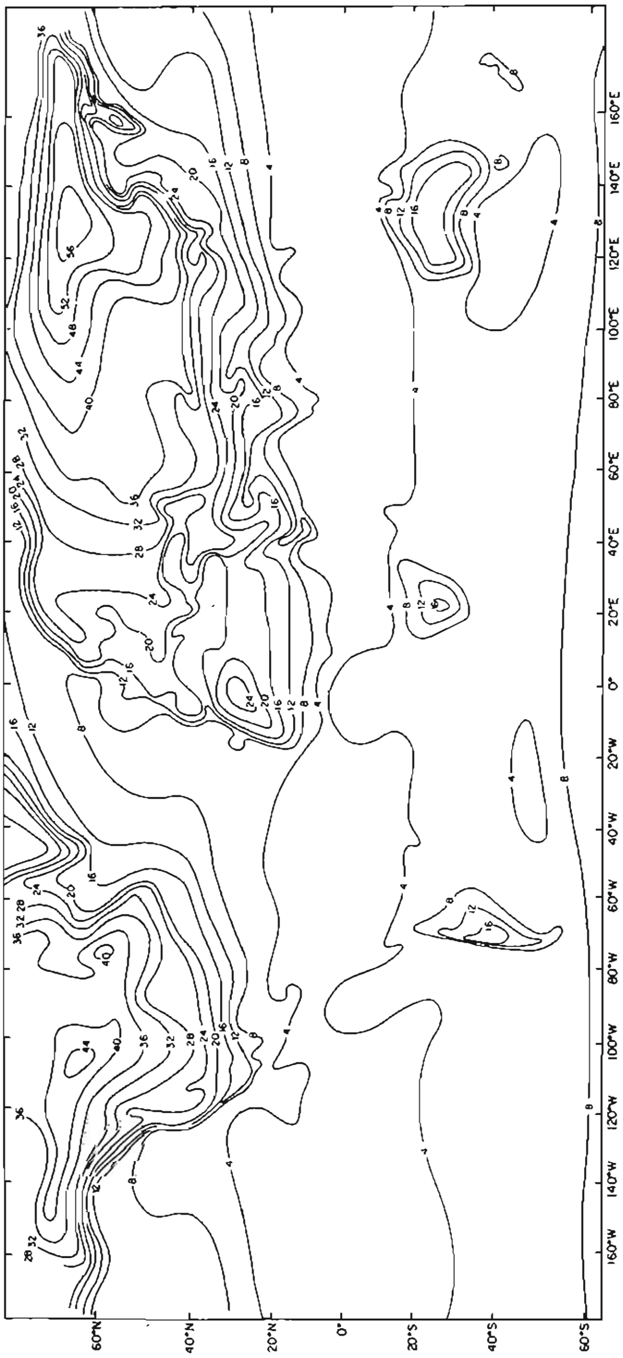


Fig. 6 The "Find the Continents" game! Annual range in temperature (°C) at the earth's surface (from Wallace & Hobbs 1977).

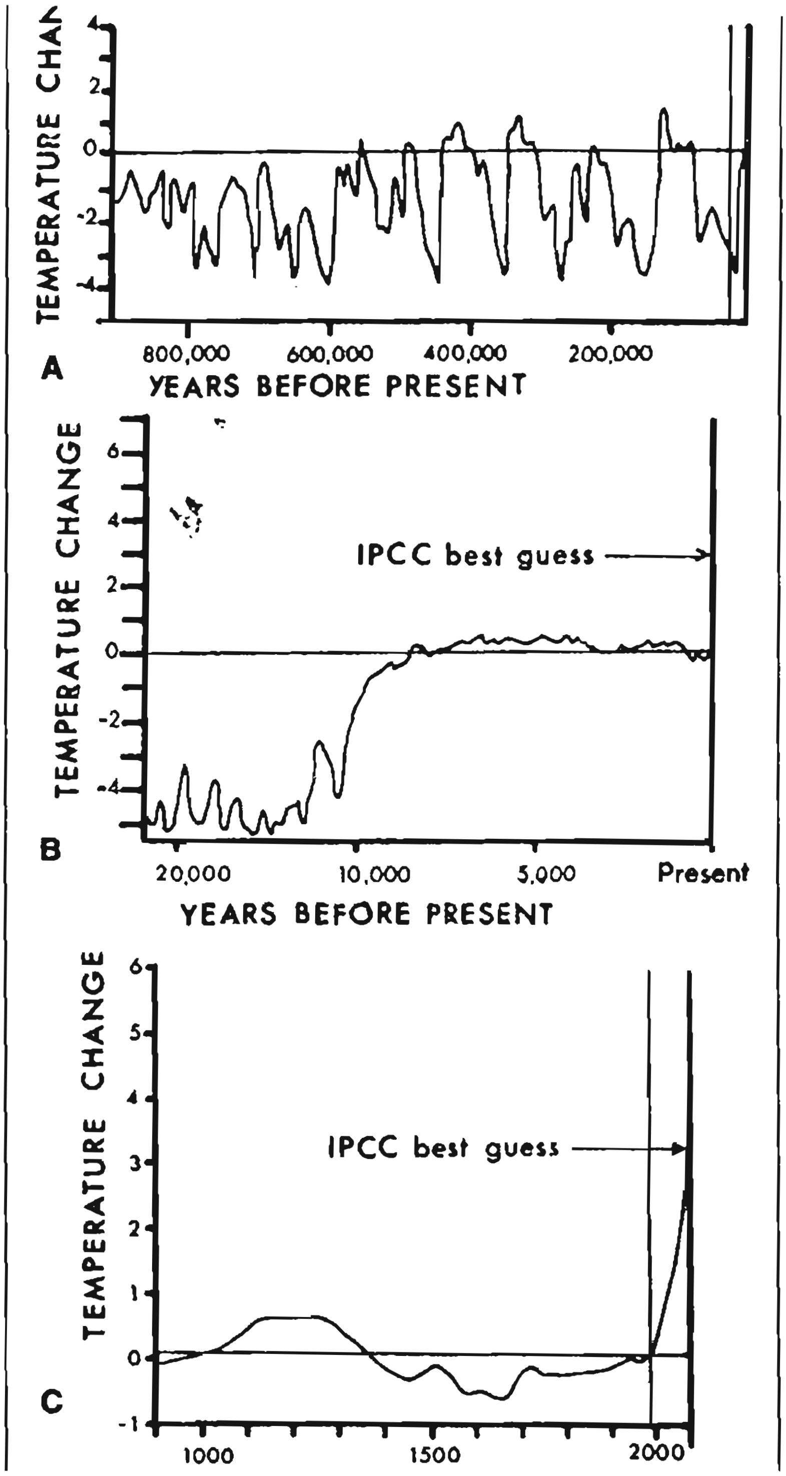


Fig. 7 Schematic diagrams of global temperature variations since the Pleistocene on three time scales, (a) the last million years, (b) the last 20 000 years, and (c) the last 1000 years. Predicted increase in temperature over the next century is included for comparison (from Callaghan 1993).

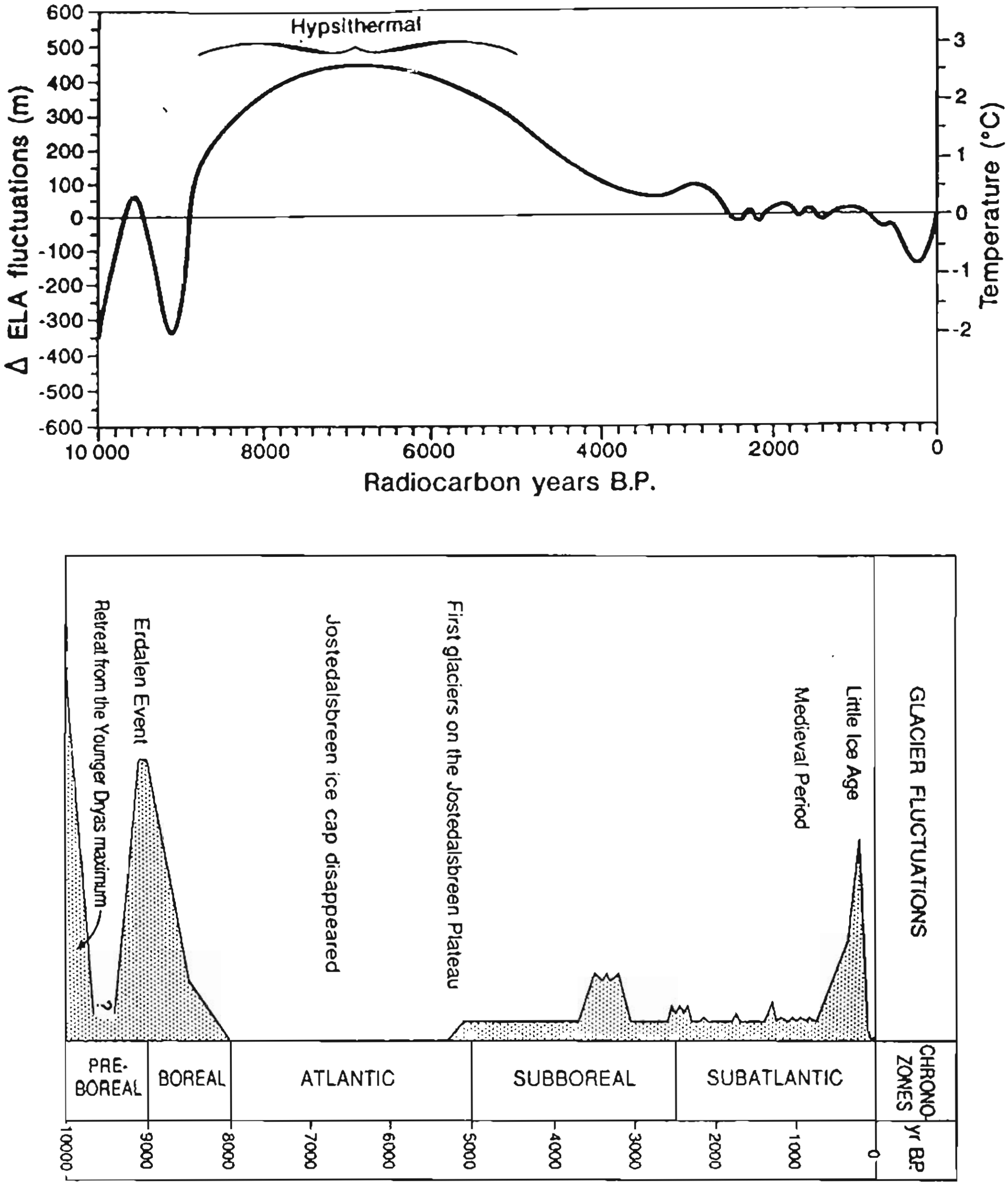


Fig.8 Holocene equilibrium-line altitude (ELA) fluctuations and temperature variations relative to present in Jostedalsbreen region (Figure 2 in Nesje & Kvamme 1991). Holocene glacier fluctuations in Jostedalsbreen region. Pattern indicates inferred extent of ice cover. Horizontal scale is schematic (Figure 3 in Nesje & Kvamme 1991).

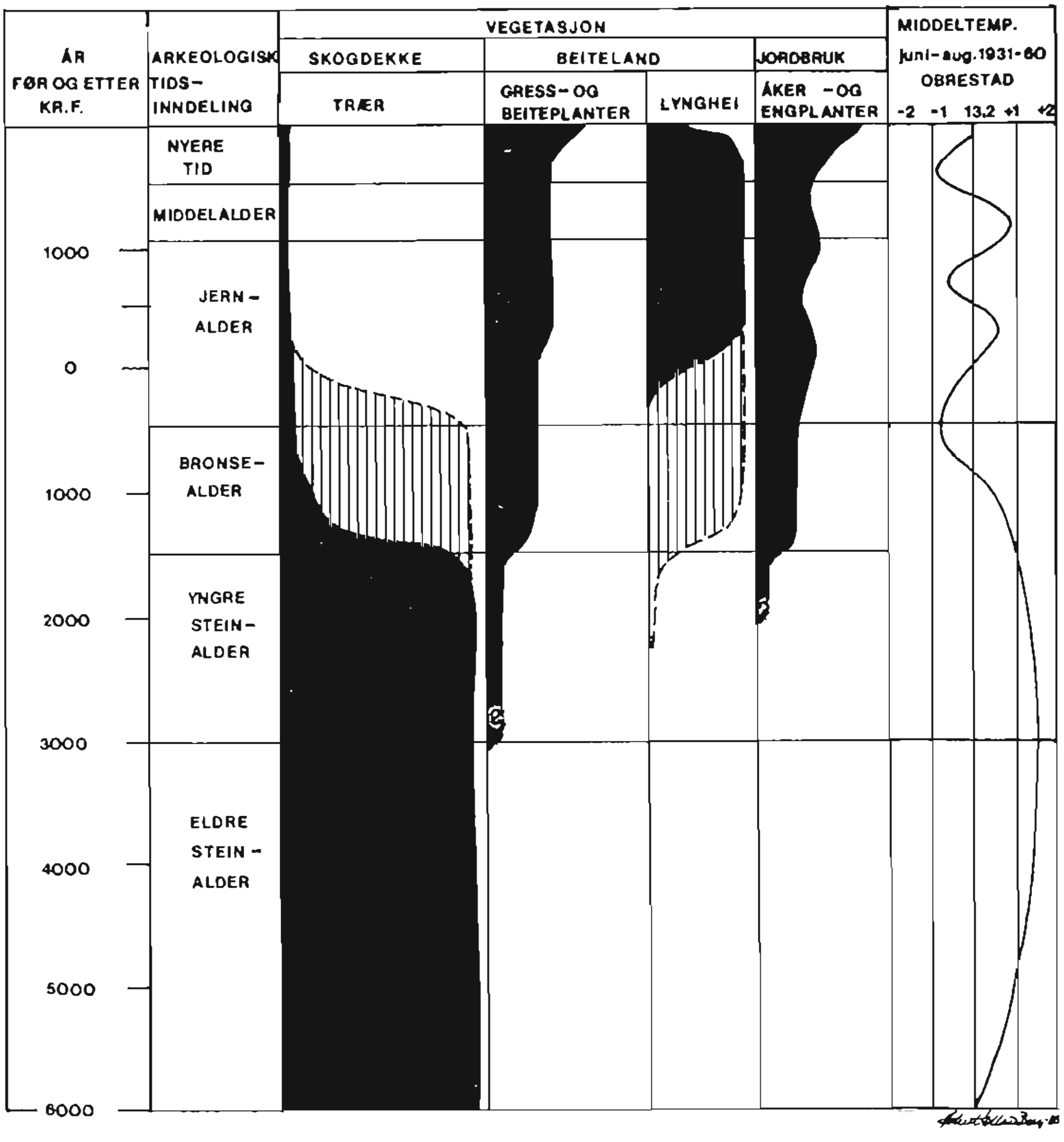


Fig.9 Vegetation composition and summer temperatures relative to present (13.2°C) at Jæren during 8000 years (from Wishman).

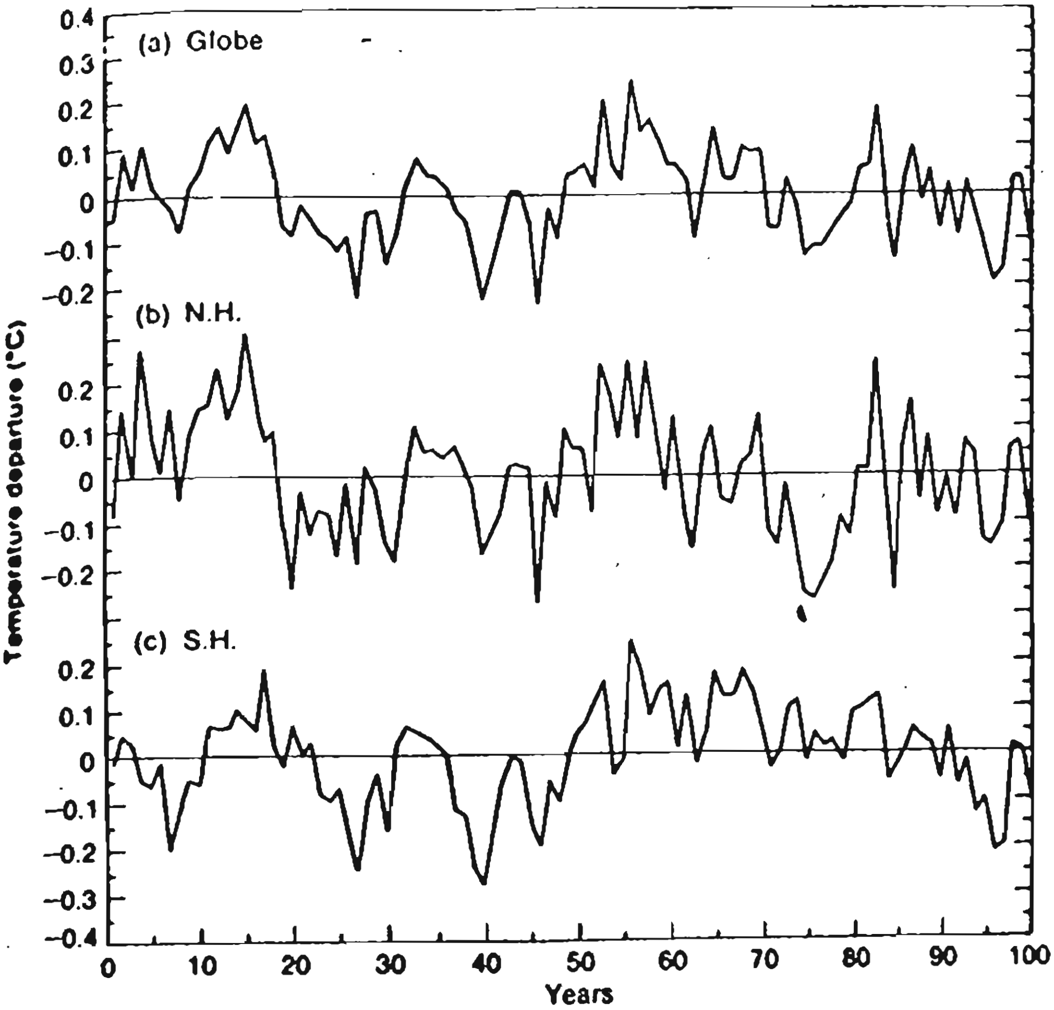


Fig. 10 Temporal variations of the area-averaged deviation of annual mean surface air temperature (°C) from the corresponding 100-year average of the control as simulated by the GFDL coupled ocean-atmosphere model for: (a) the globe, (b) the Northern Hemisphere, and (c) the Southern Hemisphere (from IPCC 1992).

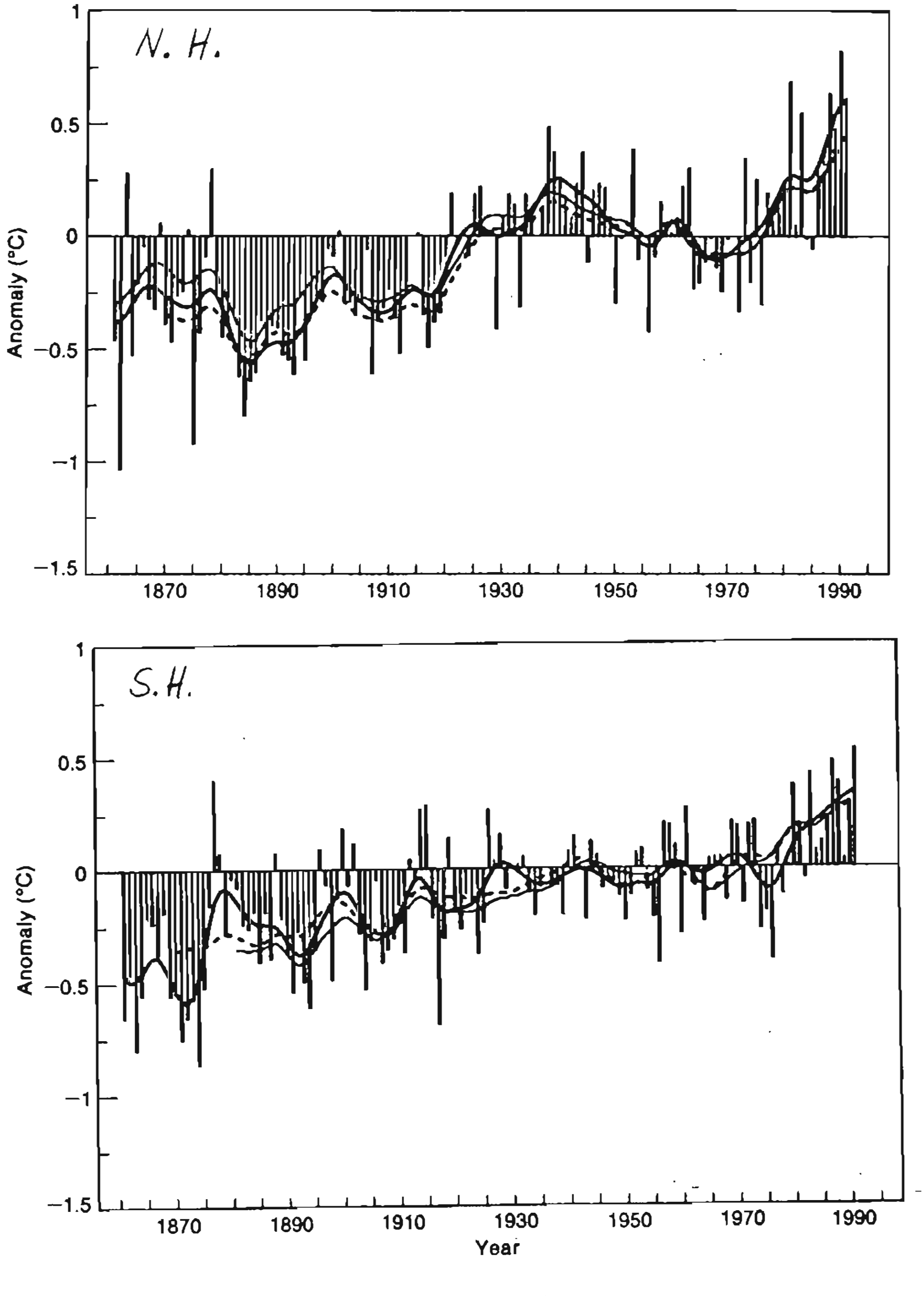
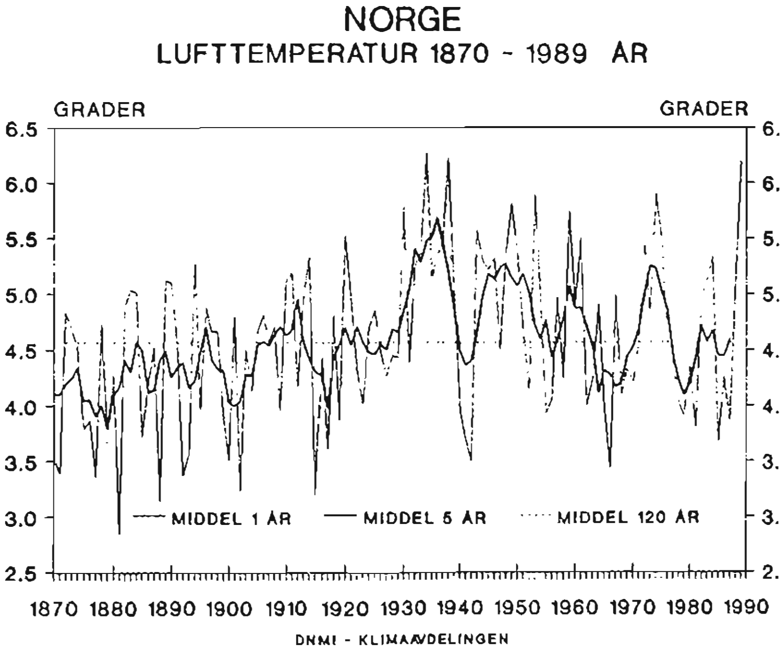


Fig.11 Land air temperature anomalies, relative to 1951-1980. (a) Northern Hemisphere; (b) Southern Hemisphere (from IPCC 1992).



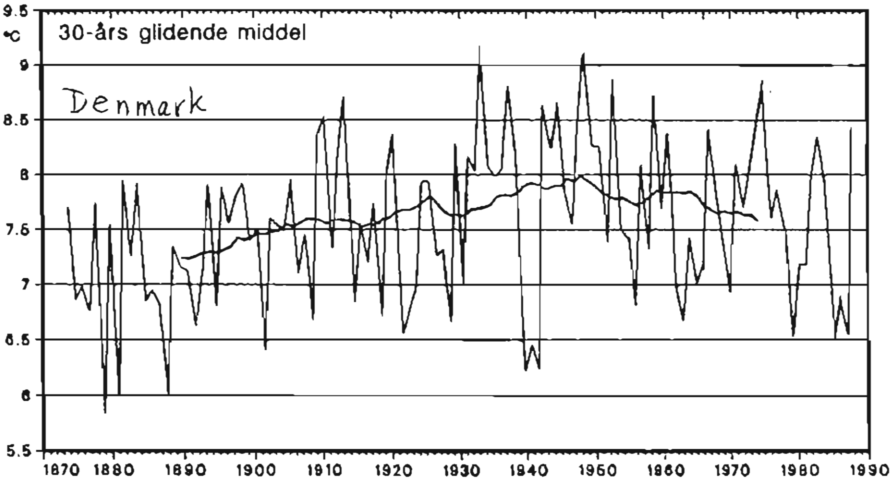


Fig.12 From Hanssen-Bauer (1990):
Norway: Moving averages of annual temperatures at 9 stations. Denmark: Annual temperatures and moving 30 year averages.

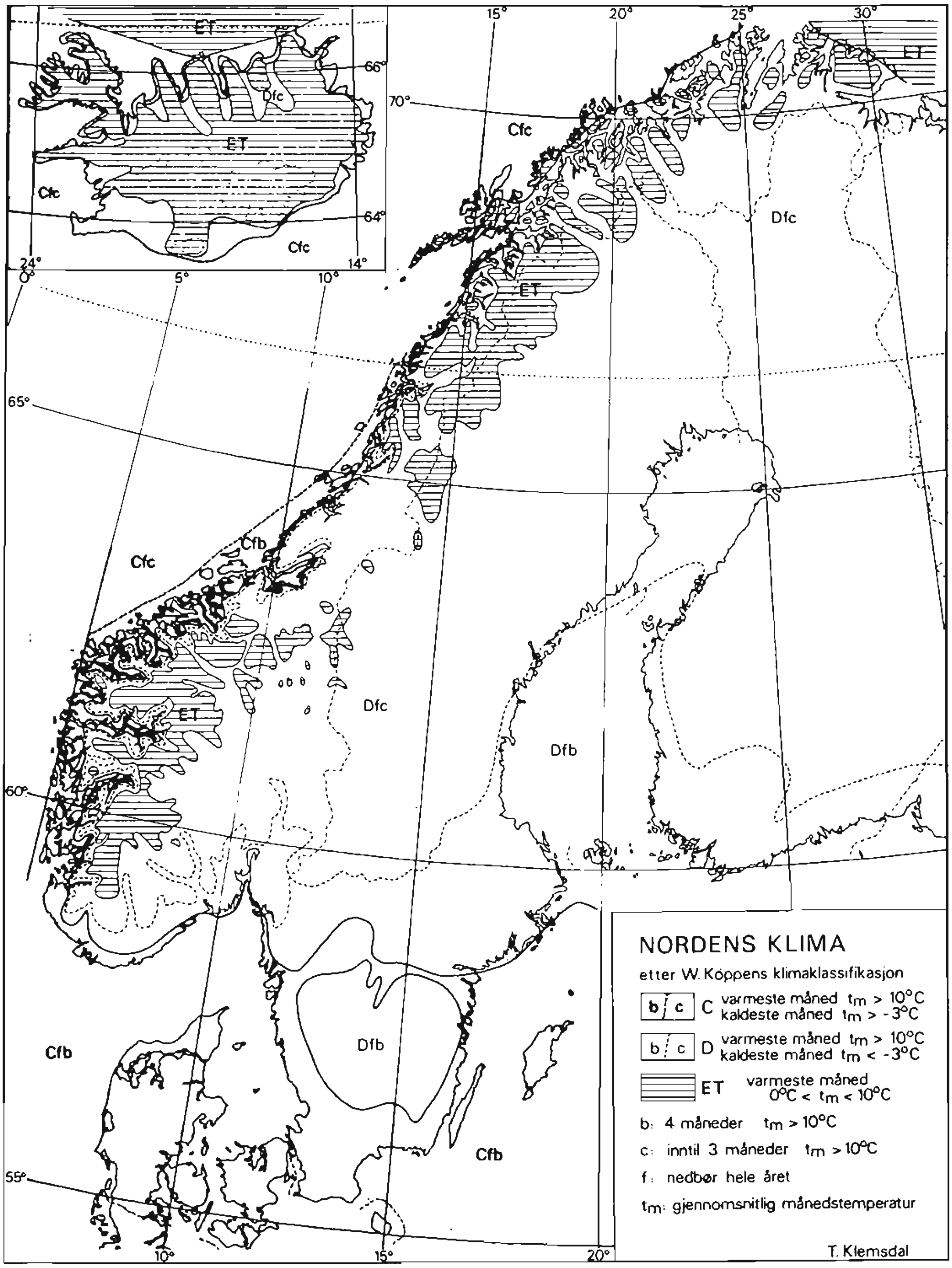


Fig. 13 The climate of the Nordic countries classified according to Köppens scheme (from Klemsdal).

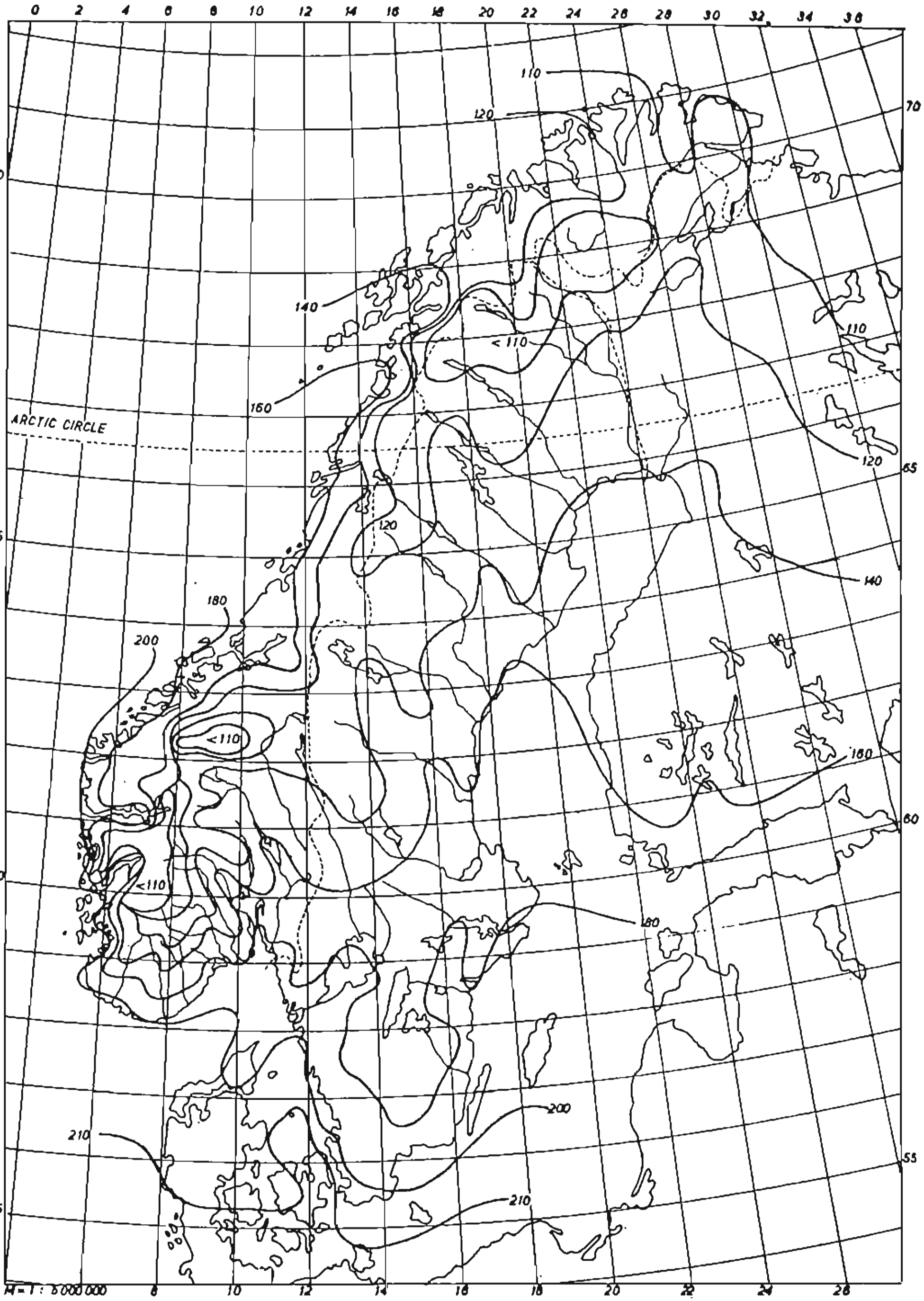


Fig. 14 From Johannessen (1970): The normal length (1931-1960) in days of the vegetation period (daily mean air temperature ≥ 6°C).

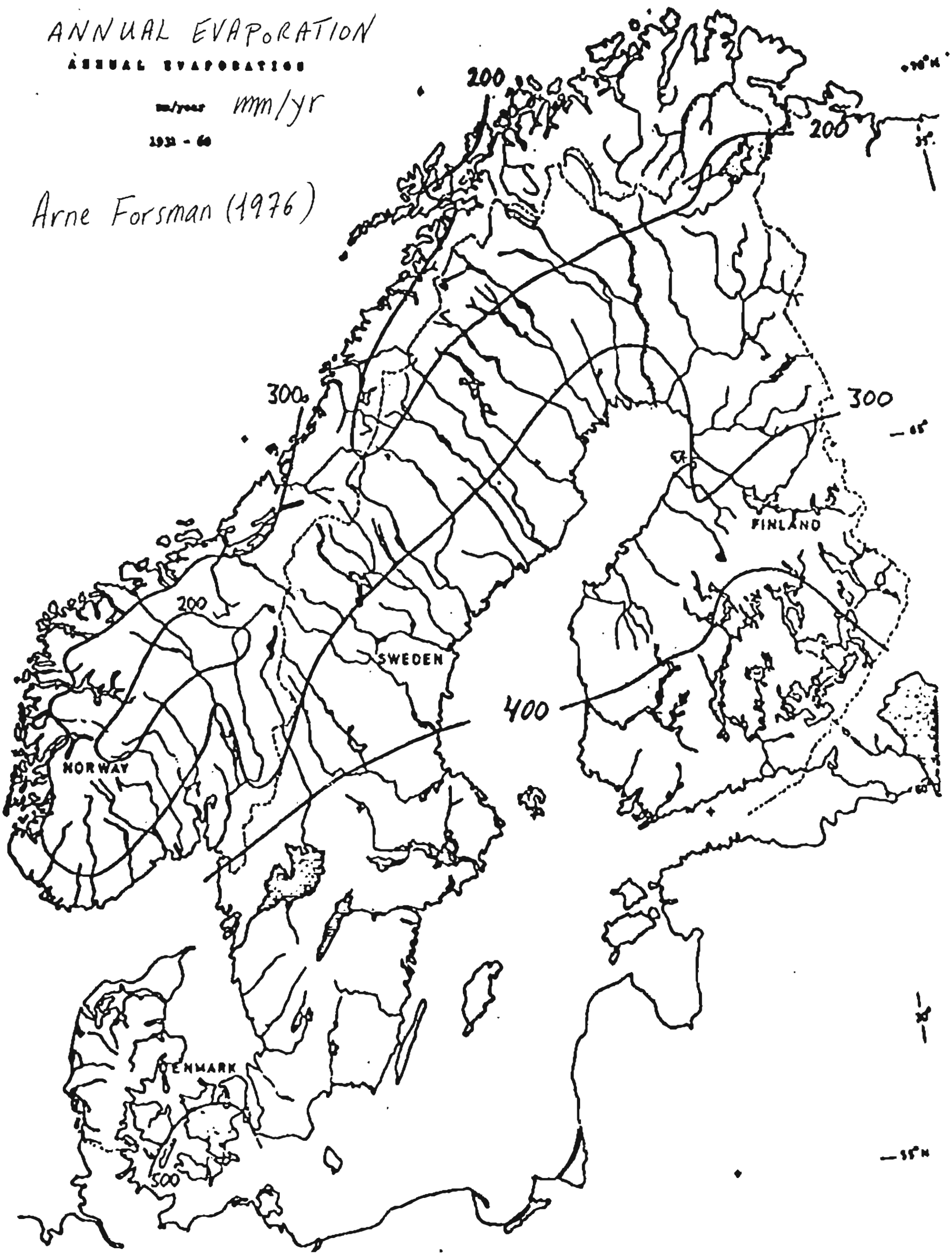


Fig. 15 From Forsman (1976): Annual evapotranspiration (mm).

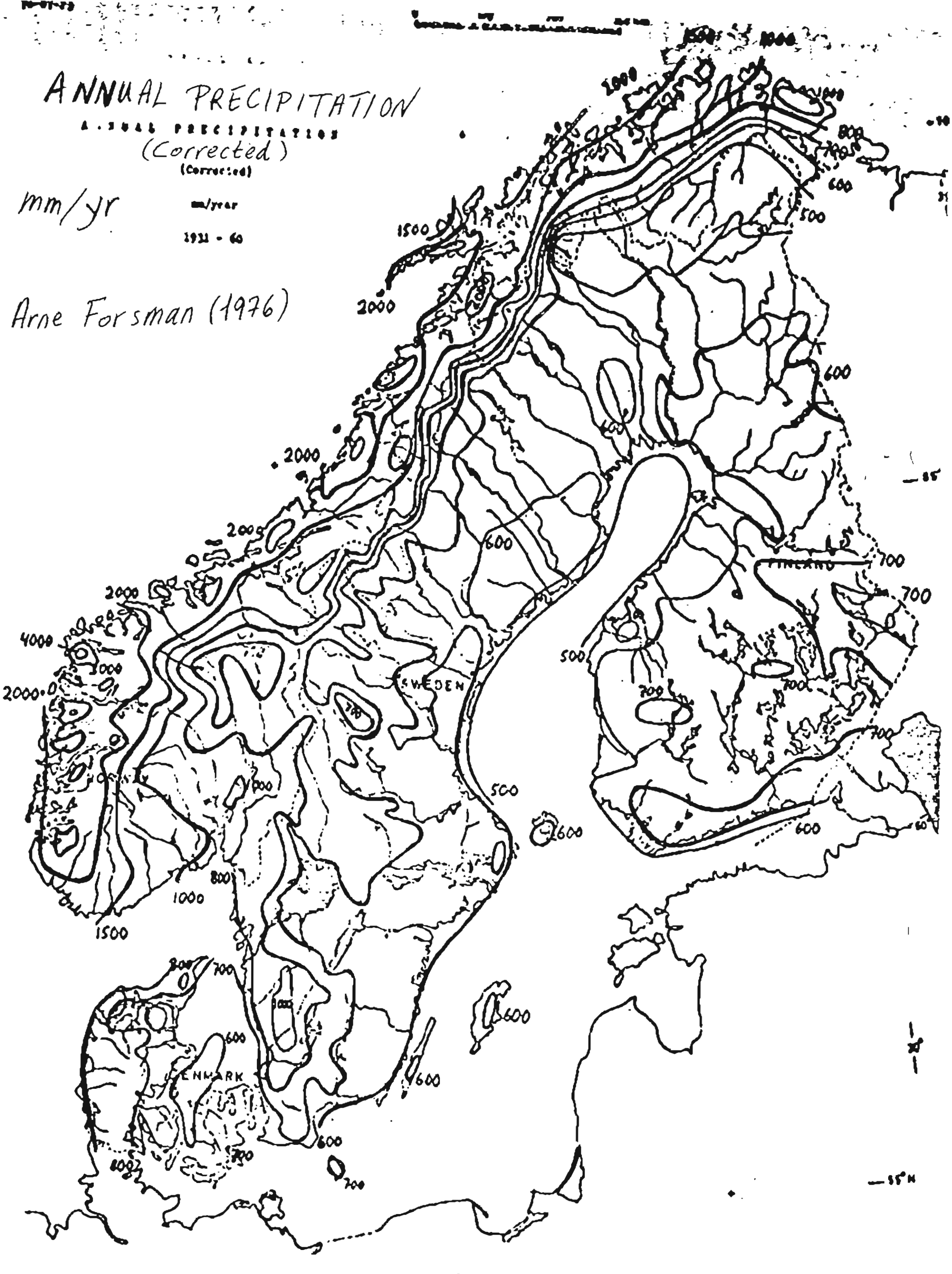


Fig.16 Annual precipitation (corrected, mmyr⁻¹) 1931-60 (from Forsman 1976).

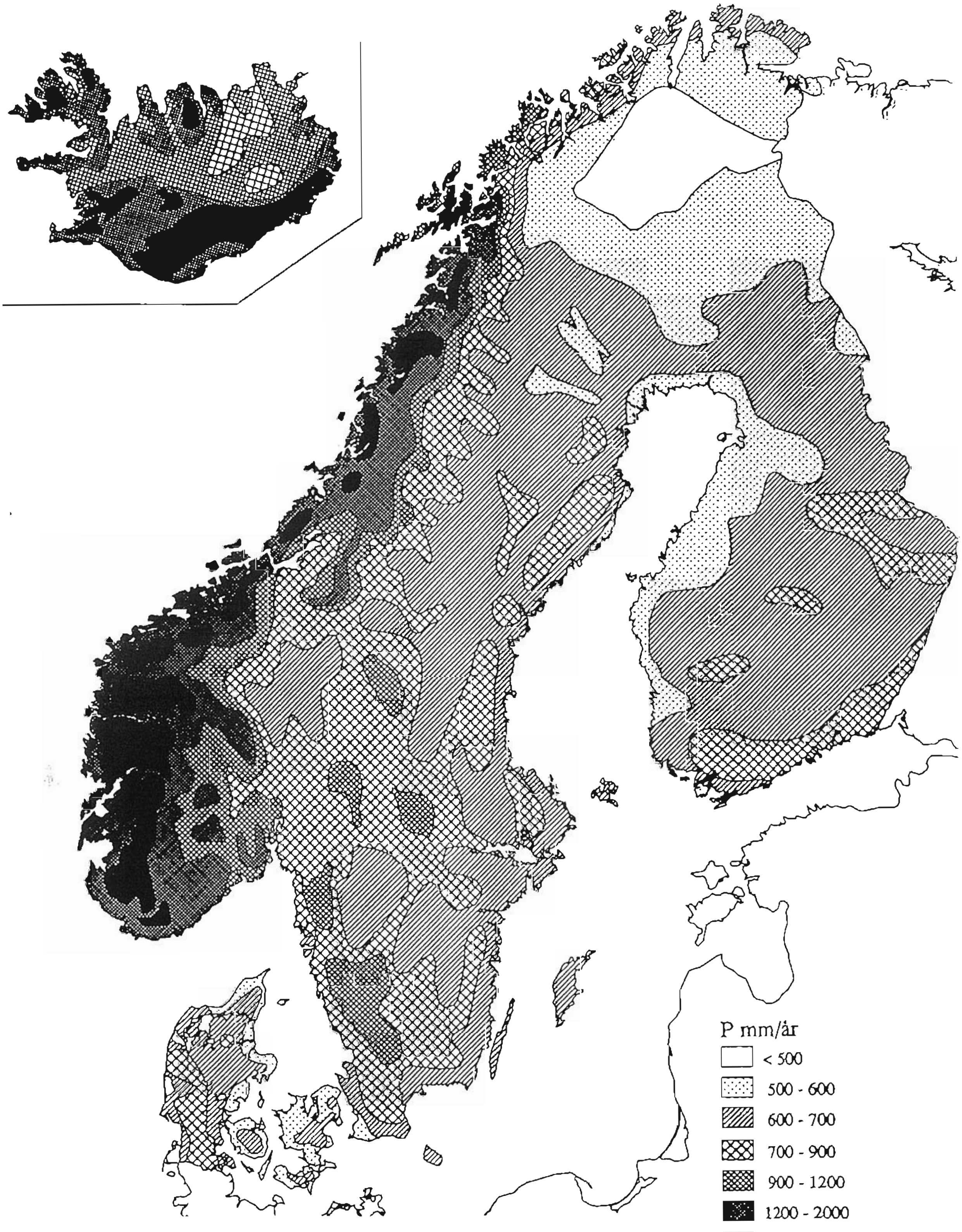


Fig. 17 Annual precipitation (mmyr^{-1}) (Lemmelä & Hyyverinen).

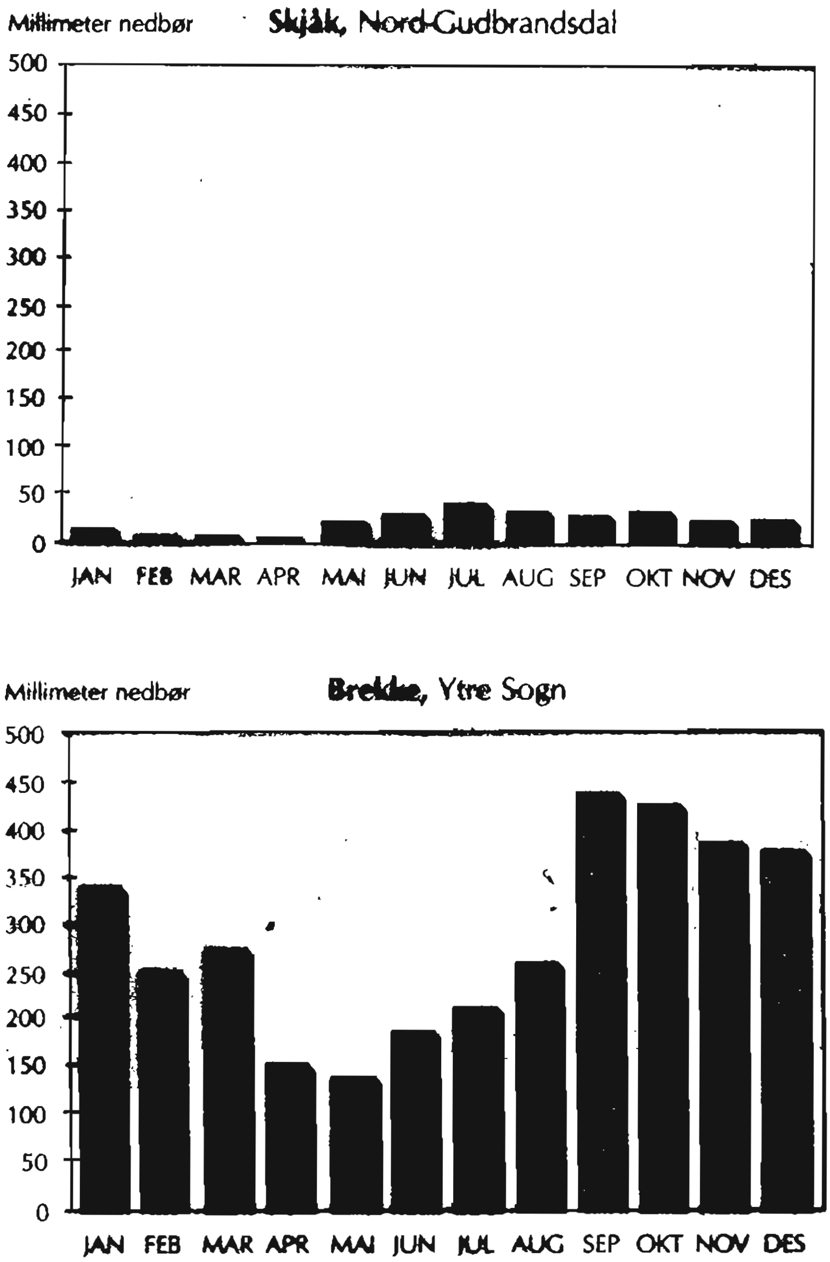
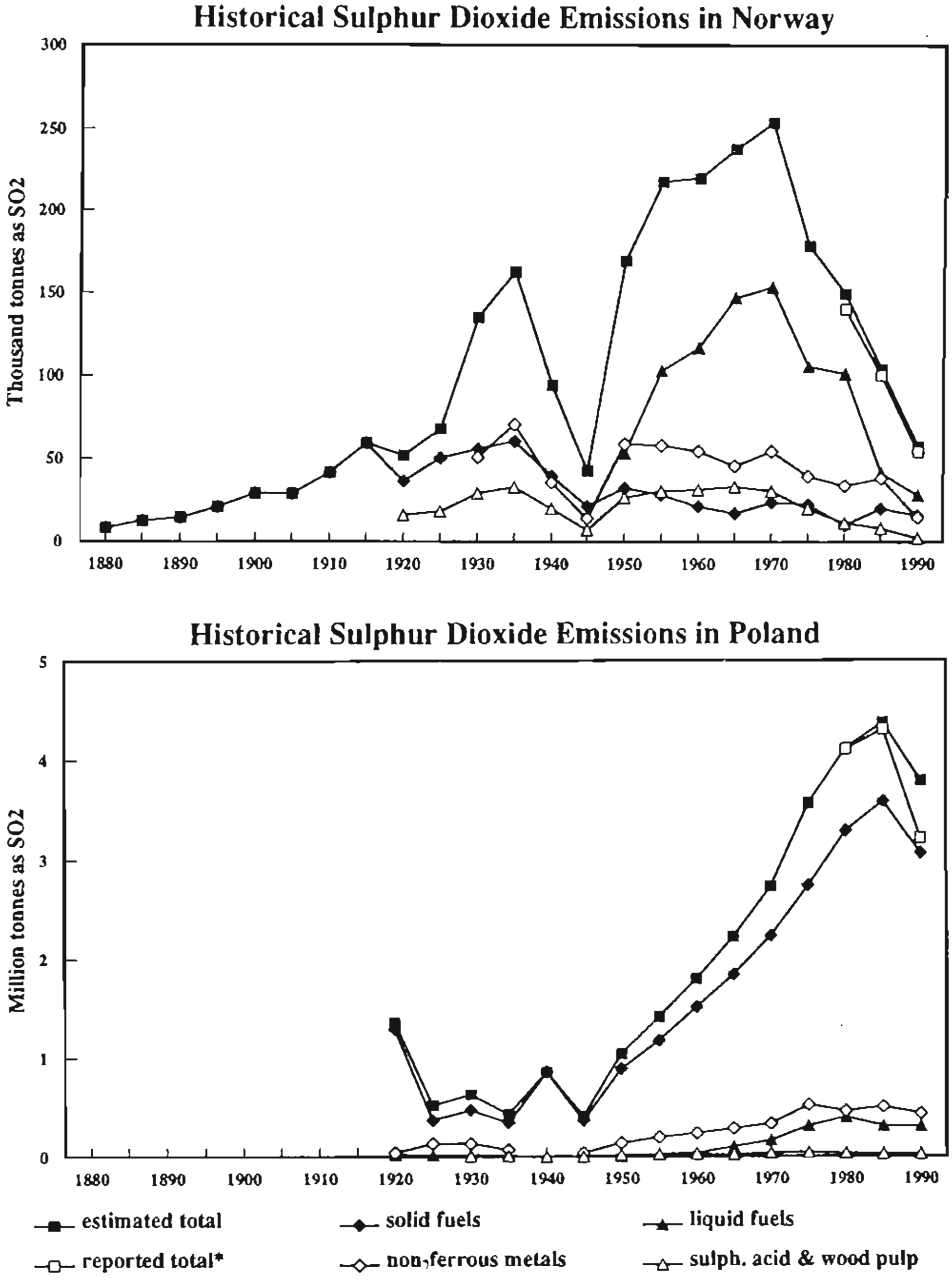


Fig.18 Monthly average precipitation at Skjåk in northern Gudbrand Valley and at Brekke (Sogn) close to the western Norwegian coast (from Aune 1993).



* Emissions submitted officially to UNECE/EMEP.

Fig.19 Historical Sulphur Dioxide Emissions in Norway and in Poland (from Mylona 1993).

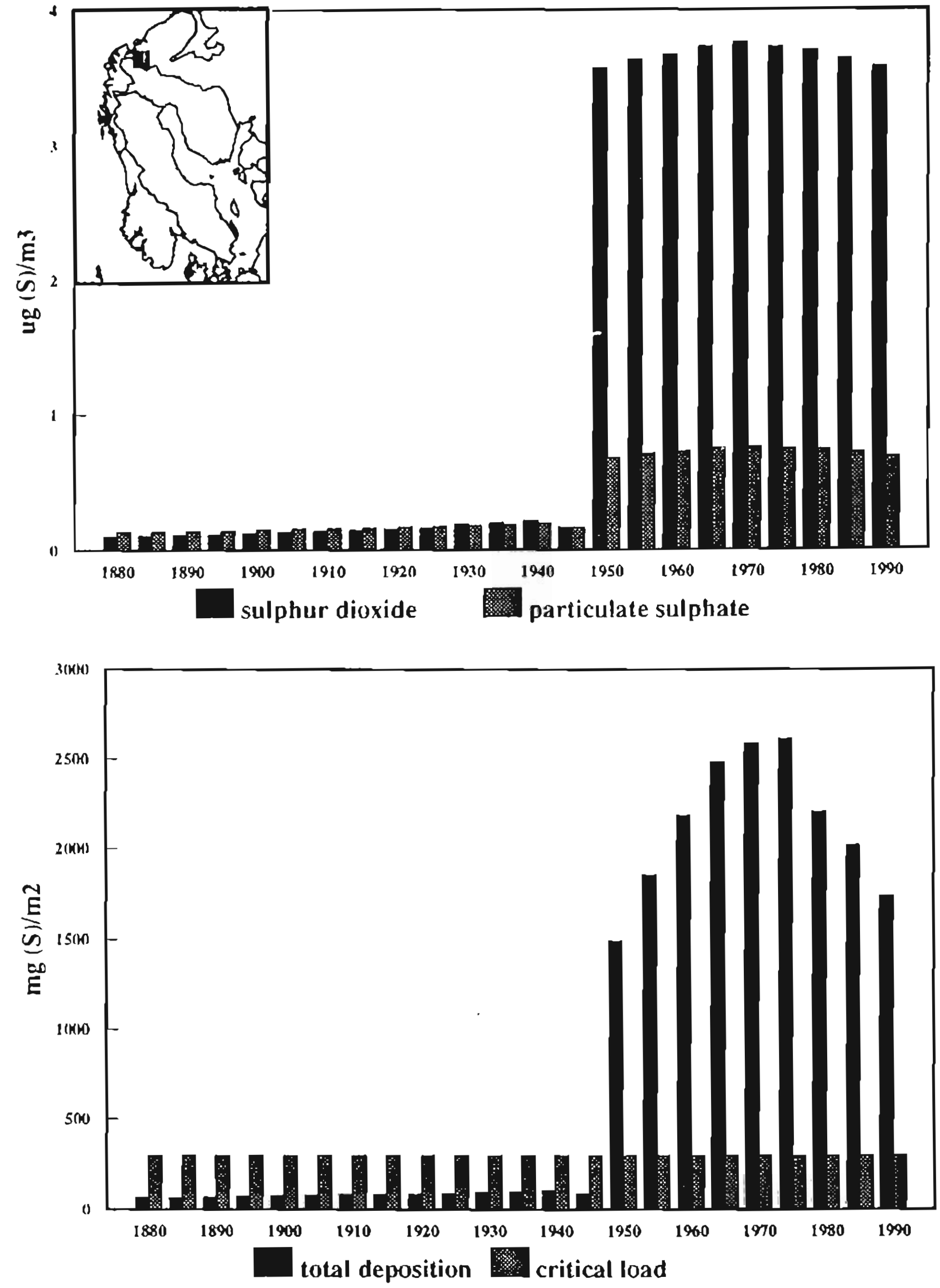


Fig. 20 Calculated historical sulphur concentrations in air and deposition of oxidized sulphur at a grid square in the Kola peninsula (from Mylona 1993).

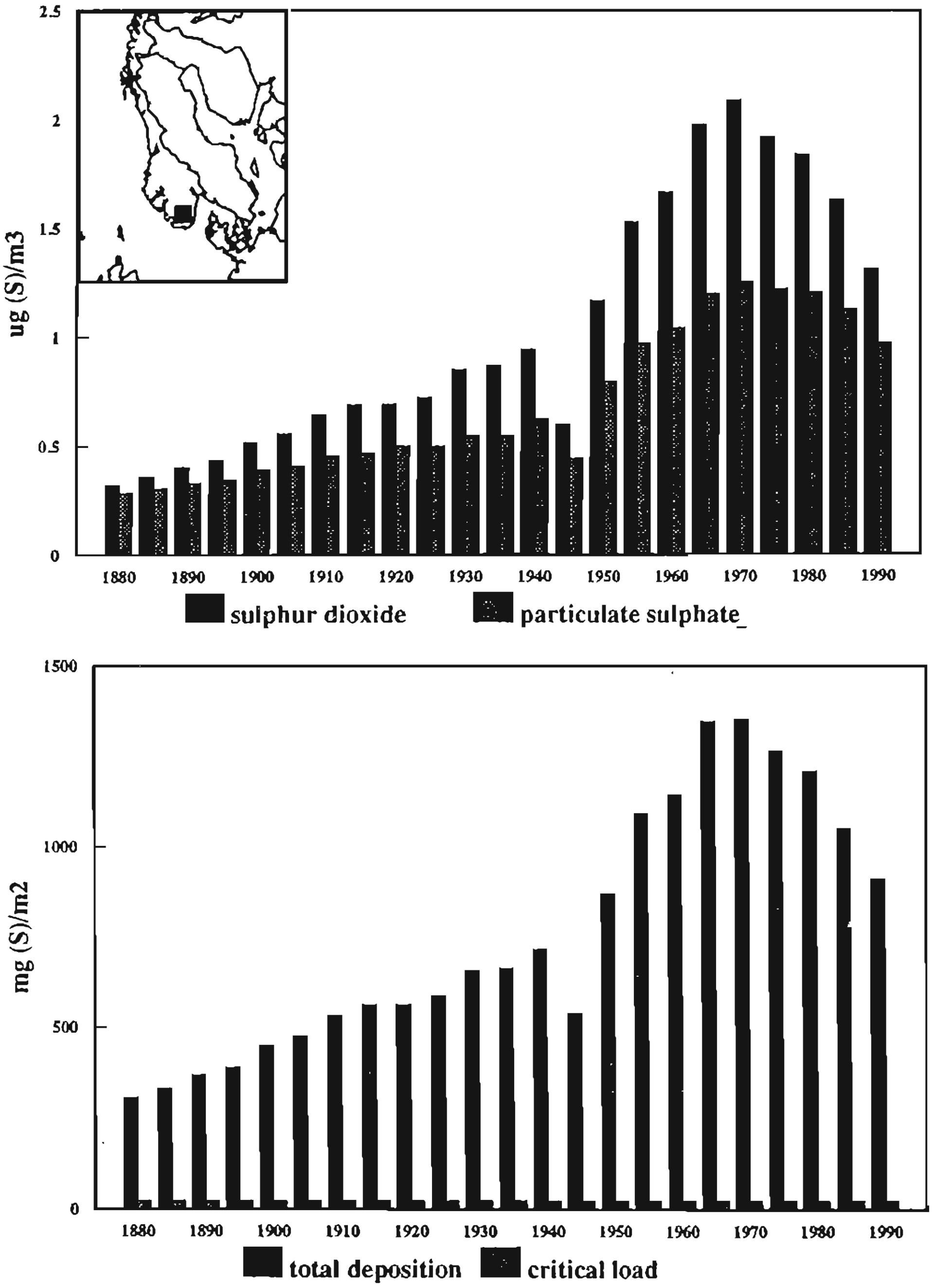


Fig.21 Calculated historical sulphur concentrations in air and deposition of oxidized sulphur at a grid square in southern Norway (from Mylona 1993).

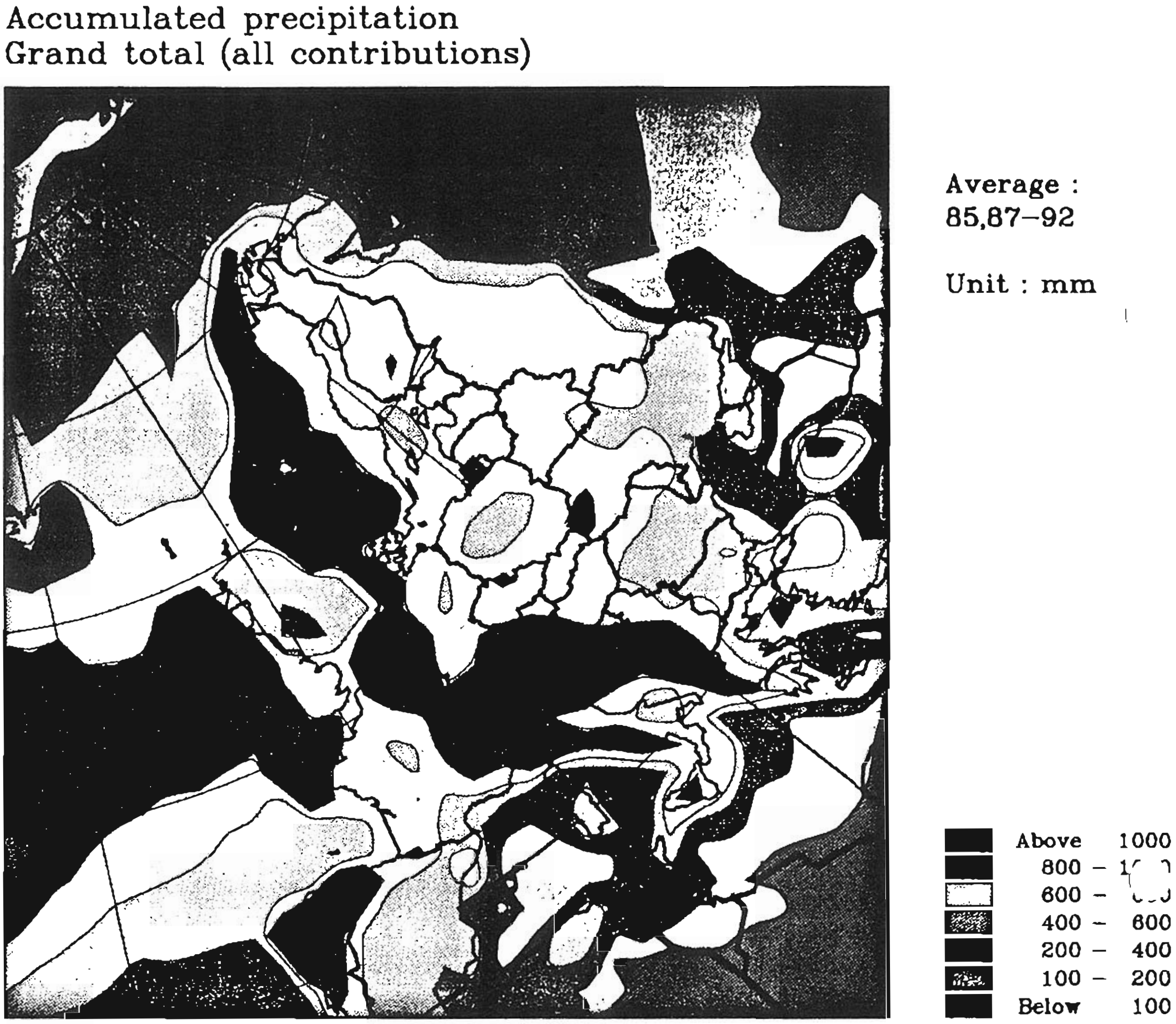


Fig. 22 Calculated (Berge, EMEP/MSC-W) annual precipitation (1985,87-92).

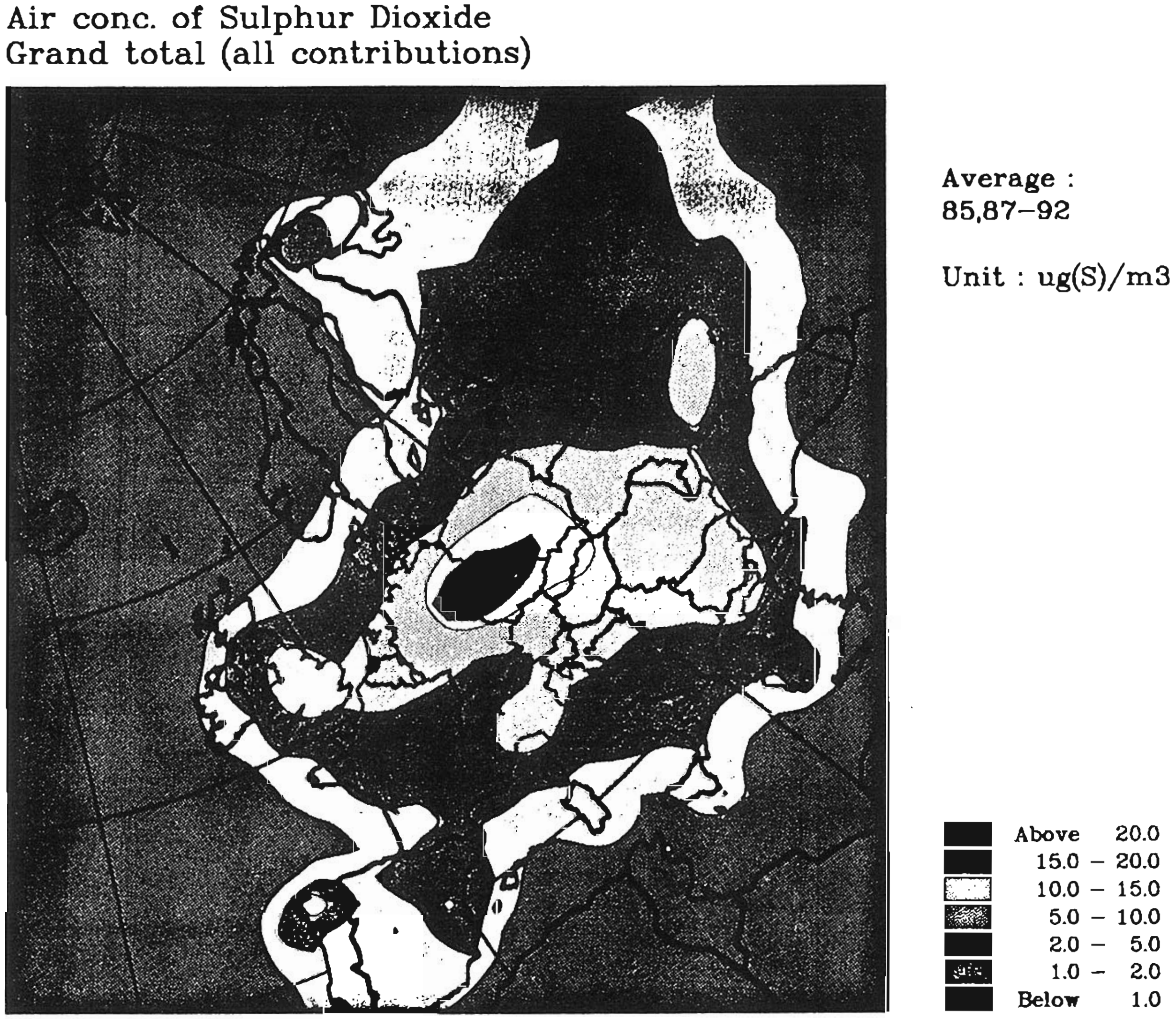


Fig.23 Calculated (Berge, EMEP/MSC-W) air concentration of SO_2 (1985,87-92).

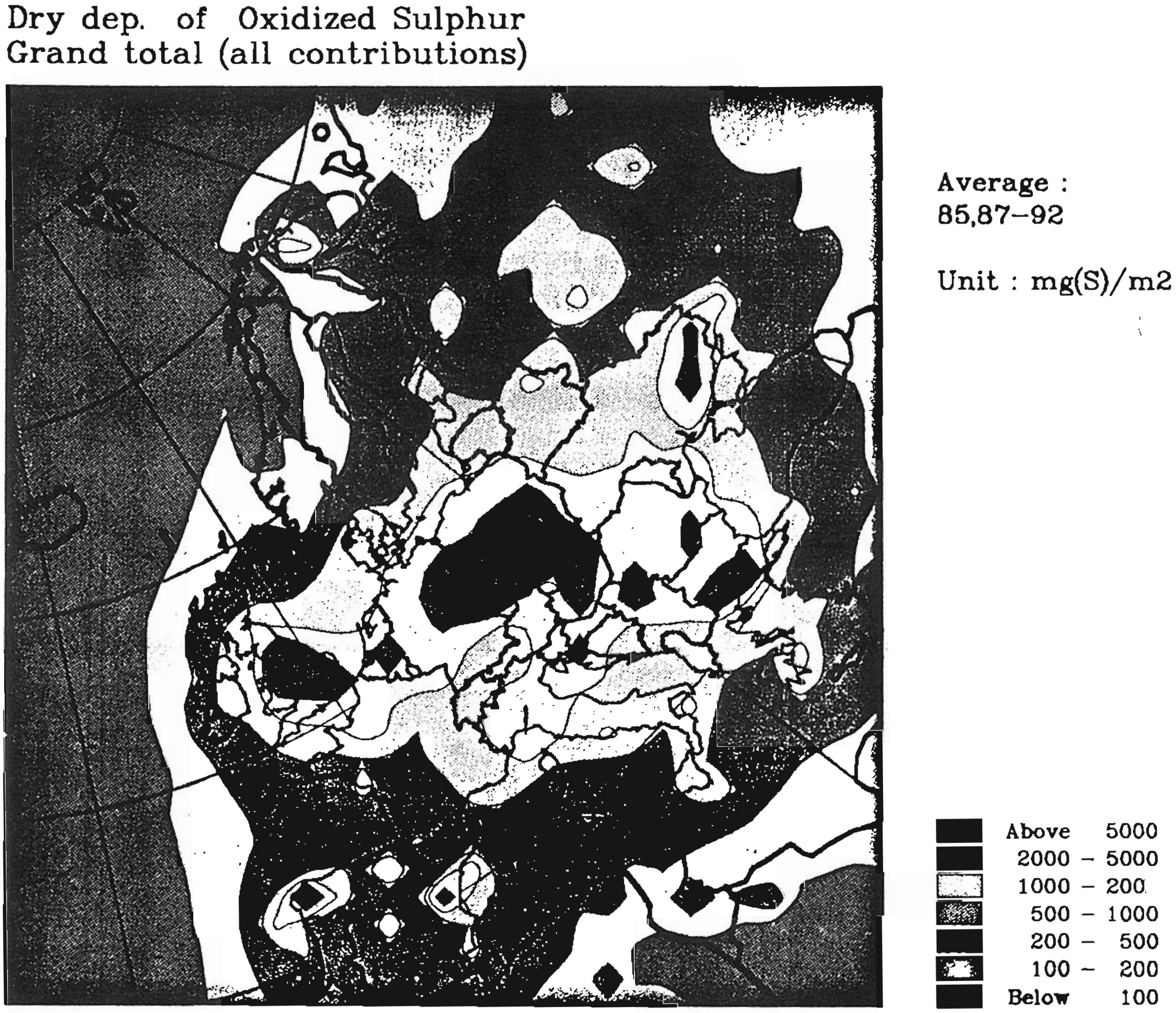


Fig.24 Calculated (Berge, EMEP/MSC-W) dry deposition of oxidized sulphur (1985,87-92).

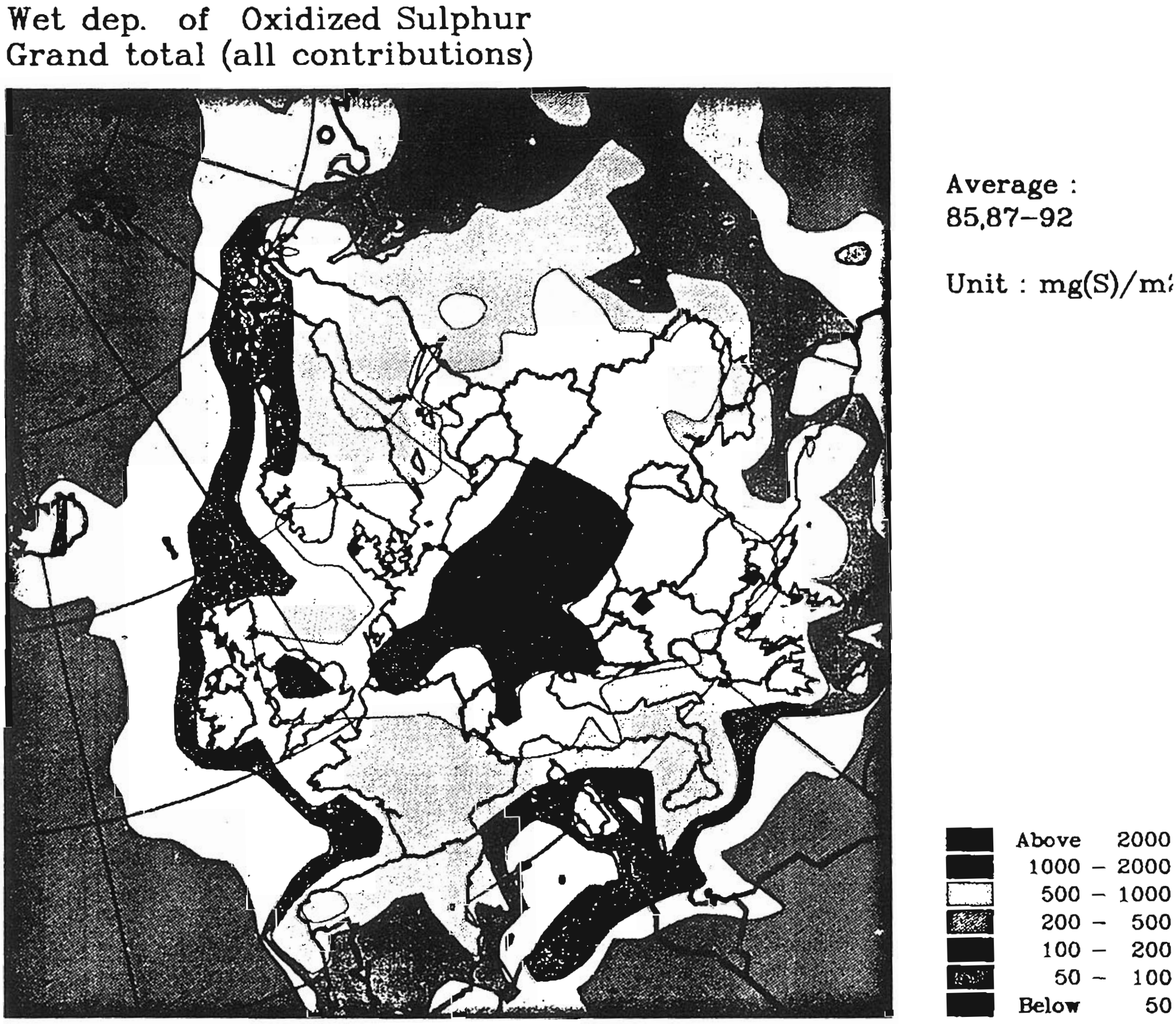


Fig. 25 Calculated (Berge, EMEP/MSC-W) wet deposition of oxidized sulphur (1985, 87-92).

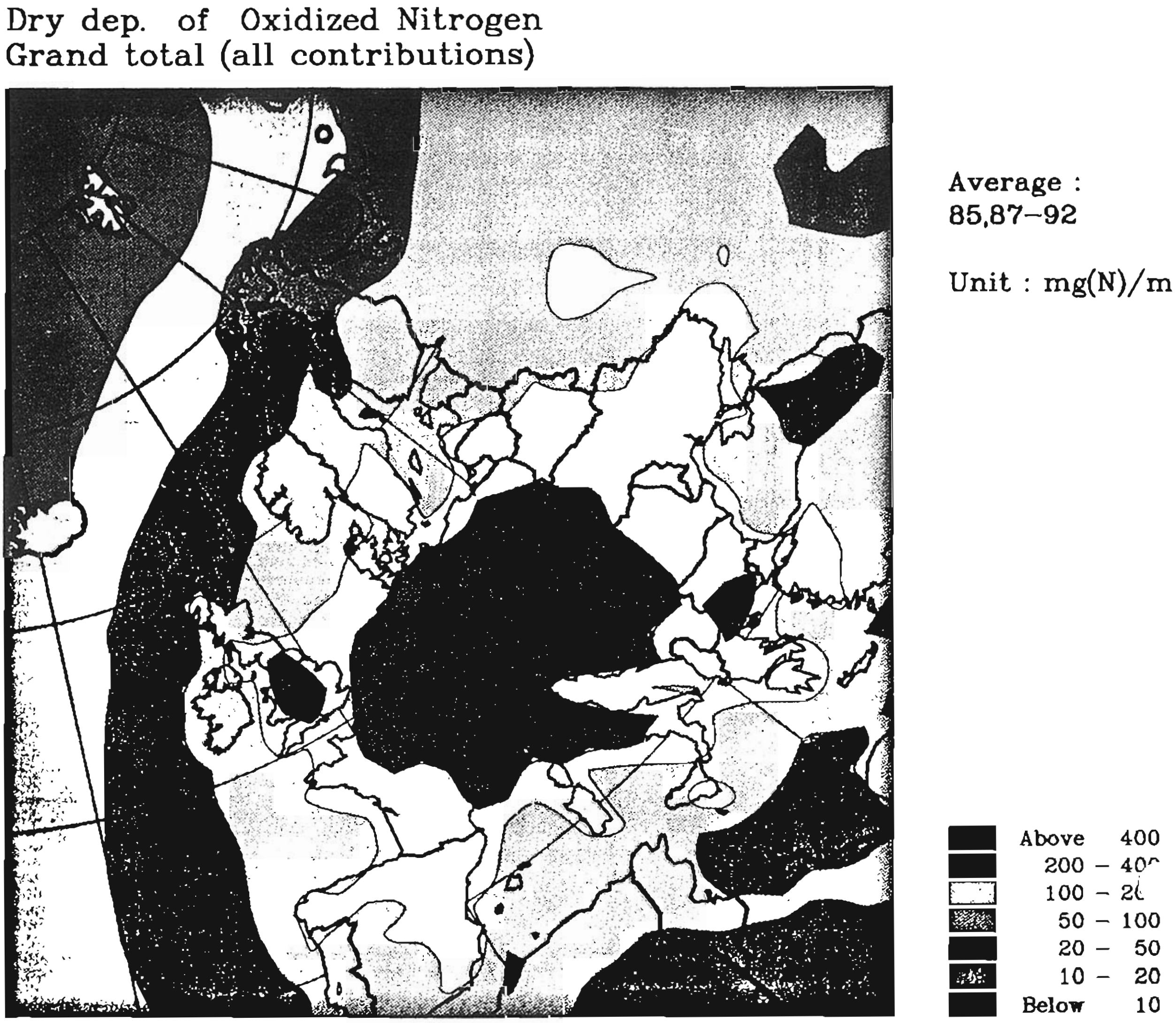


Fig.26 Calculated (Berge, EMEP/MSC-W) dry deposition of oxidized nitrogen (1985,87-92).

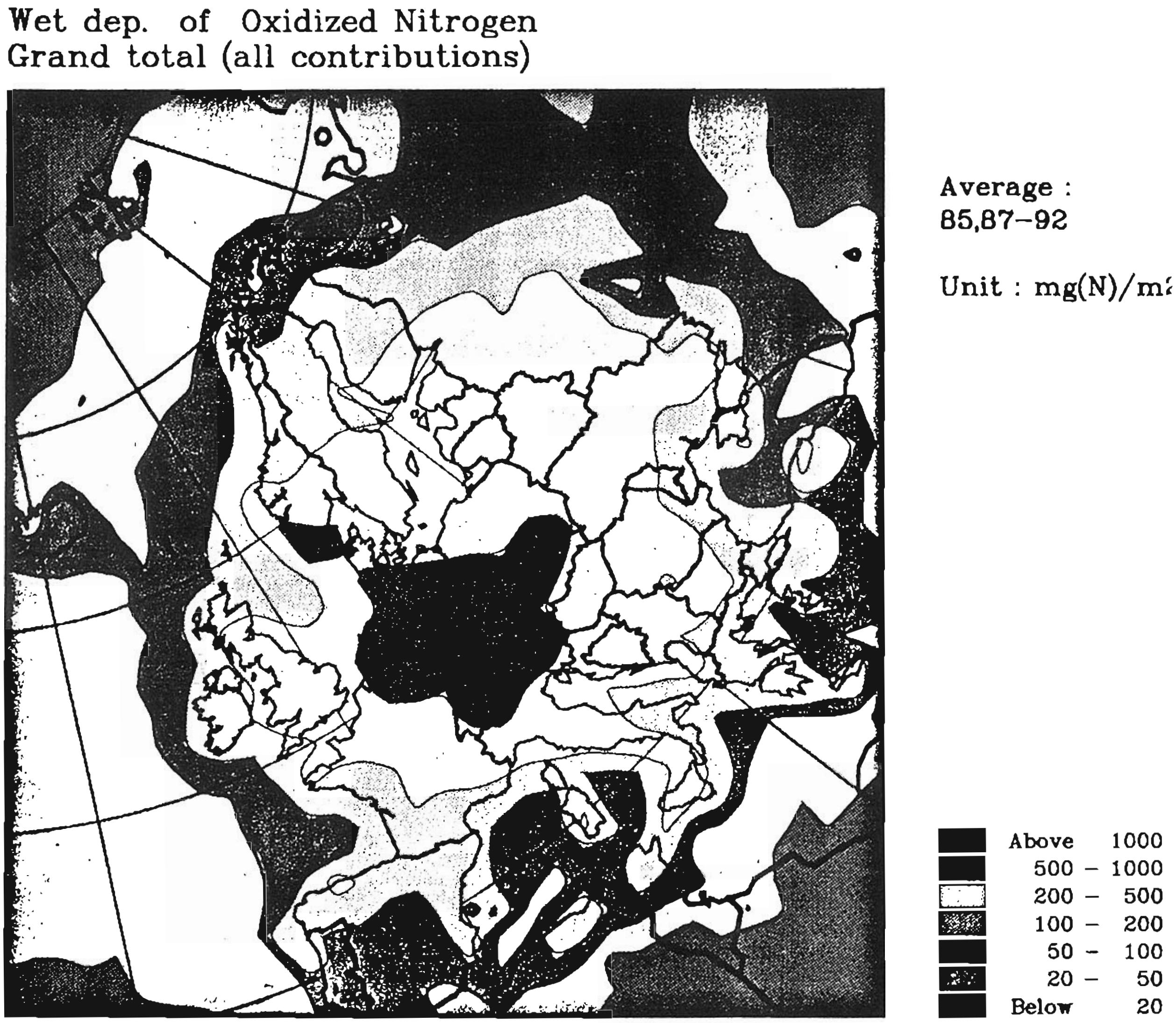


Fig.27 Calculated (Berge, EMEP/MSC-W) wet deposition of oxidized nitrogen (1985,87-92).

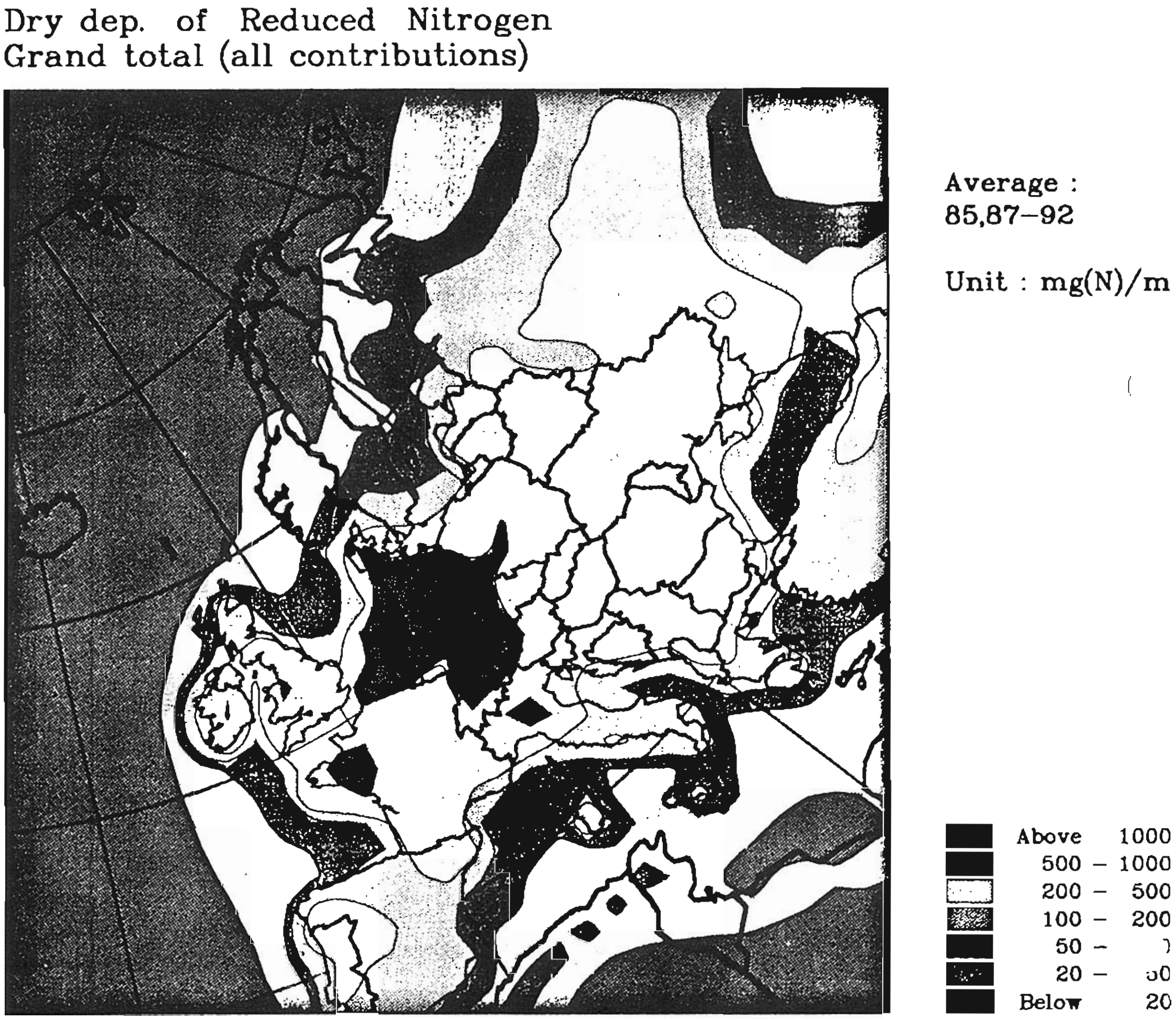


Fig.28 Calculated (Berge, EMEP/MSC-W) dry deposition of reduced nitrogen (1985,87-92).

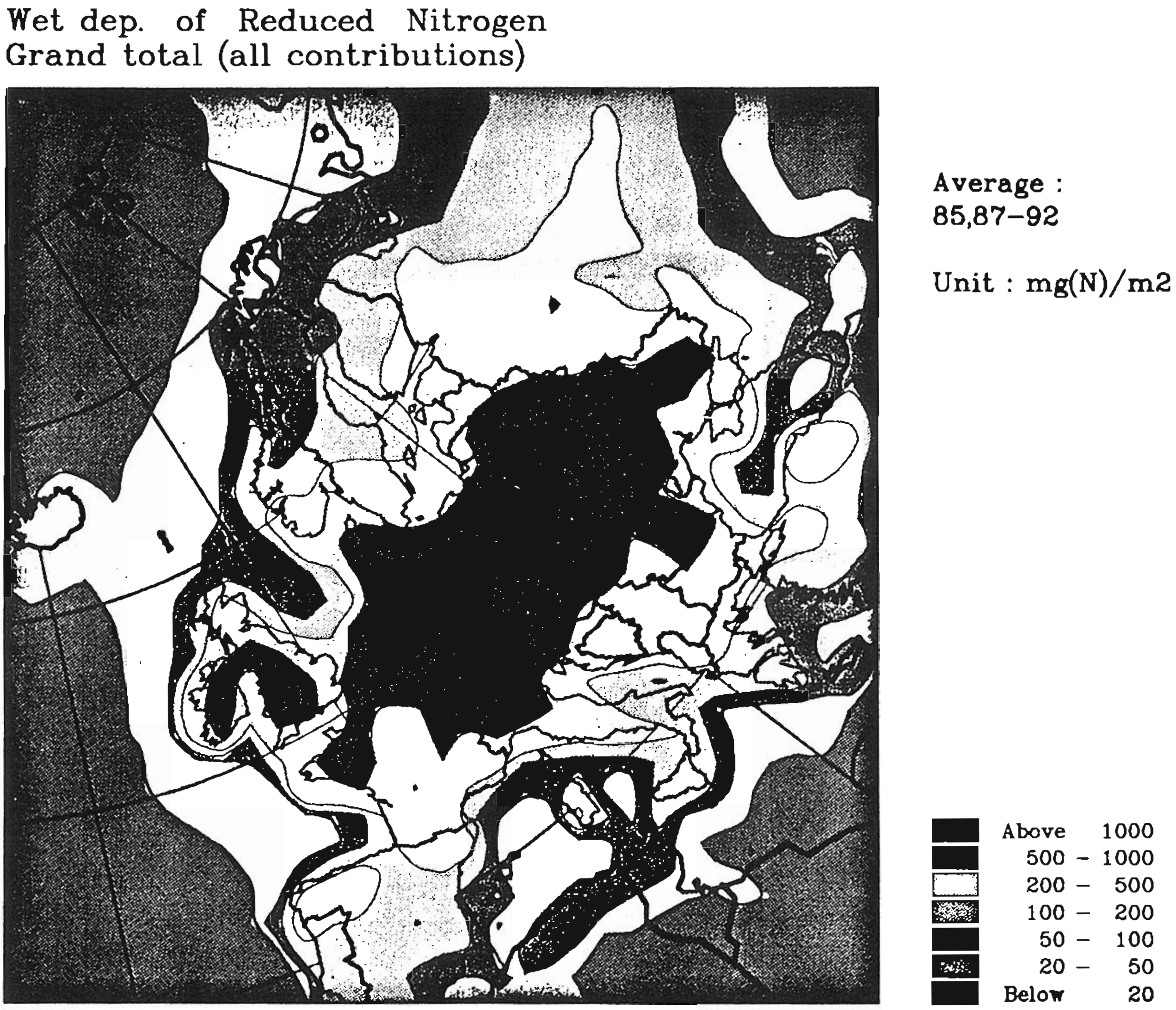


Fig.29 Calculated (Berge, EMEP/MSC-W) wet deposition of reduced nitrogen (1985,87-92).

Annual runoff
mm/yr

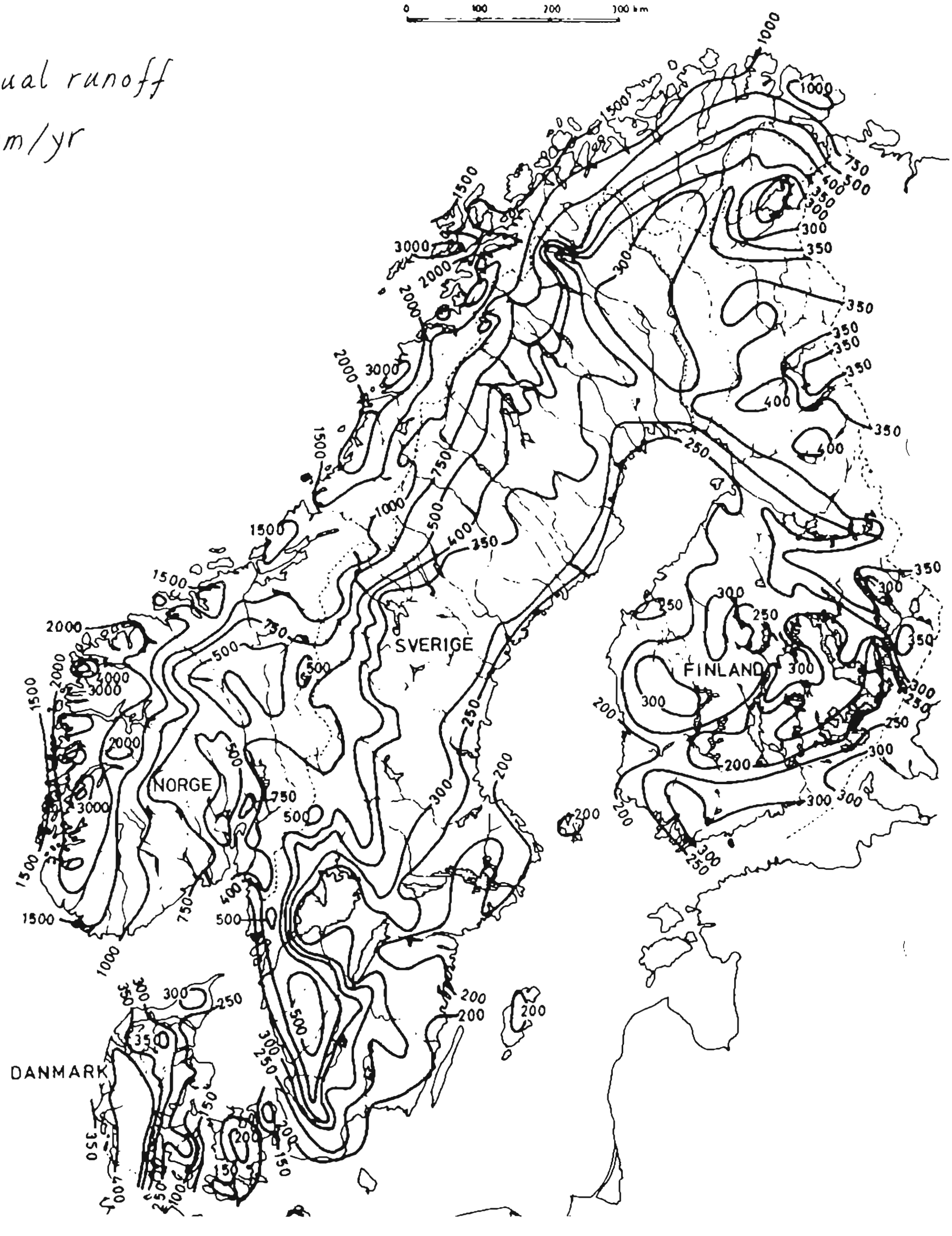


Fig. 30 Map of annual runoff in the Nordic countries (from Gottschalk et al. 1979).

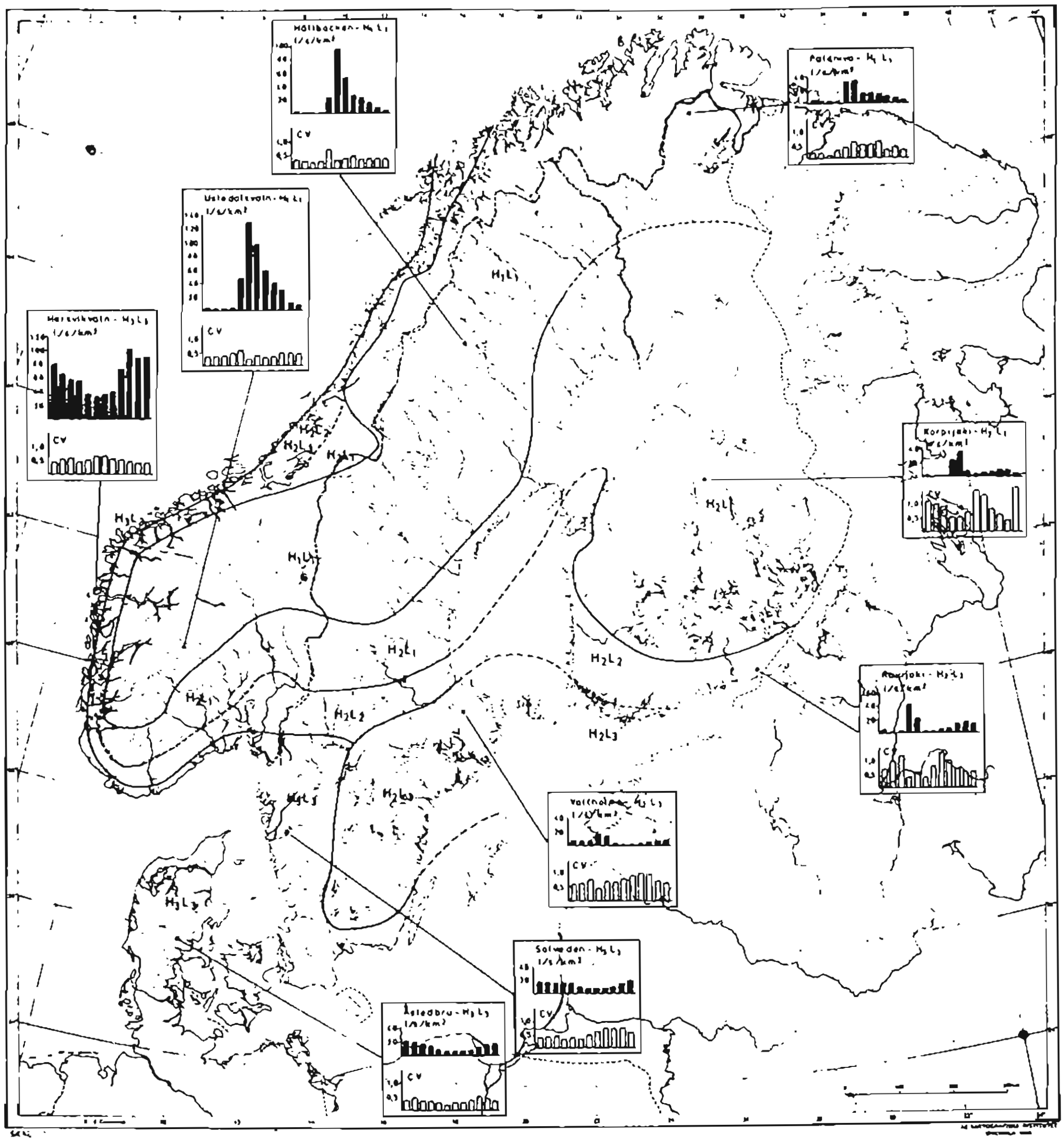


Fig.31 Map of runoff regions in the Nordic countries (from Gottschalk et al. 1979).

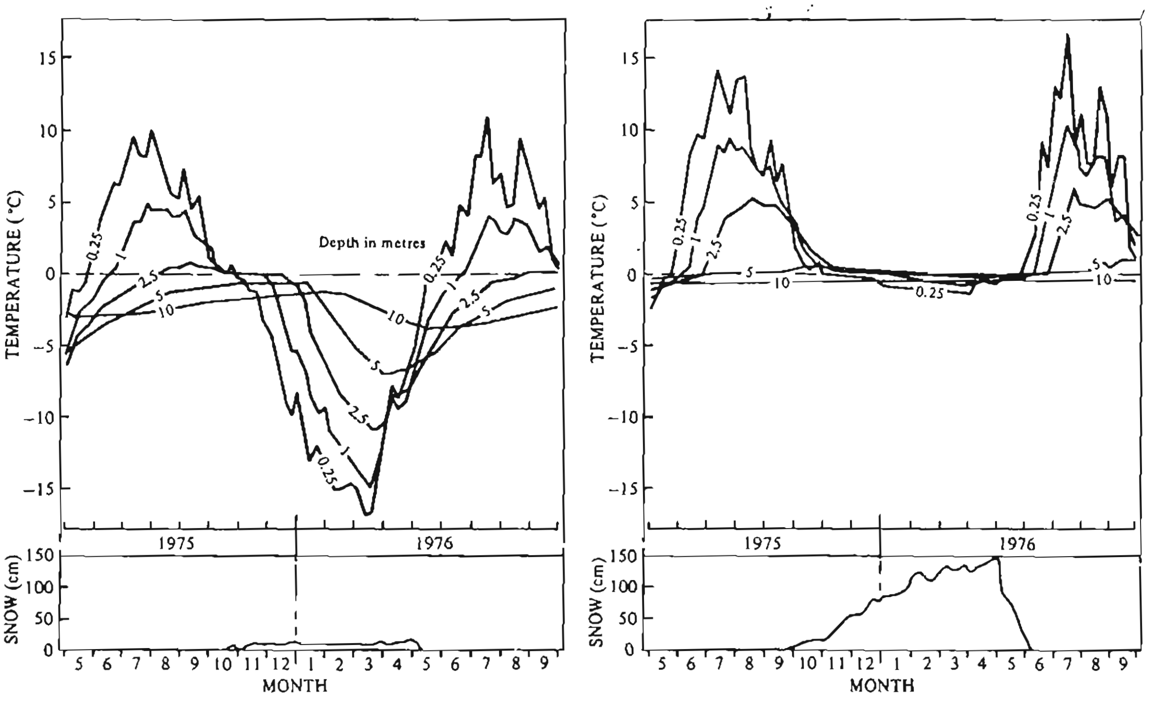


Fig. 32 Annual temperature variations at two sites of contrasting snow cover conditions (from Williams & Smith 1989).

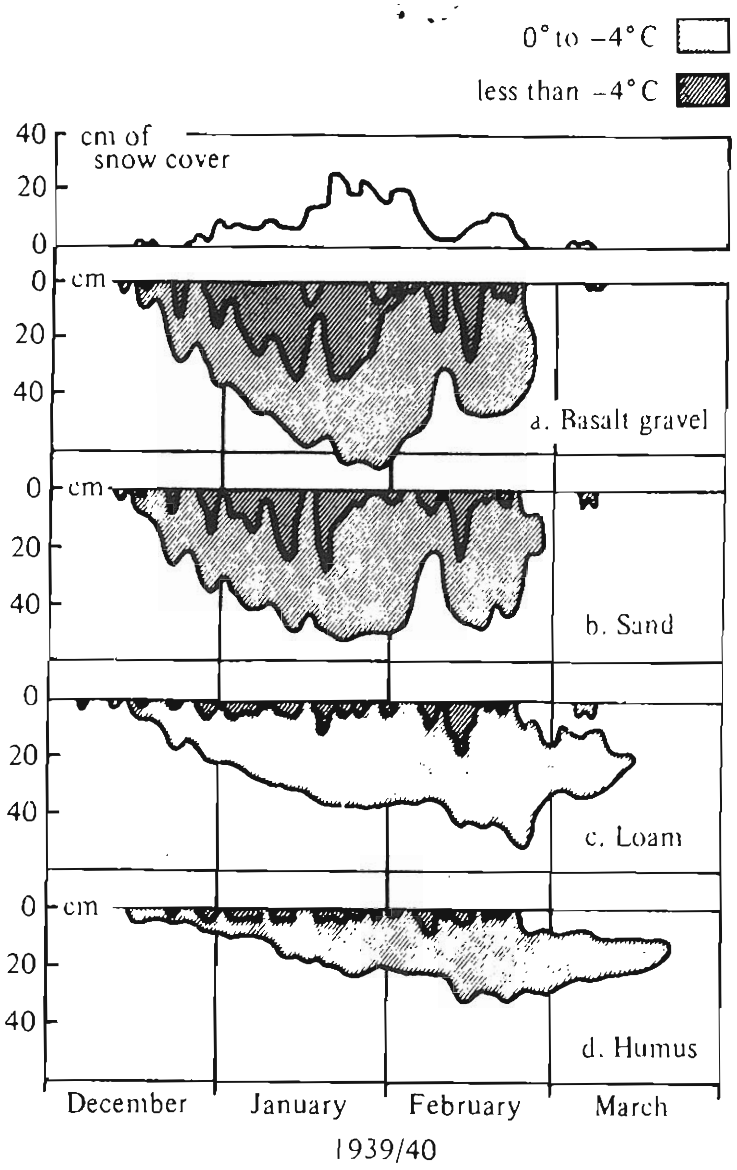


Fig.33 Duration and depth of ground frost in four different soils in winter 1939-40 at Giessen. Climate has similar effects and the situation presented in d) occurs more commonly in maritime climates, in many soils,. while that in a) is characteristic of climates with hot summers and cold winters (from Williams & Smith 1989).

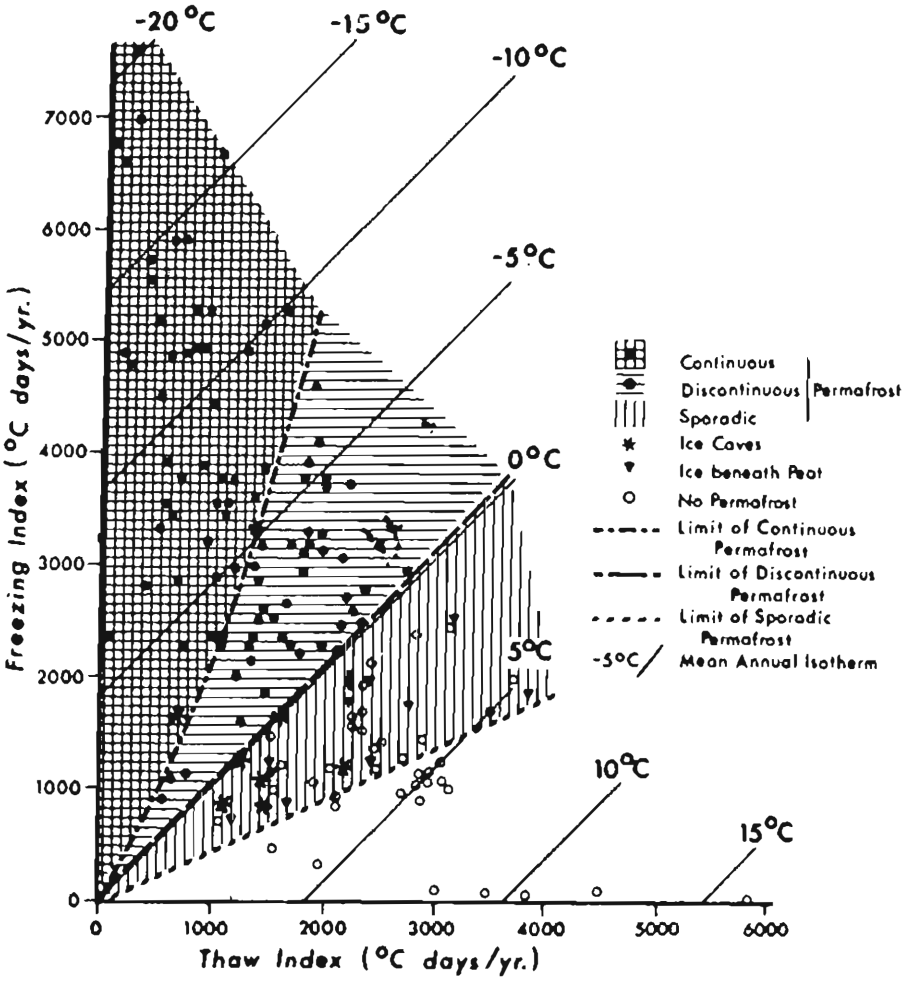


Fig.34 Relationship between occurrence of permafrost zone and the plot of freezing index versus thawing index for stations (less than 0.5 m snow depth) in Canada, Norway, Spitsbergen, Iceland and the Mongolian People's Republic (from Harris 1986).

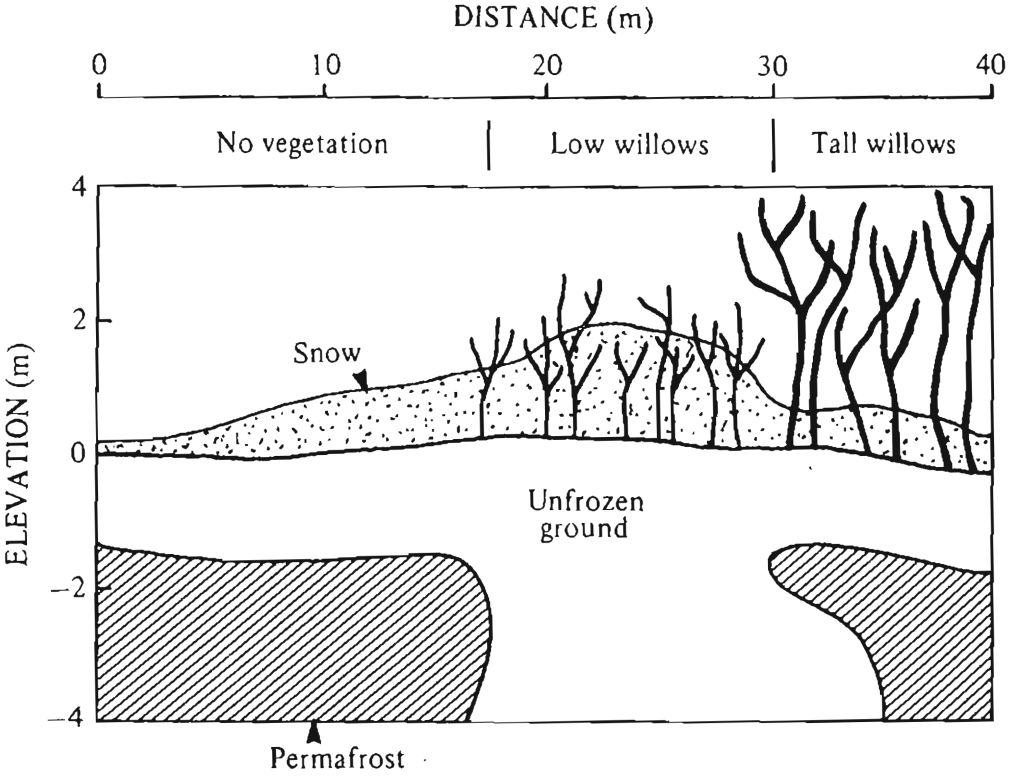


Fig.35 Details of permafrost configuration under a perennial snowbank (from Williams & Smith 1989).

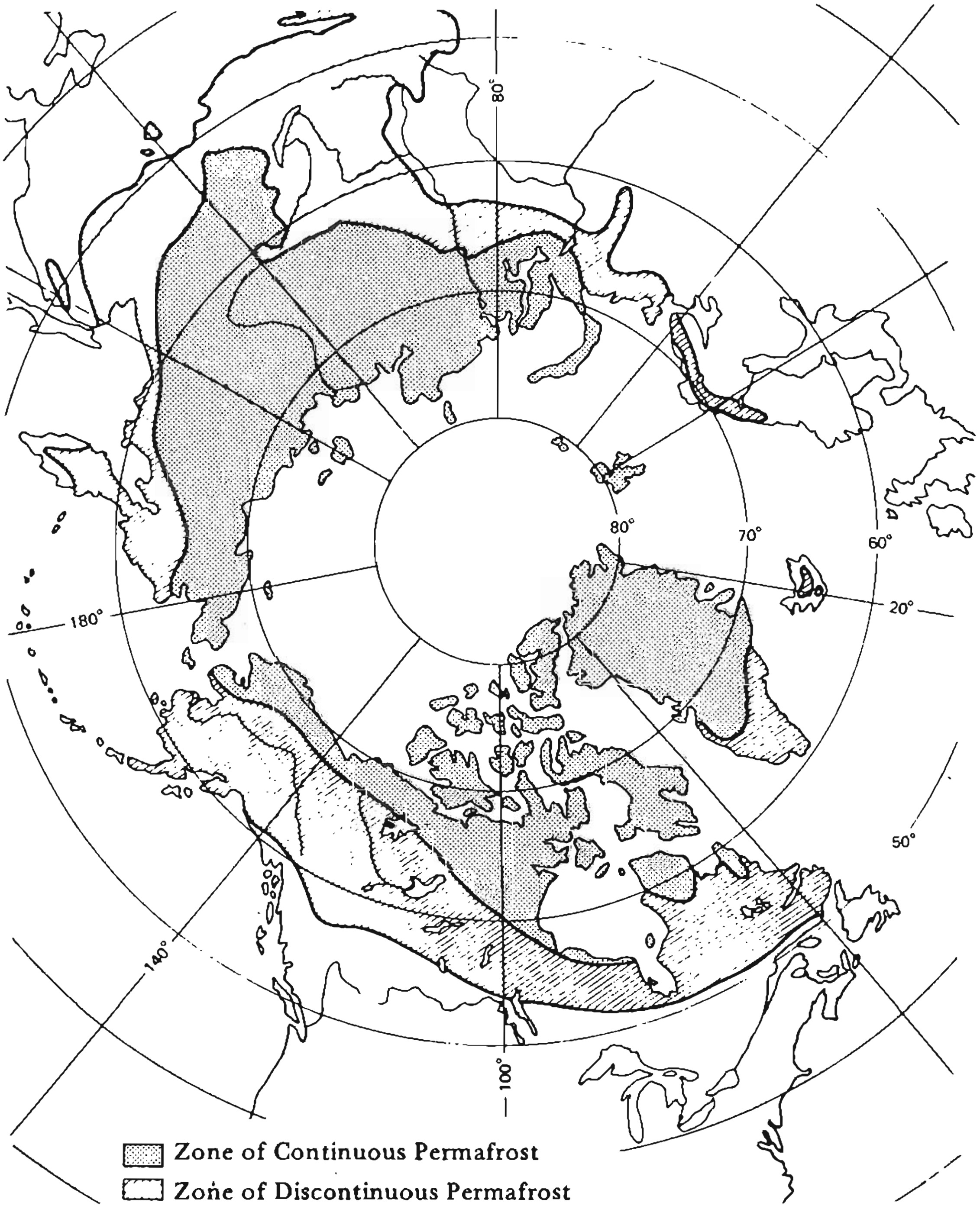


Fig.36 Extent of land permafrost zones in the Northern Hemisphere (from Péwé 1977).

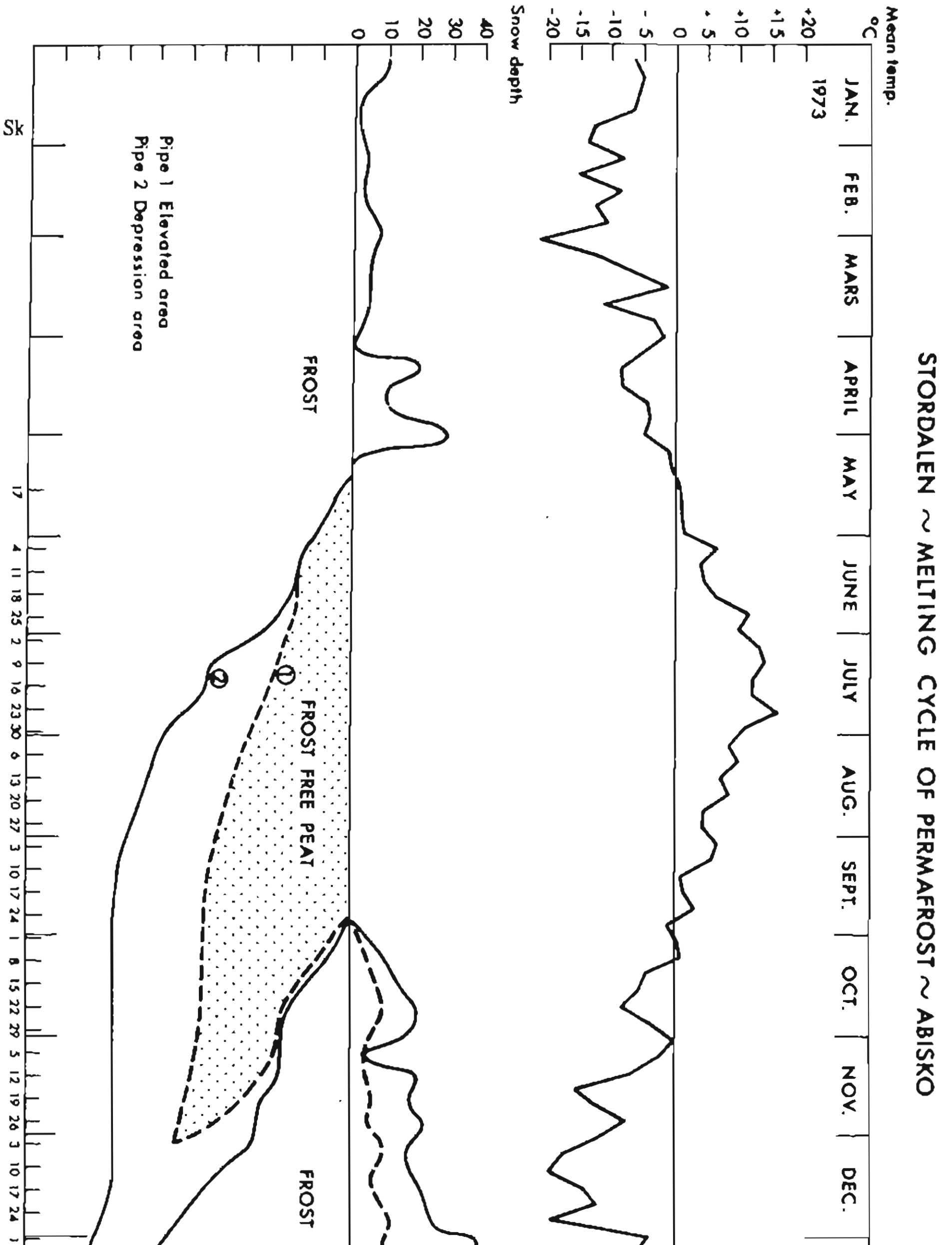


Fig.37 Seasonal thawing and freezing of peat at Stordalen, Abisko, Sweden. Frost penetration meters were installed two meters apart: in an elevated area (1) and a depression area (2), respectively. Snow depths (cm) are plotted in the same scale as frost depths, and also five-day-means of air temperature are plotted (from B.E. Rydén, in Skartveit et al. 1975).

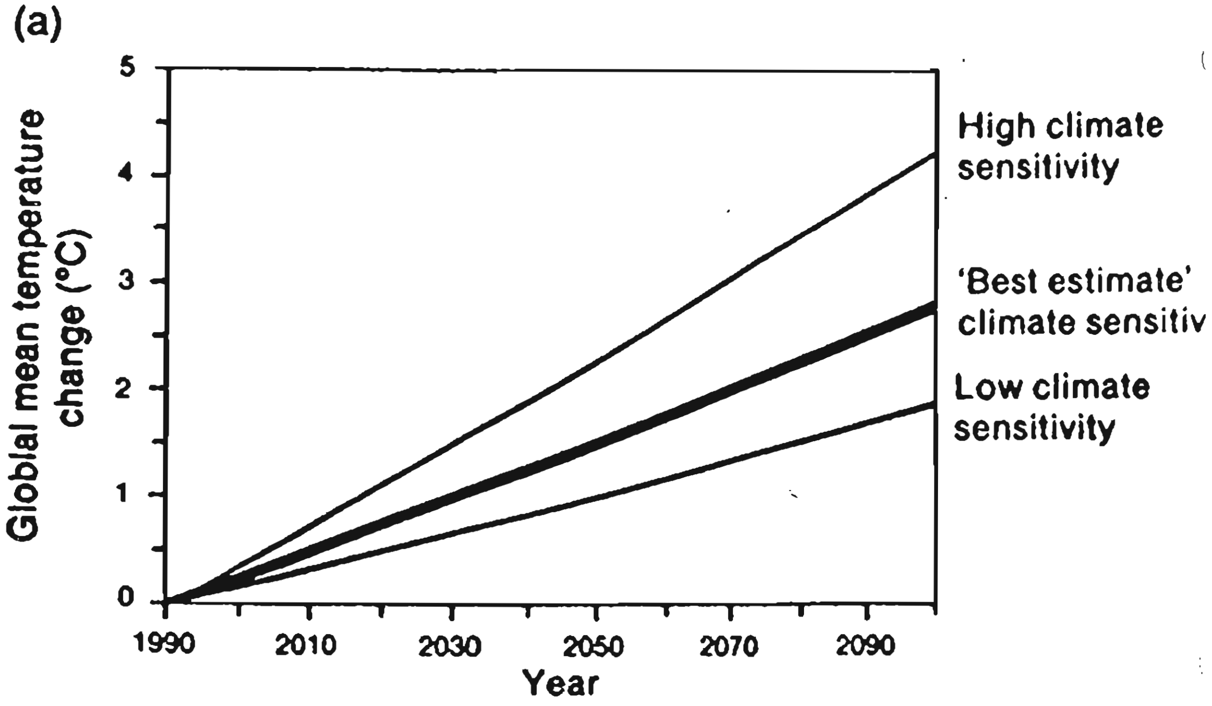


Fig.38 Global temperature change under scenario IS92a (from IPCC 1992).

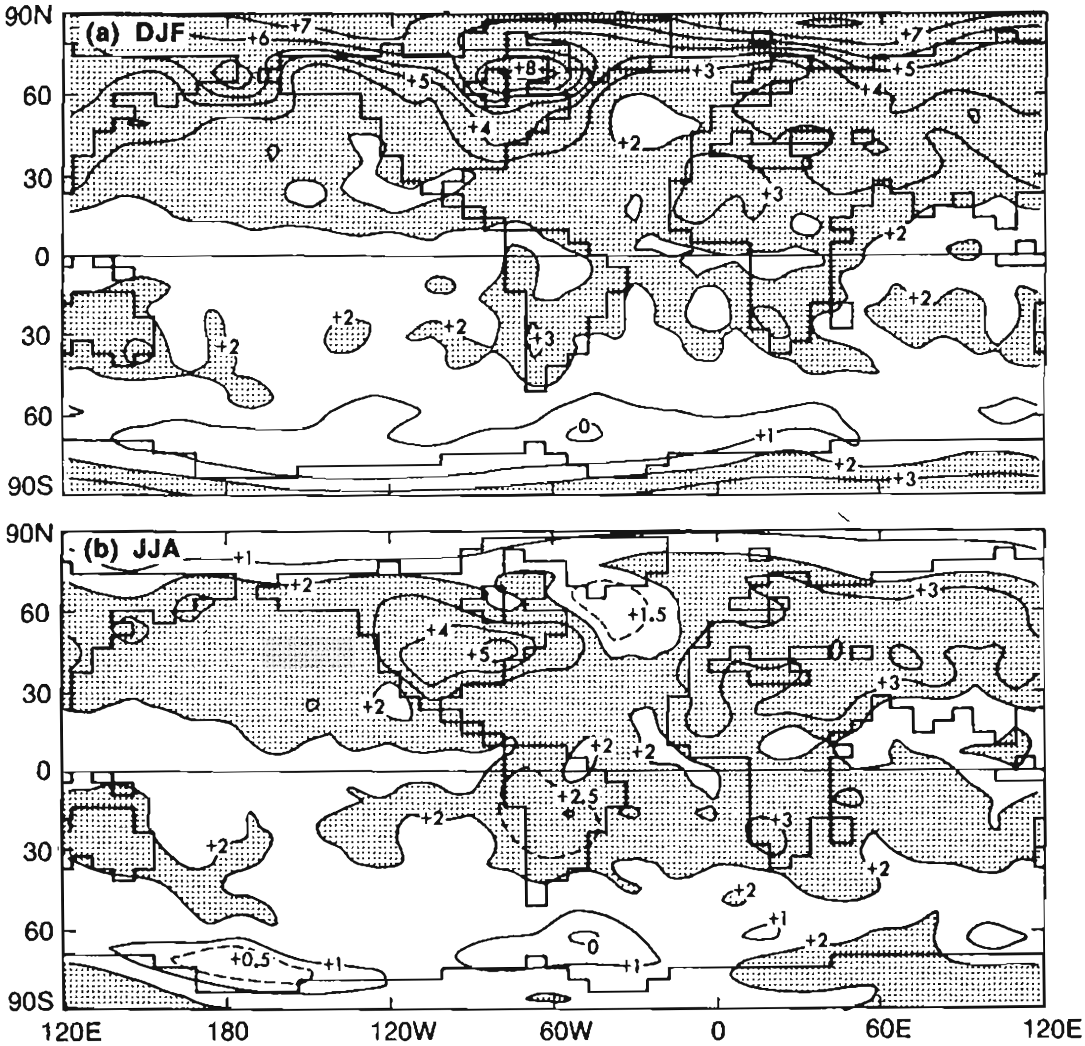


Fig.39 Distribution of the mean surface air temperature (°C) for (a) Dec-Feb, and (b) June-Aug during years 60-80 of a transient CO_2 simulation with the GFDL model, relative to the 100 year average of the control (from IPCC 1992).

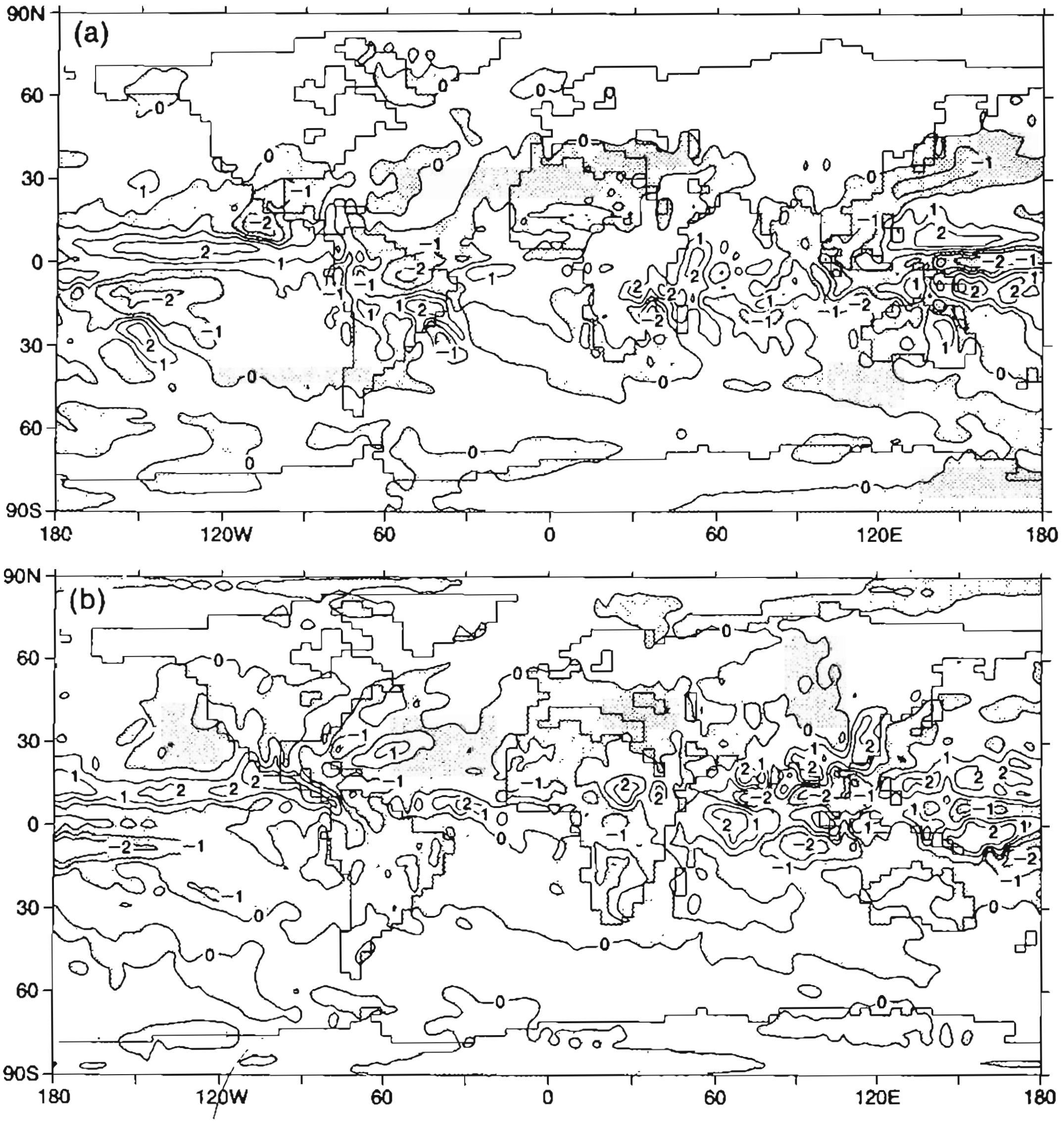


Fig. 40 Decadally averaged changes in precipitation (mm/day) for (a) DJF, and (b) JJA around the time of doubling of CO_2 in an experiment in which CO_2 was increased by 1%/year in the UKMO model. Contour intervals are every 1mm/day and areas of decrease are shaded (from IPCC 1992).

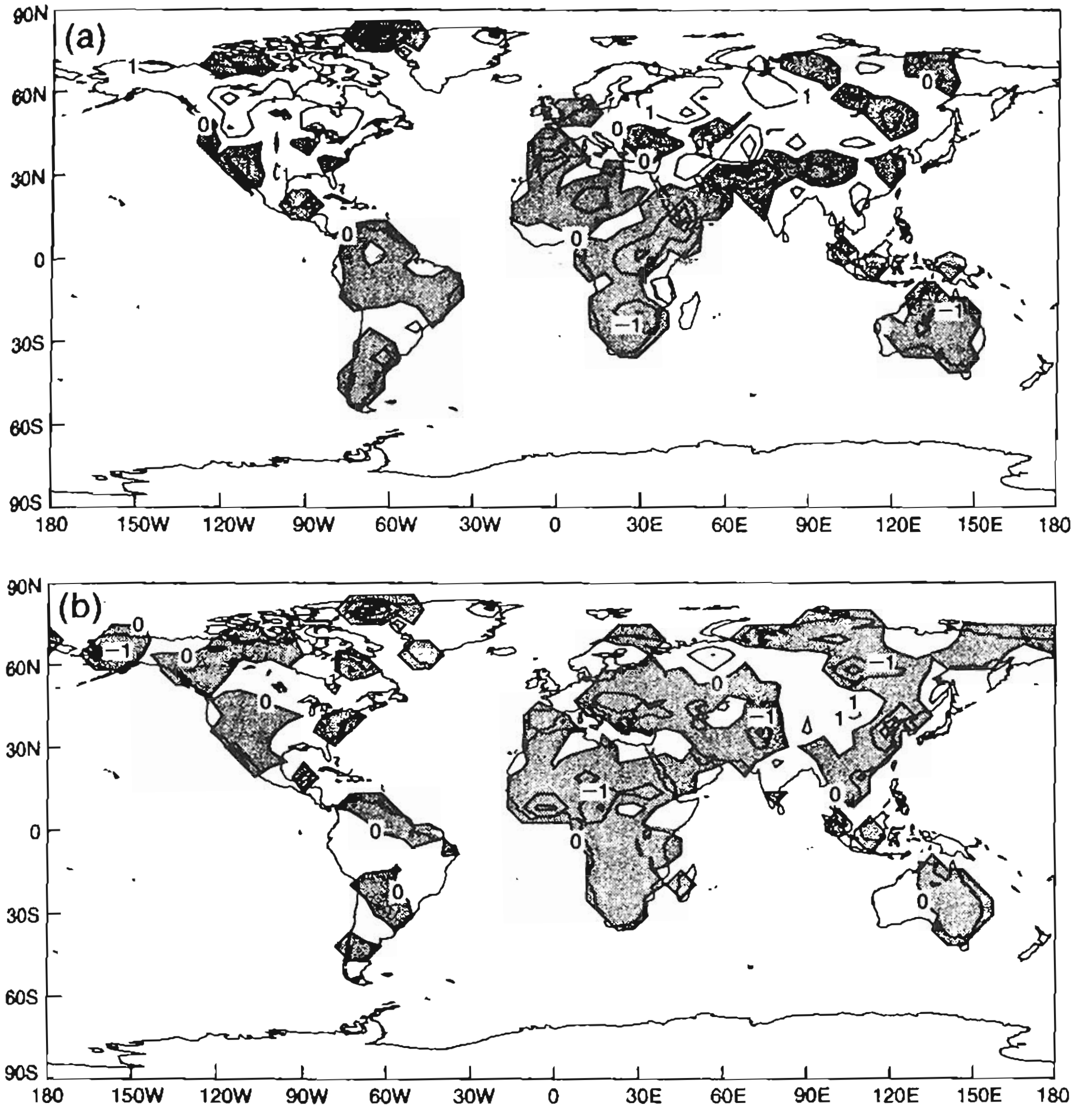


Fig. 41 Decadally averaged changes in soil moisture in the MPI model for (a) DJF, and (b) JJA around the time of doubling of CO_2*. Contour intervals are every cm and areas of decrease are shaded (from IPCC 1992).*

Hydrological context of Nordic countries

Ånund Killingtveit
Department of Hydraulic and Environmental Engineering,
Norwegian Institute of Technology

INTRODUCTION

Ground water recharge and flow is a complex process, governed by climatic and physiographic characteristics. It is difficult to measure ground water recharge directly, and most often it has to be computed from the other terms in the water balance: Precipitation, runoff and evaporation. Groundwater recharge vary both spatially and in time, and an attempt to describe general characteristics of groundwater recharge must be based on correlation between groundwater recharge and other, more easily determined hydrological elements.

A typical example of time variability in groundwater recharge is shown in Figure 1. Figure 1a shows a timeseries of annual recharge to the Gardermoen aquifer for the periode 1958-1993, computed as part of the water balance study for the new main airport under construction at Gardermoen, Norway (Killingtveit & Markhus, 1994). Figure 1b shows average seasonal distribution, with a typical maximum in the spring snow melt periode, and another secondary maximum in the autumn. Both year-to-year variability and seasonal variability is mainly governed by climatic factors, while small scale spatial distribution of groundwater recharge will mainly depend on local physiographic factors like soil type, vegetation and local topography.

This regional description of hydrology can only consider the large-scale features in the main water balance elements precipitation, evaporation and river runoff, and their relationship to climate and physiography. From this information it is also possible to estimate the volume of water available for ground water recharge, and its variability in time and space.

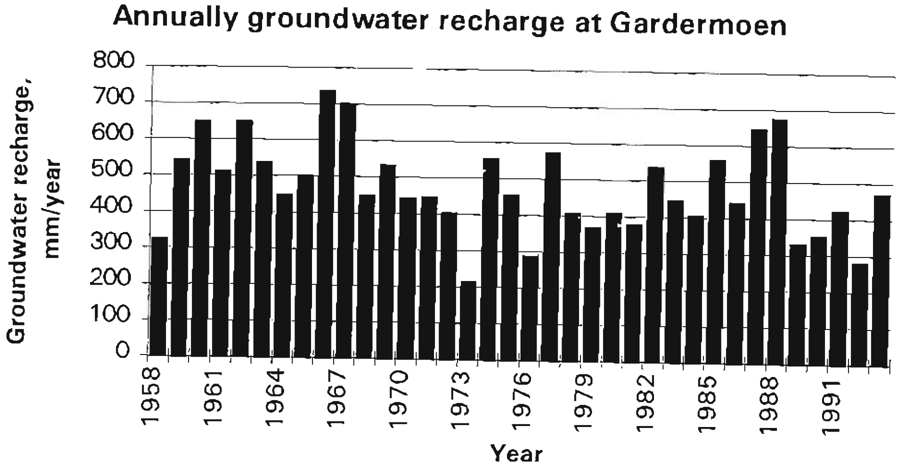


Figure 1a Groundwatere recharge to the Gardermoen aquifer
Computed from climatological data by the SINBAD-model
(From Killingtveit & Markhus, 1994)

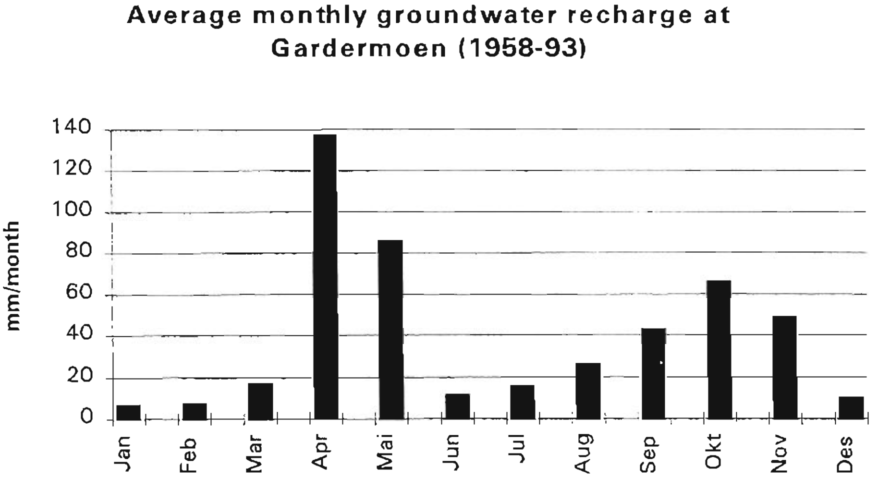


Figure 1b Seasonal distribution of computed groundwater recharge for the Gardermoen aquifer (From Killingtveit & Markhus, 1994)

HYDROLOGY, CLIMATE AND PHYSIOGRAPHY

In a recent attempt to classify river flow regimes for countries in northern and western Europe (Krasovskaia & al, 1993) one of the conclusions were: "*It was shown that there is a much greater variability in flow regime types over space in Scandinavia than in the rest of Europe, due to the complex interaction of maritime and continental climates with varied topography*". This conclusion is based on the fact that the hydrological characteristics in catchments in general are determined by climate and physiographic factors. To understand the variability in hydrology within the Nordic countries it is therefore nessecary to give a short introduction to climate and physiography, and their effect on hydrology. Generally, the input of water as precipitation in catchments is determined by climate and large scale topographical factors, while hydrological processes like groundwater recharge and river runoff formation are determined by physiographic factors such as soil type, vegetation and the local topography.

Climate

The climatic conditions in the Nordic countries are determined by their high latitude and the position on the west coast of the Eurasian continent and around the North Atlantic Ocean and the Baltic sea. The climate is dominated by a high negative annual radiation balance, and a high energy input from latent heat released by condensation (Sømme, 1968). The annual deficit in radiation balance varies from approximately 170 kJ/cm^2 in Denmark (at 55^0 N) to 280 kJ/cm^2 in Finnmark in Northern Norway (at 70^0 N). This deficit in radiation energy is partly compensated for by energy released from condensation (latent heat) and partly by advective heat (sensible heat) from air transported up from southern latitudes. This results in a pronounced seasonal variability in climate, high precipitation and an unusual high air temperature, compared with other areas at the same latitude. The climate is maritime with high annual precipitation in the western part of the region (Iceland, Norway and Denmark) and more continental with lower precipitation in the eastern part (Sweden, Finland). The degree of continentality in climate strongly affects the seasonal pattern in runoff and other hydrological processes. It can be illustrated by the difference between mean temperature in the warmest and coldest mont. The amplitude on the western coast of Norway and on the coast of Iceland is 8-10 ^{0}C, increasing up to >28 ^{0}C in Norrbotten (Sømme, 1968). The climatic effect on precipitation and evaporation distribution is described in more detail in Chapter 4.

Physiography

Topography

The large scale physiographic features are very different in the Nordic countries. The most pronounced effect of topography is related to precipitation, as mountainous areas generally receive most precipitation. Behind mountain ridges lee-zones or "rain-shadow" may occur. As the air-temperature also decreases with elevation, more precipitation falls as snow, affecting both the annual runoff regime and the groundwater recharge regime.

Denmark is a lowland, rising only to 173 masl, most of the country has a softly rolling character with relative relief mostly below 100 m.

Finland also lies mainly below 200 m, but some higher areas are found in Lappland and along the eastern and northern frontiers. The relative relief is mostly below 50 m, but marked relief of up to 500 m occurs in a few isolated mountains in the north.

Norway contrasts strongly with Denmark and Finland. More than 50% of the area lies above 500 m and 25% above 1000 m. The Scandes, the backbone of the Scandinavian peninsula, has a pronounced effect on precipitation distribution in Norway and also in parts of Sweden and Finland. The relative relief is generally high, from typically 500 m up to 1000-1500 m.

Sweden has an intermediate position, with plains similar to Denmark and Finland in the south, and mountainous regions like Norway in the north and west along the Norwegian border. The Scandes also cross into Sweden in some places, giving elevations up to and above 2000 m and relative reliefs of up to 1000 m locally. In large areas east of the Scandes the land slopes gradually to the Baltic sea in Norrland.

Iceland is mainly a bowl-shaped highland with plateaux 500-800 m in the interior. The icecapped volcanic areas in the south-east reach above 2000 m. Coastal mountains up to 1000 m are found in the east and locally in the north. In front of the large glaciers vast sand plains, *sandurs*, are formed by sediments from glacier-fed rivers and hjøkullaup.

Geology

The whole of Norden was affected by glaciations during the Quaternary period, and glacial deposits are found almost everywhere. The thickness of moraine cover is typically 5-10 m in Sweden and Finland, up to 50 m or above in Danmark and Skåne and less than 5 m in most of Norway and Iceland. Areas with thin moraine cover and exposed bedrock are found widely in southern, western and northern Norway, and generally in areas with high and steep relief (Sømme, 1968). Generally the moraine soil will have high infiltration capacity, allowing most of the precipitation and snowmelt to infiltrate into the soil and give the possibility to recharge groundwater. Locally, low infiltration capacity due to exposed bedrock, steep slopes or low hydraulic conductivity may result in overland flow and rapid runoff formation.

REGIONAL WATER BALANCE ELEMENTS

The three main water balance elements are Precipitation, Evaporation and Runoff. On a small scale it may be nessecary to distinguish between river runoff and groundwater runoff, but for most of the Nordic countries the groundwater runoff is very small and may be neglected. Due to large variability in both climatic and physiographic factors, as described briefly in the previous chapter, the main water balance element also show large spatial variability within the region.

A set of water balance maps was produced after the IHD-periode as a result of a coordinated Nordic effort to produce consistant water balance maps, and was first published in 1976 (Forsman, 1976). The maps are shown in Figure 2, 3 and 4 on the following pages. The maps show the distribution of Mean Annual Precipitation (P), Mean Annual Runoff (Q) and Mean Annual Evaporation (E) for the periode 1931-60, expressed in millimeters of water depth. The maps are consistent in the sense that P is intended to equal Q + E in any point on the map, and they do not show any abrupt changes at the borders. The precipitation and runoff maps are intended to show averages over an area at the order of 1000 km^2, while the evaporation map shows averages over approximately 10000 km^2. Some comments and additional information concerning each map is given below.

Precipitation

The precipitation map in Figure 2 shows the corrected or Actual precipitation. Most other published precipitation maps shows only the precipitation measured at standard gauges in the national precipitation station network. The actual precipitation will usually be significantly higher than measured precipitation, especially in mountainous areas and far to the north where a large part of the precipitation falls as snow. Further, the map includes the effect of precipitation increase with elevation.

The map shows very high precipitation on the west coast of Norway, and the strong effect of the Scandes. This mountain ridge, running from SW to NE along the Scandinavian peninsula, forces westernly air masses to lift over the mountain ridge and release much of the humidity as precipitation in front of and over the ridge. Average annual precipitation is around 2000 mm in the maximum zone along the coastline, and isolated areas get over 4000 mm annualy. Behind the ridge a "rain shadow" can be seen, with precipitation down to 700-800 mm annualy. The precipitation distribution in Denmark, Sweden and Finland is much less variable, most of these countries have average annual precipitation ranging from 600 mm to 800 mm.

In Iceland the average annual precipitation varies from below 500 mm/year in dry areas in the inland and to the north, up to and above 5000 mm/year at the glacier Myrdalsjøkull. Annual average for Icelend is estimated to 2000 mm/year, highest of all Nordic countries (Tollan, 1994).

Evaporation

The evaporation map, seen in Figure 3, is much less detailed than the two other maps. One reason for this is that evaporation is usually more evenly distributed spatially, another reason is that less data were (and still is) available for evaporation than for precipitation and runoff data, making a more detailed description difficult. Evaporation is high (>500 mm) in the south and decreases

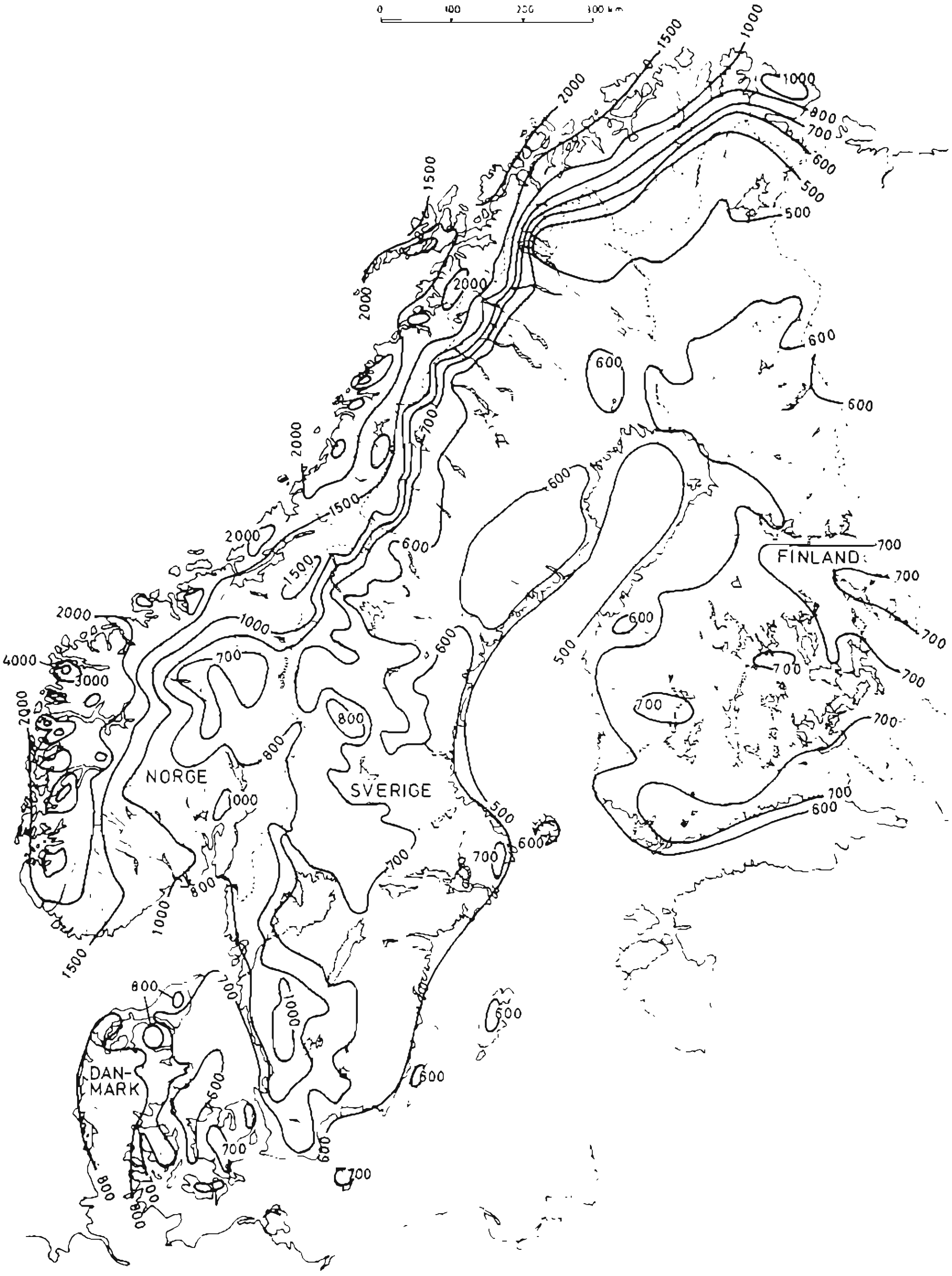


Figure 2 Average annual precipitation in Norden for the periode 1930-60 (From Otnes og Ræstad, 1978)

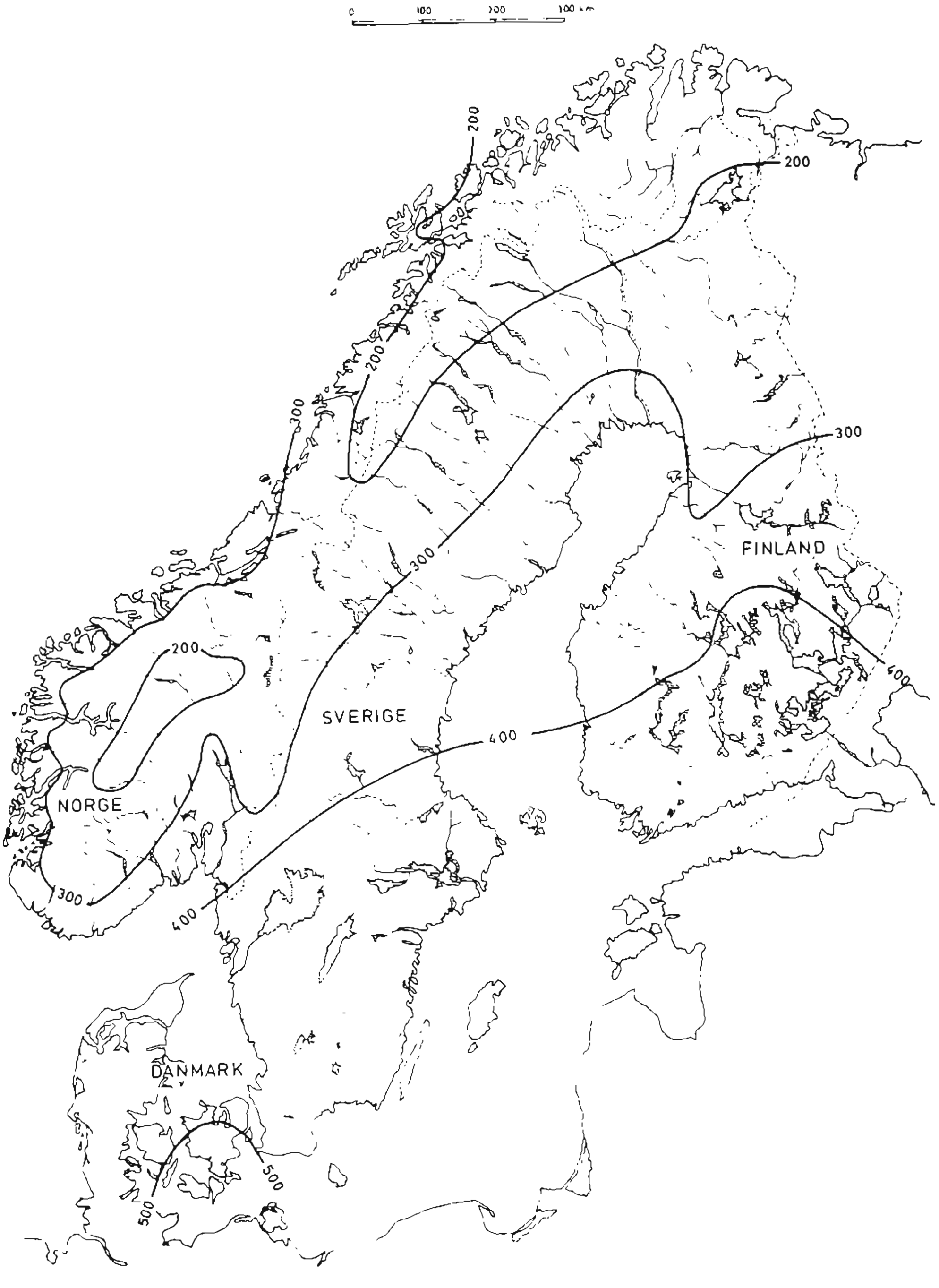


Figure 3 Average annual evaporation for Norden for the periode 1930-60 (from Otnes og Ræstad, 1978)

as a function of elevation and latitude down to 200 mm or below in the north and in mountainouss areas. It is still possible that these evaporation values may be too low, due to difficulties both with measurement techniques and computational methods (Forsman, 1976).

As an example, the actual evaporation in Sagelva research basin close to Trondheim was calculated to be 430 mm/year shortly after these maps were published (Killingtveit, 1985). This value is more than 30% over the value found in Figure 3. Still, the map probably gives a good general description of actual evaporation distribution in Norway, Sweden, Finland and Denmark.

In Iceland evaporation measurements are scarce, but has been estimated to be in the range from 300 to 400 mm/year, based on climatological data (Tollan, 1994).

Runoff

Mean annual runoff varies as shown in Figure 4, from a maximum of > 4000 mm/year in a small area near the Ålfoten glacier in western Norway, down to below 150 mm/year in Sjælland. In Denmark and Finland the mean annual runoff varies in the range 200-400 mm/year, in Sweden betweem 200 and 1000 mm/year and in Norway between 500 and >4000 mm/year. The large scale pattern of runoff is mainly determined by the precipitation distribution. It is important to realize that the actual average runoff locally may deviate from this map, due to local gradients in precipitation distribution.

In Iceland the distribution of average annual runoff follows the same pattern as the precipitation distribution. Average runoff from Iceland has been estimated to be 1586 mm/year as river runoff, and 33-62 mm/year as groundwater runoff (Tollan, 1994).

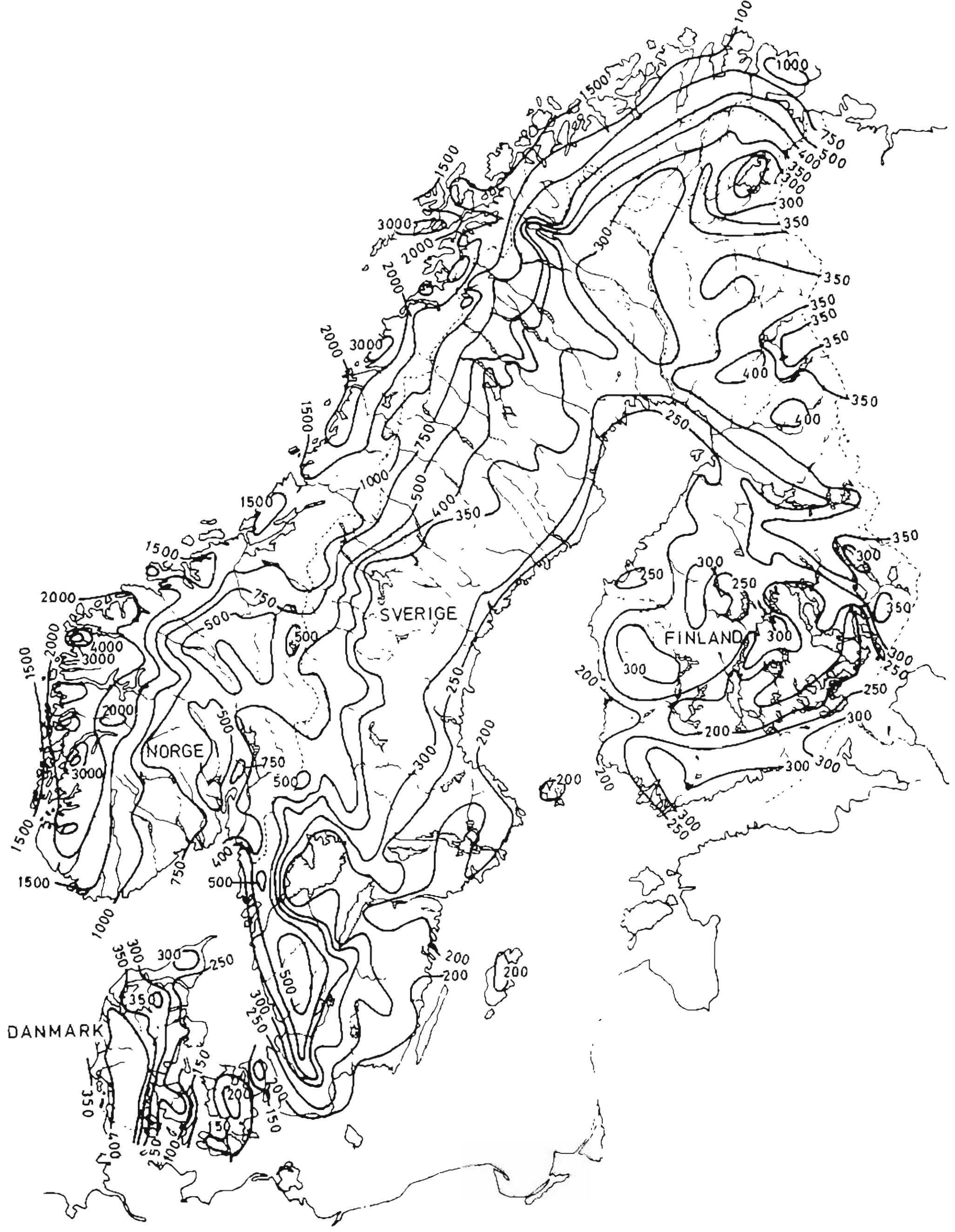


Figure 4 Average annual runoff for Norden for the periode 1930-60 (From Otnes og Ræstad, 1978)

RIVER FLOW REGIMES

The term *river flow regime* is used for the time variation of flow and especially the time of occurrence of high and low flows (Gottschalk & al, 1979). This variation is governed by climate and physiography, as described in Chapter 2. Classification of flow regimes is important in order to to describe regions with similar seasonal distribution of river flow, and it can also be used to identify areas with similar groundwater recharge distribution. The following system was defined for classification according to high and low flows (Gottschalk & al, 1979):

High water

H_1: Dominant snowmelt high water. An area is classified as H_1 if the three months with the highest average runoff belongs to spring or early summer (typically May-July).

H_2: Transition to secondary rain high water. An area is classified as H_2 when the second highest or third highest monthly runoff takes place in autumn (typically: October, November, on the Scandinavian peninsula - early in the west and late in the east and vica verca in Finland)

H_3: Dominant rain highwater. An area is classified as H_3 when the highest monthly runoff takes place in autumn or early winter (typically November-December)

Low water

L_1: Dominant low water flow in winter, caused by snow accumulation. An area is classified as L_1 when the two months with the lowest runoff belong to winter or early spring (typically: February-March)

L_2: Transition zone, when the two months with the lowest runoff do not belong to the same time of the year (typically: February and July)

L_3: Dominant summer low water caused by high evapotranspiration and/or low precipitation when the two months with the lowest belong to summer or early autumn (typically: June-August)

Five main combinations of regimes has been identified in the Nordic countries, not including Iceland (Gottschalk & al, 1979). These regions are shown on the map in Figure 5.

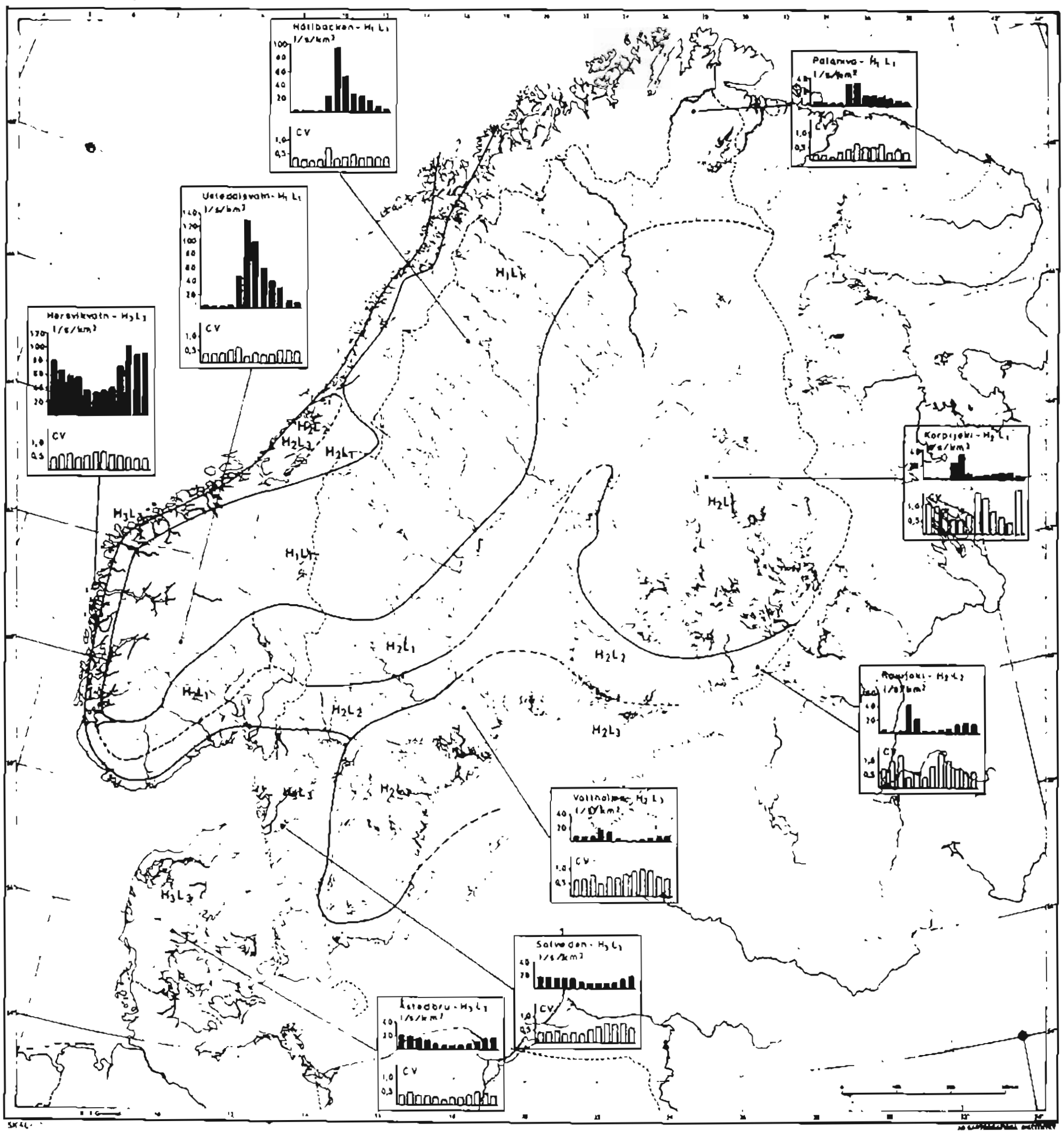


Figure 5 Runoff regimes in Norden (From Gottschalk & al, 1979)

The map shows the geographical location of these five main regimes. For each region a "typical" monthly runoff hydrograph is shown together with the coefficient for variability, C_v, for runoff in each month. A general trend is that high monthly runoff is associated with low C_v, and vice versa. In Table 1 each regime is described shortly with the name suggested by Gottschalk & al, and an description of the geographical area it covers.

Table 1: Runoff regions in the Nordic countries. (From Gottschalk & al, 1979)

Regime	Name	Geographical area covered
H_1L_1:	Mountain regime	Inner and northeast parts of Norway, Sweden and Finland
H_2L_1:	Inland regime	Fiord area along Norway's western coast, forest and coastal areas to the west of the Gulf of Bothnia, Finland except southwestern parts and northern Lapland.
H_2L_2 and H_3L_2:	Transition regime	Inlands in Østlandet (Norway), Vermland, Dalecarlia and Gestrikland (Sweden) and the western and southern coast of Finland
H_2L_3:	Baltic regime	Parts of Møre-Trøndelag (Norway), southeastern Sweden, southwestern coast and the archipelago of Finland
H_3L_3:	Atlantic regime	Frontier coastal regions of Norway, soutwestern Sweden and Denmark

Special regimes such as glacial (extreme H_1L_1) and the lindá regime in Iceland might be added to the five regimes above. Areas with glacial influence have maximum flow in late summer, and has been classified into region H_1L_1. A division into more detailed sub-regimes is also possible, such an attempt has been made in Finland by Solantie (Solantie, 1975). In this system, shown in Figure 6, the regions H_1L_1 and H_2L_1 has been divided into 6 sub-regions by taking into account the physiography in addition to runoff data.

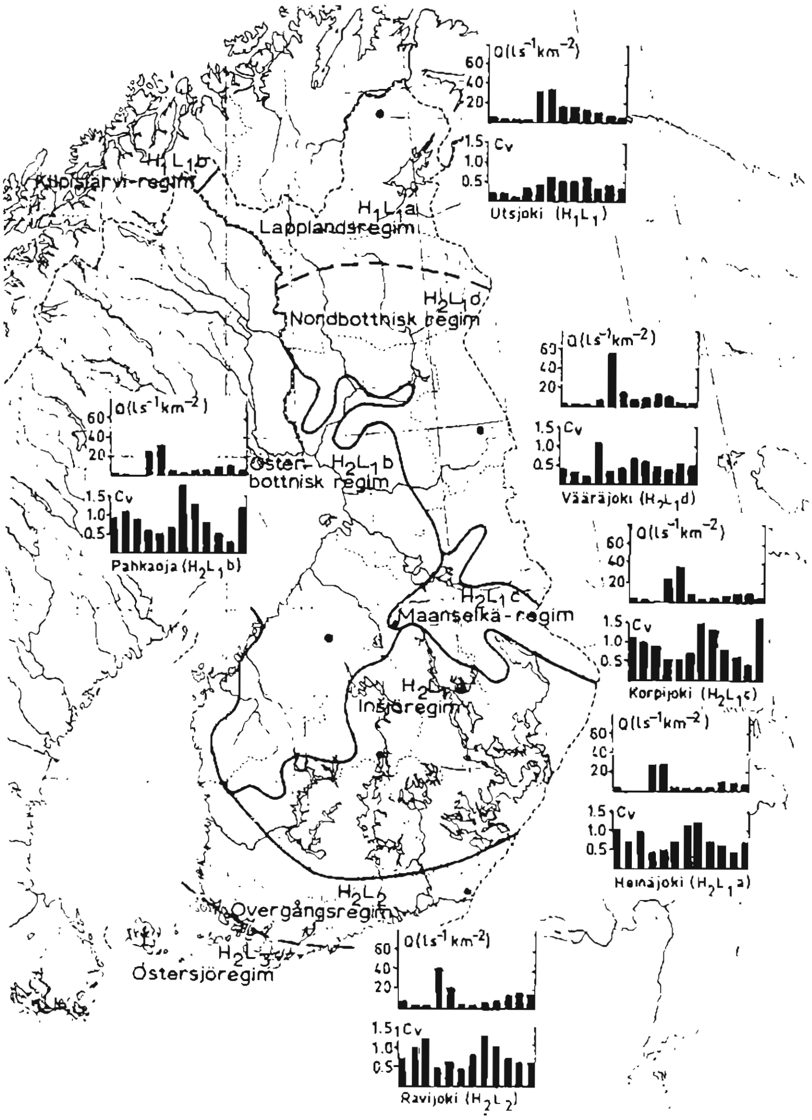


Figure 6 Detailed runoff regimes in Finland (From Gottschalk & al, 1979)

Runoff regimes in Iceland

Rivers in Iceland may be grouped into the following three classes according to Tollan:

> Rivers with direct runoff fom rain and snowmelt, *dragá* in Icelandic. Flow in these rivers is closely correlated with rainfall and snowmelt. The Icelandic dragá-rivers are located in areas with older basaltic bedrock, mostly in NV and SE of Iceland.
>
> Groundwater dominated rivers, *lindá* in Icelandic. Many of these rivers are fed by springs, and have very stable flow through the year. The lindá rivers carry small amounts of sediments, and water temperature is stable. Lindá rivers are found in areas with young and highly porous lava.
>
> Glacier streams, *jökullvatn* in Icelandic. These rivers receive glacier meltwater, and the flow is closely correlated with the glacier melt and therefore has a pronounced seasonal variation. These rivers usually has a high sediment transport, with high turbidity and brownish or gray colour. Some of the rivers deposit large amounts of sediments in front of the glaciers, forming large flat sand plains, called *sandur* in Icelandic.

Typical runoff hydrographs for these three types of rivers are shown in Figure 7. The dragá river has two high flow and two low flow periodes during the year, similar to the rivers in the H_2L_2 region in Scandinavia. The glacier stream has a minimum flow in the winter and maximum in late summer, similar to H_1L_1 rivers in Scandinavia. The effect of rainfall combined with glacier melt can be seen as "spikes" in the hydrograph in early June, October and November. The lindá river has almost constant flow during the year, with two small peaks which could be caused by rainfall or snowmelt.

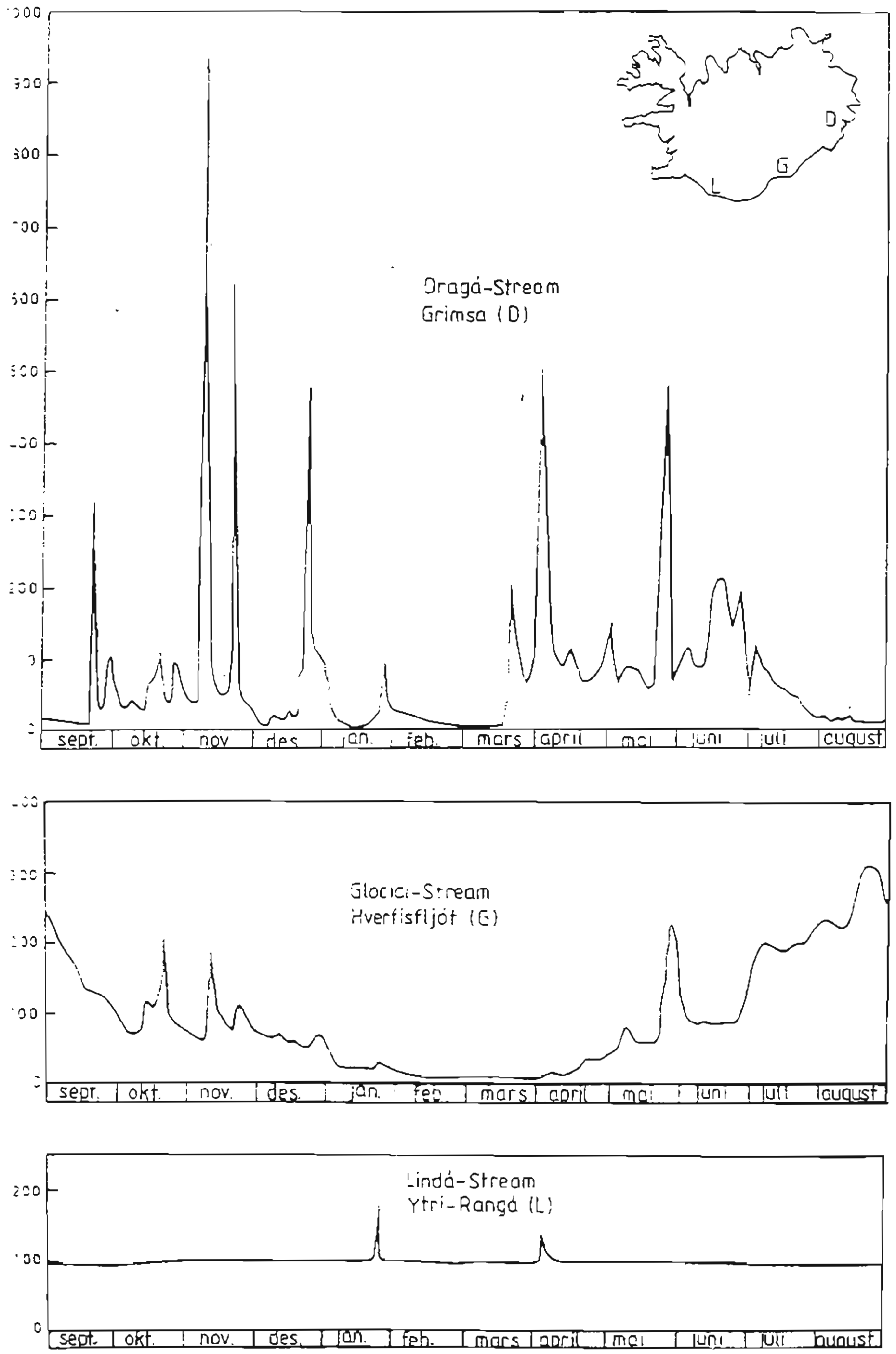


Figure 7 River flow regimes in Iceland (From Tollan, 1994)

GROUNDWATER REGIMES

The regional variations in river flow described in the previous chapter may also be used to describe the general conditions for groundwater recharge and groundwater fluctuations in Norden. Kirkhusmo identified three regions of regional variations in groundwater level in Norden (Kirkhusmo, 1988). These regions are shown in Figure 8. Typical annual groundwater variation within each region is described in Figure 9. By comparing Figure 8 and Figure 5 it can be seen that Groundwater region 1 corresponds well to the Atlantic flow region H_3L_3. Groundwater region 3 corresponds roughly to the Mountain flow region H_1L_1, while groundwater region 2 includes the Inland, Baltic and Transition flow regimes.

In the Atlantic region (I) the groundwater reaches maximum in winter or early spring, due to high precipitation, no ground frost and low evaporation. Minimum groundwater level occurs in late summer, after a long periode of low precipitation and high evaporation.

In the Mountain region (III) the groundwater also has one well defined maximum and one minimum during the year, but with a different phase than in the Atlantic region. Maximum groundwater level is found at the end of the snowmelt season, where most of the groundwater recharge occurs. From this maximum the groundwater level decreases steadily until the minimum which occurs in April-May, just before the next snowmelt season starts. This pattern clearly show the dominating role of the snow melt in this region.

In the region II the groundwater has two maxima and two minima annually, as seen in Figure 10. The two maxima occur in the two periodes with high groundwater recharge, during the spring snowmelt and during the autumn rains. The two minima occur in the two periodes with low groundwater recharge, the first during the winter when temperature is low, the second during late summer when temperature and evaporation is high. This pattern is also clearly demonstrated in Figure 1b, for the average annual groundwater recharge at Gardermoen.

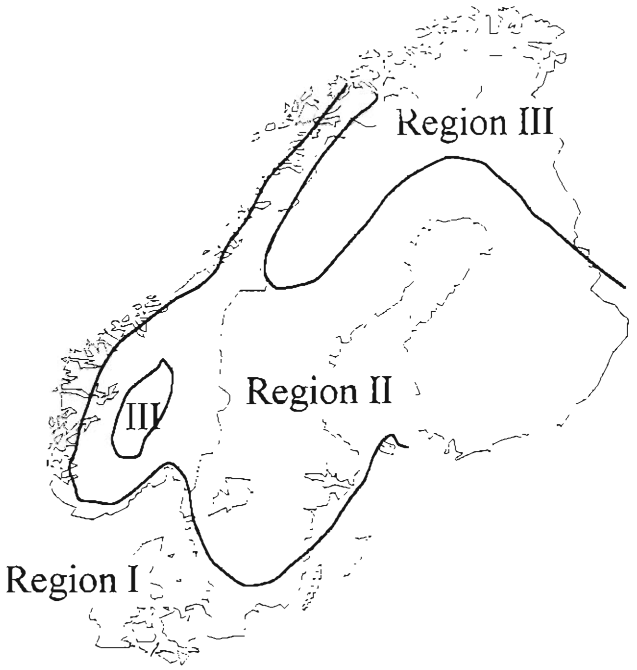


Figure 8 Three main groundwater regions in Norden (From Kirkhusmo, 1988)

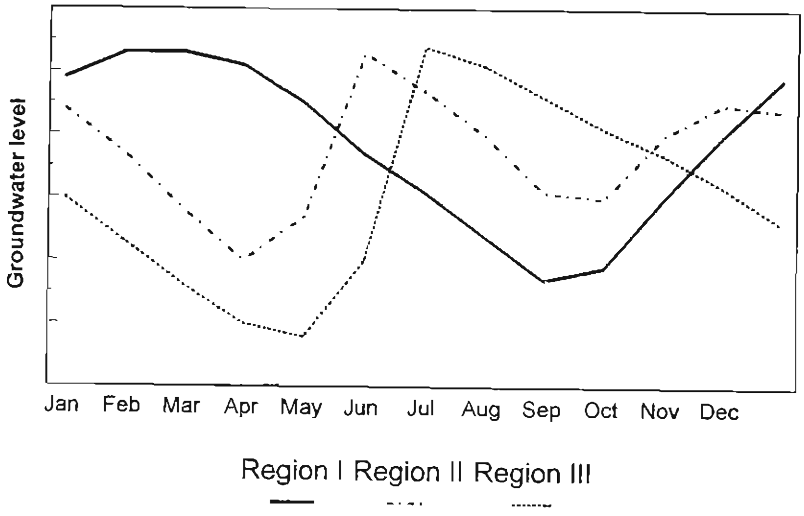


Figure 9 Typical groundwater level variation in the three main groundwater regions defined in Figure 8 (from Tollan, 1994)

HYDROLOGICAL IMPACTS OF CLIMATE CHANGE

The overview given in the previous chapters was based on the actual climate and hydrology during this century, mostly based on data for the periode 1931-60. A possible future climate change may change the stability of the climate, and change the hydrology. One of the most important consequences of future climate change may be an alteration of the hydrological balance, affecting both river runoff and groundwater recharge. The possible effects of such changes are not well understood, but many studies have been done and more are underway today. This chapter includes a brief summary of some of the most important such studies, and some of the conclusions concerning possible effects on hydrology in the Nordic countries. Only the most recent studies are considered for this summary.

Most of the studies performed in Norden is based on an assumption of an increase in CO_2 up to twice the present level. This level is typically reached in the middle or to the end of the next century if the present trends continue ("Business as usual" Scenario). From this assumption the effects on global climate is simulated using so-called Global Circulation Models (GCM). Several such models are in use, with varying degree of accuracy. It is established that the North-Atlantic and the adjoining land areas constitute a difficult region because the climate is strongly affected by the oceanic circulation and the deep water renewal in this area. The present CGMs have problems in simulating these processes correctly (Sælthun & al, 1994). Therefore, it seem that some of the results, especially from previous simulations, may not be very accurate. The GCMs typically simulate global climate with a coarse spatial resolution (typically 200X200 km), but the spatial resolution is increasing with each new generation of the models.

A study in Norway in 1990 and 1991 was based on analysis and recommendations from UN's Intergovernmental Panel on Climate Change (IPCC), but adjusted to Nordic conditions. A typical scenario in this study is an increase in global mean temperature of 1.5-4.5 ^{0}C as a result of 2xCO_2. This level may be reached in 2025-2050. The temperature increase will be largest near the poles, and smaller near the equator. For Norway the following effects were assumed:

Air temperature increases 3-4 ^{0}C in winter and 2 ^{0}C in summer
Precipitation increases 10% in spring and summer, and 5% in fall and winter.
Soil water increases 10-20 mm in winter, spring and fall, no changes in summer.

Typical hydrological changes were computed by using a precipitation-runoff model (The HBV-model) in selected catchments. The results indicated a strong change in seasonal runoff regime, with more winter-runoff and less runoff in spring and summer. In some catchments the computed changes were dramatic, with winter runoff many times todays values. (Miljøverndep., 1991)

Vehvilainen and Lohvansuu studied the possible effects on river discharge and snow cover in Finland (Vehvilainen & Lohvansuu, 1991), and based their analysis on a similar methodology. Based on a GCM from Goddard Institute for Space Studies (GISS) a temperature increase of 2-6 ^{0}C was computed for Finland, together with a precipitation increase of 10-30 mm/month and evaporation increase of 5-30 mm/month. The hydrological effects were evaluated by simulations in 12 catchments using the HBV-model. According to their results the hydrological changes would be very large: Average mean runoff increased 20-50%, the mean minimum discharge in winter increased considerably, average snow cover periode was reduced by 2-3 months and mean maximum discharge decreased due to diminished maximum snow water equivalents.

Krasovskaia and Gottschalk studied the frequency of extremes and the stability of flow regimes and the relation to climate change. They concluded that a possible consequence could be a change in frequency and magnitude of extreme floods (Krasovskaia & Gottschalk, 1992). They also concluded that the flow response to a temperature change of only ±1 ^{0}C and annual precipitation change of ±100 mm was specially pronounced in southern Scandinavia (Krasovskaia & Gottschalk, 1993).

The Nordic project "Climate Change and Energy Production" has been defined to study the effect of climate change on hydrology and water power production. The program is near completion, and a summary of findings were recently presented at NHK'94 (Sælthun &al, 1994). This study utilizes results from the newest GCM simulations, and climate change scenarios are established for all Nordic countries, including the Faroe Islands and Greenland. The temperature increases

123

predicted is in the order of 0.3-0.45 ^{0}C per 10 years, and a precipitation increase of 1.5% per 10 years. There are marked regional and seasonal differences in temperature predictions, with largest increase in winter and in the continental areas. The precipitation predictions from the GCM simulations vary considerably with no particular spatial pattern. A Nordic scenario was computed by a Nordic expert group, with results as shown in Table 2 for the 10 prediction points selected, as shown in Figure 10. The predictions given in Table 2 should occur over 70 years, as a result of a 2xCO_2 scenario.

Some preliminary results are given, based on simulations with an improved HBV-model. The simulations confirm the results from other studies: A possible future climate change will have a pronounced effect on hydrology in the Nordic countries. Spring and summer runoff (and groundwater recharge) will typically decrease, while winter runoff and winter floods will increase substantially. Also glaciated catchments will be strongly affected, and some glaciers may decrease and finally disappear if the predicted trend continues.

Table 2: **Precipitation and temperature changes predicted and estimated Changes are after 70 years (From Sælthun & al, 1994)**

Station	Temperature, °C Winter - Summer	Precipitation,%
Nuuk	3.1 - 1.8	11
Kirkjubæjarklauster	2.5 - 1.8	11
Torshavn	2.5 - 1.8	11
Kastrup	3.3 - 1.9	11
Bergen	3.1 - 1.8	14
Tromsø	3.3 - 2.0	14
Østersund	3.8 - 2.0	11
Stockholm	3.7 - 1.7	11
Sodankyla	4.1 - 2.3	11
Kuopio	4.0 - 2.2	11

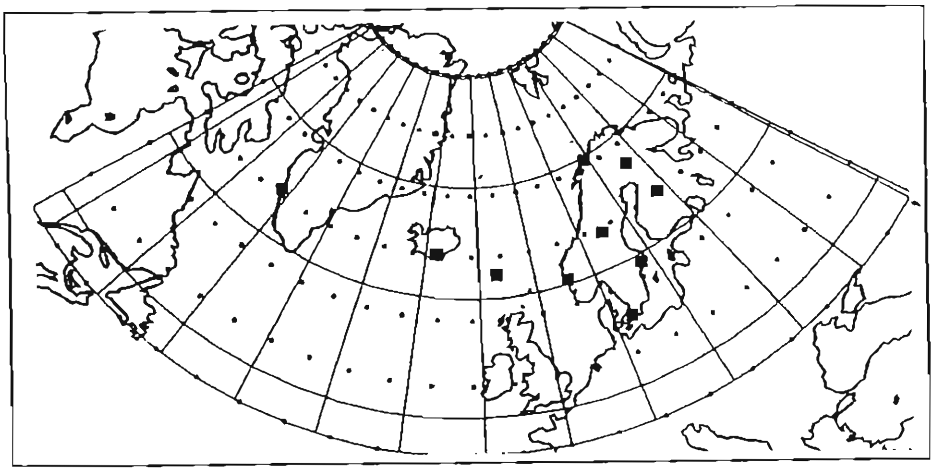


Figure 10 Prediction points (•) and climatic stations (■) used in the analysis (From Sælthun & al, 1994)

REFERENCES

Forsman, A. (1976): Water Balance Maps of the Nordic Countries
Vannet i Norden, nr. 4

Gottschalk, L., Lundager Jensen, J., Lundquist, D., Solantie, R. and Tollan, A. (1979): Hydrological regions in the Nordic Countries
Nordic Hydrology, Vol 10 No. 5

Killingtveit, Å. and Markhus, E. (1994): Grunnvannsdannelse på Gardermoen
NGU Grunnvannseminar Trondheim 24/10, 1994

Killingtveit, Å. (1985): Computation of mean areal evapotranspiration by the water-balance method
In: *Casebook on Operational Assessment of Areal Evaporation*
WMO Operational Hydrology Report No. 22

Kirkhusmo, L. (1988): Groundwater fluctuation patterns in Scandinavia
In: *Studies of Groundwater Recharge in Finland, Norway and Sweden*
NHP-Report No. 23

Krasovskaia, I. & Gottschalk, L. (1992): Stability of Flow Regimes
Nordic Hydrology, Vol 23(3)

Krasovskaia, I. & Gottschalk, L. (1993): Frequency of Extremes and its Relation to Climate Fluctuations
Nordic Hydrology, Vol 24

Krasovskaia, I, Gottschalk, L., Arnell, N.W. and van der Vateren-de.Hoog, B. (1993): Variations in regime class over time.
In: *Flow Regimes from International Experimental and Network Data (FRIEND) Volume I Hydrological Studies*

Miljøverndepartementet (1991): *Drivhuseffekten, virkninger og tiltak* Rapport fra den interdepartementale klimagruppen.
Miljøverndepartementet, T-841, Oslo, ISBN 82-7243-851-8

NVE (1987): *Avrenningskart over Norge*
Norges Vassdrags- og Energiverk, Vassdragsdirektoratet, Hydrologisk Avdeling

Otnes, J. og Ræstad, E. (1978): *Hydrologi i praksis*
Ingeniørforlaget A/S, Oslo 1978 (ISBN 82 524 0036 1)

Solantie, R. (1975): A proposal for the definition of Finnish hydrological zones from a consideration of small, almost lake free basins.
Vannet i Norden, No. 4

Sælthun, N.R., Bergstrøm, S., Einarsson, K., Thomsen, T. & Vehvilainen, B. (1994): Simulation of Climate Change Impact on Runoff in the Nordic Countries
Nordic Hydrological Conference 1994, Torshavn, Faroe Island 2-4 August 1994

Sømme, A. (1968): *A Geography of Norden*
J.W.Cappelens Forlag, Oslo

Tollan, A. (1975): Hydrologiske regioner i Norden
Vannet i Norden, No. 1

Tollan, A. (1994): Norden som hydrologisk region
In: *Kompendium i Hydrologi*, Inst. for Geofysikk, Oslo, August 1994

Vehvilainen, B. & Lohvansuu, J. (1991): The effects of climate change on discharges and snow cover in Finland.
Hydrological Sciences Journal Vol 36, No. 2

Introduction to weathering processes

James I. Drever
University of Wyoming

DEFINITIONS OF WEATHERING

The concept of weathering is relatively simple. Minerals, particularly minerals formed at high temperatures and pressures in the earth's interior, are unstable in contact with dilute waters at the earth's surface and tend to dissolve or transform into minerals that are more stable in the environment of the earth's surface. In geochemical and hydrologic studies, however, our interest is not so much in what happens to the minerals themselves, it is rather the effect of weathering reactions on water chemistry. The focus of this chapter will therefore be on the production of solutes rather than on mineralogy. The solutes of particular interest are the *base cations*, Ca^{2+}, Mg^{2+}, Na^+ and K^+. These are important in part because of their role as plant nutrients, but primarily because of their importance in neutralizing acidity or generating alkalinity. Weathering of primary (meaning present in the bedrock) minerals can be described by the general equations:

$$\text{Primary mineral} + n\,H^+ = \text{Uncharged secondary products} + \text{cation}^{n+}$$

or

$$\text{Primary mineral} + n\,CO_2 + m\,H_2O = \text{Uncharged secondary products} + \text{cation}^{n+} + n\,HCO_3^-$$

Uncharged secondary products may include secondary minerals such as clay minerals or $Al(OH)_3$, and may include uncharged solutes, particularly H_4SiO_4. Alkalinity, and hence chemical weathering, is a major control on the pH (and hence Al concentration and suitability as fish habitat) of surface waters (Fig. 1). I shall define *rate of weathering* for the purpose of this chapter as rate of production of dissolved base cations, in units of equivalents per unit area of catchment per unit time, commonly keq ha^{-1} y^{-1}. The base cations need not be transported in runoff: they may also be taken up by plants or adsorbed on a pre-existing exchange site. I would stress, however, that this is a very limited definition of weathering, and other definitions will be appropriate when weathering is viewed from other perspectives.

Generally speaking, rates of weathering are of interest in environmental studies when they are low—that is to say when acidification or depletion of nutrient cations is a potential problem. This occurs where the bedrock is composed of silicates rather than carbonates, so the following discussion will focus on silicate rather than carbonate rocks.

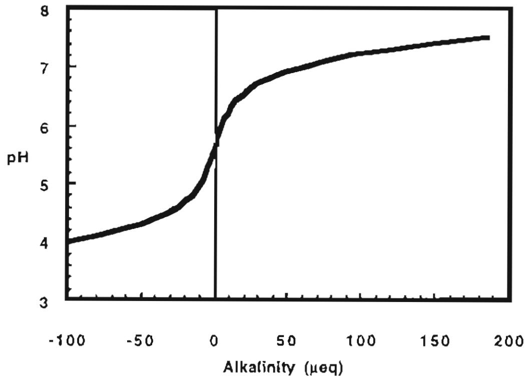


Figure 1: Relationship between pH and alkalinity (= $[HCO_3^-] - [H^+]$) for waters in equilibrium with atmospheric P_{CO_2} at 10 °C. Note the abrupt decrease in pH as alkalinity changes from positive to negative values.

TYPES OF WEATHERING REACTION

Congruent solution

Congruent dissolution means that the entire solid dissolves, leaving no secondary solid phase. An obvious example would be calcite

$$CaCO_3 + 2H^+ = Ca^{2+} + H_2O + CO_2$$

or forsterite olivine

$$Mg_2SiO_4 + 4H^+ = 2\,Mg^{2+} + H_4SiO_4$$

Aluminosilicates such as feldspars may dissolve congruently under strongly acidic or extremely dilute conditions

$$NaAlSi_3O_8 + 4H^+ + 4H_2O = Na^+ + Al^{3+} + 3H_4SiO_4 \qquad \text{(acidic conditions)}$$

$$NaAlSi_3O_8 + 8H_2O = Na^+ + Al(OH)_4^- + 3H_4SiO_4 \qquad \text{(neutral conditions)}$$

More commonly, because of the low solubility of aluminum hydroxide (Fig. 2) and hydrous aluminum silicates, feldspars dissolve incongruently with the formation of a secondary phase such as gibbsite or kaolinite.

Incongruent solution

Incongruent solution occurs when dissolution of the primary phase is accompanied by formation of a secondary solid phase. For example the alteration of albite to gibbsite

$$NaAlSi_3O_8 + H^+ + 7\,H_2O = Na^+ + Al(OH)_3 + 3\,H_4SiO_4$$

or kaolinite

$$2\,NaAlSi_3O_8 + 2H^+ + 9\,H_2O = 2Na^+ + Al_2Si_2O_5(OH)_4 + 4\,H_4SiO_4$$

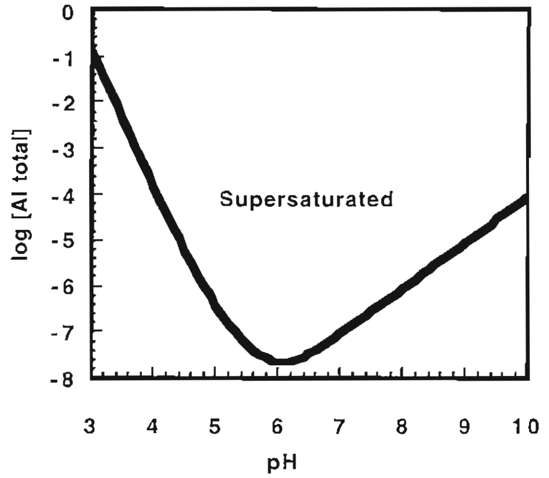


Figure 2: Solubility of gibbsite [$Al(OH)_3$]. Note the low solubility at near-neutral pH and the rapid increase in solubility with decreasing pH. The solubility curve for kaolinite has a similar shape; its vertical position depends on the activity of silica in solution (see Drever, 1988).

The identity of the secondary phase can be predicted approximately by thermodynamic equilibrium calculations (Fig. 3). At low silica activities an aluminum hydroxide phase is likely to form (gibbsite or a disordered or amorphous equivalent). At higher silica activities, a kaolinite-type mineral is likely to form. In more concentrated solutions (high silica activity, relatively high pH), a smectite-type mineral may form. Smectite is a general term for expandable clay minerals, equivalent to montmorillonite in some older literature. A simple example would be beidellite

$$(X)_{0.3}Al_2(Si_{3.7}Al_{0.3})O_{10}(OH)_2.nH_2O$$

where X represents exchangeable cations. Natural smectites are highly variable in composition, commonly containing some Mg.

Although natural systems generally conform qualitatively to these predictions, exceptions do occur, and the precise positions of the lines on Fig. 3 are subject to considerable uncertainty.

We can distinguish two types of incongruent dissolution: in the first type, exemplified by the alteration of albite given above, the primary phase dissolves completely and the secondary products are completely different phases that are precipitated from solution. The second type occurs when ions are leached from a solid phase, but the structure of the solid phase is retained, with a slightly different chemistry. An example would be the alteration of biotite to vermiculite. Biotite can be represented by the formula

$$K(Mg_{1.5}Fe_{1.5})AlSi_3O_{10}(OH)_2$$

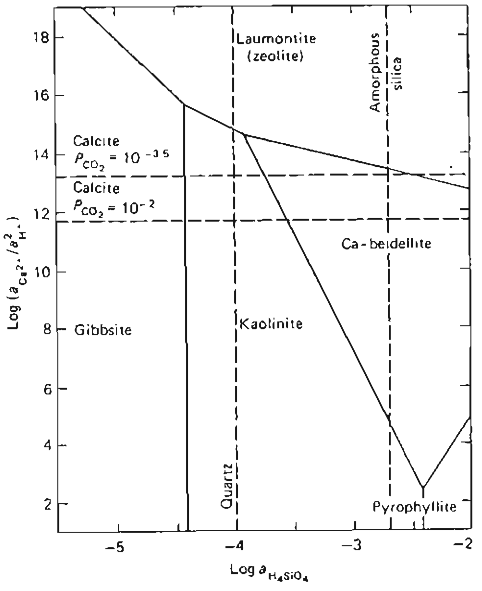


Figure 3: Stability relationships among some minerals in the system CaO-Al_2O_3-SiO_2-H_2O at 25 °C. Vertical dashed lines are solubilities of silica phases; horizontal dashed lines are solubility of calcite at the P_{CO_2} values indicated.

When it weathers, some ferrous iron is commonly oxidized to Fe(III). To maintain charge balance, K^+ is lost to solution. The process opens up the mica layers and additional K^+ is lost, being replaced by other ions (typically Ca^{2+} or Mg^{2+}) from solution, resulting in a vermiculite. The fundamental layer structure of the mica is retained, but its chemical composition is changed. As a general statement, incongruent reactions of the first type are relatively easy to model quantitatively, reactions of the second type are much more difficult because thermodynamic equilibrium is rarely attained and stoichiometry is usually not well defined.

CATION EXCHANGE

Clay minerals and oxides undergo cation exchange reactions. Ions loosely held by electrostatic forces at the surface of the solid are readily exchanged

$$2Na^+ + Ca\text{—clay} = 2Na\text{—clay} + Ca^{2+}$$

When a soil is "acidified", base cations on the exchange sites are replaced by hydrogen ions or aluminum ions. For hydrogen ions, the reaction can be written

$$2H^+ + Ca\text{—clay} = 2H\text{—clay} + Ca^{2+}$$

Aluminum uptake is usually a two-stage process: hydrogen ions are consumed and Al ions released as acid water interacts with an aluminum hydroxide phase

$$Al(OH)_3 + 3H^+ = Al^{3+} + 3H_2O$$

and the Al^{3+} displaces base cations from exchange sites

$$2Al^{3+} + 3Ca\text{—clay} = 2Al\text{—clay} + 3Ca^{2+}$$

Soils in high-latitude forests are very commonly acidified in the sense of having exchange sites in the soil dominated by H^+ and Al^{3+}. Acidification can occur as a result of natural processes. It does not have to be related to anthropogenic emissions.

In this discussion, I shall maintain a distinction between ion exchange reactions and chemical weathering and specifically exclude ion exchange reactions from my definition of weathering. Ion exchange reactions are typically fast, but limited by the number of exchange sites available in a soil. Weathering reactions are typically slow, but the amount of material available for reaction (minerals in the bedrock) is typically very large. From a modeling perspective, ion exchange can be treated as an equilibrium system, whereas weathering is modeled in terms of kinetics.

MINERAL DISSOLUTION KINETICS

Relative rates

There has been a large amount of work done in the laboratory on the rates at which minerals dissolve (for reviews, see Sverdrup, 1990; papers in Stumm, 1990). A simple way (Lasaga et al., 1994) of comparing rates is to calculate the time it would take for a 1 mm sphere of each mineral to dissolve completely in a dilute solution at pH 5 (Table 1). The rates vary enormously. Minerals such as quartz and kaolinite dissolve extremely slowly; feldspars and mafic minerals at intermediate rates, and carbonates very rapidly.

One implication of these differences is that weathering should be highly selective. If a rock contains several minerals, they will not all dissolve at the same rate. Any carbonate minerals will dissolve first, followed calcic feldspars and ferromagnesian minerals, followed in turn by alkali feldspars. We will see examples of this in the discussion of catchment mass balance(chap. IV.1)

Table 1: Approximate time for a hypothetical sphere of various minerals to dissolve in dilute solution at pH 5 (modified from Lasaga et al., 1994)

Mineral	Lifetime (y)
Quartz	34,000,000
Kaolinite	6,000,000
Muscovite	2,600,000
Epidote	923,000
Microcline	921,000
Albite	575,000
Sanidine	291,000
Gibbsite	276,000
Enstatite	10,100
Diopside	6,800
Forsterite	2,300
Nepheline	211
Anorthite	112
Wollastonite	79
Dolomite	1.6
Calcite	0.1

Also, weathering reactions in a catchment will evolve through time. When weathering starts, commonly at the end of glaciation, rapidly-weathering minerals such as calcite will be the dominant source of solutes, and the overall weathering rate will be relatively rapid. Over time, these reactive minerals will be depleted in the near-surface environment, and weathering will be dominated by less reactive but more abundant phases such as feldspars.

Effect of solution composition on dissolution rates of silicate minerals

pH

Far from equilibrium, the dissolution rates of most silicate minerals have the form shown in Figure 4. In dissolution experiments, the units are moles per unit surface area of the mineral per unit time. In the acid region, the rate increases exponentially with increasing hydrogen ion concentration, or

$$(\mathrm{Rate})_H = k_H[H^+]^n \qquad \text{(acid region)}$$

where k_H is a rate constant and n an exponent that is different for different minerals. n is typically about -0.5; some values are shown in Table 2. An exponent of -0.5 means that for a change in pH of one unit, the rate should change by a factor of about 3 ($=10^{-0.5}$).

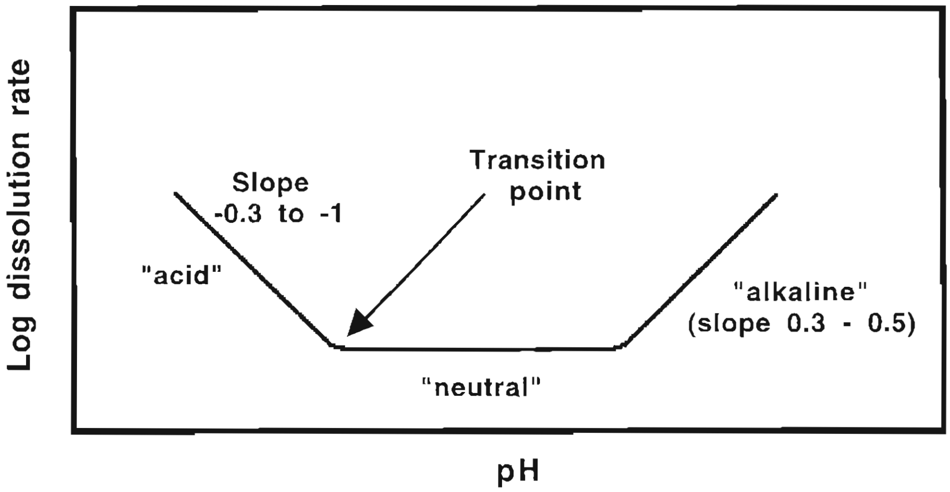


Figure 4: Schematic representation of silicate dissolution rates far from equilibrium as a function of pH (after Drever, 1994).

Table 2: pH of the transition points (See Fig. 4) and slope of log (Rate) vs. pH graph in the acid region for silicate minerals (after Drever, 1994).

Mineral	Transition pH (acid)	Slope in acid region	Transition pH (basic)	Source
Albite	4.5	-0.5	7.5	Sverdrup (1990)[1]
Oligoclase	4.5	-0.5		Oxburgh et al, (1994)
Andesine	4.5	-0.5		Oxburgh et al, (1994)
Bytownite	5	-0.75		Oxburgh et al, (1994)
Anorthite	4.5	-3		Amrhein & Suarez (1988)
K-feldspar	5	-0.5		Schweda (1990)
Forsterite	7[2]	-0.9	7[2]	Sverdrup(1990)[1]
Forsterite	4.5 (?)	-0.6	8(?)	Blum & Lasaga (1988)
Garnet	5.5	-0.9	8	Sverdrup (1990)[1]
Amphiboles	5.5	-0.8		Sverdrup (1990)[1]
Amphibole	<3			Mast & Drever (1987)
Diopside, Augite	ca. 6	-0.7 to -0.9	ca. 8	Sverdrup (1990)[1]
Phlogopite, Biotite		-0.4		Sverdrup (1990)[1]
Kaolinite	4	-0.4		Wieland & Stumm (1992)

[1] Sverdrup (1990) represents both original work and a compilation of data from the literature.

[2] No pH-independent region.

Above some transition pH, typically 4-5 (Table 2), dissolution rates are generally independent of pH,

$$(\text{Rate})_{\text{neutral}} = k_{\text{N}} \qquad \text{(neutral region)}$$

and then at pH values above 8 or so, they increase with increasing pH

$$(\text{Rate})_{\text{OH}} = k_{\text{OH}}[\text{OH}^-]^{m} \qquad \text{(basic region)}$$

where m typically has a value of about 0.3 to 0.5 (Brady and Walther, 1989).

Rates of dissolution are commonly explained by *transition state theory*. This approach was originally developed to explain the kinetics of gas-phase reactions involving simple molecules. Its extension to silicate dissolution kinetics involves a certain amount of faith. In brief, the rate of dissolution is assumed to be proportional to the concentration of an activated complex at the surface of the solid. The precursor to the activated complex is a *surface complex*, which is formed when an ion (or neutral molecule) is adsorbed from solution onto the mineral surface, forming a complex involving a metal in the solid and the adsorbed species. Adsorption of the solute weakens the bonds between the metal and the bulk of the solid, accelerating its detachment. The rate of dissolution is proportional to the concentration of the surface complex. In the discussion of pH dependence above, adsorption of protons results in a protonated surface complex, and the rate of dissolution is proportional to the concentration of this species. The exponent n reflects either a non-linear dependence of adsorbed protons on proton concentration in solution or the stoichiometry of the surface complex. A similar logic applies to formation of a surface complex involving OH^- (or deprotonation of a surface –OH group) under basic conditions.

Following Stumm and Wieland (1990), the total rate can be expressed as the sum of the individual rates

$$(\text{Rate})_{\text{total}} = k_H[H^+]^n + k_N + k_{OH}[OH^-]^m$$

Organic ligands

Organic ligands may accelerate the dissolution rates of silicates by forming complexes with metals, particularly aluminum on the mineral surface, and thereby weakening the bonds between the metal and the solid. The effect can be described by the same equations as were used for proton adsorption

$$(\text{Rate})_{\text{ligand}} = k_L[\text{ligand}]^p$$

and

$$(\text{Rate})_{\text{total}} = k_H[H^+]^n + k_N + k_{OH}[OH^-]^m + k_L[\text{ligand}]^p$$

The exponent p again reflects the stoichiometry of the complex and the relationship between adsorbed concentration and concentration in solution. The overall effect of organic ligands on dissolution rate will depend on how the magnitude of the term $k_L[\text{ligand}]^p$ compares to those of the other terms in the equation. There has been some controversy over the quantitative importance of organic ligands in nature (Drever, 1994). High concentrations of chelating ligands such as oxalate certainly accelerate the dissolution rates of silicates far from equilibrium, but the effect of organic ligands *at the concentrations commonly observed in soil solutions* on dissolution rate is probably minimal. The effects in microenvironments adjacent to roots and fungal hyphae, however, may be much greater.

The effect of organic ligands on state of saturation is a separate issue discussed below.

Chemical affinity and other ions in solution

As chemical reactions approach equilibrium the rate of the back reaction, here reprecipitation of the original solid phase, become finite. Thus the net rate of reaction decreases as equilibrium is approached and ultimately becomes zero at equilibrium, where the rate of the back reaction exactly equals the rate of the forward reaction. The question is how close to equilibrium to be before this effect becomes significant. This is a controversial question which has not been fully resolved in the literature. Transition state theory gives, *for a simple, elementary reaction*, (Aagaard and Helgeson, 1982; Velbel, 1989)

$$\text{Rate} = k_a[1 - \exp(\Delta G_r/\sigma RT)]$$

where k_a is the apparent forward rate constant, σ is a "stoichiometric number", conventionally set equal to 1, R is the gas constant and T temperature.(K) ΔG_r is the departure from equilibrium, or *chemical affinity*, in units of kJ or kcal per mole. It is related to the activity product for the reaction by

$$\Delta G_r = RT \ln (Q/K_{eq})$$

where Q is the activity product and K_{eq} the equilibrium constant. Thus for dissolution of albite in acid solution

$$NaAlSi_3O_8 + 4H^+ + 4H_2O = Na^+ + Al^{3+} + 3H_4SiO_4$$

the activity product, Q, is given by

$$Q = \frac{[Na^+][Al^{3+}][H_4SiO_4]^3}{[H^+]^4}$$

where [ ] represent activities in solution. For undersaturated solutions, the value of Q is less than the equilibrium constant, K_{eq}.

The term $\exp(\Delta G_r/\sigma RT)$ represents the "back reaction"–in this case reprecipitation of the primary mineral. It will be 5% of the forward reaction if

$$\exp(\Delta G_r/\sigma RT) = 0.05 \qquad \Delta G_r/RT = -3 \text{ (assuming } \sigma = 1)$$

i.e. if $\Delta G_r = -7.45$ kJ (-1.78 kcal), or $\log_{10}(Q/K_{eq}) = -1.3$ at 25°C

If the departure from equilibrium is greater than this, the back reaction should be negligible and rate should be independent of affinity of reaction. Most solutions in the weathering environment should be much farther from saturation with respect to most important primary minerals other than potassium feldspars than $\log_{10}(Q/K_{eq}) = -1.3$ (Velbel, 1989) so, according to this analysis, chemical affinity should have no effect on the dissolution rates of most primary minerals in the weathering environment. [The difference in weathering rate between K-feldspars and albite in nature may in fact reflect decreased dissolution of K-feldspar resulting from the chemical affinity effect.] The assumption has often (naively) been made that dissolution rates of plagioclase and ferromagnesian minerals in the field depend only on pH, and other aspects of solution composition could be ignored.

The problem with this analysis is that it assumes that mineral dissolution behaves as a simple elementary reaction. In experiments conducted by Burch et al (1993) at 80 °C, the dissolution rate of albite seemed to be reduced by chemical affinity effects at much higher degrees of undersaturation than $\log_{10}(Q/K_{eq}) = -1.3$. Oelkers et al. (1994), in experiments at 150 °C, showed that the dissolution rates of albite at pH 9 and kaolinite at pH 2 were strongly affected by dissolved Al concentrations. This could be interpreted in terms of a chemical affinity effect or as a specific inhibition caused by Al. It is difficult to separate the two effects. Dissolved Al concentration may affect reaction rate either through its effect on chemical affinity or through some other mechanism that is independent of chemical affinity. A possible mechanism would be competition with H^+ for adsorption sites (Murphy and Helgeson, 1987; Schweda, 1990; Brantley and Stillings, 1994). Conceptually, adsorption of Al^{3+} would decrease adsorption of H^+. If dissolution rate were a strong function of adsorbed H^+ concentration, then increased Al^{3+} concentrations would result in decreased dissolution rates. Similarly, at high pH, $Al(OH)_4^-$ might compete with $(OH)^-$.

There is also evidence that the results of Oelkers et al. (1994) and Burch et al. (1993) may not be generally applicable to dissolution reactions at or below 25 °C. Mast and Drever (1987) showed that oxalate ion at a concentration of 10^{-3} M had no effect on the dissolution rate of oligoclase feldspar. [Higher concentrations of oxalate do affect oligoclase dissolution rate, and 10^{-3} M oxalate has some effect on dissolution rates of more calcic feldspars (Welch and Ullman, 1993; Sun, 1994).] Oxalate is a strong chelator of Al. The presence of 10^{-3} M oxalate will have an enormous effect on the activity of free Al species and hence also on the affinity of reaction. The lack of a strong oxalate effect indicates a lack of sensitivity to dissolved Al activity or chemical affinity under the conditions of these experiments. In batch experiments on albite dissolution at room temperature, Murphy (1993) observed no significant decrease in rate until $\log_{10}(Q/K_{eq})$ had a value of about -4.7. This is farther from equilibrium than the simple model would predict, but is much closer to equilibrium than suggested by the results of Burch et al. and Oelkers et al..

In summary, the effect of solution composition, particularly chemical affinity and Al concentration, on silicate dissolution rates is not well understood. Models of weathering rates in the field have generally assumed that pH was the only important variable. This may be adequate in very dilute solutions but is certainly not adequate as solutions become more concentrated and equilibrium with primary minerals is approached.

COMPARISONS BETWEEN FIELD AND LABORATORY DISSOLUTION RATES

In order to apply laboratory-based dissolution rates to weathering in the field, we need to know the surface areas of all minerals exposed to weathering in the catchment (or soil profile) under study. This is not an easy number to obtain. In catchment-scale studies, Paces (1983) and Velbel (1985) came up with areas based on plausible fracture patterns and an assumed grain-size for minerals in the soil. They estimated that rates in the field were 1 to 3 orders of magnitude slower than predicted by laboratory experiments. Although their estimates of surface area were crude, they were conservative and probably underestimated the actual wetted surface area of minerals in the catchment and hence overestimated the weathering rate in the field. Swoboda-Colberg and Drever (1993) attempted to measure weathering rates in plot-scale ($2m^2$) experiments at a forested site in Maine (USA). They artificially irrigated the plots with acid and collected soil solutions at 25 and 50 cm depths. They also measured the mineralogy and grain-size distribution of the soil to get an accurate estimate of surface area, and they measured the dissolution rates of minerals from the soil in laboratory reactors. From these experiments they concluded that

1. Rates in the field were a factor of about 200 slower than rates measured in the laboratory with the same minerals.

2. Untreated minerals from the soil dissolved at rates similar to those of crushed fresh "pure" minerals. Aging of mineral surfaces in the soil was thus not the cause of the discrepancy between field and lab.

Clow (1992) conducted somewhat similar experiments in a small high-elevation catchment on granite in Colorado (USA), using natural precipitation and distilled water rather than acid. He came to the same conclusions: field rates were much slower than laboratory rates, and differences between fresh and weathered minerals was not the cause of the discrepancy.

Schnoor (1990) reviewed published studies on field and laboratory dissolution rates and made the interesting observation that silica concentrations, a measure of silicate weathering, were generally very similar (within an order of magnitude), about 60 μM. Water fluxes and exposed mineral surface areas varied over a far wider range. This suggests that dissolution rates in the field are not controlled by surface reactions at high degrees of undersaturation, as is the case in laboratory experiments, but by either some sort of saturation effect or by a transport process. Drever et al (1994) visualized the control in terms of two end-members, based on the idea that the porosity in soils and shallow aquifers was very heterogeneous. Under rapid flow conditions, solutions would move largely through high-permeability channels. Solutions in low-permeability channels would have a relatively long residence time; pH values would be higher and affinities of reaction lower than in the large channels. Weathering rate would be determined by reactions in the large channels and by diffusional transport between the large channels and the soil matrix. The other end-member would be represented by slow flow conditions. Concentrations would be more spatially uniform but, because of the implied long contact time, solutions would be more concentrated and closer to saturation with respect to primary phases. Weathering rate would be limited by the affinity of reaction or a related chemical effect. The situation is more complicated in the unsaturated zone because high flow rates are also associated with a flushing effect.

MODELING APPROACHES

The main efforts to model chemical weathering quantitatively have been in conjunction with integrated models for predicting the effect of acid deposition on surface water chemistry. Some sort of integrated model is essential in order to calculate the pH of soil solutions, which is generally not the same as the pH of incoming precipitation. Chemical weathering is very important in this context because it is the dominant long term (decades or more) control on whether or not a water body becomes acidified under a given acid loading. The short-term variations in the chemistry of a stream or lake, however, are related to changes in input, adsorption processes and biological processes rather than to chemical weathering. What this means is that short-term data (which tend to be the only detailed data available) provide very little information that can be used to calibrate a weathering model. The most common approach (e.g. the MAGIC model, Cosby et al., 1985; 1990) has been to assume that "weathering" is a single, lumped process that can be described by an equation such as

$$\text{Rate} = [H^+]^n$$

where n is selected by the operator. This expression is plausible in light of the experimental results discussed above. If, however, rates in the field are effectively determined by transport processes and not surface reaction, the equation (and/or the choice of n) may not be appropriate. One model that does treat mineral dissolution kinetics explicitly is the PROFILE model (Sverdrup and Warfvinge, 1988; Altveg, Warfvinge, and Sverdrup, 1993). In this model, the mineralogy of each soil layer is input, as is an equation describing the dependence of dissolution rate for each mineral on pH and dissolved Al concentration. The only major "free parameter" is the specific surface area of the minerals in each soil layer, which must be supplied as an input. The model has been quite successful in describing the chemistry of surface waters in Scandinavia.

The models of Henriksen (Henriksen , 1984; Brakke et al., 1990) and Kirchner (1993) do not consider weathering explicitly, weathering is included implicitly as part of alkalinity balance.

The problem with essentially all these models is that they are calibrated using short-term data. They should thus, in principle at least, describe well the short term behavior of the system and the underlying controlling processes. However, because the processes controlling the long-term behavior of the system (chemical weathering) is not really constrained by the short-term calibration procedure, there is no way of knowing whether or not the long-term predictions are accurate.

REFERENCES

AAGAARD P. AND HELGESON H.C. (1982) Thermodynamic and kinetic constraints on reaction rates among minerals and aqueous solutions, I. Theoretical considerations. *Am. J. Sci.* **282**, 237-285.

ALTVEG M., WARFVINGE P., AND SVERDRUP H. (1993) Profile 3.2: User's guidance for the Apple Macintosh version. Unpublished manuscript, Dept. of Chemical Engineering II, Chemical Center, P.O. Box 124, S-221 00 Lund, Sweden.

AMRHEIN C. and SUAREZ D.L. (1988) The use of a surface complexation model to describe the kinetics of ligand-promoted dissolution of anorthite. *Geochim. Cosmochim. Acta* **52**, 2785-2793.

AMRHEIN C., AND SUAREZ D.L. (1992) Some factors affecting the dissolution kinetics of anorthite at 25°C. *Geochim. Cosmochim. Acta* **56**, 1815-1826.

BLUM A.E. and LASAGA A.C. (1988) The role of surface speciation in the low-temperature dissolution of minerals. *Nature* **331**, 431-433.

BRADY P.V, AND WALTHER J.V. (1989) Controls on silicate dissolution rates in neutral and basic pH solutions at 25°C. *Geochim. Cosmochim. Acta*, **53**, 2823-2830.

BRAKKE D.F., HENRIKSEN A., AND NORTON S.A. (1990) A variable F-factor to explain changes in base cation concentrations as a function of strong acid deposition. *Verh. Internat Verein. Limnol.* **24**, 146-149.

BRANTLEY S.L. AND STILLINGS L. (1994) An integrated model for feldspar dissolution under acid conditions. *Mineralogical Magazine* **58A**, 117-118.

BURCH T.E., NAGY K.L. AND LASAGA A.C. (1993) Free energy dependence of albite dissolution kinetics at 80°C and pH 8.8. *Chemical Geology* **105**, 137-162.

CLOW D.W. (1992) *Weathering rates from field and laboratory experiments on naturally weathered soils.* Unpublished PhD dissertation, Univ. of Wyoming, Laramie, WY, USA.

COSBY B.J., HORNBERGER G.M., GALLOWAY J.N., AND WRIGHT R.F. (1985) Modeling the effects of acid deposition: assessment of a lumped parameter model of soil water and stream chemistry. *Water Resources Res.* **21**, 51-63.

COSBY B.J., JENKINS A., FERRIER R.C., MILLER J.D., AND WALKER T.A.B. (1990) Modelling stream acidification in afforested catchments: long-term monitoring at two sites in central Scotland. *J. Hydrol.* **120**, 143-162.

DREVER J.I. (1988) *The Geochemistry of Natural Waters*, 2nd Edition. Englewood Cliffs, Prentice-Hall, 437 p.

DREVER J.I. (1994) The effect of land plants on the weathering rates of silicate minerals. *Geochim. Cosmochim. Acta* **58**, 2325-2332.

DREVER J.I., MURPHY K.M., AND CLOW D.W. (1994) Field weathering rates versus laboratory dissolution rates: an update. *Mineralogical Magazine* **58A**, 239-240.

HENRIKSEN A. (1984) Changes in base cation concentrations due to freshwater acidification. *Verh. Internat. Verein Limnol.* **22**, 692-8.

KIRCHNER J.W. (1992) Heterogeneous geochemistry of catchment acidification. *Geochim. Cosmochim. Acta* **56**, 2311-2327.

LASAGA A.C., SOLER J.M., CANOR J., BURCH T.E., AND NAGY K.L. (1994) Chemical weathering rate laws and global geochemical cycles. *Geochim. Cosmochim. Acta* **58**, 2361-2386.

MAST M.A. and DREVER J.I. (1987) The effect of oxalate on the dissolution rates of oligoclase and tremolite. *Geochim. Cosmochim. Acta* **51**, 2559-2568.

MURPHY K.M. (1993) *Kinetics of albite dissolution: the effect of grain size.* Unpublished M.S. thesis, Univ. of Wyoming, Laramie, WY, USA.

MURPHY W. M. AND HELGESON H. C. (1987) Thermodynamic and kinetic constraints on reaction rates among minerals and aqueous solutions. III. Activated complexes and the pH-dependence of the rates of feldspar, pyroxene, wollastonite, and olivine hydrolysis. *Geochim. Cosmochim. Acta* **51**, 3137-3153.

OELKERS E.H., SCHOTT J., AND DEVIDAL J-L. (1994) The effect of aluminum, pH, and chemical affinity on the rates of aluminosilicate dissolution reactions. *Geochim. Cosmochim. Acta* **58**, 2011-2024.

OXBURGH R., DREVER J.I., AND SUN Y-T. (1994) Mechanism of plagioclase dissolution in acid solution at 25°C. *Geochim. Cosmochim. Acta* **58**, 661-669.

PACES T. (1983) Rate constants of dissolution derived from measurements of mass balance in hydrologic catchments. *Geochim. Cosmochim. Acta* **47**, 1855-1863.

SCHNOOR J.L. (1989) Kinetics of chemical weathering: A comparison of laboratory and field weathering rates. In Stumm W., ed., *Aquatic Chemical Kinetics.* New York, Wiley-Interscience, pp. 475-504.

SCHWEDA P. (1990) *Kinetics and mechanisms of alkali feldspar dissolution at low temperatures.* PhD dissertation, Dept. of Geology and Geochemistry, Stockholm University, Sweden.

STUMM W. (1990) *Aquatic Chemical Kinetics.* New York, Wiley-Interscience, 545 p.

STUMM W. and WIELAND E. (1990) Dissolution of oxide and silicate minerals: Rates depend on surface speciation. In Stumm W., ed., *Aquatic Chemical Kinetics.* New York, Wiley-Interscience, pp. 367-400.

SUN Y. (1994) *The effect of pH and oxalate ion on the dissolution rates of plagioclase feldspar.* Unpublished M.S. thesis, Univ. of Wyoming, Laramie, WY, USA.

SVERDRUP H.U. (1990) *The Kinetics of Base Cation Release due to Chemical Weathering.* Lund University Press, 246 p.

SVERDRUP H. AND WARFVINGE P. (1988) Weathering of primary silicate minerals in the natural soil environment in relation to a chemical weathering model. *Water, Air, and Soil Pollution* **38**, 387-408.

SWOBODA-COLBERG N.G AND DREVER J.I. (1993) Mineral dissolution rates in plot-scale field and laboratory experiments. *Chemical Geology* **105**, 51-69.

VELBEL M.A. (1985) Geochemical mass balances and weathering rates in forested watersheds of the southern Blue Ridge. *Am. J. Sci.* **285**, 904-930.

VELBEL M.A. (1989) Effect of chemical affinity on feldspar hydrolysis rates in two natural weathering systems. *Chemical Geology* **78**, 245-253.

WELCH S.A. AND ULLMAN W.J. (1993) The effect of organic acids on plagioclase dissolution rates and stoichiometry. *Geochim. Cosmochim. Acta* **57**, 2725-2736.

WIELAND E. and STUMM W. (1992) Dissolution kinetics of kaolinite in acid aqueous solutions at 25°C. *Geochim. Cosmochim. Acta* **56**, 3339-3363.

Introduction to soil development processes

Ole K. Borggaard
Chemistry Department, Royal Veterinary and Agricultural University,
40 Thorvaldsensvej, DK-1871 Frederiksberg C, Denmark

LIST OF CONTENTS

INTRODUCTION

Soils are dynamic, open systems with influxes and effluxes of matter and energy. Many soil-forming (and destroying) processes proceed simultaneously, and the resulting soil reflects the balance of these processes. Five factors (climate, organisms, relief, parent material and time) drive the internal processes within the soil. Distinct regimes or combinations of processes produce distinctive soils with distinct compositions and properties. True equilibrium is nearly never attained but soil development may reach a steady-state. The soil processes form soil from not-soil but the processes are themselves functions of the soil change, because they go on in a continuing changing environment (soil). Some processes may initiate other processes. Processes involving very reactive materials are fast, while more inert materials react slowly.

The composition and properties of a soil can be measured or estimated but it will, of course, only give a snapshot of the soil due to its dynamic nature. Since present days soils carry the imprint of processes that was active in the past, such information is fundamental for the interpretation and understanding of soil formation (genesis) and of the processes involved. Knowledge about soil development and processes is, in turn, very important in forecasting future soil development, and hence composition and properties, under specified but maybe changing environmental conditions. Such information is basic to sustainable soil use and management.

Soil classification aims at organizing (and mapping) different soils in different groups or classes according to their composition and properties. The soils belonging to one class have specified, alike composition and properties different from those of the soils in the other classes. Since soil composition and properties, and hence classification, are functions of soil genesis, understanding of and knowledge about soil genesis is very useful for interpretation and applicability of the results of soil classification. On the other hand, soil classification systems should, however, not be based entirely on genesis (as in older soil classification systems), because soil formation processes can seldom be observed and because the processes operating within a soil have been subject to changes over time.

This paper aims at giving a short overview on soil classification, composition, formation, processes, and properties with emphasis on aspects relevant to soils in the Nordic countries, Denmark, Finland, Norway and Sweden. The soil classification will be according to the American Soil Taxonomy and FAO/Unesco systems (FAO/Unesco, 1990; Soil Survey Staff, 1992). More information on soils in general can be found in numerous textbooks (e.g. Brady, 1990; Foth & Turk, 1972; Schachtschabel et al., 1989; Singer & Munns, 1987; Tan, 1994), whereas the books of Bohn et al. (1985) and Sposito (1989) cover soil chemistry and those of Buol et al. (1989), Duchaufour (1982), Fanning & Fanning (1989) and Wilding et al. (1983) deal with pedology (soil genesis and classification). Those interested in soil analysis and sampling may find appropriate information in FAO (1990), Hodgson (1978), Klute (1986), Page et al. (1982), Rowell (1994) and Westerman (1990).

DEFINITIONS AND TERMINOLOGY OF SOILS

Soil is a term understood by almost everyone but the meaning of it varies among different people, e.g. the farmer, the engineer, the forester and the gardener think of soil in different ways and for different purposes, a thinking that deviates from that of the soil scientist. From a soil science point of view soil is, according to the American soil classification system (Soil Survey Staff, 1992), defined in the following way:

Soil is the collective term used in this text for the natural bodies, made up of mineral and organic materials, that cover much of the earth's surface, contain living matter and can support vegetation out of doors, and have in places been changed by human activity. The upper limit of soil is air or shallow water. Its horizontal boundaries are where it grades to deep water or to barren areas of rock or ice. The lower boundary that separates soil from the not-soil underneath is most difficult to define. Soil consists of the horizons near the earth's surface which, in contrast to the underlying rock material, have been altered by the interactions, over time, between climate, relief, parent materials, and living organisms. In the few places where it contains thin cemented horizons that are impermeable to roots, soil is considered to be as deep as the deepest cemented horizon. More commonly, soil grades at its lower boundary to hard rock or to earthy materials virtually devoid of animals, roots, or other marks of biologic activity. Thus the lower limit of soil is normally the lower limit of biologic activity, which generally coincides with the common rooting depth of native perennial plants. If, however, either biological activity or current pedogenic processes extend to depths greater than 200 cm, the lower limit of the soil that we classify is arbitrarily set at 200 cm. For certain soil management purposes, layers deeper than the lower boundary of the soil that we classify must also be described if they affect the content and movement of water and air in the soil of the root zone.

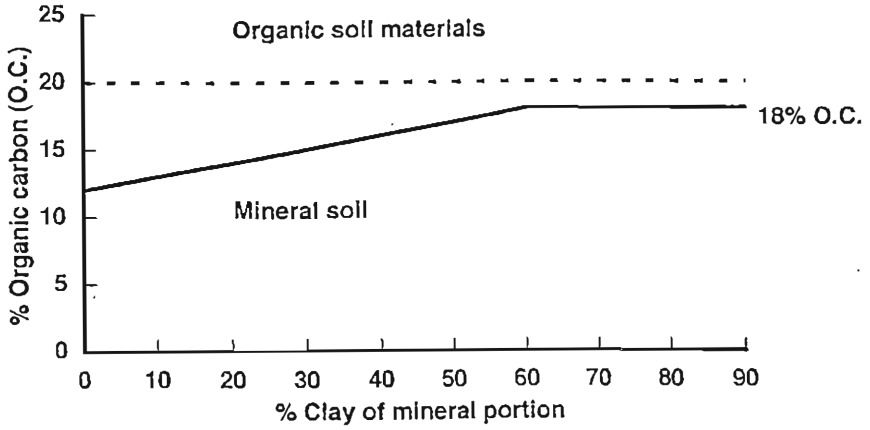


Figure 1. Definitions of mineral and organic soil materials by Soil Survey Staff (1992). Solid lines are for materials usually saturated with water (or artificially drained) and the dashed line is for usually dry materials.

The distinction between mineral and organic materials can be seen in Figure 1. Active sand dune, glaciers, bare rock and the like are not-soil. For a land surface to be recognized as a soil it must be capable of supporting vegetation but plants need not actually be present (Fanning & Fanning, 1989). Soils buried deeper than 50 cm are considered buried soils (paleosols, fossil soils).

A soil is a 3-dimensional component of the landscape. The pedon is the smallest volume that can be called a soil and covers an area of 1-10 m^2 depending on the variability of the soil layers or horizons (Buol et al., 1989; Fanning & Fanning, 1989). The 2-dimensional vertical surface of a pedon from top to bottom is called a soil profile (Figure 2). Contiguous pedons all falling within the defined range of the lowest category of soil classification (see section 3) is a true individual soil body and is termed a polypedon.

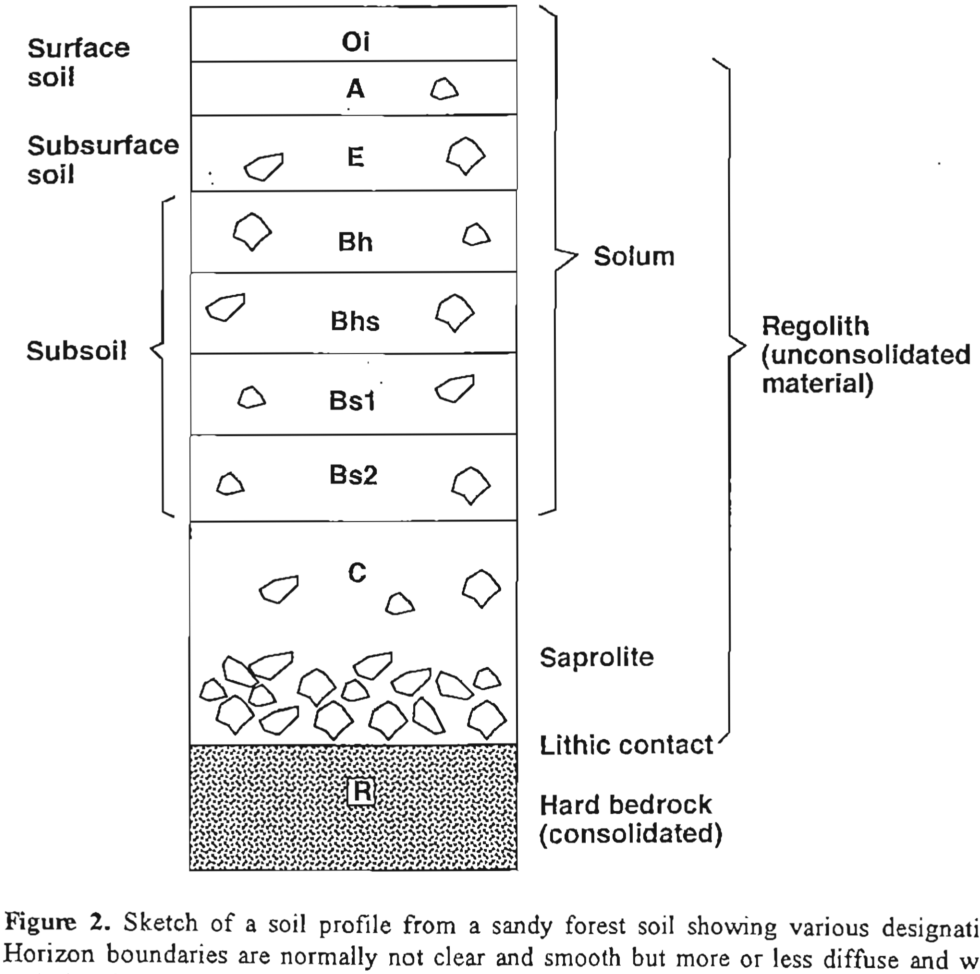


Figure 2. Sketch of a soil profile from a sandy forest soil showing various designations. Horizon boundaries are normally not clear and smooth but more or less diffuse and wavy. Strictly, the term surface soil (or topsoil) is restricted to the plow layer (Ap) in cultivated soils.

Table 1. Selected soil processes.

Term	Brief definition
Eluviation	Movement of material out of a portion of a soil profile as in an albic horizon.
Illuviation	Movement of material into a portion of a soil profile as in an argillic or spodic horizon.
Leaching	General term for washing out or eluviating soluble materials from the solum.
Enrichment	General term for addition of material to a soil body.
Erosion	Removal of material from the surface layer of a soil.
Cumulization	Aeolian and hydrologic additions of mineral particles to the surface of a soil solum.
Decalcification	Reactions that remove calcium carbonate from one or more soil horizons.
Lessivage	The mechanical migration of small mineral particles from the A and E to the B horizons of a soil, producing in B horizons relative enrichment in clay (argillic horizons).
Pedoturbation	Biologic, physical (freeze-thaw and wet-dry cycles) churning and cycling of soil materials, thereby homogenizing the solum in varying degrees.
Podzolization	The chemical migration of aluminum and iron and/or organic matter from the A and E to the B horizons, producing in B horizons enrichment in these materials (spodic horizons).
Littering	The accumulation on the mineral soil surface of organic litter.
Humification	The transformation of raw organic material into humus.
Mineralization	Complete decomposition of organic matter.
Gleization	The reduction of iron under anaerobic "waterlogged" soil conditions, with the production of bluish to greenish-gray matrix colours, with or without yellowish-brown, brown and black mottles, and concretions of iron and manganese oxides.

The soil definition points out the five soil forming factors, which according to Jenny (1941) determine the state of the soil:

Soils = f(cl, o, p, r, t)

This relation is called the "clorpt" concept, where cl is climate, o is organisms, p is parent material, r is relief and t is time. The fundamental processes including (i) humus formation, (ii) weathering and (iii) transport leading to formation of soil from not-soil and to changes of soil development are ruled by these factors. These fundamental processes are integrated in the processes, some of which are shortly described in Table 1.

Precise information about the soil forming factors at a certain soil site should be based on detailed climatic, topographic and geological maps and on carefull observations about land-use, vegetation etc. at the site. In relation to Nordic soils the soil forming factors can shortly be summarized in the following way:

Climate (cl): The socalled temperature regimes are mesic (mean annual soil temperature 8^0-15^0C) and frigid or cryic (mean annual soil temperature 0^0-8^0C, where summer temperature

is higher under frigid than cryic conditions). The socalled moisture regime is in most well-aerated soils udic (the soil is not dry in 90 cumulative days per year and is dry less than 45 consecutive days in the 4 months following the summer solstice), whereas in water-logged soils it will be aquic (a reducing regime that is virtually free of oxygen as shown by redoximorphic features). The complete definitions of temperature and moisture regimes are given by the Soil Survey Staff (1992).
Organisms (o): Can roughly be divided into two groups according to land-use as woodland and as arable land. Soils in the latter group are seriously affected by liming and fertilization. Different plants such as conifers and deciduous undoubtedly affect soils differently.
Parent material (p): Most soils are developed on glacial deposits originating from the last glaciation (Weichsel Glaciation). The deposits include glacial till and water sorted materials of different texture ranging from sand (gravel) to clay. Many shallow lakes have turned into peatland. Due to land uplift, where previously submerged material is now above sea level, the parent material in some parts, especially in northern regions, has been reworked by the sea. More information about geology of the Quarternary period may be found elsewhere (e.g. Ehlers, 1983; Rankama, 1965).
Relief (r): The relief exhibits great variability ranging from level or almost level areas through undulated land to very steep slopes on mountains.
Time (t): Soils in the Nordic countries may be considered to have developed after termination of the Weichselian Glaciation, i.e. within less than 10000-15000 years. Even the soils in the south-western part of Denmark that was not covered by ice during the last ice age are considered to have formed since its termination but on previously leached materials, which have been affected by freeze-thaw actions. Soils in northern Scandinacia and Finland, particularly those on uplifted land, are younger than 10000 years.

Genetic horizons

Two kinds of soil horizons (layers) are considered, genetic horizons and diagnostic horizons. They are not equivalent. While genetic horizon designations used in the socalled ABC system express qualitative judgments about believed changes (genesis), diagnostic horizons are quantitatively defined features used in soil classification according to the Soil Taxonomy (Soil Survey Staff, 1992) and the FAO/Unesco systems (FAO/Unesco, 1990). The definitions of the diagnostic horizons, which will be considered in more details in a later section, relay on combinations of field information (soil description) and laboratory measurements.

The ABC system is especially useful in making soil (profile) descriptions, e.g. according to the FAO guidelines (FAO, 1990). In the ABC system, horizon designations are based on field examination alone.

For the genetic master O, A, E, B and C horizons and R layers the designations are shortly described as follows:

O horizons: Layers dominated by organic material.
A horizons: Mineral horizons that formed at the surface or below an O horizon and are characterized by an accumulation of humified organic matter intimately mixed with the

mineral fraction or have properties resulting from cultivation, pasturing, or similar kinds of disturbance.
E horizons: Mineral horizons in which the main feature is loss of silicate clay, iron, aluminum, or some combination of these, leaving a concentration of sand and silt particles of quartz or other resistant materials.
B horizons: Horizons that formed below an O, A or E horizon and are dominated by: (1) evidence of removal of carbonates; (2) concentrations of sesquioxides or silicate clay; (3) formation of granular, blocky, or prismatic structure; or (4) combination of these.
C horizons: Horizons or layers, excluding hard bedrock, that are little affected by pedogenic processes and lack properties of O, A, E, or B horizons.
R layers: Hard bedrock including granite, basalt, quartzite and indurated limestone or sandstone that i sufficiently coherent to make hand digging impractical (hardness > 3 on Mohs scale).

Two kinds of transitional horizons occur. In one, the properties of an overlying or underlying horizon are superimposed on properties of the other throughout the transition zone (i.e. AB, BC). In the other, distinct parts that are characteristic of one master horizon are recognizable and enclose parts characteristic of a second recognizable master horizon (i.e. E/B, B/C). Examples: AB - a horizon with characteristics of both an overlying A horizon and an underlying B horizon, but which is more like the A than the B. E/B - a horizon comprised of individual parts of E and B horizon components in which the E component is dominant and surrounds the B materials.

An obvious change in the mineral material, e.g. as seen in a soil on glacial till covered by aeolean sand, is considered a lithological discontinuity and the underlying horizons are designated by a number prefix, e.g. A, E, 2B, 2C, if the discontinuity occurs af the E to B boundary. Lower-case letters are used as suffixes to designate specific kinds of master horizons and layers.

These subordinate distinctions are shortly described as follows:
a: highly decomposed organic matter
b: buried soil horizon
c: concretions or nodules
e: intermediately decomposed organic matter
f: frozen soil
g: strong gleying
h: illuvial accumulation of organic matter
i: slightly decomposed organic matter
k: accumulation of carbonates
m: strong cementation
n: accumulation of sodium
o: residual accumulation of sesquioxides
p: plowing or other disturbance
q: accumulation of silica
r: weathered or soft bedrock
s: illuvial accumulation of sesquioxides
t: accumulation of clay
v: plinthite

w: colour or structural B
x: fragipan character
y: accumulation of gypsum
z: accumulation of salts

An example of various horizon designations is shown in Figure 2.

Soil profile description

Soil classification according to the Soil Taxonomy and FAO/Unesco systems is based on field data (soil profile description) and laboratory data (FAO/Unesco, 1990; Soil Survey Staff, 1992). Apart from information about the soil forming factors at the site in question (geology, climate etc.) a soil profile description contains a careful description of the different soil horizons including colour, texture, structure, consistence, voids, cutans, roots, cementation, nodules/concretions, stoniness, faunal activity and horizon boundary (FAO, 1990; Hodgson, 1978; Table 2).

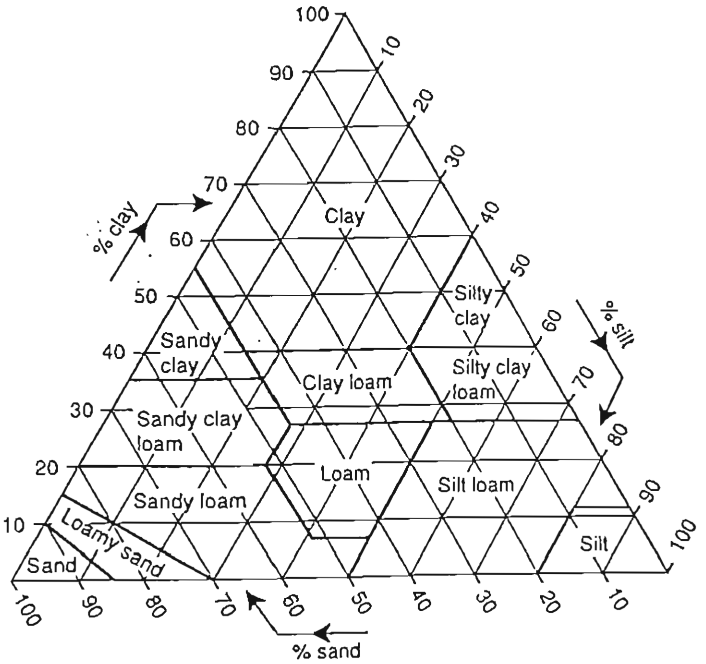


Figure 3. Soil textural classes according to Soil Taxonomy. Clay <2 μm, silt 2-50 μm, very fine sand 50-100 μm, fine sand 100-250 μm, coarse sand 250-2000 μm.

Table 2. Description of a profile (Klosterhede) from a Danish sandy soil in western Jutland.

Location: **Klosterhede** plantage (dep. 316).
UTM coordinates:: 32V MH 630589 (GI: Klosterhede Plantage 1115 IV SØ).
Moisture regime: Udic.
Temperature regime: Mesic.
Classification:
USDA: Typic Haplorthod, sandy, mixed/siliceous, mesic, noncemented.
FAO/Unesco: Haplic Podzol.
Elevation: 27 m.a.s.
Land-form: Almost flat.
Vegetation-Landuse: Picea abies (planted 1923).
Parent material: Glacial fluvial sand.
Drainage: Well drained.
Evidence of erosion: None.

Description of individual soil horizons
Oi (0-10)*: Very dark brown (10YR 2/2) mainly undecomposed needles and lichens.

A (10-25): Very dark gray (10YR 3/1, moist) sand; fine weak crums; non sticky, non plastic, friable, loose; many fine and medium interstitial voids; many, very fine, fine and medium roots; few, fine angular, slightly weathered stones (flint, granite); clear wavy boundary.

E (25-30): Dark gray (10YR 4/1, moist) sand; porous massive; non sticky, non plastic, friable; few fine and medium interstitial voids; few fine and medium roots; few, fine angular, slightly weathered stones (flint, granite); clear smooth and wavy boundary.

Bh (30-35): Black (5YR 2/1, moist) sand; porous massive; non sticky, non plastic, friable; broken thin humus cutans between and on mineral grains; few, fine and medium interstitial voids; many , very fine, fine and medium roots;fine angular, slightly weathered stones (flint, granit); gradual smooth boundary.

Bhs (35-45): Dark reddish brown (5YR 3/2, moist) matrix with few, medium to coarse, distinct, clear brownish yellow (10YR 6/8, moist) mottles; sand; porous massive; non sticky, non plastic, loose; broken thin humus and sesquioxide cutans between mineral grains; few fine and medium interstitial voids; common, very fine and fine roots; few, fine angular and rounded slightly weathered gravelsized particles (flint, gneiss); diffuse smooth boundary.

Bs1 (45-70): Dark brown (7,5YR 3/3, moist) matrix with few, medium, faint, diffuse, very dark grayish brown (10YR 3/2, moist) mottles; sand; porous massive; non sticky, non plastic, loose; broken thin sesquioxide cutans on mineral grains; few, fine interstitial voids; few, very fine roots; few, fine angular slightly weathered gravelsized particles (flint, gneiss); diffuse smooth boundary.

Bs2 (70-100): Yellowish brown (10YR 5/4, moist) sand; porous massive; non sticky, non plastic, loose; broken thin sesquioxide cutans on mineral grains; few, very fine interstitial voids; few, fine angular slightly weathered gravelsized particles (flint, gneiss); diffuse smooth boundary.

C (100-): Pale brown (10YR 6/4, moist) sand; porous massive; non sticky, non plastic, loose; few, very fine interstitial voids; few, fine angular slightly weathered stones (flint, gneiss granite).

---

*Very often the upper boundary is set at the boundary between the organic layer and the mineral soil, i.e. Oi(-10-0), A(0-15), E(15-20) etc.

Soil colours are mainly determined by free or non-silicate bound iron compounds, which give yellow, orange and red colours under aerobic conditions and green and blue colours in waterlogged soils together with the black and brown colours of manganese oxides and organic matter (Schachschabel et al., 1989). The colours of the soil matrix and mottles (if present) are measured by comparison with Munsell (1975) colour charts, which quantifies the colour by its hue (red, yellow, green, blue), value (lightness) and chroma (strength) arranged as (hue value/chroma), e.g. the notation of a yellowish brown coloured soil layer will be (10YR 5/6). The moisture content (dry, moist) is given since it affect the colour.

Soil texture classes of the fine earth fraction (<2 mm) are shown in Figure 3. Soil structure refers to the natural organization of soil particles into discrete soil units (aggregates, peds), which are separated from each other by persistent surfaces of weakness. Grade, class and type of these units is used in describing soil structure. Soil consistence is recorded in dry, moist and wet states and is as structure determined by abundance and composition of the organic and clay fractions.

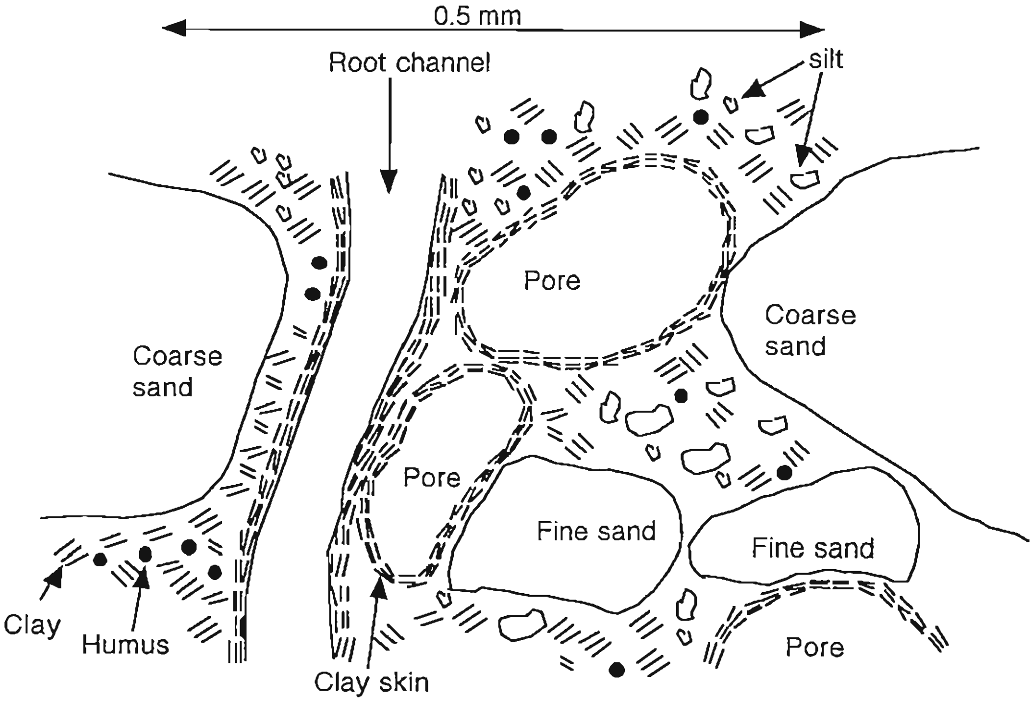


Figure 4. Thin section of soil showing various micromorphological features. Sand and silt (+ coarse organic matter) form the skeleton grains, whereas clay, oxides and humus form the plasma.

Voids (pores) include all space in the soil and is described in terms of type, size and abundance. Cutans (coatings) consisting of illuvial clay silicates, humus and aluminium and iron oxides (sesquioxides), or mixtures thereof, deposited on ped and pore surfaces are described according to abundance, contrast, nature and location. Abundance, size and orientation of roots are recorded. The occurrence of cementation or compaction, in pans or otherwise, is described according to continuity, structure, nature of cementing agent and degree. Nodules and concretions are described in terms of abundance, kind, size, shape, hardness, nature and colour. Stones and rock fragments are described according to abundance, size, shape, weathering state and nature. Abundance and kind of faunal activity is recorded.

Micromorphology

Soils are heterogenous as indicated in Figure 2 and shown in Table 2 by the different horizons. However, even within each horizon voids and different solid components are easily recognizable. The heterogeneity can be seen by the naked eye but it is much more comprehensively studied under the microscope using thin sections prepared by cutting and polishing slices of soil impregnated with a resin (Douglas, 1990; FitzPatrick, 1984). The spatial arrangement of solids and voids (pores) is called the soil fabric, the coarser (>2 μm) organic and inorganic fragments form the soil skeleton and the less than 2 μm organic and inorganic materials constitute the soil plasma (Figure 4).

Table 3. Some characteristics of the Klosterhede soil.

Horizon, depth (cm)		%C	pH	Exchangeable cations (cmol(+)/kg)						Oxalat-extr. (μmol/g)	
				Ca	Mg	K	Na	Al	CEC_7	Al	Fe
Oi	(0-10)		2.98	6.31	6.47	0.87	1.59	3.86	65.6	30.8	20.5
A	(10-25)	3.64	3.32	0.09	0.14	0.05	0.07	1.67	6.14	9.6	6.8
E	(25-30)	1.49	3.53	0.05	0.07	0.03	0.05	1.84	3.65	10.7	12.9
Bh	(30-35)	6.99	3.71	0.04	0.13	0.08	0.13	11.0	23.6	172	267
Bhs	(35-45)	2.69	4.01	0.02	0.03	0.04	0.05	4.38	8.22	218	214
Bs1	(45-70)	0.30	4.09	0.01	0.00	o.01	0.02	0.59	1.06	83.4	21.1
Bs2	(70-100)	0.15	4.07	0.01	0.00	0.01	0.01	0.29	0.42	31.5	13.8
C	(100-)	0.09	4.03	0.01	0.00	0.01	0.01	0.20	0.21	20.4	9.4

CLASSIFICATION

Soil classification according to the FAO/Unesco and Soil Taxonomy systems is based on socalled diagnostic surface horizons (epipedons), subsurface horizons and diagnostic soil characteristics or properties, which are quantitatively defined features determined by field and laboratory data (FAO/Unesco, 1990; Soil Survey Staff, 1992). For the Klosterhede soil the profile description is shown in Table 2 and the laboratory data in Table 3. The two classification systems are briefly outlined in Figure 5. Since both systems are (and have been) under current revision, it is important to state which edition the soil classification corresponds to.

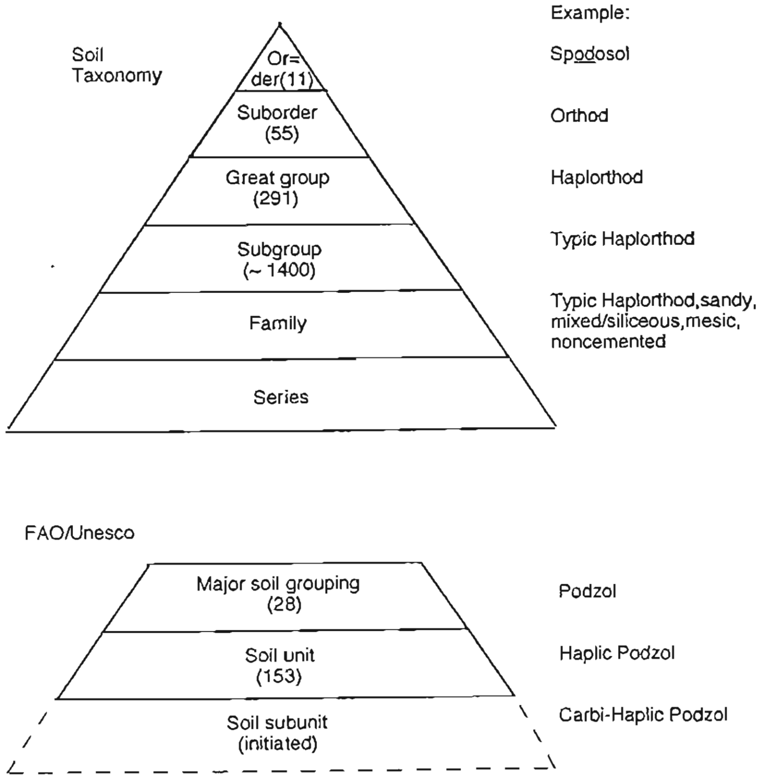


Figure 5. The categories of the Soil Taxonomy (Soil Survey Staff, 1992) and of the FAO/Unesco (FAO/Unesco, 1990) soil classification systems and number of units in each category. Detailed guidelines for classification at the subunit level in the FAO/Unesco system is still under consideration.

Table 4. Simplified key to epipedons according to the definitions of the Soil Survey Staff (1992). In addition to those listed, two more epipedons may occur, i.e. the melanic (organic matter - rich, developed on vulcanic materials) and the plaggen (man-made, > 50 cm thick) epipedons. The definitions agree with those of the FAO/Unesco system (FAO/Unesco, 1990), where the anthropic and plaggen epipedons are called fimic.

Property	Mollic	Umbric	Anthropic	Histic	Ochric
1. SOM	+	+	+	+++	+/-
2. Colour	+	+	+	+	+/-
3. Structure	+	+	+	+	+/-
4. Base Sat.	+	-	+/-	+/-	+/-
5. Depth	+	+	+	+/-	+/-
6. Moisture	+	+	+/-	+	+/-
7. P content	+	+	-	+/-	+/-
8. n value	+	+	+	+/-	+/-

1: More than 0.6% organic carbon.
2: Dark coloured.
3: Not hard and massive at the same time.
4. More than 50% base saturation.
5: More than 25 cm (18 cm) deep.
6: Not aridic moisture regime.
7: Less than 250 mg P_2O_5/kg (<110 mg P/kg).
8: n value less than 0.7.

Diagnostic horizons

The epipedons are shortly defined in Table 4. Some diagnostic subsurface horizons are briefly described (Soil Survey Staff, 1992) as follows:
Agric horizon: This horizon is an illuvial horizon which occurs directly under the plow layer, has formed under cultivation and contains significant amounts of illuvial silt, clay and humus.
Albic horizon: This horizon is an eluvial horizon (typically E) consisting of albic materials, i.e. light coloured materials with high colour value and low chroma.
Argillic horizon: This horizon is an illuvial B horizon formed by illuviation of clay silicates (for mechanism of formation see section 7.2). In general, it contains at least 1.2 times as much clay as an overlying eluvial horizon or 3% or more clay if the overlying horizon contains <15% clay or 8% or more if it contains >40% clay. It should be at least 0.1 times as thick as all overlying horizons. Clay cutans may occur on peds.
Cambic horizon: This horizon is a structured B horizon of very fine sand, loamy very fine sand or finer texture (Figure 3) being too weakly developed to meet the requirements for argillic and spodic horizons.
Fragipan: This horizon is a subsoil layer with hard to very hard consistence when dry and brittle when moist.

Glossic horizon: This horizon is ≥5 cm thick and occurs between an overlying albic horizon and an underlying argillic horizon. It developes as a results of degradation of an argillic horizon, where tongues of albic materials extend into the argillic horizon.
Placic horizon: This horizon is a 2-10 mm thick, strongly cemented (by iron-humus complexes) layer commonly occurring within 50 cm of the soil surface.
Spodic horizon: This horizon is a ≥2.5 cm thick, illuvial B horizon consisting of spodic materials, i.e. illuvial active materials composed of humus and aluminium, and sometimes iron oxides.
Sulfuric horizon: This horizon is a ≥15 cm thick mineral or organic layer that has a pH <3.5 and is toxic to plant roots.

According to the FAO/Unesco system (FAO/Unesco, 1990) these horizons are defined similarly, although small differences occur, e.g. the argillic (termed argic) and the cambic horizons contains ≥8% clay according to the FAO/Unesco system.

In addition to those listed several more diagnostic horizons exists including the calcic, gypsic, kandic, natric, oxic, salic and sombric horizons (Soil Survey Staff, 1992) but since their occurrence in Nordic soils is highly unlikely they will not be considered further.

Soil Taxonomy system

The system contains six categories (Figure 5). The nature of the kinds of differentiating characteristics employed in the various categories may summarized as follows:
Order: This division is based largely on soil-forming processes as indicated by presence or absence of diagnostic horizons.
Suborder: The subdivision of orders emphasizes properties that suggest genetic homogeneity including wetness, soil temperature and moisture regimes and major parent material; for Histosols organic fiber decomposition stage.
Great group: The subdivision of suborders is according to kind, arrangement and degree of expression of horizons. Presence of placic horizons and fragipans is considered at this level as is the climate, if not accounted for at a higher category.
Suborder: This subdivision of great groups arranges soils in relation to the central concept of the great group and to properties indicating intergradation to other great groups, suborders and orders.
Family: This subdivision of subgroups is according to texture, mineralogy, soil temperature, depth and other properties affecting plant growth.
Series: This subdivision of families is mainly based on kind and thickness of horizons and other properties, particularly those affecting soil use, not accounted for at higher levels. Series are not very much used outside USA.

Soils in Nordic countries will be classified as Histosols, Spodosols, Ultisols, (Mollisols), Alfisols, Inceptisols or Entisols. The main characteristics of these soil orders may be seen in Table 5.

Table 5. Simplified key to soil orders.

	If soil has:	Order	Formative element
1.	Organic material to a depth of more than 40(60) cm	Histosols	ist
2.	Other soils with a spodic horizon within 2 m	Spodosol	od
3.	Other soils with > 35 cm of andic soil properties	Andisols	and
4.	Other soils with an oxic horizon within 1.5 m and no kandic horizon, or contain ≥ 40% clay in the surface 18 cm and have a kandic horizon within 1.5 m	Oxisols	ox
5.	Other soils with more than 30% clay in all horizons; some cracks when dry at 50 cm	Vertisols	ert
6.	Other soils that are dry more than 50% of the year and have an ochric or anthropic epipedon	Aridisols	id
7.	Other soils that have an argillic or kandic horizon but a *B.S. at pH 8.2 less than 35% at a depth of 1.8 m	Ultisols	ult
8.	Other soils that have a mollic epipedon	Mollisols	oll
9.	Other soils that have an argillic or kandic horizon	Alfisols	alf
10.	Other soils that have an umbric, mollic, or plaggen epipedon, or a cambic horizon	Inceptisols	ept
11.	Other soils	Entisols	ent

*B.S. (base saturation) of 35% at pH 8.2 corresponds to B.S. of 50% at pH 7 (ammonium acetate).

The classification is carried out according to the key. A simplified key to soil orders is shown in Table 5. It is very important to follow through the key in order from the beginning (from the top) in attempting to "key out" the identification and name of an unknown soil, else misclassification will result. Classification at the suborder, great group and subgroup levels is carried out in the same way, i.e. by keying, whereas classification of soil families is indicated by listing texture, temperature, etc.

Naming of soils according to the Soil Taxonomy system is carried out by putting together formative elements for the order, suborder and great group but in the opposite order as examplified by the Klosterhede soil in Figure 5. The formative elements of the orders are included in Table 5. Subgroups are identified by adjectives and families by texture, mineralogy, etc. as outlined in Figure 5.

FAO/Unesco system

The system contains only two categories, i.e. the major soil grouping and soil unit, with a third category, the soil subunit, under consideration (Figure 5). With only 153 units at the soil unit level it is obviously less elaborated than the Soil Taxonomy system with >1000 units already at the subgroup level. Classification according to the FAO/Unesco system is based on similar criteria, i.e. on presence and absence of diagnostic horizons and properties, as classification according to the Soil Taxonomy system.

The major soil groupings most relevant for Nordic soils may briefly be summarized in the following way:
Histosols: Include organic soils, i.e. soils containing organic materials to a depth of at least 40 cm (or 60 cm if fibers).
Anthrosols: Include soils profoundly modified by human activities.
Leptosols: Include shallow soils formed on rocks and lack diagnostic subsurface horizons.
Fluvisols: Include soils with fluvic properties, i.e. exhibit stratification due to water deposition of the materials.
Gleysols: Include soils with gleyic properties, i.e. exhibit evidence of reduction processes (these processes are considered in section 7.1).
Arenosols: Include poorly developed soils containing <8% clay and an ochric epipedon.
Regosols: Include poorly developed soils containing >8% clay.
Podzols: Include soils with a spodic horizon.
Podzoluvisols: Include soils with an argic B horizon showing an irregular or broken upper boundary resulting from deep tonguing of the E into the B horizon.
Alisols: Include soils with an argic B horizon and a CEC ≥24 cmol(+)/kg clay and a base saturation <50%.
Luvisols: Include soils with an argic B horizon and a CEC ≥24 cmol(+)/kg clay and a base saturation ≥50%.
Cambisols: Include soils with a cambic B horizon.

Classification is carried out by means of a key in the same way as outlined above for classification according to the Soil Taxonomy system. The soil name consists of the major grouping designation and an adjective indicating the subdivision at the soil unit level as shown by the example in Figure 5.

An updated description of soils of the various major soil groupings (and some modification of groupings) is under launching (Spaargaren, 1994).

Finally, it must be emphasized that soil classification should be based on correct use of the entire classification key (FAO/Unesco, 1990; Soil Survey Staff, 1992) and never on short excerpts like those outlined above.

SOIL COMPOSITION

Soil is a mixture of inorganic and organic solids, air, and water as indicated in Figure 4. The solids are of different size and shape. The spatial arrangement or layering of the solids determines the volume (space) of voids (pores) occupied by air and water; the closer the solids are arranged the smaller will the porespace be. A mineral soil showing good physical conditions for plant growth will have a solid:void ratio of ~1 as illustrated in Figure 6. Decreased porespace due to increased soil compaction affects plant growth negatively.

The bulk density of a soil measures the mass of a unit volume (voids and solids) of dry soil in natural layering. Since soil particle (solid) density exhibit very limited variation around 2.6 Mg/m^3 in most mineral soils, bulk density is a measure of porespace or compaction of soil. Bulk densities of various clayey to sandy soils are in the range 1.1-1.8 Mg/m^3 but may exceed 2 Mg/m^3 in very compacted (sub)soils, where root development is seriously impaired. Organic soils have much lower particle and bulk densities.

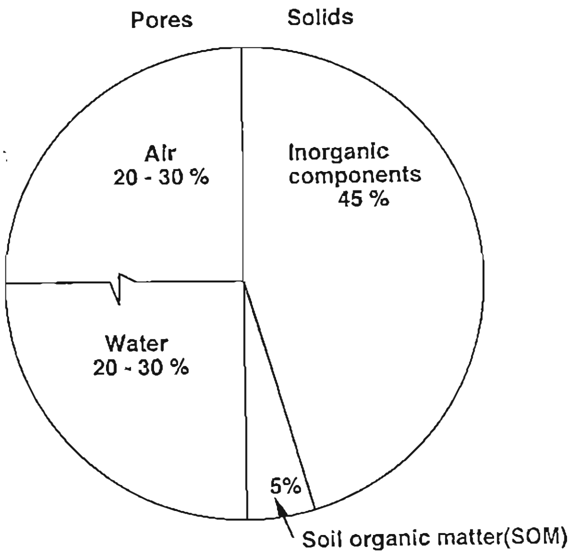


Figure 6. Volume composition of a loamy surface soil with good physical plant growth conditions. The air and water in a soil are extremely variable.

Most reactions in soils occur at the surfaces of the solid particles. The surface area, and hence reactivity, increases at decreasing particle size as shown in Figure 7 for mineral particles. Texture (Figure 3) accounts for the particle size distribution in soils.

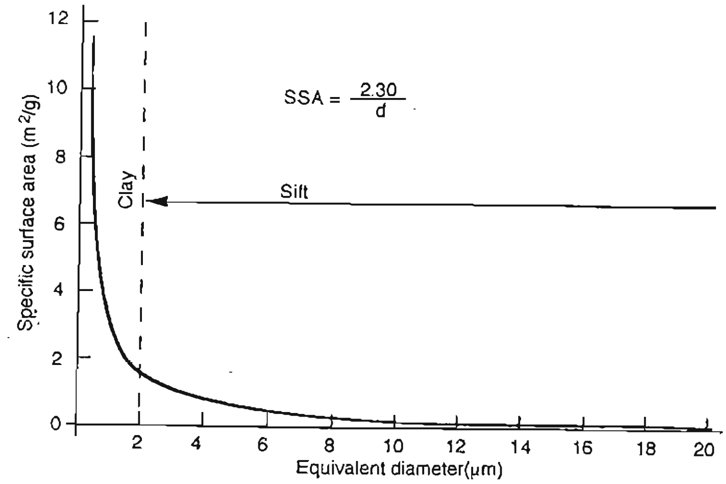


Figure 7. The specific surface area of mineral soil materials as a function of particle size.

Soil air

The components of soil air are typically the same as in atmospheric air but the concentrations are different because of biological activity (Brady, 1990; Schachschabel et al., 1989; Singer & Munns, 1987). Well-aerated soils contain ~20% of O_2 but this figure may drop to ~2% in isolated microenvironments near plant roots. On the other hand, soil air is enriched in CO_2. Often it contains 0.33-3% but at the bottom of the root zone the concentration of CO_2 may approach 10% in the vicinity of roots. Under anaerobic conditions soil air may contain N_2O, CH_4 and H_2S.

Soil water

Apart from being vital for the plants in itself, soil water is a repository for dissolved solids and gases, which are immediately available for plants and for soil processes (Bohn et al., 1985; Schachschabel et al., 1989; Singer & Munns, 1987; Sposito, 1989) . The mixture of water and dissolved solids and gases is called the soil solution. In well-aerated soils the main gaseous soil solution components include O_2 and CO_2. The dissolved solids include a hugh number of different cations (Al^{3+}, Ca^{2+}, Mg^{2+},K^+, Na^+, NH_4^+, H^+, etc.), anions (HCO_3^-, Cl^-, NO_3^-, SO_4^{2-}, etc.) and molecular species such as various organic compounds.

The dissolution of gases in water is determined by Henry's law:
$[gas] = K_H P_{gas}$,
where K_H is a constant. The constants for O_2 and CO_2 are 1.26 and 34.06 mol/(m³atm), respectively, at 25 °C. A partial pressure of 0.2 atm (atmospheric air) will result in a O_2 concentration of 0.25 mM, whereas the considerably higher CO_2 constant suggests soil solution concentrations up to ~1 mM (P_{CO2} = 0.03 atm) or even higher.

Water potentials

Soil water is affected by gravitational forces, capillary and adsorptive forces and osmotic force (Brady, 1990; Schachschabel et al., 1989; Singer & Munns, 1987). Due to these forces energy is consumed or released when water is moved (flows) from one zone to another. Water flow can occur as a saturated flow when all pores are full of water or as unsaturated flow when the larger pores have been emptied. The corresponding potentials can shortly be described in the following way:

Gravitational potential (Ψ_z): This potential is a function of the height (z) of water above a reference point (e.g. the groundwater table).

Matrix potential (Ψ_m): This potential is a function of the attraction of surfaces to water molecules and is most important to unsaturated flow. Under saturated conditions Ψ_m is replaced by Ψ_h, the pressure potential.

Osmotic potential (Ψ_o): This potential is a function of the soil solution composition.

The sum of these potential (called water potential, Ψ) is a quantitative expression for the tendency of water to move, which will be from zones of high water potential to zones where the water potential is lower. For pure water the osmotic potential is zero giving the hydraulic potential:

$\Psi_H = \Psi_z + \Psi_m$ (or Ψ_h).

The variation of the potentials at equilibrium and under non-equilibrium conditions at raising (upward) water due to uptake or drying and descending (downward) water after rain is shown in Figure 8.

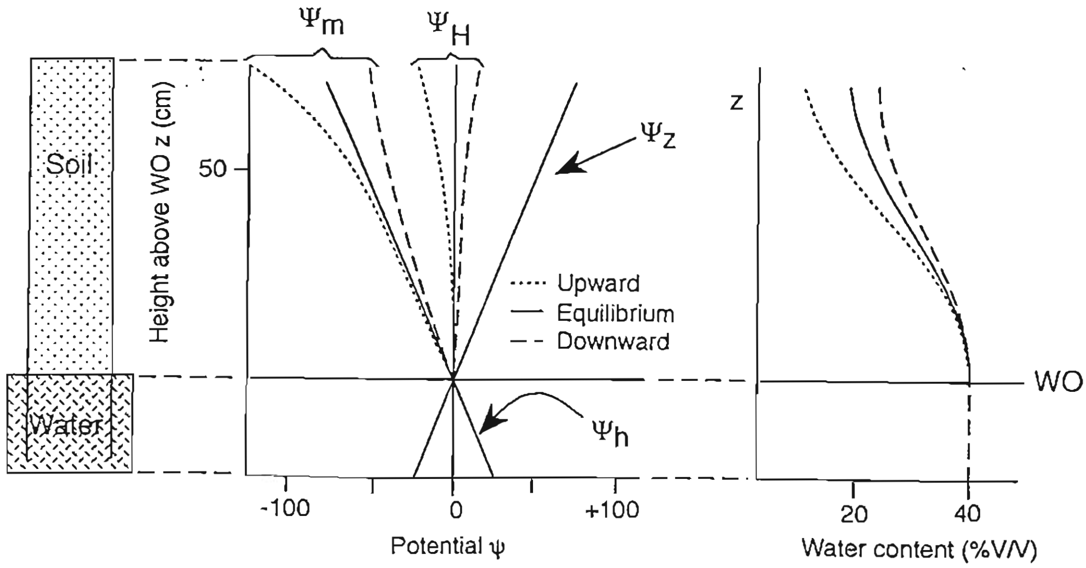


Figure 8. Various potentials for pure water under equilibrium ($\Psi_z = \Psi_m$) and non-equilibrium conditions.

The gravitational potential can be measured with a piezometer and the osmotic potential can be determined from the decrease in water vapor pressure or calculated by the expression:

$\pi = MRT$, where M is the salt concentration, R the gas constant and T the temperature. The matrix potential, i.e. the negative pressure retaining water in the pores under unsaturated conditions, can be measured by a tensiometer or by the membrane-plate method. Details of these methods may be found elsewhere (Klute, 1986; Rowell, 1994).

Field capacity and wilting point represent the approximate upper and lower limits of water that can be held in soils and is available for plants. At field capacity the medium and small pores are filled with water but most of the larger pores are nearly empty, whereas at the wilting point only the smallest pores contain water. The corresponding approximate matrix potentials are app. -0.1 bar or -100 cm water column (field capacity) and -15 bars or -15000 cm water column (wilting point). Normally the pF values ($pF = -\log(\text{cm water column})$) are used, i.e. $pF = 2$ at the field capacity and $pF = 4.2$ at the wilting point. In water retention curves pF is plotted against the corresponding water content (Figure 9). Plant available water is taken as the difference between the water contents at pF 2 and pF 4.2.

Obviously clayey soils retain much more water than silty and sandy soils but the available water contents show little difference because micropores dominate in the clayey soils. In fact, for Danish soils plant available water (PAW) can be calculated by the expression:

$PAW = 1.79 \times \%\text{organic matter} + 0.07 \times \%\text{clay} + 0.29 \times \%\text{silt} + 0.18 \times \%\text{fine sand} + 2.56$,
where silt is 2-20 µm and fine sand is 20-200 µm (Madsen et al., 1992). This expression emphasizes the importance of the silt and fine sand fractions as reservoirs of the plant available water.

The retention curves may also give information about the pore size distribution. Thus the pore diameter (in µm) is related to pF in the following way:
$\log(d) = 3.5 - pF$.
The pores are divided into fine pores with $d < 0.2$ µm ($pF > 4.2$), medium pores with $d = 0.2$-30 µm (pF 4.2-2) and coarse pores with $d > 30$ µm ($pF < 2$). Accordingly, plant available water is retained in the medium pores.

Water flows through soils in response to differences in the hydraulic potentials over the distance s, the hydraulic gradient ($d\Psi_H/ds$). Under saturated conditions water moves in the entire pore volume. The stationary, one-dimensional water movement per unit time (V) can be expressed as a function of the hydraulic gradient (Darcy's law):

$V = K(d\Psi_H/ds)$,
where K is a constant, the hydraulic conductivity of the soil, which depends on size and shape of the pores. Soil texture and structure are closely related to the hydraulic conductivity. Thus sandy soils and well-structured soils have higher hydraulic conductivities than clayey soils, particularly if their structure is weak.

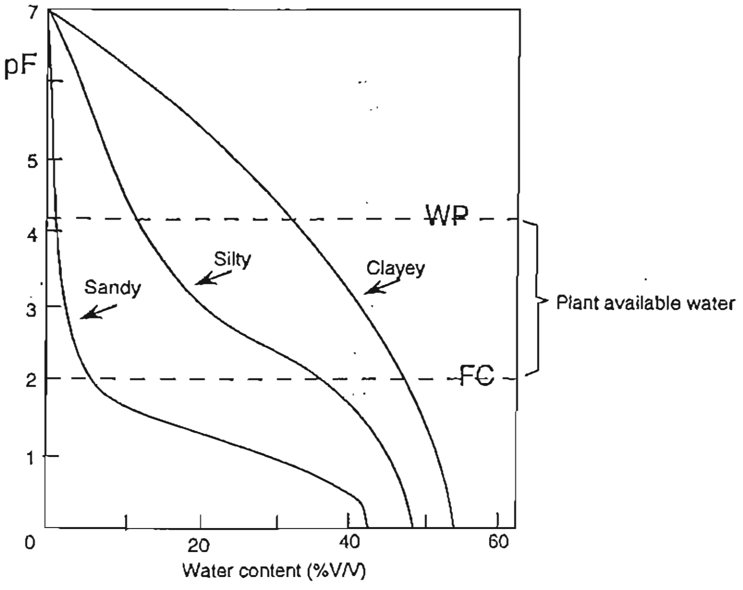


Figure 9. Water retention curves for three soils of different texture.

In unsaturated soils, water moves much slower and only in the finer pores because the coarser pores are filled with air. Unsaturated flow is governed by the same general principles as saturated flow, i.e. the hydraulic conductivity and the hydraulic gradient (mainly matrix potential differences). The hydraulic conductivity strongly depends on the water content, since when the soil becomes dried only the finest pores will be filled with water. Due to differences in pore size distribution the hydraulic conductivity is more affected by the moisture content in sandy soils than in clayey soils. Thus, at high moisture contents (high matrix potential levels) the hydraulic conductivity is higher in sandy soils than in clayey soils, whereas at low moisture contents the opposite is true.

Since the water flow under natural conditions is normally not constant and occurs in more than one direction, its description often necessitates application of mathematical expressions more complicated than Darcy's law shown above (Schachschabel et al., 1989). Furthermore, it has been assumed that soils are uniform in texture and structure, which is nearly never the case. Different soil horizons and layers will often have different hydraulic conductivities because of inhomogeneity in texture, structure and compaction.

Leaching

Nordic soils are subject to water leading because precipitation exceeds evapotranspiration. Textural and structural inhomogeneities will affect the leaching of water through the soils. A sand layer underlying a clayey soil will retard the vertical water movement, since the macropores of the sand offer less attraction for the water than does the finer textured materials. Rainwater may then cause saturation with water of the upper horizons, which can lead to oxygen depletion and reduction. On the other hand, some clayey soils have fractures running from upper layers to a depth of several meters, which are filled with more sandy materials (Jørgensen & Fredericia, 1992). The water movement through such fractures can be considerably faster than through the clayey matrix. Big, vertical continous pores starting at or near the soil surface may have the same effect. Consequently, in many, particularly clayey, soils, water will leach through different parts of the soil at different rates, which is termed dual flow, bypass or preferential flow (Booltink & Bouma, 1991; Jørgensen & Fredericia, 1992). After drying organic and inorganic soil components are difficult to wet (water repellency) forcing water and solutes to flow via preferential paths through the unsaturated soil.

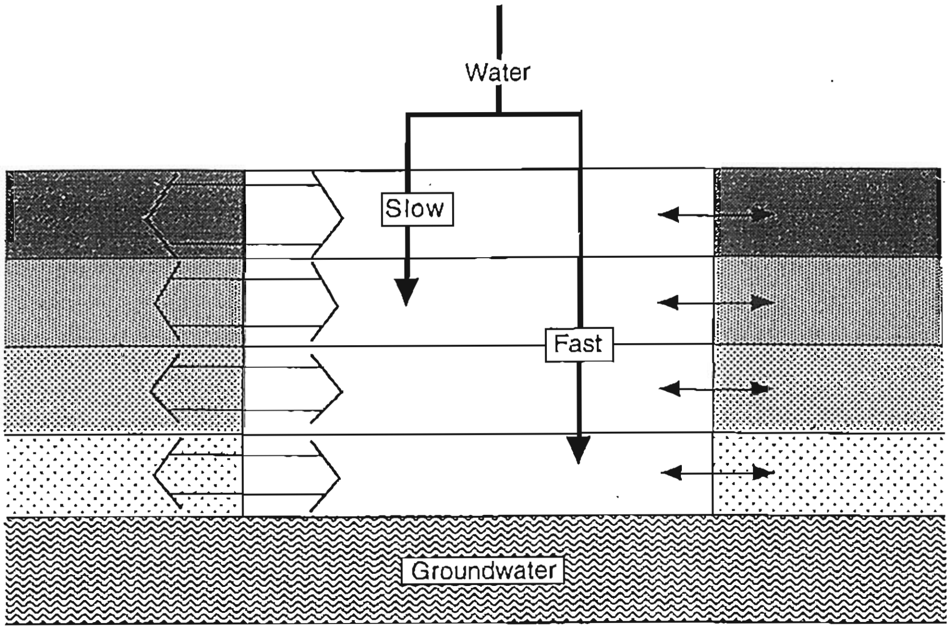


Figure 10. Schematic diagram of a soil with a macropore system allowing preferential flow. The thickness of the horizontal arrows indicates extent of reactions.

The flow of water through a system of larger pores and fractures that allows fast flow rates and bypasses the unsaturated soil matrix is of great environmental and soil development importance. Preferential flow may lead to transportation of solutes and colloids by the leaching water in soils, where the soil matrix has the capacity to retain such compounds. Thus, clayey soils were formerly considered impermeable or slowly permeable for various pollutants but recent investigations have suggested readily transport of pesticides (and clay minerals) to deep soil layers through such a macropore system (Jørgensen & Fredericia, 1992). The presence of preferential flow ways in clayey soils has been demonstrated directly by staining experiments (Booltink & Bouma, 1991).

The equilibria of most reactions occurring between compounds in the soil solution and the soil solids are not established instantaneously. Therefore, the faster the water flows, and hence the compounds therein, the further away from equilibrium will the reactions be. Under fast leaching through a macropore system the exchange between the water phase and the solid phase is limited, whereas under slow leaching through the finer pores the processes may have time to proceed. This is schematically shown in Figure 10.

Soil organic matter (SOM)

The soil organic matter (SOM) pool consists of living organisms and dead plant and animal residues and it is determined by formation and decay processes involving autotrophs and heterotrophs (Figure 11). The overall process is:

$CO_2 + H_2O \rightleftarrows CH_2O + O_2$

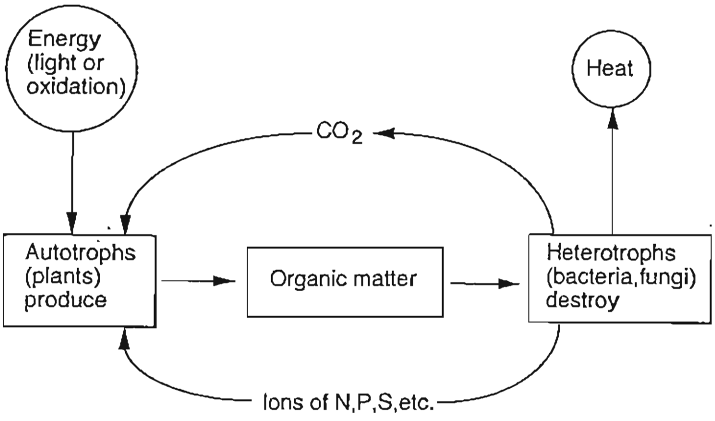


Figure 11. Overall energy, heat and element flows during organic matter formation and decomposition.

From left to right, the process shows the net result of photosynthesis, i.e. production by autotrophs of carbohydrates and oxygen from carbon dioxide and water, whereas from the opposite direction it shows the complete decomposition by heterotrophs of organic matter, i.e. mineralization. The balance between organic matter formation and mineralization determines the SOM content in soils.

Organic matter decomposition is composed of a hugh number of chemical and microbiological subprocesses. Partial decomposition leads to formation of humus, which is a dark, heterogeneous mass of colloidal organic compounds. The complex of degradative and synthetic processes leading to humus formation is called humification. SOM is an extremely complex mixture. In addition to fresh organic materials and humus, SOM contains nonhumic compounds consisting of carbohydrates, lipids, proteins, etc., which have been released by decay of plant, animal and microbial tissue. Soils also contain simple aliphatic and aromatic compounds such as formic acid, acetic acid, oxalic acid, citric acid and benzoic acid in amounts up to several mmol/kg soil.

Humus is often divided into three fractions according to solubility as follows:
Humin: This fraction is insoluble also in alkali and therefore probably consists of high molecular weight compounds with few functional groups.
Humic acid: This fraction is soluble in alkali but insoluble in acid. It is relatively rich in oxygen-containing groups such as carboxyl and phenol groups.
Fulvic acid: This fraction is soluble in alkali and acid. It is the low molecular weight fraction and it is considered the most reactive portion of SOM because of its high content of functional groups.
Typically, humic acid and fulvic acids contain 1.5-6.0 and 5.2-11.2 mol/kg of carboxyl groups and 2.1-5.7 and 0.3-5.7 mol/kg of phenol groups (Schnitzer & Khan, 1978). Fulvic acid appears to be less aromatic than humid acid. Typically, fulvic acid exhibits molecular weights in the range 700-2300 daltons, while molecular weights of 2.3×10^3 to 1.4×10^6 daltons are common for humic acid.

The rate of organic matter decomposition increases at increasing soil temperature but decreases at very low and very high moisture contents exhibiting maximum decomposition at a moisture content, which at the same time ensures optimal microbial activity and good oxygen access. In poorly drained and swampy soils, dead plant material can be covered with water and oxygen excluded leading to peat, and hence Histosol formation. Compared to fresh organic matter residues with half-lives of only a few months, humus, or at least part of it, is very resistant to microbial desomposition. The resistant fraction can be >1000 years. In association with clay silicates and aluminium and iron oxides SOM seems to be stabilized.

The main elements in SOM include C, O, N, S, P and H. In well-aerated, temperate soils the mass ratios for the C/N/S/P concentrations are about 100/10/1/1. Since the C/N in fresh plant material such as straw is much higher, humification leads to a relative enrichment of N (Figure 12). The elements are linked together in numerous complex alifatic and aromatic compounds. Oxygen is present in carboxyl, phenol, quinone and other groups. During humification the oxygen content increases as shown by the increased carboxyl content in Figure 12. Since SOM appears to contain ~58% C, the soil SOM content is taken as (100/58) times the % C content determined by wet or dry combustion.

SOM is per unit mass the most active portion of soils. It is the reservoir for a large pool of carbon and for various essential elements, promotes good soil structure, is a source of cation exchange capacity and soil pH buffering and is important for complexation processes and mineral weathering.

More information about humus formation, composition and properties may be found elsewhere (Aiken et al., 1985; Bohn et al., 1985; Brady, 1990; Hayes et al., 1989; Schachschabel et al., 1989; Singer & Munns, 1987; Sposito, 1989; section 5).

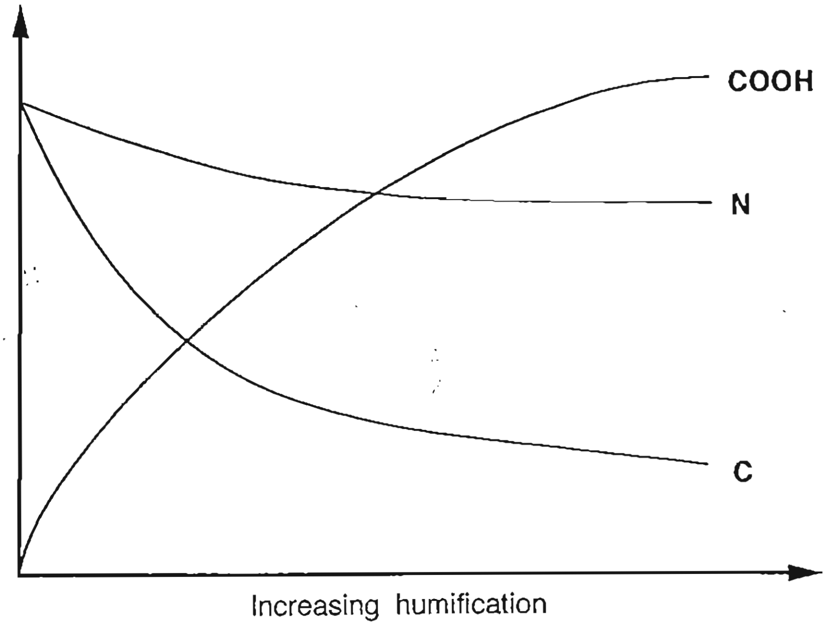


Figure 12. Decreases of N and C and increase of carboxyl groups, and hence CEC, in SOM during decomposition of organic matter to humus.

Soil microorganisms

In addition to the macroorganisms such as plant roots, earthworms, moles, ants etc. important for mixing of the various soil components, soils contain microorganisms. Soil microorganisms, which include aerobic and anaerobic bacteria, fungi, actinomycetes and algae, are essential for organic matter decomposition and many other soil processes (Brady, 1990; Schachtschabel et al., 1989; Tan, 1994). While the plow layer (Ap horizon) may contain several Mg per ha of microbial biomass, the B and C horizons are very low in microorganisms.

Soil bacteria are either autotrophic or heterotrophic. The autotrophs obtain their energy from the oxidation of inorganic substances such as ammonium, sulfur and iron compounds and most of their carbon from carbon dioxide. Most bacteria are, however, heterotrophic, that is both energy and carbon come from organic matter. Thus, heterotrophic bacteria, along with fungi and actinomycetes, account for the general breakdown of organic matter in soils. This degradation involves many different bacteria (and other organisms), where specific species are active in different steps of the process.

While fungi are rather unaffected by soil pH, SOM decomposing bacteria exhibit maximum activity at pH 6-8 and high calcium concentration. SOM degradation therefore depends on soil pH. Under neutral to slightly alkaline pH conditions with good calcium availability, bacteria are the main SOM decomposers leading to mull formation, which is well humified SOM with a low C/N ratio. Mull is well mixed with the inorganic soil fraction because of earthworm activity. In nutrient-poor soils of low pH SOM is mainly decomposed by fungi leading to formation of mor, i.e. poorly humified SOM with a high C/N ratio. Mor accumulates at the soil surface due to absence of soil mixing organisms like earthworms. Mor formation is most likely to occur on sandy soils due to their low buffering capacity against acidification and under certain kinds of vegetation such as spruce and calluna.

Under anaerobic conditions bacteria can decompose organic matter under formation of methan. Nitrate, if present, may serve as electron acceptor leading to its reduction to N_2 (denitrification).

Nitrogen fixation either symbiotic or nonsymbiotic, where atmospheric nitrogen is converted to plant available nitrogen compounds, is another very important microbial process. Symbiotic nitrogen fixation under ammonium formation is carried out by heterotrophic bacteria living in root nodules of legumes. Nonsymbiotic nitrogen fixation can be carried out by certain heterotrophic bacteria and by autotrophic blue-green algae.

Other bacteria are involved in the oxidation of ammonium to nitrate (nitrification), which occurs in two main steps, i.e. oxidation of ammonium to nitrite followed by oxidation of nitrite to nitrate. Normally the first step is rate limiting preventing accumulation of high concentrations of poisonous nitrite in soil.

Soil minerals

The inorganic soil fraction consists of primary and secondary minerals. The silt and sand fractions consist largely of primary minerals, i.e. minerals formed at elevated temperatures and inherited from igneous and metamorphic rocks. Some resistant, primary minerals also occur as minor constituents of the clay fraction. The main components of the clay fraction are, however, the secondary minerals, which are formed by low temperature reactions and either inherited from the parent material or formed by weathering (section 6). Various silicate minerals constitute the most abundant primary minerals, whereas the secondary minerals in soils include layer silicates and accessory minerals such as carbonates and oxides.

Silicate minerals

The composition, structure and formation of various silicate minerals have been compiled in numerous textbooks (e.g., Dixon & Weed, 1989; Deer et al., 1962, 1963; Gieseking, 1975). Briefly, the fundamental building block in the structure of these minerals is the silica tetrahedron consisting of four oxygen ions forming covalent bonds to Si^{4+}, giving the ionic unit SiO_4^{4-}. Silica tetrahedra can occur as isolated, single units, in single or double chains linked together by shared corners, in sheets and in fully 3-dimensional frameworks. The various structures are shown schematically in Figure 13. In silicate minerals the silicate structures are held together with divalent cations like Mg^{2+} and Fe^{2+} or trivalent cations including Al^{3+} and Fe^{3+} in octahedral coordination. In the layer silicates the octahedra are linked together in sheets. Socalled isomorphous substitution, where one cation is substituted by another cation of comparable size but often of different charge, occurs in many silicates.

166

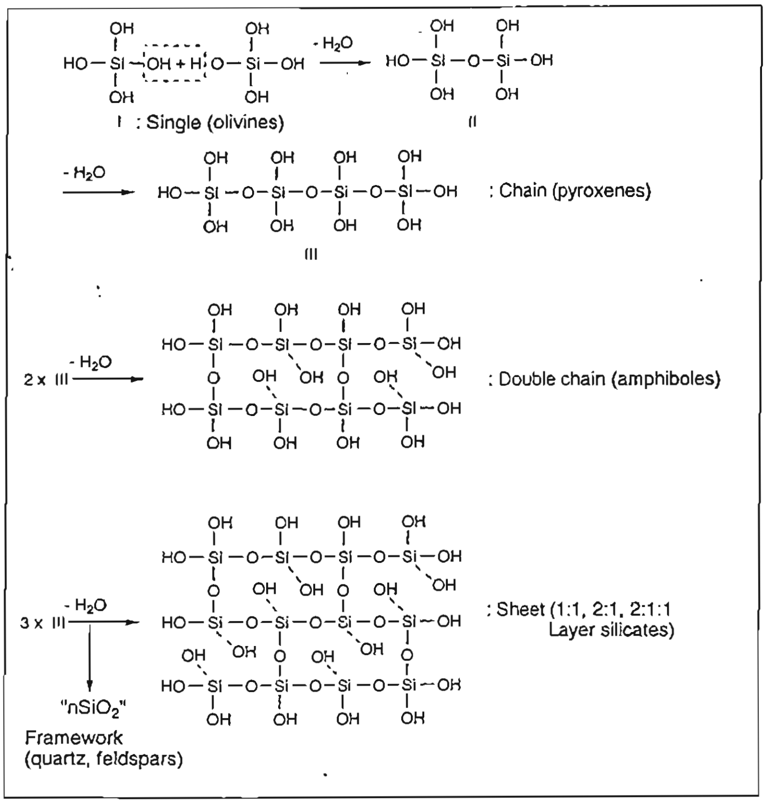


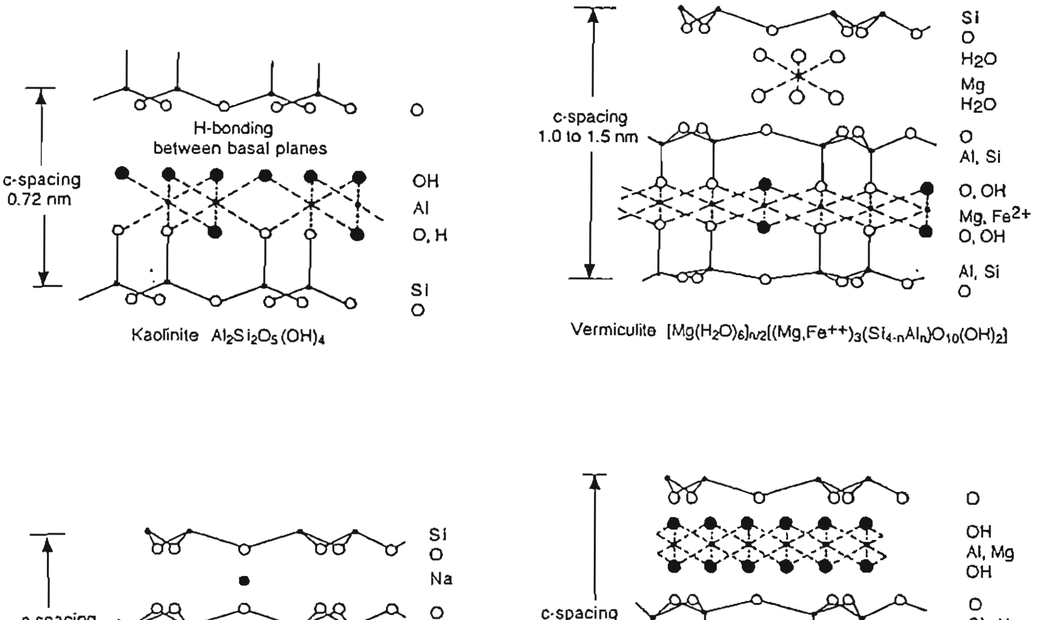
Figure 13. Sketch of di- and polymerization of silicic acid tetrahedra forming, as anions, the basic Si(IV) structures in various classes of soil silicates.

The primary silicates occurring in soils may shortly be described in the following way:
Olivines: These minerals consists of single silica tetrahedra linked together by divalent cations such as Mg^{2+} and Fe^{2+} in octahedral coordination. Example: $FeMgSiO_4$.
Pyroxenes: These minerals contain single chains of silica tetrahedra forming repeating units of $Si_2O_6^{4-}$ linked together by a variety of mainly divalent cations in octahedral coordination. Example: $FeSiO_3$.
Amphiboles: These minerals contain double chains of silica tetrahedra forming repeating units of $Si_4O_{11}^{6-}$ linked together by a variety of mainly divalent cations in octahedral coordination. Example: $NaCaMg_3FeAlSi_8O_{22}(OH)_2$.
Micas: These minerals are built up from two sheets of silica tetrahedra fused to each planar side of a sheet of metal cation octahedra, typically containing Al, Mg and Fe ions. The formula of layer silicates is often based on 22 negative charges corresponding to $O_{10}(OH)_2^{22-}$. If the metal cation is trivalent, only two of three possible cationic sites in the octahedral sheet can be filled to achieve charge balance and the sheet is termed dioctahedral. If the metal cation is divalent, all three possible sites are filled and the sheet is trioctahedral. Isomorphic substitution of Al for Si, Fe(III) for Al and Fe or Al for Mg occurs typically in the micas.

Substitution of an ion of higher charge such as Si^{4+} by an ion of lower charge such Al^{3+} will result in a negative charge, which must be balanced by a positive ion. Examples: $K(Al_2)[AlSi_3]O_{10}(OH)_2$ (muscovite); $K(Mg_2Fe)[AlSi_3]O_{10}(OH)_2$ (biotite). Muscovite is dioctahedral, whereas biotite is trioctahedral. In both mineral one Si per formula unit in the tetrahedral layer has been substituted by Al resulting in a negative charge, which is balanced by K^+.

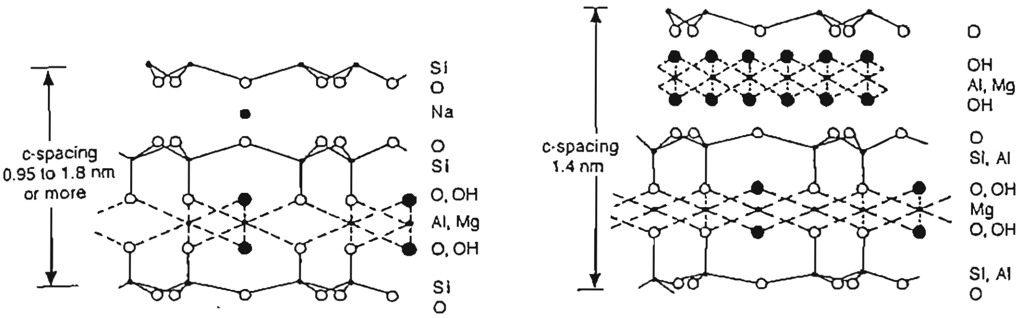
Framework silicates: These minerals consist of a continous, 3-dimensional framework of silica tetrahedra sharing corners. The general formula of the minerals is:
$M_x[Al_{0-2}Si_{2-4}O_8]$, which includes Si_4O_8 (or SiO_2, quartz), $KAlSi_3O_8$ (orthoclase, K-feldspar), $NaAlSi_3O_8$. (albite, Na-feldspar), $CaAl_2Si_2O_8$ (anorthite, Ca-feldspar) and mixed Na/Ca-feldspars (plagioclase).

The secondary silicates (often called clay silicates) are layer silicates and consist, like the micas, of sandwiches of tetrahedral and octahedral sheets (Figure 14). The bonding between tetrahedral and octahedral sheets occurs through the apical oxygen ions in the tetrahedral sheet and always produces, because of mismatch, distortion of the anion arrangement in the final layer structure. The minerals are differentiated by (1) the number and sequence of tetrahedral and octahedral sheets; (2) the layer charge per formula unit; (3) the type of interlayer bonding and interlayer cations; (4) the cations in the octahedral sheet; (5) the type of stacking along the c-dimension.



Kaolinite $Al_2Si_2O_5(OH)_4$

Vermiculite $[Mg(H_2O)_6]_{n/2}[(Mg,Fe^{++})_3(Si_{4-n}Al_n)O_{10}(OH)_2]$



Smectite $Na_x[(Al_{2-x}Mg_x)Si_4O_{10}(OH)_2]$

Chlorite $[AlMg_2(OH)_6][(Mg)_3(Si_3Al)O_{10}(OH)_2]$

Figure 14. Schematic structures of common soil layer silicates.

Three layer types occur, distinguished by the number of combined tetrahedral and octahedral sheets, which is further divided into five mineral groups according to kinds of isomorphic substitution. The kaolin group consists of 1:1 layer silicates with a tetrahedral sheet and an octahedral sheet. The illite, vermiculite and smectite groups are 2:1 layer silicates and are built of an octahedral sheet sandwiched between two tetrahedral sheets. The chlorites are 2:1:1 layer silicates consisting of a 2:1 layer structure as in the micas, but instead of K^+ it contains an extra octahedral sheet.

The three 2:1 groups differ from another in two ways: (1) by the layer charge decreasing in the order illite > vermiculite > smectite and (2) by vermiculite exhibiting higher substitution in the tetrahedral sheets than smectite. The smectites with layer charge of 0.25-0.6 per formula unit due to isomorphous substitution in mainly the octahedral sheet are freely expanding, that means, the cations balancing the layer charge are readily exchangeable. The higher layer charge of the vermiculites, 0.6-0.9 per formula unit, which is caused by substitution in the tetrahedral sheets, results in limiting expansion of these minerals, and hence of cation exchangeability. In illites the layer charge can be up to 1.0 per formula unit as in micas from which they are formed. The high charge originating from isomorphous substitution in the tetrahedral sheets results in a strong coulombic attraction for charge-compensating cations. Therefore, the K^+ occurs in a nonexchangeable or fixed form between 2:1 layers in illite. In chlorites the negative charge of the 2:1 layer is balanced by the positively charged extra octahedral (hydroxy) interlayer.

Table 6. Selected properties of common soil layer silicates.

Property	Kaolinite	Smectite	Vermiculite	Illite	Chlorite
Mineral type	1:1	2:1	2:1	2:1	2:1:1
Layer charge	~0	0.25-0.6	0.6-0.9	0.7-1.0	~1.0
CEC (cmol(+)/kg)	1-10	80-120	120-150	20-50	10-40
Spec. surf. area (m^2/g)	1-20	600-800	600-800	70-200	70-150
c-spacing (nm)	0.7	variable	1.0-1.5	1.0	1.4
Expansible	No	Yes	Partly	No	No
pH-dependent charge	Most	Little	Little	Medium	Most

Certain soil horizons, particularly upper Spodosol horizons, often contain socalled hydroxy interlayered minerals (HIM), where the charge-balancing cations in smectites and vermiculites are replaced by more or less polymerized hydroxyaluminium and -iron(III) compounds. Since the minerals resemble chlorite, they are sometimes called pseudochlorites.

Apart from isomorphous substitution, which determines the permanent, or pH-independent, charge, charge of silicate particles (and other soil constituents) can develope in response to changes of pH, socalled pH-dependent charge. This charge can be positive or negative depending on pH. On silicates pH-dependent charge originates mainly from amphoteric AlOH and SiOH groups occurring on edges of the minerals. Consequently, soil charge is the sum of permanent and pH-dependent charges (Figure 15). Some properties of the layer silicates are summarized in Table 6.

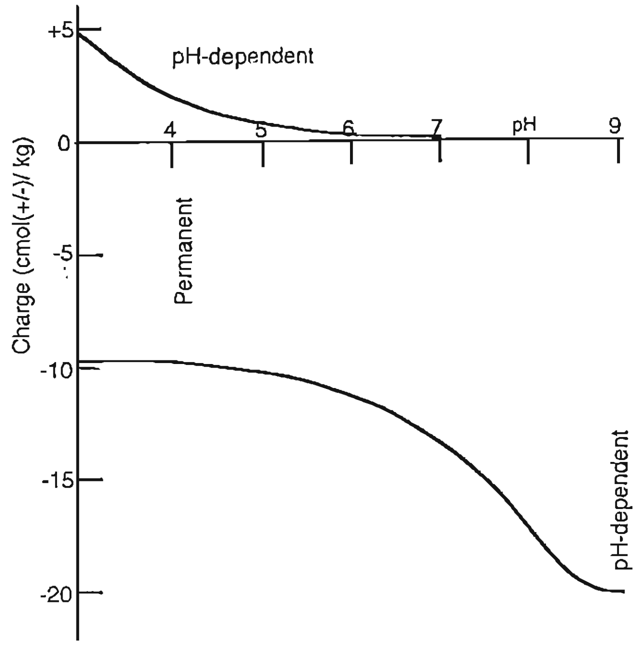


Figure 15. Sketch of soil charge versus pH. The positive charge, and hence the anion exchange capacity (AEC), is considerably smaller than the negative charge corresponding to the cation exchange capacity (CEC).

Accessory minerals

Soils also contain minor amounts of oxides, oxyhydroxides and hydroxides (collectively called oxides) of aluminium, iron, manganese and titanium, which are weathering products retained in the soil due to very low solubility (Deer et al., 1962; Dixon & Weed, 1989; Gieseking, 1975). These solids range in degree of organization from amorphous to crystalline. The most poorly ordered forms will have the largest specific surface area and will therefore be the most reactive in soils. The oxides have only pH-dependent charge properties but they are effective adsorbents for many organic and inorganic compounds (section 5.2).

While anatase (TiO_2) is the most commonly occurring soil titanium oxide, gibbsite ($Al(OH)_3$) is the most abundant form of free aluminium in soils. Both oxides are white. In contrast, manganese oxides (MnO_2) are black.

In Nordic soils the yellow goethite (α-FeOOH), the brown ferrihydrite ($5Fe_2O_3$, $9H_2O$) and the orange lepidocrocite (γ-FeOOH) are the most commonly occurring iron oxide forms, whereas the red hematite (α-Fe_2O_3) is not likely to form in cold climates and the black, magnetic magnetite (Fe_3O_4) appears to be inherited from the parent rock (Borggaard, 1990; Stucki et al., 1988). Since lepidocrocite only forms by oxidation of Fe(II) it serves as a pedogenic indicator of temporarily poor drainage (water-logging). In addition to these iron(III) oxides, zones at redox boundaries, e.g. around the groundwater table, probably contain small amounts of very reactive iron(II)-iron(III)-hydroxide compounds (green rust) consisting of two positively charged octahedral sheets with exchangeable charge-compensating anions between the sheets (Hansen et al., 1994).

Allophane and imogolite are structurally disordered aluminium silicates commonly found in clay fractions from soils formed on volcanic ash materials but have also been suggested to occur in certain Spodosols. The Al/Si ratio is often 1-2. While occurs as hollow sperules, imogolite has a tubular morphology.

Soils, at least in southern Scandinavia, formed on glacial till often contain calcium carbonate because of its occurrence in the parent material. Calcium carbonate is very important in counteracting soil acidification. Pyrite (FeS_2) occurs in many subsoils. It will only form and be stable under very reductive conditions as found under anaerobic organic matter decomposition in marsh areas. Under aerobic conditions pyrite will oxidize under formation of sulfuric acid.

SOIL PROPERTIES

The most important soil properties are the adsorption properties, i.e. the capacity of organic and inorganic solid soil components to act as sinks in nonspecific and specific adsorption of organic and inorganic anions, cations and uncharged molecules (Bohn et al., 1985; Brady, 1990; McBride, 1994; Schachtschabel et al., 1989; Sposito, 1989). Adsorption is sometimes called sorption or retention. Nonspecific adsorption (or outer-sphere complexation) is due to coulombic attraction between adsorbate (the solute) and adsorbent (the solid) with at least one water molecule interposed, whereas in specific adsorption (or inner-sphere complexation) the adsorbate is directly bound to the adsorbent with no water molecules interposed (Sposito, 1989). Several mathematical models have been developed in order to describe nonspecific and specific adsorption (Goldberg, 1993; Stumm, 1987).

Cation exchange

Nonspecific adsorption reduces the loss of ions by leaching while keeping them available for plant uptake. Nonspecifically adsorbed ions can be replaced or exchanged by other similarly charged ions. More precisely, ion exchange is the capability of soil particles to reversibly take up nonspecifically adsorbable anions and cations from solution. Typically, nonspecifically adsorbable ions include the anions of strong acids, e.g. Cl^- and NO_3^- and cations of strong bases, e.g. Na^+ and K^+.

The ion exchange properties in soils are attributed to the permanent and pH-dependent charge on soil particles. While layer silicates have both permanent and pH-dependent charge properties, SOM and oxides only have pH-dependent charge. Ion exchange reactions are extremely fast. Thus, cation exchange on clay silicates without narrow interlayer regions such as kaolinite and smectite appears to be instantaneous. At a negatively charged surface, cations are attracted while anions are repelled. Therefore regions close to the surface will be enriched in cations and depleted in anions (Figure 16).

Cation exchange prevails in soils because the negative charge on soil particles normally is considerably higher than the positive charge at naturally occurring soil pH (Figure 15). Therefore, unless the soil is very rich in pH-dependent charge components such as aluminium and iron oxides, which acquire a positive charge at low pH, nonspecific anion adsorption will be of limited importance (Borggaard, 1990).

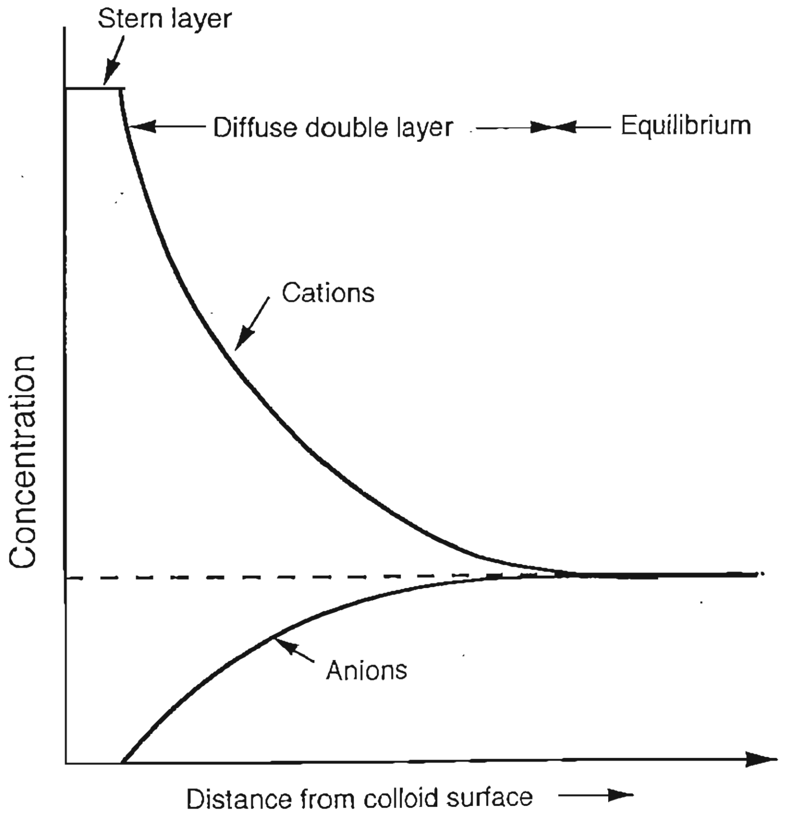


Figure 16. Concentrations of cations and anions at the surface of a negatively charged particle. The thickness in nm of the diffuse double layer is $0.303 \times I^{-0.5}$ at 25 °C, where I is the ionic strength.

The most important exchangeable cations in soils include the base cations, Na^+, K^+, Ca^{2+} and Mg^{2+}, and the acidic cations, H^+, Al^{3+}, $Al(OH)^{2+}$ and $Al(OH)_2^+$. In soils rich in ammonium, NH_4^+ will also contribute to the base cations, while in very acid soils Fe^{3+} and its mono- and dihydroxylated forms need consideration. The composition of the cation pool at the soil particle surfaces is determined by the environment, that is the cation composition and concentrations of competing cations in solution. When the soil meets a new environment, the composition will change. This is illustrated by the reaction schemes in Figure 17.

The cations are held at the negative particle surface by coulombic forces suggesting the composition of adsorbed cations to be a function of concentration only for ions of the same valence. However, the attraction is also affected by the size of the hydrated cation, the larger the hydrated radius of the ion the less tightly will it be held. This is shown by the relative replaceability in the socalled lyotropic series:

$Na^+ > K^+ \sim NH_4^+ > Mg^{2+} > Ca^{2+} > Al^{3+}$.

Compared to the other exchangeable cations, H^+ behaves somewhat different. Adsorption sites on pH-dependent charge components must de deprotonated ("titrated") before they are available for exchange reactions. Clay silicates saturated with H^+ are unstable and rearrange within rather short time to Al-saturated clay. Therefore H^+ will occur mainly as titratable acidity at normal soil pH.

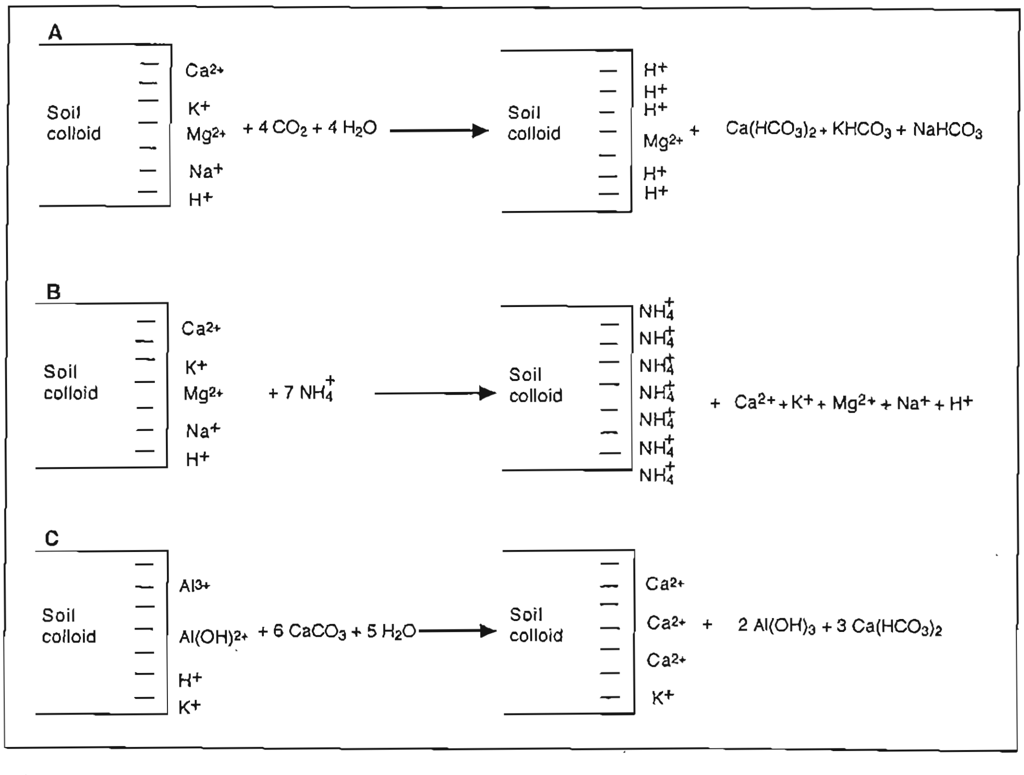


Figure 17. Various cation exchange reactions. A: Acidification by CO_2; B: Saturation with ammonium ions; C: Exchange of acidic cations by liming.

The cation exchange capacity (CEC) of a soil is the amount in cmol(+)/kg of base and acidic cations that can be adsorbed by the soil under given conditions:

$CEC = \Sigma(\text{base cations}) + \Sigma(\text{acidic cations})$.

Since CEC is a function of soil particle charge, which is partly pH-dependent, CEC will also be dependent on pH, that is, CEC increases at increasing pH. In soils dominated by permanent charge components such as smectite the increase is limited, whereas in soils rich in SOM and other pH-dependent charge components CEC will exhibit strong pH dependency.

In the laboratory soil CEC is often determined at soil pH (ECEC) using a neutral salt solution, e.g. KCl, at pH 7 (CEC_7) using ammonium acetate or at pH ~8 (CEC_8) using a buffer solution at that pH. Description of methods for CEC determination may be found elsewhere (Page et al., 1982; Rowell, 1994). In most Nordic soils the cation exchange capacities will increase in the order $ECEC < CEC_7 < CEC_8$.

Soil CEC is largely determined by the small soil particles including the colloidal SOM and clay minerals. It will therefore be dependent on abundance and composition of these fractions, which can show high variability in different soils. However, for Danish soils the following equation often gives a fair estimate of CEC:

$CEC_8 = 4x\%\text{organic carbon} + 0.5x\%\text{clay}$.

The percentage of CEC that is satisfied by base cations is termed percentage base saturation (BS):

$BS = 100x\Sigma(\text{base cations})/CEC$.

BS is closely related to soil pH. Often a linear or almost linear relation is found between BS and pH. Thus, in many mineral soils BS=0, which corresponds to complete base depletion, occurs at pH ~4 and BS=100, which corresponds to fully base saturation, will be at pH ~8 (if CEC_8 is used). In organic soils pH <4 at BS=0.

Under natural conditions soils are subject to acidification because of many external sources of acidity (atmospheric CO_2, HNO_3, H_2SO_4) and because of internal proton sources in the soil (CO_2 produced by root and microbial respiration, organic acid production during organic matter decomposition, nitrification leads to H^+ formation). Acidification leads to a decrease in BS, and hence in pH, as shown in Figure 17. At the same acid load, soils such as sandy soils with low CEC, and therefore low buffer capacity, will turn acidic faster than soils with high CEC. To counteract the negative environmental and vegetational effects of acidification, soils can be limed. The reaction resulting from liming an acid soil is shown in Figure 17.

Specific adsorption

Although trace (or heavy) element cations such as Co^{2+}, Cu^{2+}, Cd^{2+}, etc. can be adsorbed by cation exchange in soils, their strong adsorption even in the presence of high concentrations of exchangeable base cations neccessitates an alternative mechanism, that is specific adsorption. Rather than the permanently charged sites on planar surfaces of layer silicates, it is the amphoteric AlOH and FeOH groups on aluminum and iron oxides and edges of silicates together with carboxyl and phenol groups on SOM, i.e. the pH-dependent charge sites, which are the adsorption sites for these metals. Manganese oxides are important adsorbents for Co^{2+}.

At low pH the amphoteric groups are positively charged, whereas at alkaline pH they will be negatively charged. The pH at which the charge is zero is called the zero point of charge (ZPC). Pure aluminium and iron oxides will have ZPC near 7 (Borggaard, 1990). The ZPC is affected by specific adsorption (Sposito, 1989). An important characteristic of specific cation (and anion) adsorption is that adsorption can occur at pH values, where adsorbent and adsorbate have similar charge. Hence heavy metal cations are also adsorbed at pH below adsorbent ZPC.

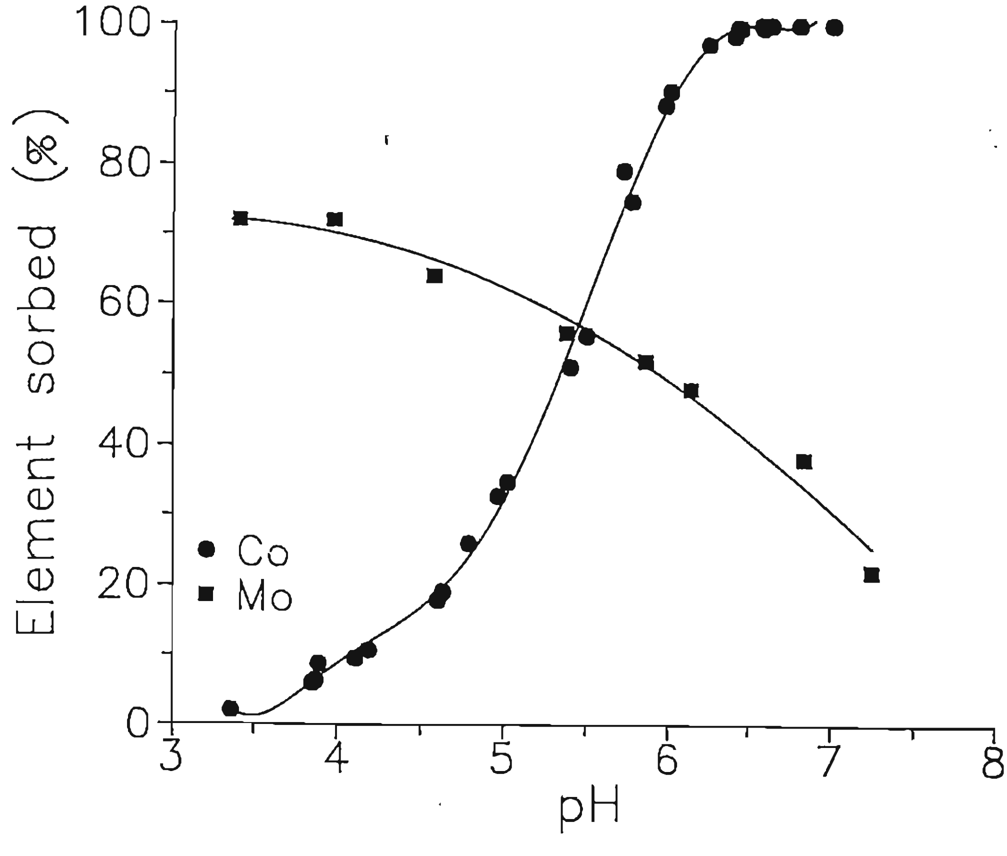


Figure 18. Adsorption edge for Co^{2+} and adsorption envelope for MoO_4^{2-}. The adsorbent is a soil sample from a Danish Humic Haplorthod. The initial cobalt concentration was 1 μM, whereas the initial molybdate concentration was 50 μM. Background electrolyte 0.1 M $NaNO_3$.

Specific adsorption of these metal cations is strongly dependent of pH raising within a narrow pH range from negligible adsorption at low pH to high adsorption at higher pH. The curve resulting from plotting amounts adsorbed versus pH is called an adsorption edge. This is shown for adsorption of cobalt by a Spodosol sample in Figure 18. The position of the adsorption edge along the pH axis depends on the metal and on the adsorbent. It is also affected by concentration and reaction time. At increasing reaction time the adsorption edge is shifted to lower pH corresponding to increased adsorption, which has been interpreted as diffusion of the metal ions into micropores in the adsorbent (Bibak et al., 1994).

Adsorption of heavy metal ions is followed by a release of protons. For iron oxides the proton release corresponds to 1-2 protons per metal ion adsorbed (Borggaard, 1990). This is shown schematically by the reactions in Figure 19.

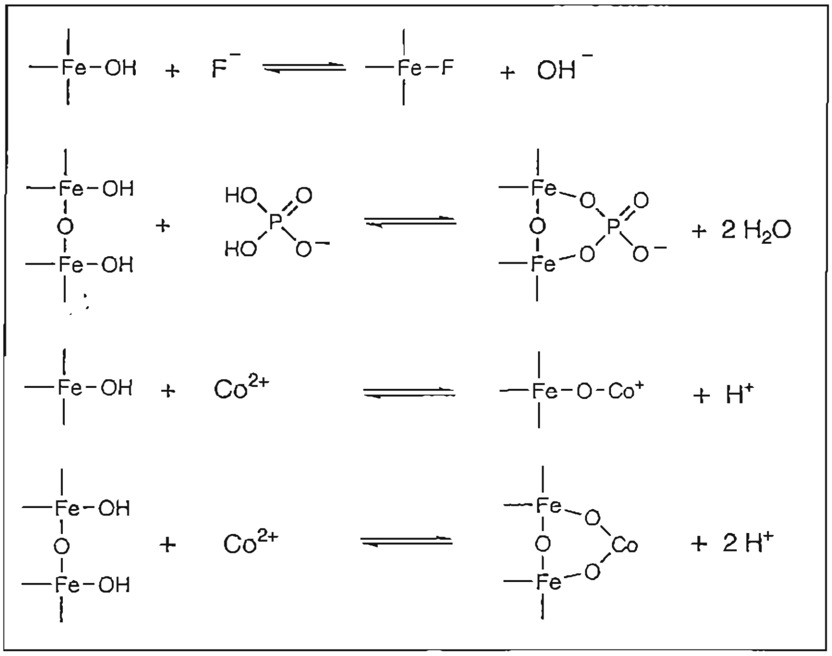


Figure 19. Examples of specific anion and cation adsorption onto an iron oxide surface.

In contrast to their weak nonspecific adsorption properties, most soils including those in Nordic countries often have significant specific anion adsorption properties. Thus, anions of weak organic and inorganic acids such as citrate, oxalate, fluoride, molybdate and phosphate can be retained by specific adsorption (Borggaard, 1990; Parfitt, 1978). Adsorption occurs as a result of ligand exhange as shown in Figure 19. The pH dependency of specific anion adsorption exhibits the socalled adsorption envelope with maximum adsorption around pH corresponding to the pK of the adsorbate and decreasing adsorption at lower and higher pH values (Figure 18).

Aluminium and iron oxides are considered the most important components acting as specific anion adsorbents in Nordic soil. SOM affects specific anion adsorption but often indirectly, because it retards development of crystalline aluminium and iron oxides resulting in poorly crystalline oxides with high specific surface area and hence high reactivity (Borggaard, 1990).

However, metallic anions such as molybdate has been shown to react directly with SOM, probably by complexation of Mo to organic ligands (Bibak & Borggaard, 1994).

WEATHERING PROCESSES

Weathering includes disintegration by physical forces and chemical decomposition. Disintegration or mechanical weathering is accompanied by a variety of physical processes such as freeze-thaw action, exfoliation and plant wedging, all of which break the rock materials into smaller pieces that retain the chemical composition of the parent rock. Although mechanical weathering is an important prerequisite for soil formation, only chemical weathering will be shortly reviewed here. More information about weathering may be found elsewhere (Birkeland, 1974; Dixon & Weed, 1989; McBride, 1994).

Primary minerals formed at high temperature in the absence of oxygen are unstable under the humide, oxygen-rich and often acidic conditions at the earth surface. Therefore primary minerals will be transformed or weather to secondary minerals. The resistance of igneous rock minerals to weathering is the same as the order of crystallization from cooling magmas (Figure 20). Minerals such as olivines, pyroxenes and Ca-feldspar that are most stable at high temperatures, and that therefore crystallize first from the molten magma, are the least stable in soils.

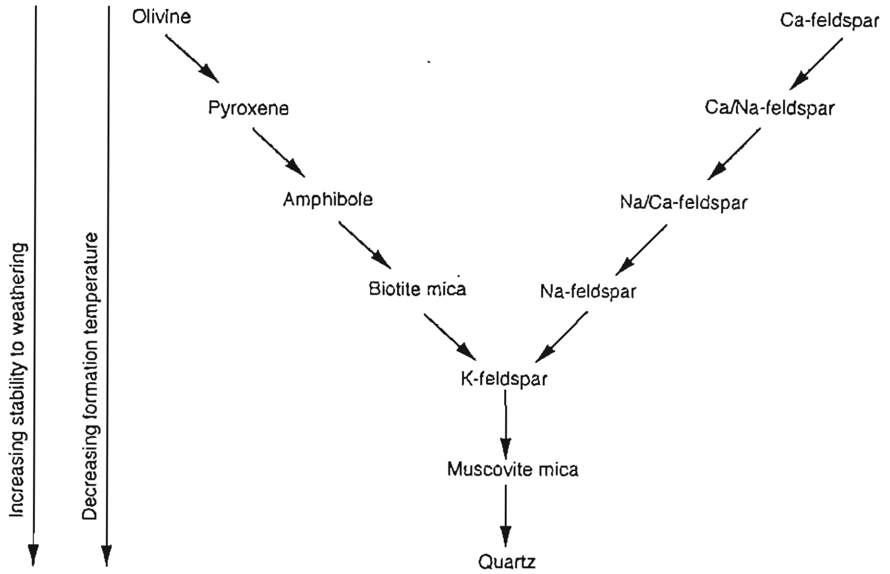


Figure 20. Series of primary minerals occurring in soils. Left and right sequences are ferromagnesian and nonferromagnesian minerals, respectively.

Overall weathering processes are irreversible even though individual steps may be considered reversible. The main chemical mechanisms of weathering include (1) exchange reactions, where structural base cations are replaced by solution cations, primarily H^+; (2) hydration reactions, where water addition leads to metal-oxygen bond hydrolyzation; (3) oxidation reactions, where electrons are transferred from the mineral to an electron acceptor, often oxygen. The importance of these mechanisms are illustrated by weathering of olivine, biotite, K-feldspar and Ca-feldspar:

$$Fe_2SiO_4 + 0.5O_2 + 3H_2O \rightarrow 2FeOOH + H_4SiO_4$$

$$2K(Mg_2Fe)[AlSi_3]O_{10}(OH)_2 + 0.5O_2 + 4.5H^+ + 6.5H_2O \rightarrow 1.25Mg^{2+} + 2K^+ + Mg_{0.25}(Mg_{2.5}Fe(III)_{0.5})[AlSi_3]O_{10}(OH)_2 + Al(OH)_3 + 1.5Fe(OH)_3 + 3H_4SiO_4$$

$$2KAlSi_3O_8 + 2H^+ + 9H_2O \rightarrow Al_2Si_2O_5(OH)_4 + 4H_4SiO_4 + 2K^+$$

$$3CaAl_2Si_2O_8 + 0.4Mg^{2+} + 4.8H^+ + 4.2H_2O \rightarrow 2.8Ca^{2+} + Al_2Si_2O_5(OH)_4 + Ca_{0.2}(Al_{1.6}Mg_{0.4})[Si_4]O_{10}(OH)_2 + 2.4Al(OH)_3$$

In addition to silicates many (young) soils contain calcium carbonate, which is easily weathered (decalcification):

$$CaCO_3 + CO_2 + H_2O \rightarrow Ca^{2+} + 2HCO_3^-$$

The reactions indicate the general characteristics of primary mineral weathering including (1) loss of tetrahedral coordinated Al (octahedral Al is favoured at low temperature); (2) oxidation of Fe(II); (3) participation of water; (4) consumption of H^+; (5) release of silica and base cations (Ca^{2+}, Mg^{2+}, K^+, Na^+). Often organic and inorganic ligands are participating in the reactions.

The released silica and base cations will be available for plant uptake or they will be removed by leaching. When the elements are released into solution in proportion to their mole fractions in the mineral, the dissolution is congruent while selective release of one or more elements into solution is called incongruent dissolution.

Initially formed weathering products are subject to further weathering leading ultimatively to the final weathering products. This is indicated by the weathering sequence in Table 7. In general, mineral weathering rates will be determined by several factors including (1) intrinsic structural stability; (2) pH; (3) presence of complexing ligands (organic and inorganic acids and anions); (4) temperature; (5) specific surface area of the mineral; (6) efficiency of removal by leaching, precipitation, etc. of soluble weathering products.

Table 7. Sequence of clay mineral distribution with increasing soil development.

Relative degree of soil development	Prominent minerals in soil clay fraction
1	Gypsum
2	Calcite, dolomite, apatite
3	Olivines, amphiboles, pyroxenes
4	Biotite
5	Feldspars
6	Quartz
7	Muscovite
8	Vermiculites, illites
9	Smectites
10	Kaolinite
12	Gibbsite
12	Iron oxides
13	Titanium oxides, zircon

The mineralogy of Nordic soils is often very complex, particularly because of the cold climate and the relative short time of development, which have resulted in limited weathering. The content of weatherable minerals (primary minerals except quartz) in silt and sand fractions can be high and the clay fraction often consists of many minerals. Thus, although illites appear to be the dominant group of clay silicate minerals, smectites, vermiculites, chlorites, HIMs and kaolines are also commonly occurring but in highly variable amounts (Møberg, 1991; Møberg et al., 1988; Olsson & Melkerud, 1989). In addition, the soils contain minor amounts of aluminium, iron, manganese and titanium oxides.

SOIL PROCESSES

Soils in Denmark, Finland, Norway and Sweden can be classified as Histosols, Spodosols, Ultisols, (Mollisols,) Alfisols, Inceptisols and Entisols according to the Soil Taxonomy system and as Histosols, Anthrosols, Leptosols, Fluvisols, Gleysols, Arenosols, Regosols, Podzols, Podzoluvisols, Alisols, Luvisols and Cambisols according to the FAO/Unesco system (Breuning-Madsen & Jensen, 1992; Møberg et al., 1988; Møberg & Breuning-Madsen, 1991; Rasmussen et al., 1991). The most important processes in these soils (except the Histosols) include decalcification, gleization, lessivage and podzolization. While lessivage and podzolization are very complex processes integrating the fundamental processes of humification (section 4.3), weathering (section 6) and transport (section 4.2.2), decalcification and gleization are more simple. These processes are shortly described below. More information can be found elsewhere (Birkeland, 1974; Boul et al., 1989; De Coninck, 1980;

Duchaufour, 1982; Fanning & Fanning, 1989; Petersen, 1976; Wilding et al., 1983).

Gleization

Gleization processes are considered to explain the origin of soil features, especially colour patterns, formed in response to wetness-induced reduction of iron and manganese in some or all parts of a soil. By reduction iron and manganese are solubilized because of formation of Fe(II) and Mn(II), which can leach or diffuse from one zone to another within the soil or they may be leached from the soil. Typically Fe(II) and Mn(II) move to zones of higher redox potential (aerated zones), where they precipitate as iron and manganese oxides. The processes tend to give low chroma (typically gray) colours to those parts of the soil from which iron has been removed and high chroma (typically brown, orange, yellow) where the iron oxides have accumulated. Manganese oxides are black. The different colouring is termed mottling.

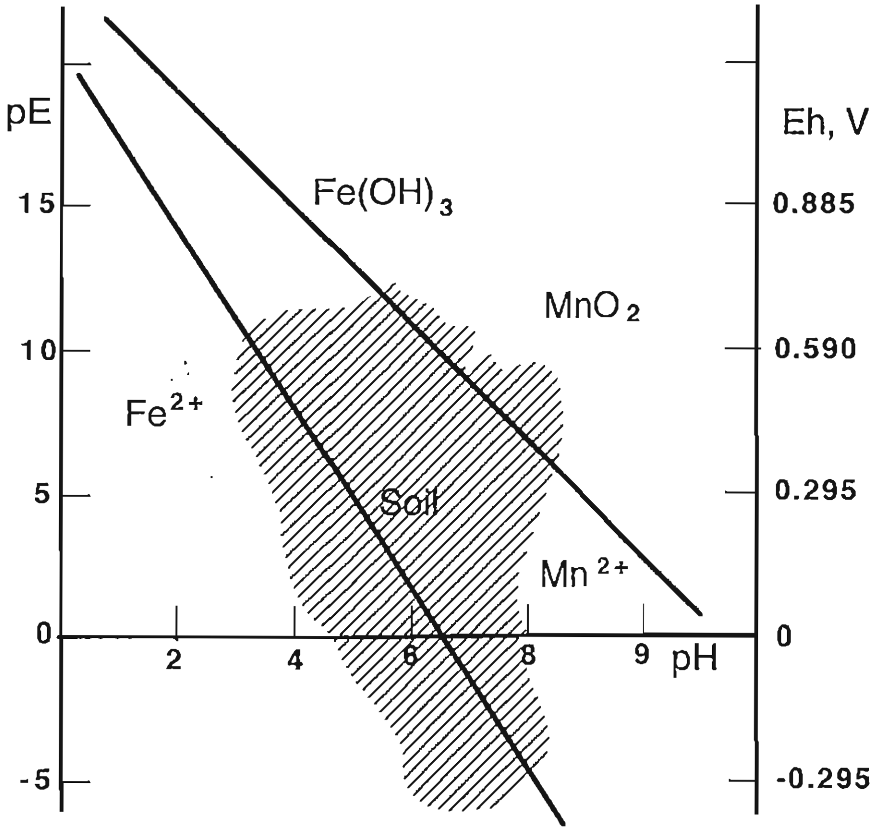


Figure 21. Stability fields for the Fe(III)/Fe(II) and Mn(IV)/Mn(II) systems. Shading indicates soil measurements. The straight lines correspond to: $pE = 16.0 - 3pH + pFe(II)$ and $pE = 20.7 - 2pH + 0.5pMn(II)$, where $pFe(II) = pMn(II) = 4$.

Reduction of iron and manganese oxides are followed by proton consumption and oxidation of Fe(II) and Mn(II) by proton release as shown for the iron system:

$$FeOOH + e + 3H^+ \Leftrightarrow Fe^{2+} + 2H_2O$$

Consequently, the redox potentials for the Fe(III)/Fe(II) and Mn(IV)/Mn(II) systems depend on pH as shown in Figure 21. It should be noticed that the straight lines for the two systems are not coinciding suggesting that oxidation and reduction occur at different redox potentials, and hence different environments. At the same pH, lower redox potentials are needed to reduce iron oxides than manganese oxides. On the other hand, at redox potentials high enough to cause iron oxide formation, manganese may still exist as Mn(II). Therefore, iron oxides and manganese oxides enrichments in soils often occur separated. This is schematically shown in Figure 22 by separated rims and nodules of the two oxides. However, as outlined in section 5.2 iron oxides are effective adsorbents for heavy metal ions including Mn^{2+}. Soil iron oxide nodules are therefore often enriched in manganese because, once formed, they will adsorb Mn(II).

Exclusion of oxygen and microbial activity leads to oxygen depletion and low redox potentials. Such anaerobic conditions will occur when the soil is water-logged either permanently or temporarily. Soils, which are permanently water-logged because of a shallow groundwater table, will be gley soils, whereas pseudogley soils are formed due to surface water accumulation (episaturation) over a slowly permeable soil horizon (perched or stagnant water table). Slowly permeable soil horizons or layer can be dense argillic (argic) horizons and various pans such as fragipans. The main difference between gley soils and pseudogley soils is that while the soil matrix is reduced in gley soils, it is oxidized in pseudogley soils (Figure 22).

Gleization processes are important in many Nordic soils. Clayey soils are often mottled due to low hydraulic conductivity, and hence temporare water saturation (pseudogley). Soils at low places in the landscape often suffer from shallow groundwater tables suggesting that they may be gley soils. The extent of gleization is quantified in the FAO/Unesco and Soil Taxonomy systems (FAO/Unesco, 1990; Soil Survey Staff, 1992) by socalled gleyic and stagnic properties and aquic conditions. Soils with gleyic properties will be classified as Gleysols according to the FAO/Unesco system, whereas the Soil Taxonomy system addresses aquic properties at the suborder level, e.g. Aqualfs and Aquods.

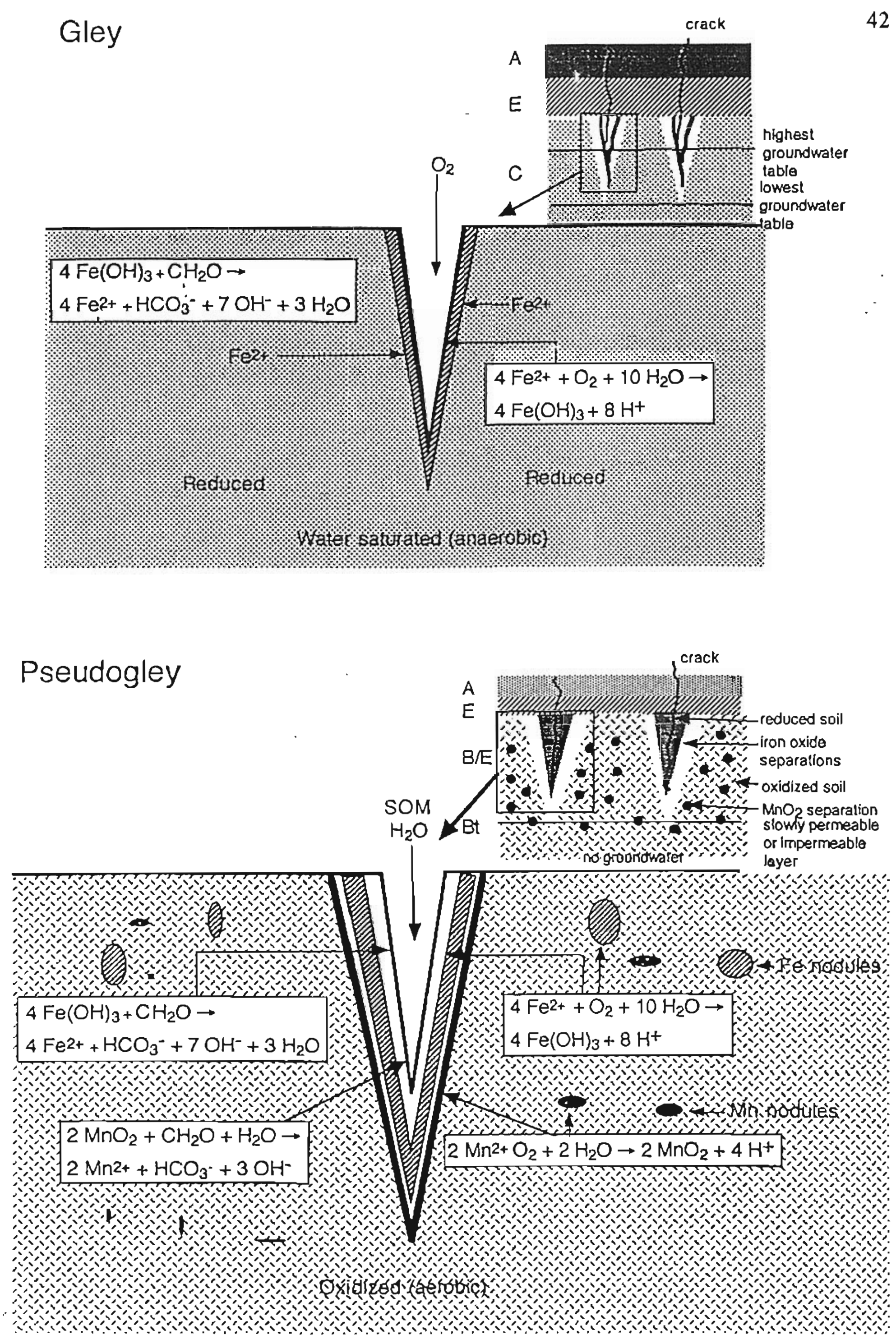


Figure 22. Schematic diagram showing iron and manganese oxide depletions and concentrations in gley and pseudogley soils.

Lessivage

Lessivage includes all the processes leading to migration by eluviation and illuviation of clay minerals from the A and E horizons to the B horizon. If the clay accumulation exceeds certain limits the B horizon will fulfil the requirements to an argillic (argic) horizon (section 3). Lessivage consists of two main steps, that is eluviation by leaching water after clay dispersion and deposition by flocculation or other mechanisms of the clay particles.

In general, two similarly charged particles such as two negatively charged smectites platelets will affect each other by opposite acting forces,.,mass attraction and charge repulsion. The result of these forces is determined by the thickness of the electric double layer (Figure 16). Since mass attraction is due to short-range forces, repulsion will prevail if double layers of the particles are thick. Flocculation results if the particles are attracted to each other, whereas repulsion causes dispersion. Thickness of the elctric double layer is a function of ionic strength, i.e. the concentration and charge of the ions. The clay minerals will disperse in solutions with low concentrations of, in particular, monovalent ions but will flocculate at high concentrations, especially if the ions are di- or trivalent.

Saturation with monovalent ions, particularly Na^+, promote dispersion, whereas Ca^{2+}- and Al^{3+}-saturation causes flocculation. Therefore, clay minerals in calcareous soils will not be dispersed before calcium carbonate has been removed (by decalcification) and part of the Ca^{2+} on the clay particles has been replaced by monovalent ions. On the other hand, since Al^{3+}-saturation leads to flocculation, lessivage will not be important in acid soils. SOM by complexation of di- and trivalent ions or because of adsorption onto clay particles and iron oxides due to their positive charge can affect clay silicate dispersion, and hence lessivage.

The composition of the clay fraction and the amount of precipitation (leaching) stongly affects lessivage. Smectites, which have a high charge and are the smallest clay particles, will most easily be dispersed, and hence transported, whereas kaolinite migrates only slowly. Lessivage appears to be most pronounced in a climate with dry and wet seasons. By drying cracks will form in the soils. When the rain comes, the water with its suspended clay will move into these cracks. Since the soil matrix may still be dry, the water will be removed by capillary withdrawal into the matrix leading to formation of clay cutans (argillans) on void walls. Therefore, argillic horizons will not form under very wet conditions (perudic), where precipitation exceeds evapotranspiration throughout the year.

Clay migration will stop due to flocculation when the clay suspension run into calcium carbonate containing soil layers. Accordingly, clay accumulation is often found in or just above a calcareous horizon. The development sequence for soils on calcareous till under udic moisture conditions will therefore start with removal of calcium carbonate by decalcification and leaching from upper soil layers followed by lessivage. Since natural and anthropogenic acids will cause decalcification to increasing depth over time clay migration will proceed. However, at a certain stage of development the upper boundary of the argillic horizon tends to degrade, which may lead to formation of a glossic horizon with tongues of albic material from the E horizon extending into the argillic horizon. This is schematically shown in Figure 23. Apart from decalcification and decreasing base saturation, acidification is also important for formation by weathering of migrating clay minerals.

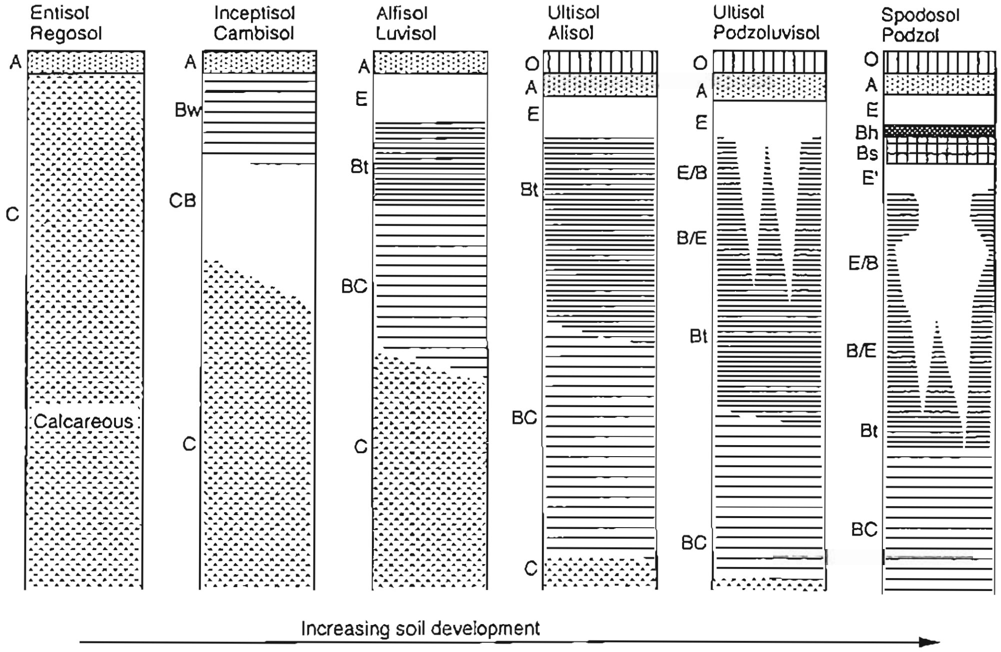


Figure 23. Likely development of soils formed on calcareous till in a humid climate.

Soils formed by lessivage are commonly occurring particularly in southern Scandinavia, where many soils developed on calcareous till are Alfisols (Luvisols), whereas cleyey soils developed on more acidic deposits will be Ultisols (Alisols). The deveolopment sequence shown in Figure 23 indicates main features in the formation of such soils.

Podzolization

Podzolization is the process leading to the formation of Spodosols or Podzols, which are typically formed on coarse textured materials in a cold and humid climate. In fact, it is not a single process but rather a bundle of processes bringing about translocation by leaching water, under the influence of protons and organic compounds, of aluminium with or without iron from upper to deeper soil layers. This transport can lead to the formation of a spodic horizon. The horizontation and horizon composition of a typical Spodosol are shown in Figure 2 and Tables 2 and 3.

Several theories have been suggested to explain podzolization but many aspects of Spodosols and their formation are still unresolved. According to the organic complexation theory negatively-charged, water-soluble organic compounds (fulvic acid) formed in base cation depleted topsoil layers take up aluminium and iron from inorganic soil components during its downward movement in leaching water. Increasing uptake of aluminium and iron ions results in a gradual reduction of the negative charge of the complex (or maybe more correctly, the association) leading eventually to saturation with aluminium and iron and then precipitation of the association. Continued production and leaching of organic compounds with varying composition (molecular weight and number of functionalities) will result in formation of various spodic subhorizons (Bh, Bhs, Bs) and the gradual descending of the spodic horizon. Other theories are focusing on microbial degradation of the organic compounds as the cause for precipitation and the occurrence of imogolite-like compounds, found in some Spodosols, as the active carrier of aluminium.

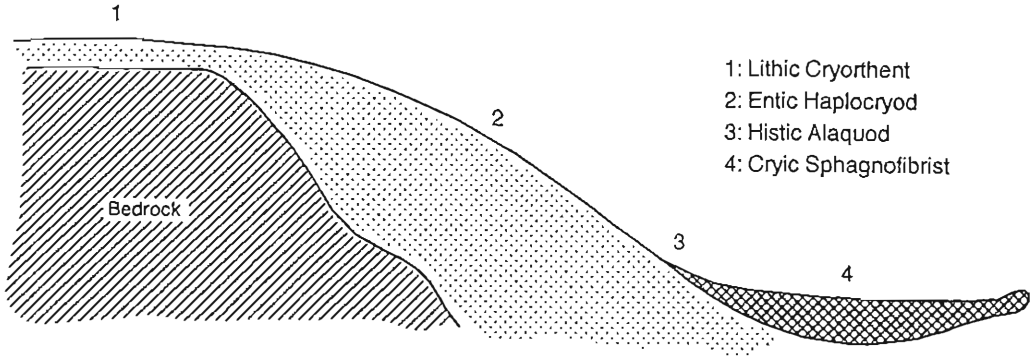


Figure 24. Sketch of a toposequence near Joensuu, Finland (Møberg, personal information).

While aluminium can not be transported unless it is associated or complexed to water-soluble organic (or inorganic) compounds, iron movement, and hence its depletion and concentration, can occur as a result of redox processes (section 7.1). This may, at least partly, be the reason for the great variability in the composition of various Spodosols. Some spodic horizons are strongly enriched in aluminium while iron is dominating in others. Spodic horizons high in aluminium are often more or less indurated, whereas a high iron content is not always followed by hardening. Spodosols with loose B horizons riched in iron often appears in soils formed on iron-rich primary silicates.

On the other hand, in soils obviously affected by gleization hardening is common. The placic horizon (section 3.1) is a special occurrence of an iron-indurated soil layer, presumably formed by iron reduction followed by downward diffusion and oxidation at a sharp redox boundary between an upper, reduced spodic subhorizon and a lower, oxidized subhorizon. Placic horizons seem to occur only in well developed ("mature") Spodosols with rather stongly cemented spodic horizons of low hydraulic conductivity.

Podzolization is restricted to strongly leached (acidic) soils poor in plant nutrients. Such conditions occur in soils formed on sandy materials in a cold and humid climate (cf. also final stage in Figure 23). Formation of poorly humified SOM (mor, section 4.3.1) appears to be closely linked to podzolization, possibly because this transformation leads to production of the organic compounds involved in the transport. Mor formation, and hence podzolization, is favoured by certain plants such as spruce and calluna. Other plants, particularly oak, will retard podzolization or maybe even revert it (depodzolization). Certainly, liming, fertilizing and plowing will cause depodzolization of Spodosols.

Many Nordic soils occur on coarse textured materials. Since the vegetation is often spruce or other conifers and the climates are cold and humid, podzolization will be the dominating process in many Nordic soils, at least in the more well-drained soils (Figure 24). Accordingly, Spodosols or more or less podzolized soils are commonly occurring in Denmark, Finland, Norway and Sweden.

Acknowledgement

The author thanks Per Jensen for making the drawings and Maria Strandberg for writing and organizing the text.

REFERENCES

Aiken, G.R., McKnight, D.M., Wershaw, R.L. & MacCarthy, P. (Eds.) 1985. Humic Substances in Soil, Sediment, and Water. Wiley-Interscience, New York.

Bibak, A. & Borggaard, O.K. 1994. Molybdenum adsorption by aluminium and iron oxides and humic acid. Soil Sci., in press.

Bibak, A., Gerth, J. & Borggaard, O.K. 1994. Retention of cobalt by pure and foreign-element associated goethites. Clays Clay Miner., in press.

Birkeland, P.W. 1974. Pedology, Weathering, and Geomorphological Research. Oxford,

186

New York.

Bohn, H.L., McNeal, B.L. & O'Connor, G.A. 1985. Soil Chemistry. 2nd Edition, Wiley, New York.

Booltink, H.W.G. & Bouma, J. 1991. Physical and morphological characterization of bypass flow in a well-structured clay soil. Soil Sci. Soc. Am. J. 55, 1249-1254.

Borggaard, O.K. 1990. Dissolution and Adsorption Properties of Soil Iron Oxides. DSR-Forlag, Copenhagen.

Brady, N.C. 1990. The Nature and Properties of Soils. 10th Edition, Macmillan, New York.

Breuning-Madsen, H. & Jensen, N.H. 1992. Pedological regional variations in well-drained soils, Denmark. Geografisk Tidsskrift 92, 61-69.

Breuning-Madsen, H., Nørr, A.H. & Holst, K.Aa. 1992. Atlas over Denmark. The Danish Soil Classification. Series 1, Vol. 3, The Royal. Danish Geographical Society, Reitzel, Copenhagen.

Buol, S.W., Hole, F.D. & McCracken, R.J. 1989. Soil Genesis and Classification. 3rd Edition, Iowa State University Press, Ames.

De Coninck, F. 1980. Major mechanisms in formation of spodic horizons. Geoderma 24, 101-128.

Deer, W.A., Howie, R.A. & Zussman, J. 1962, 1963. Rock Forming Minerals. Vol. 1: Ortho- and Ring Silicates; Vol. 2: Chain Silicates; Vol. 3: Sheet Silicates; Vol. 4: Framework Silicates; Vol. 5: Non-Silicates. Longmans, London.

Dixon, J.B. & Weed, S.B. (Eds.) 1989. Minerals in Soil Environments. 2nd Edition, Soil Science Society of America, Madison.

Douglas, L.A. (Ed.) 1990. Soil Micromorphology: A Basic and Applied Science. Elsevier, Amsterdam.

Duchaufour, P. 1982. Pedology. (Translated by T.R. Paton), Allen & Unwin, London.

Ehlers, J. (Ed.) 1983. Glacial Deposits in North-west Europe. Balkema, Rotterdam.

Fanning, D.S. & Fanning, M.C.B. 1989. Soil. Morphology, Genesis, and Classification. Wiley, New York.

FAO, 1990. Guidelines for Soil Description. 3rd Edition, Food and Agriculture Organization of the United Nations, Rome.

FAO/Unesco, 1990. Soil Map of the World. Revised legend, World Soil Resources Report 60, Food and Agriculture Organization of the United Nations, Rome.

FitzPatrick, E.A. 1984. Micromorphology of Soils. Chapman and Hall, London.

Gieseking, J.E. (Ed.) 1975. Soil Components. Vol. 2: Inorganic Components. Springer, New York.

Goldberg, S. 1993. Use of surface complexation models in soil chemical systems. Adv. Agron. 47, 233-329.

Hall, D.G.M. 1993. An amended functional leaching model applicable to structured soils. I. Model description. J. Soil Sci. 44, 579-588.

Hansen, H.C.B., Borggaard, O.K. & Sørensen, J. 1994. Evaluation of the free energy of formation of Fe(II)-Fe(III) hydroxide-sulphate (green rust) and its reduction of nitrite. Geochim. Cosmochim. Acta 58, 2599-2608.

Hayes, M.H.B., MacCarthy, P., Malcolm, R.L. & Swift, R.S. (Eds.) 1989. Humic Substances II. In Search of Structure. Wiley-Interscience, Chichester.

Hodgson, J.M. 1978. Soil Sampling and Soil Description. Clarendon, Oxford.

Jenny, H. 1941. Factors of Soil Formation. McGraw-Hill, New York.

Jørgensen, P.R. & Fredericia, J. 1992. Migration of nutrients, pesticides and heavy metals in fractured clayey till. Géotechnique 42, 67-77.

McBride, M.B. 1994. Environmental Chemistry of Soils. Oxford, New York.

Klute, A. (Ed.) 1986. Method of Soil Analysis. Part 1: Physical and Mineralogical Methods. 2nd Edition, Soil Science Society of America, Madison.

Munsell, 1975. Soil Color Charts. Macbeth, a division of Kollmorgen Corp., 2441 North Calvert Street, Baltimore, Maryland 21218.

Møberg, J.P. 1991. Formation and development of the clay fraction in Danish soils. Folia Geographica Danica 19, 49-61.

Møberg, J.P., Petersen, L. & Rasmussen, K. 1988. Constituents of some widely distributed soils in Denmark. Geoderma 42, 295-316.

Møberg, J.P. & Breuning-Madsen, H. (Eds.) 1991. Soil Research in Denmark. Folia Geographica Danica 19, Reitzel, Copenhagen.

Olsson, M. & Melkerud, P.-A. 1989. Chemical and mineralogical changes during genesis of a Podzol from till in southern Sweden. Geoderma 45, 267-287.

Page, A.L., Miller, R.H. & Keeney, D.R. (Eds.) 1982. Method of Soil Analysis. Part 2: Chemical and Microbiological Properties. 2nd Edition, Soil Science Society of America, Madison.

Parfitt, R.L. 1978. Anion adsorption by soils and soil materials. Adv. Agron. 30, 1-50.

Petersen, L. 1976. Podzols and Podzolization. DSR-Forlag, Copenhagen.

Rankama, K. (Ed.) 1965. The Geological Systems. Vol.1: The Quaternary. Interscience, New York.

Rasmussen, K., Sippola, J., Urvas, L., Låg, J. Troedsson, T. & Wiberg, M. 1991. Soil Map of Denmark, Finland, Norway and Sweden, Scale 1:2000000. Landbruksforlaget, Oslo.

Rowell, D.L. 1994. Soil Science: Methods and Applications. Longman, Essex.

Schachtschabel, P., Blume, H.-P., Brümmer, G., Hartge, K.-H. & Schwertmann, U. 1989. Lehrbuch der Bodenkunde. 12th Edition, Ferdinand Enke, Stuttgart.

Schnitzer, M. & Khan, S.U. 1978. Soil Organic Matter. Elsevier, Amsterdam.

Singer, M.J. & Munns, D.N. 1987. Soils. An Introduction. Macmillan, New York.

Soil Survey Staff, 1992. Keys to Soil Taxonomy. 5th Edition, SMSS technical monograph No. 19, Pocahontas, Blacksburg.

Spaargaren, O.C. (Ed.) 1994. World Reference Base for Soil Resources. ISSS, ISRIC, FAO, Wageningen/Rome.

Stucki, J.W., Goodman, B.A. & Schwertmann, U. (Eds.) 1988. Iron in Soils and Clay Minerals. Reidel, Dordrecht.

Stumm, W. W. (Ed.) 1987. Aquatic Surface Chemistry. Wiley, New York.

Tan, K.H. 1994. Environmental Soil Science. Marcel Dekker, New York.

Westerman, R.L. (Ed.) 1990. Soil Testing and Plant Analysis. 3rd Edition, Soil Science Society of America, Madison.

Wilding, L.P., Smeck, N.E. & Hall, G.F. (Eds.) 1983. Pedogenesis and Soil Taxonomy. I. Concepts and Interactions. II. The Soil Orders. Elsevier, Amsterdam.

Introduction to catchment hydrology

Allan Rodhe

Department of Earth Sciences, Uppsala University

In1674 Pierre Perrault published "Sur l´origine des fontaines". From measurements of precipitation and estimates of stream runoff he showed that the volume of precipitatation over the River Seine catchment was more than enough to feed the river with water. He was the first ot establish mass budget for water in a catchment, which has become the basis for modern hydrology: input - output = change in storage. But to-day, 300 years after Perrault, many basic questions still reamin to be answered concerning how the water is flowing through the catchment: which paths it follows and how long time the water stays in various parts. With some general experience from the behaviour of catchments, the mass budget together with some mathematics makes it possible to calculate the mass output of water (runoff) from the input (precipitation). It is not necessary to know in detail what happens within the catchment, which can be treated almost like a black box. But when it comes to understanding, or forecasting, the chemical changes taking place in the water during its flow through a catchment, we have to know the flow-paths and transit times for water in various biogeochemical surroundings. And these questions are little known and the focus for much of to-day´s hydrological research.

Model		*Catchment information*
$Q = f(P)$	with calibration	black box
$Q = f(P)$	without calibration	from catchment physiography, necessary to know the response of different types of sub-areas, where the flow impulses are given
$c_{out} = f(c_{in})$		need to know bio-geochemical environment, flow, flow-paths and transit times

Q is stream discharge, P precipitation and c concentration of a dissolved chemical constituent.

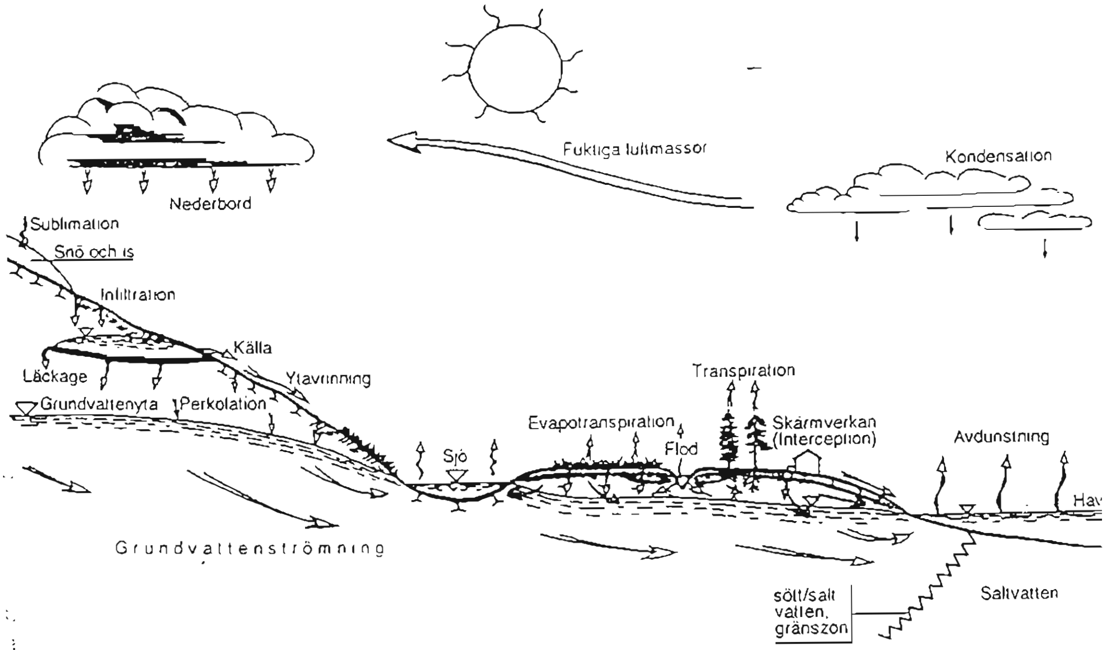


Fig.1. *The water cycle*: evaporation, transpiration, condensation, precipitation, interception, snowmelt, infiltration, percolation, groundwater flow, groundwater discharge, overland flow, stream runoff.

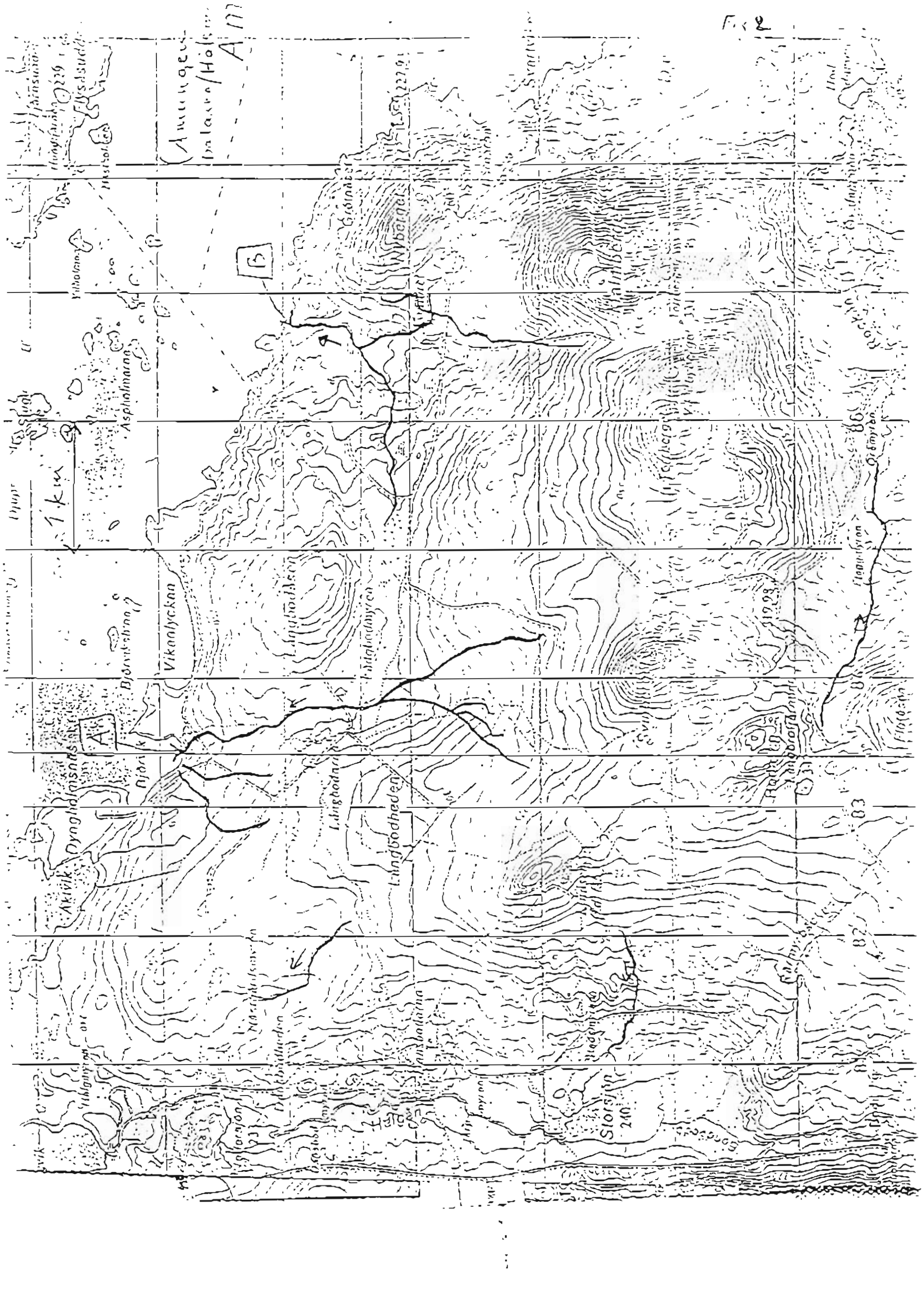
The catchment: water divide (surface water or topografic water divide and groundwater divide) (Fig 2,3 for exercice). The water balance equation reads

$$P = E + R + \Delta S \tag{1}$$

where P is precipitation, E evapotranspiration, R runoff and ΔS is change in storage, often expressed in mm/time where 1 mm water depth = 1 litre/m^2. ΔS can conveniently be divided into change in storage in surface water (ponds, streams, lakes), soil water, groundwater, snow and glaciers. Much of hydrology is concerned with the determination of these flows and changes in storage in very diffierent time and spacial scales. The effect of the storage term is to damp the time variability of the outflow terms (E and R). The rainfall lasts for a few hours, whereas the runoff and evapotranspirtion may continue for several days, fed by the declining storage.

For long time periods (several years) ΔS will be small as compared with the flow terms and the equation becomes

$$P = E + R \tag{2}$$


1 km
A
B
Storsjön

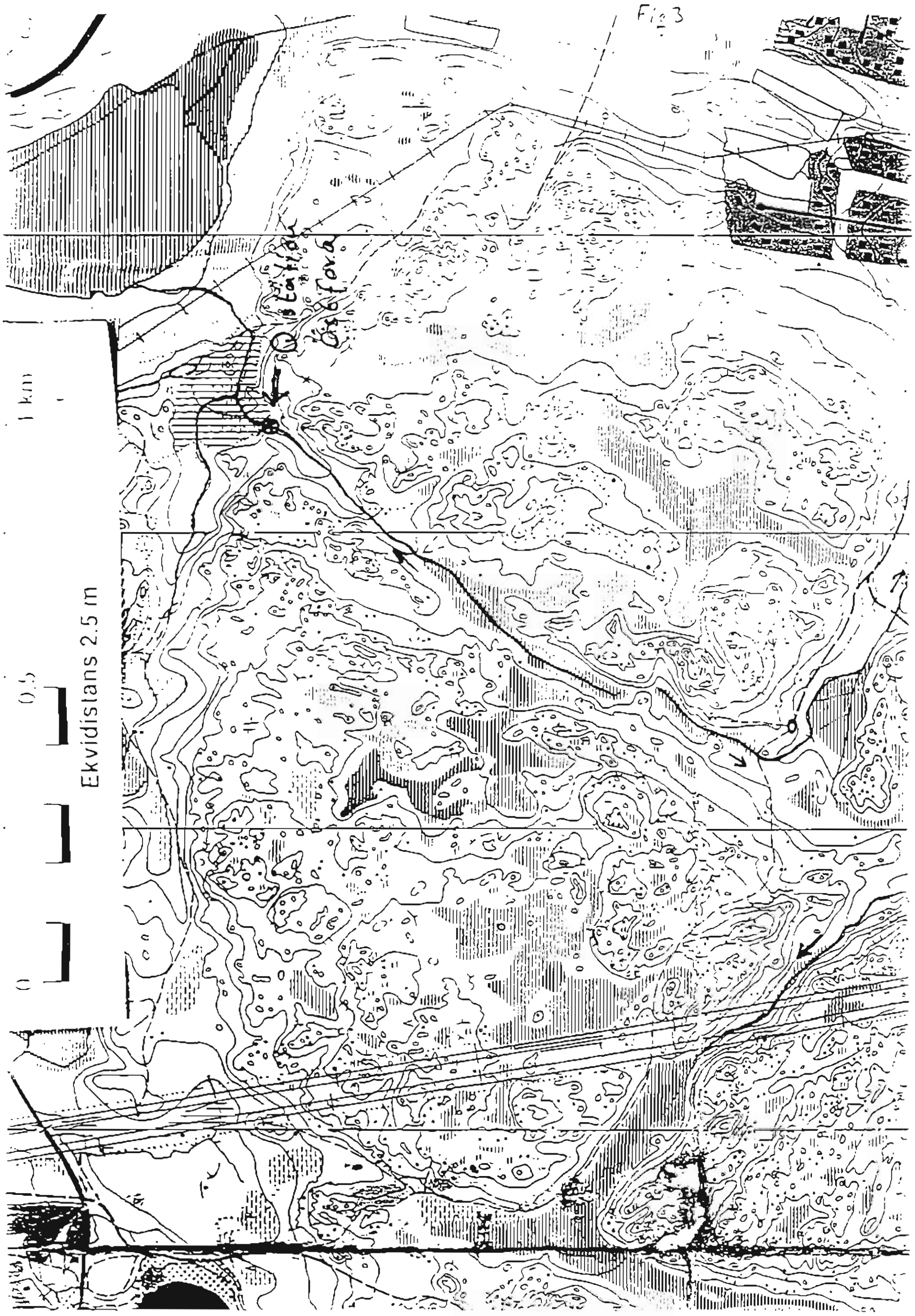


Fig 3
Östfora
Ekvidistans 2.5 m
0
0.5
1 km

Soil water storage and flow

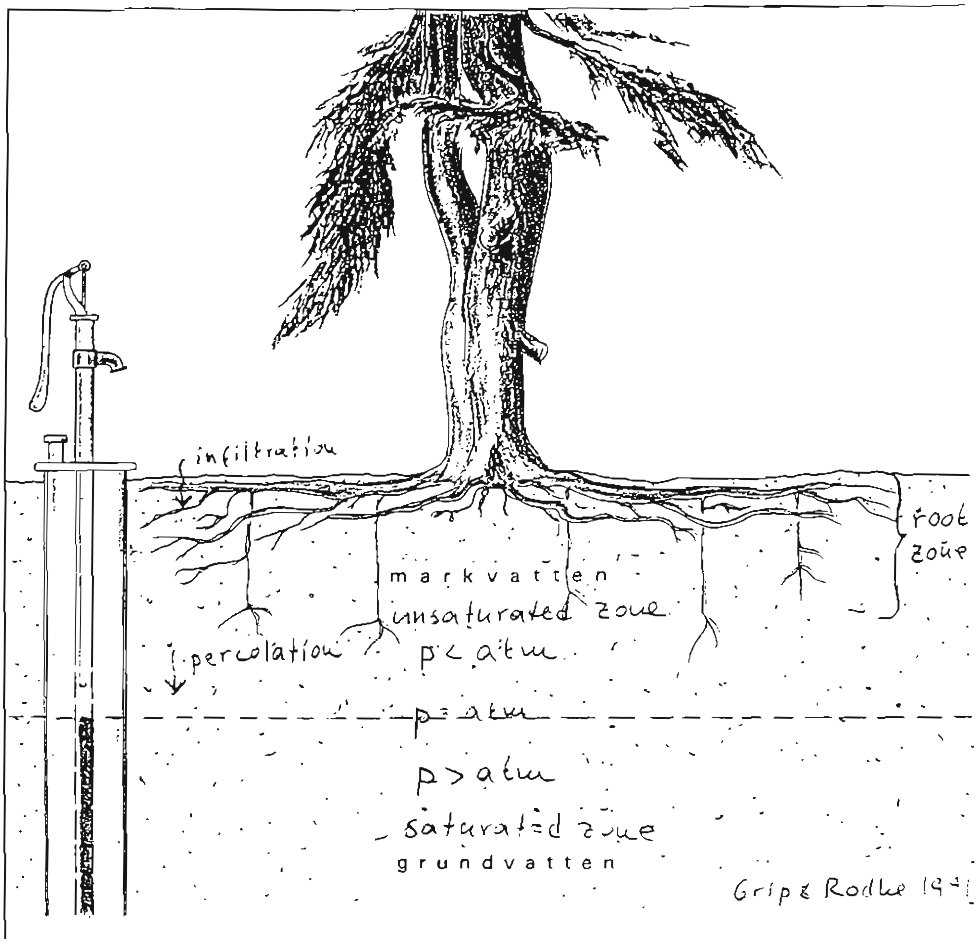


Fig 4: unsaturated and saturated zones, definitions and flow.

Demonstration:
- * pressure conditions (unsaturated < p_{atm}, saturated > p_{atm}. Let $p_{atm}=0$)
- *pF
- * field capacity - wilting point

$$\Phi = \Psi + z \qquad (3)$$

total potential = pressure potential + elevation potential
(J/m^3. Division by ρg gives J/N, i.e., m water depth)

The total potential is the height level of the water in the observation tube of a tensiometer (unsaturated) or piezometer (saturated). The flow takes place from high to low total potential.

Darcys law (1856), originally established for saturated flow, but later developed to include unsaturated flow

$$Q = -KA\frac{d\Phi}{dx} \qquad (4)$$

where Q = discharge (m^3/s)
K = hydraulic conductivity (m/s)
A = cross section area (m^2)
$d\Phi/dx$ = total potential gradient (m/m)

K increases rapidly with increasing radius of the (water-filled) pores.

Drainage equilibrium

If there is no vertical flow, i.e., $Q_z = 0$ then $d\Phi/dz = 0$ giving Φ = constant.
Thus $\Psi + z$ = constant, or $-\Psi = z$ + constant.
Choose the groundwater level ($\Psi = 0$) as reference height for z, then the constant = 0 and

$$-\Psi = z \qquad (5)$$

i.e., the suction (negative pressure) is equal to the height above the water table. This relationship holds at drainage equilibrium to a certain level above the water table, say a few dm in sand and up to a few m in more fine-grained soils. (At larger distances the hydraulic conductivity has declined so much, due to emptying of the larger pores, that the capillary contact is broken. Then Ψ tends to be constant and $d\Phi/dz$ is 1 m/m.

Demonstration *influence of the groundwater table on the soil moisture content.

The above influence of the groundwater table on the pressure potential in the unsaturated zone, and the fact that water cannot leave the soil unless $\Psi \geq 0$, may make the soil moisture in a freely draining lysimeter very different from that of the natural soil profile which it is supposed to represent. Consequently the conditions for biogeochemical processes may differ considerably.

(2) $\Phi = \psi + z$

(3) $\frac{dz}{dz} = 1$, dvs $\frac{d\Phi}{dz} = \frac{d\psi}{dz} + 1$

(4) b) $\frac{d\psi}{dz} = -1$ vilket ger $\frac{d\Phi}{dz} = 0$, dvs $Q = 0$, *inget flode*

(5) a) $\frac{d\psi}{dz} < -1$ vilket ger $\frac{d\Phi}{dz} < 0$, dvs $Q > 0$, *uppåtriktat flode*

(6) c) $\frac{d\psi}{dz} > -1$ vilket ger $\frac{d\Phi}{dz} > 0$, dvs $Q < 0$, *nedåtriktat flode*

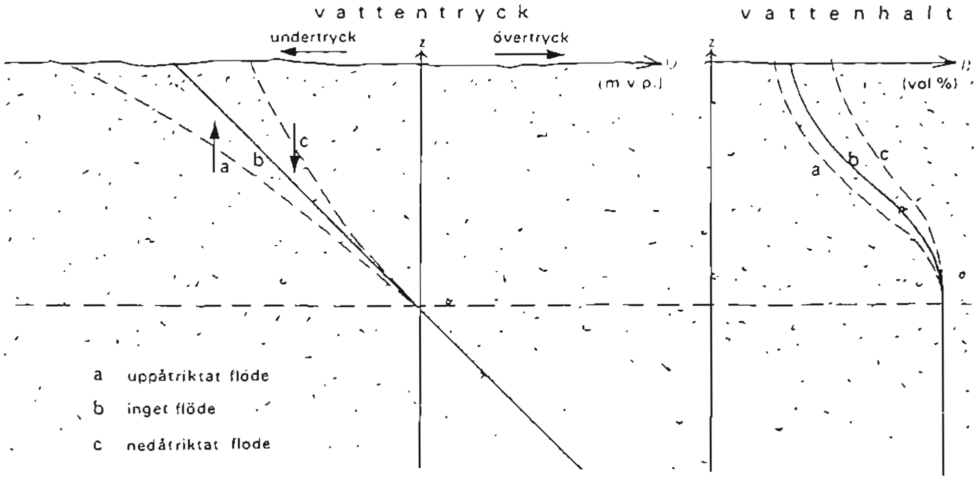


Fig 5. Watter content-water pressure, relationships with flow direction

Demonstration *A coarse layer links off unsaturated flow. Only water-filled pores can conduct water! The problem of water sampling in the unsaturated zone.

Groundwater

Unconfined aquiefer: aquifer with free water table, the groundwater table is the upper boundary of the groundwater zone.
Confined aquifer: the aquifer is overlain by a comparatively impermeable layer (aquitard). The groundwater level, defined as the water level in an observation well, is above the upper boundary of the aquifer.

Storage coefficient $M = \Delta S/\Delta h$ where S is storage (m^3/m^2 =m) and h is groundwater level (m).

Examples of M:

confined aquifer	0.001
coarse till soil	0.05
coarse sand	0.20

Δh = ΔS/M, i.e., small M gives large temporal variations in the groundwater level (Fig 8). For an unconfined auifer M = effective porosity = porosity - field capacity

$\phi = \psi + z$

at the groundwater table $\psi = 0$

$\Rightarrow \phi = z$

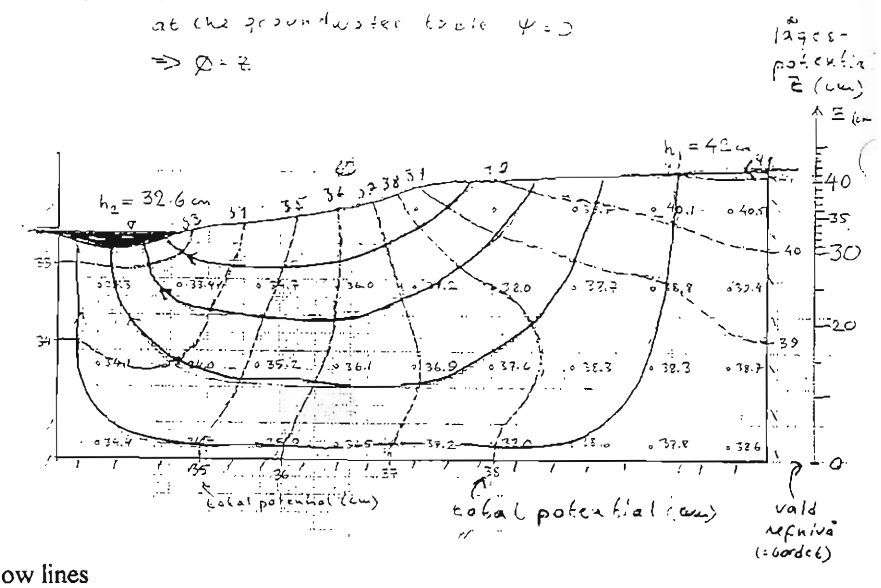


Fig. 6 Hillslope with flow lines

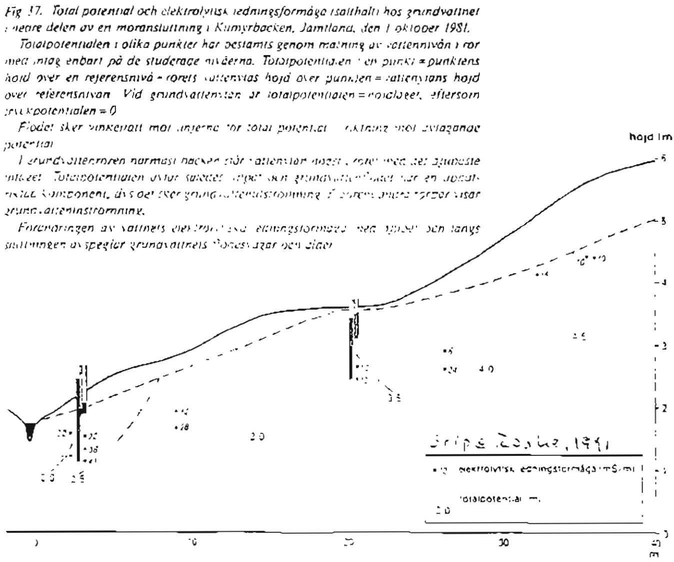


Fig 37. Total potential och elektrolytisk ledningsformåga (salthalt) hos grundvattnet i nedre delen av en moränsluttning i Kumyrbacken, Jämtland, den 1 oktober 1981.

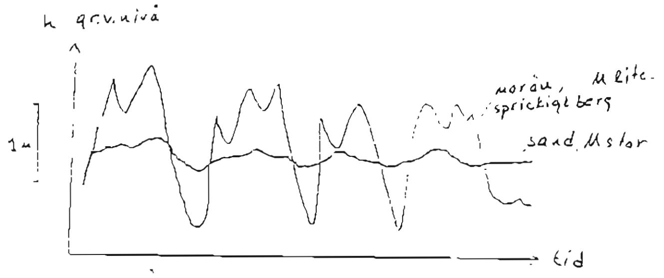
Totalpotentialen i olika punkter har bestämts genom mätning av vattennivån i rör med intag enbart på de studerade nivåerna. Totalpotentialen i en punkt = punktens höjd över en referensnivå + rörets vattenytas höjd över punkten = vattenytans höjd över referensnivån. Vid grundvattenytan är totalpotentialen = höjden, eftersom tryckpotentialen = 0.

Flödet sker vinkelrätt mot linjerna för total potential [illegible] mot avtagande potential.

I grundvattenzonen närmast bäcken [illegible]

Förändringen av vattnets elektrolytiska ledningsförmåga [illegible]

Fig. 7 Hillslope with flow lines.



magasins-koefficient $M = \frac{\Delta S}{\Delta h} = \frac{\Delta mag}{\Delta gr.v.nivå}$

Fig. 8 $\Delta h = \Delta S/M$, i.e., small M gives large temporal variations in the groundwater level.

The Dupuit assumption simplifies Darcys law for groundwater flow:
Φ is assumed to be constant with depth (equivalent with horisontal flow lines). Then

$$d\Phi/dx = dh/dx \qquad (6)$$

where h is the height of the groundwater table. This is the normal way to apply Darcys law.

$$Q = -K \cdot A \cdot dh/dx = -T \cdot b \cdot dh/dx \qquad (7)$$

Here b is the width (m) and T is the transmissivity (m^2/s) of the water conducting formation.

$$T = \int_{bottom}^{top} K(z)dz \qquad (8)$$

The direction of the groundwater flow is easily determined from the groundwater levels, being perpendicular to the height contours of the groundwater table.

Three "velocities" (m/s)

1. Darcian velocity $v_{Darcy} = \frac{Q}{A}$
2. Particle velocity $v_{particle} = \frac{Q}{A \cdot \theta}$

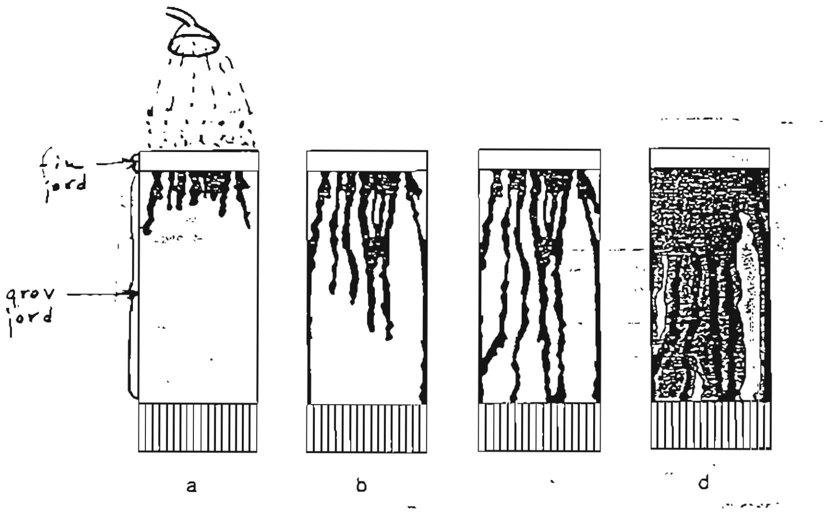
   where θ is volumetric water content (or rather content of mobile water)
3. Pressure propagation velocity $v_{pressure}$

Demonstration: * $v_{pressure} >> v_{particle} > v_{Darcy}$

Typical particle velocities:	soil water	1 - 2 m/year
	groundwater	loose deposits 0.01-0.1 m/day
		fractured rock 0.1-10 m/day
	forest stream	0.2 m/s
	river	1 m/s

Preferential flowpaths (macropore flow or fingering). Some water molecules by-pass the molecules in the soil matrix:

1. Only water-filled pores can conduct water. Large pores are empty in unsaturated soil.
2. If the hydraulic conducticityof the soil matrix is small as compared to the water input to the the pore, then the soil might get saturated around the pore, which then will conduct water.
3. In the groundwater zone a system of large pores will conduct water very efficiently.
4. Unstable unsaturated flow may occur when a coarse layer underlies a fine layer. Small areas with temporary almost saturation build up, giving local breakthroughs into the coarse layer (fingering). Fig.9.



Streamflow generation

How is rainfall or snowmelt over a catchment transformed into stream runoff?

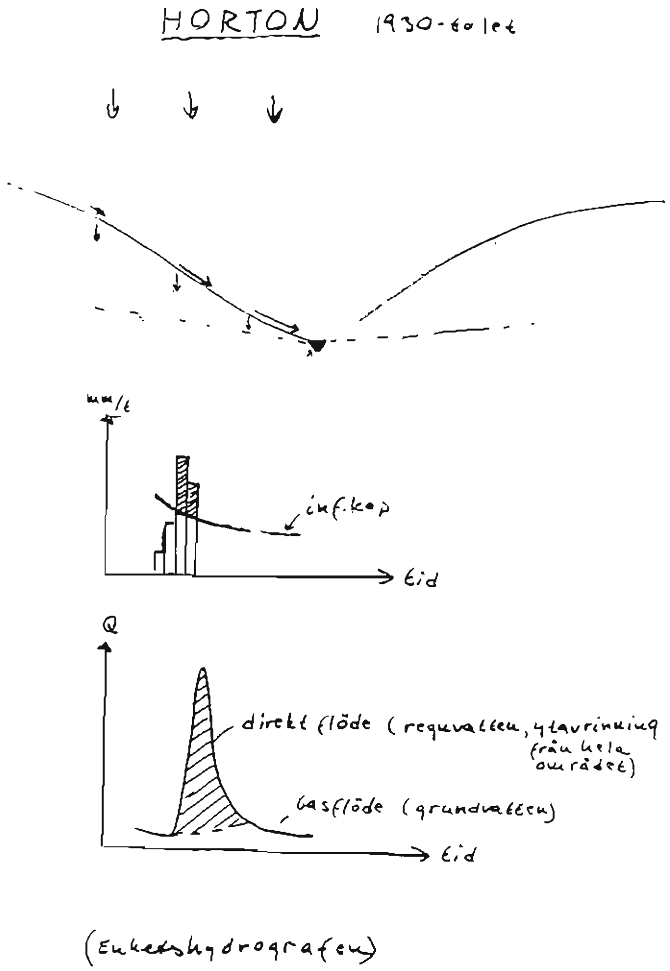
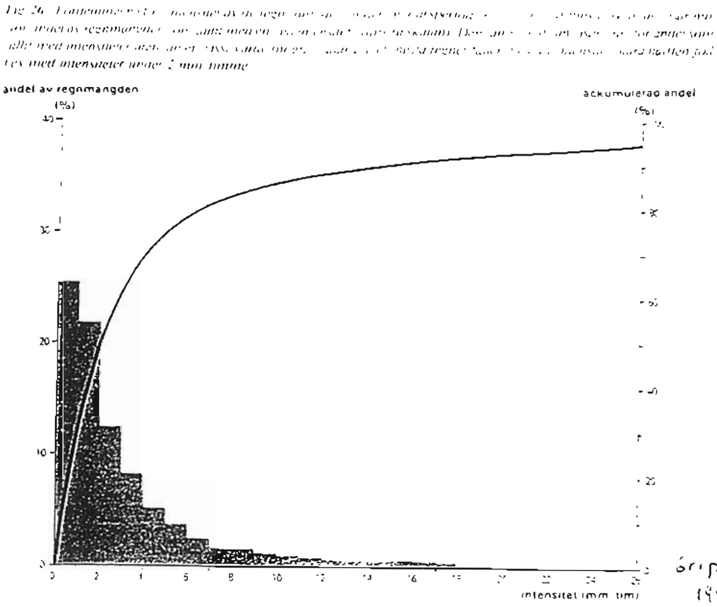


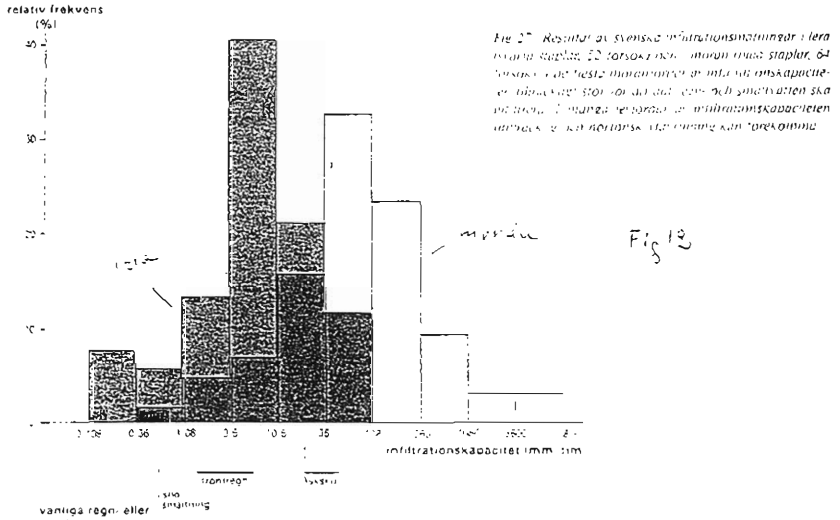
Fig 10. According to Horton runoff events in streams occur when the intensity of rainfall exceeds the infiltration capacity of the soil. Infiltration excess generates overland flow over the whole catchment. The hydrograph is supposed to be a sum of a slowly changing base flow, fed by groundwater and little influenced by rainfall episodes, and a rapidly changing direct runoff from overland flow. This concept has had a very large impact on the view of streamflow generation and applied hydrology since the 1930ies. The "Unit hydrograph" is a simple model based on Horton's view. Since it is based on linear reservoirs it can describe stream runoff, but when it comes to water chemistry it fails completely.

Objections against Horton in humid areas: Overland flow is seldom seen, the infiltration capacity is normally larger than the intensity of rainfall or snowmelt, maximum streamflow is more related to the total rainfall volume than to the rainfall intensity.



Gripe Rodhe 1991

Fig. 11. Rainfall intensity and infiltration capacity (Swedish sites)



Gripe Rodhe, 1991

Fig. 12. Rainfall intensity and infiltration capacity (Swedish sites)

USA 1960-ies

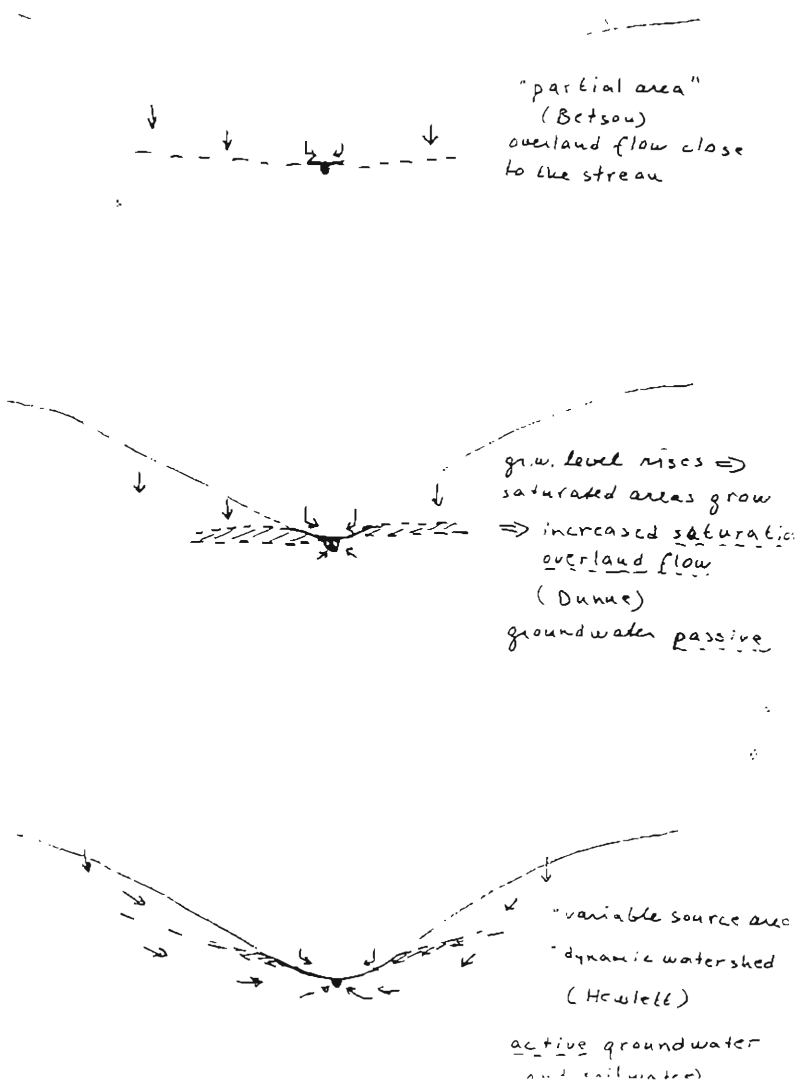


Fig. 13 Development of a new view on streamflow generation

Fig. 14. (black/whiteboard) Hypothesis for streamflow generation

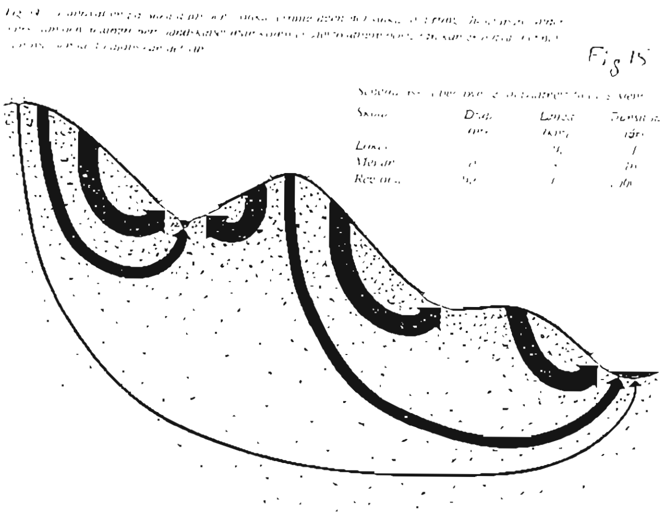


Fig 15

Grip & Rodhe, 1991

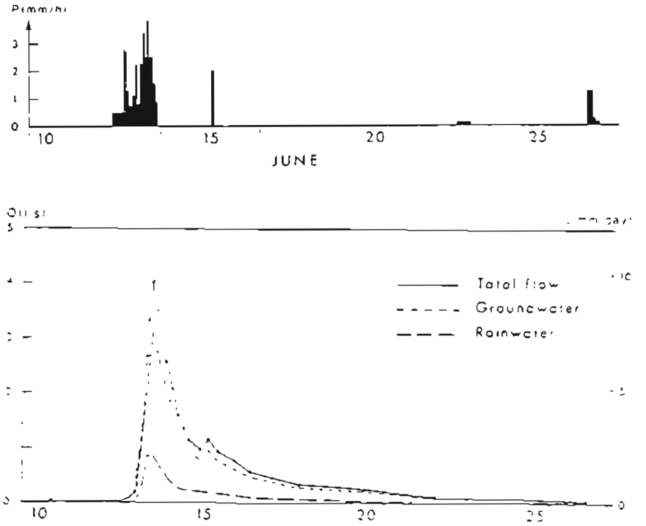
Fig. 15. Flow systems for groundwater

Recharge area: Area in which there is a flow towards the groundwater zone
" downward total potential gradient in the groundwater
Discharge area: " flow from the groundwater zone
" upward total potential gradient in the groundwater zone

The discharge area may be saturated or unsaturated. Unsaturated discharge areas may occur if there is a very large increase in the hydraulic conductivity towards the ground surface or if the transpiration from the soil equals (is larger than) the groundwater flow from below (Fig 17)
Hortonian overland flow: Overland flow on unsaturated areas (unsaturated except for the saturated layer close to the ground surface)
Saturation overland flow: Overland flow on saturated areas.

Isotope studies (see lecture on .19 october) have shown a large and rapid response of the flow of pre-event water in the stream (water that existed in the catchment before the rainfall) to rainfall input over the catchment. If there are no lakes or other surface water reservoirs, the pre-event water is by definition groundwater, since water can leave the ground only if the pressure is ≥ that of the atmosphere, i.e., as groundwater. The discharged groundwater may, on the other hand, well be soil water that was transformed into groundwater during the event, by infiltration and percoaltion from above or by a rising water table from below.

Gårdsjön, F1 1982



Rodhe, 1987

Fig. 16. Separated hydrograph

The response of the groundwater outflow to rainfall infiltration can be divided into two steps:
a) the response of the groundwater table
b) the response of the groundwater discharge to a raised groundwater table.

Two mechanisms are discussed here:

1. There is a large increase in the hydraulic conductivity towards the ground surface, which is typical for till soils. The slope of the groundwater may increase only by a few percent as the water table rises, but a moderate rise of the groundwater table generates a very large increase in the transmissivity and thus in the flow, when superficial layers with high hydraulic conductivity start to contribute to the lateral groundwater flow. Over the year there will be a large flow volume in the superficial layers, although the flow occurs only during short time periods. These "spate specific" flow-paths may have great impact on the hydrochemistry.
Fig K=f(z)

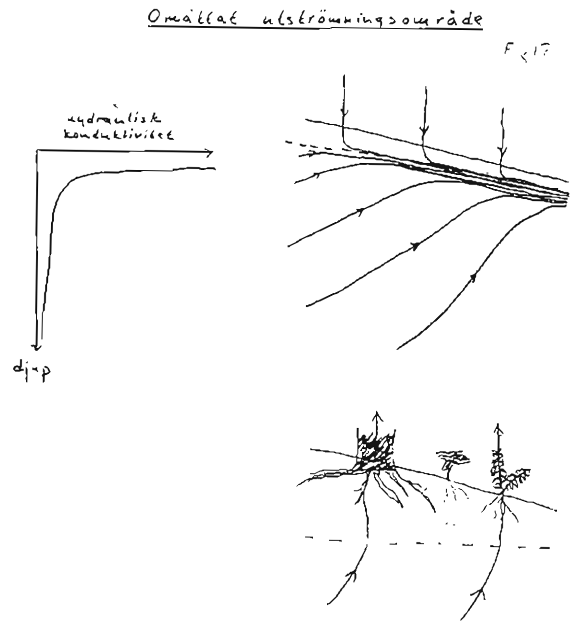


Fig. 17.

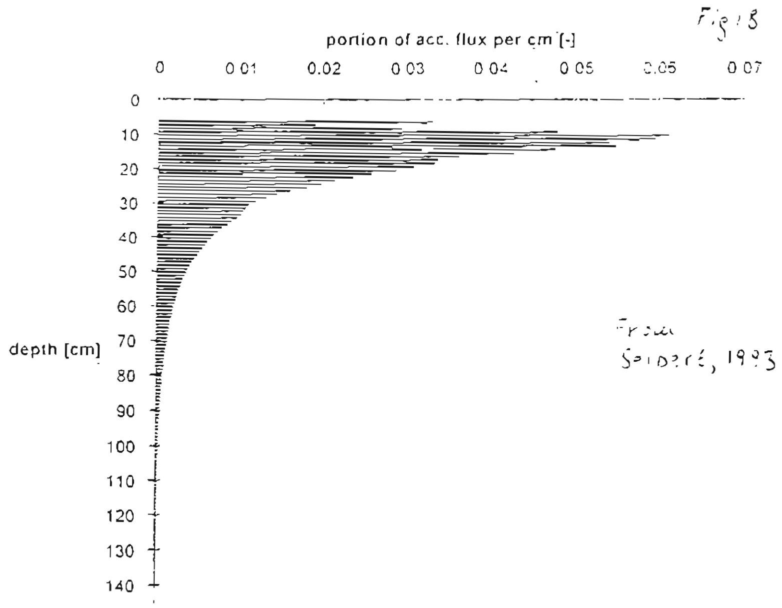


Fig 18

From Spongré, 1993

Fig. 18. Fraction of annual flow versus depth.

2. The capillary fringe model. Close to the stream, where the groundwater table is close to the ground surface, the storage coefficient may be very small since he capillary fringe, in which almost all pores are water-filled, reaches the ground surface. A small water input generates a large and rapid increase of the groundwater level and thus on the potential gradient and flow towards the stream. This could explain the rapid response in soils where the hydraulic conductivity does not vary vertically, but the effect is probably of little importance in Nordic till soils. (Fig 19)

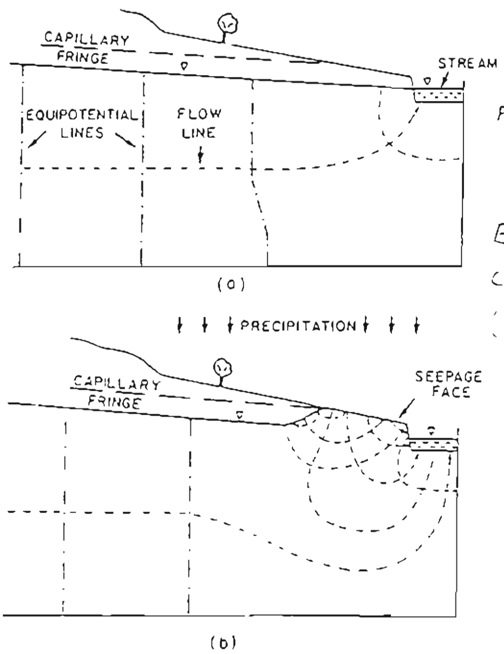


Fig 19

Effect of the capillary fringe (Gillham, 1984)

Fig. 19

The role of topography

The wetness at a certain location within a catchment (expressed as soil moisture content and groundwater level) depends on the water budget of the location, i.e., on the catchment area of the location (determines the rate of inflow) and the slope and transmissvity at the location (determines the rate of outflow). Saturated areas occur when the capacity of the soil to transport away the water downhill is exceeded by the inflow from the upper part of the hillslope. The capacity is given by

$$Q_{max} = -T \cdot b \cdot \frac{dh}{dx} \qquad (9)$$

where dh/dx is the slope of the ground surface (=maximum slope of the groundwater table). Consider a two dimensional hillslope. The mean flow at a certain location is proportional to the distance to the water divide, giving

$$R \cdot x_d \cdot b = -T \cdot b \cdot \frac{dh}{dx} \qquad (10)$$

wher R is the groundwater recharge, x_d is the distance from the water divide to the discharge area and b is the width under consideration. In an upward concave hillslope, the slope of the ground surface decreases downhill, giving large discharge area (small x_d) and in a convex hillslope, where the slope increases downhill, the discharge area will be small or absent. Human activities may increase (clear-cutting) or decrease (forest plantation) R, giving increase and decrease in the extent of the discharge area respectively. Looking in 3 dimensions, the above considerations are complemented by differences in the inflow to the point caused by different catchment areas. Hollows are wet and noses are dry.

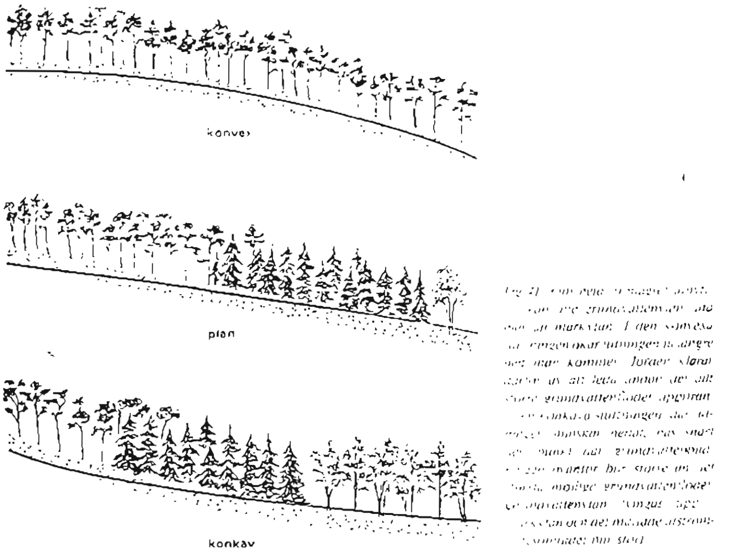


Fig. 20

Grip & Rodhe, 1991

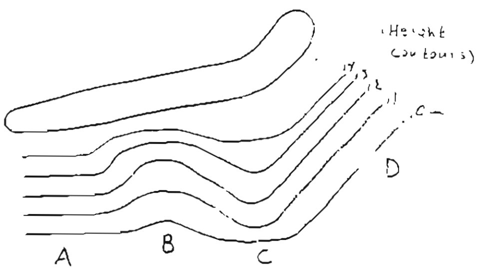


Fig. 21

Fig. 20, m21 Extension of discharge area 2-dim and 3-dim respectively

Components of the water budget and river regimes in the Nordic countries

Otnes & Raestad,
Hydrologi i Praksis 1973

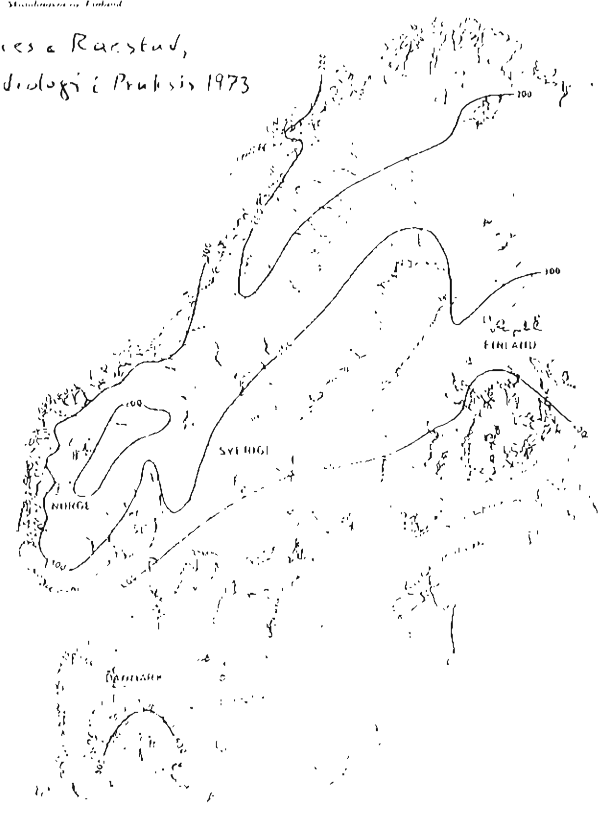


Fig. 22. Annual P, E, R in the Nordic countries

Otnes et al

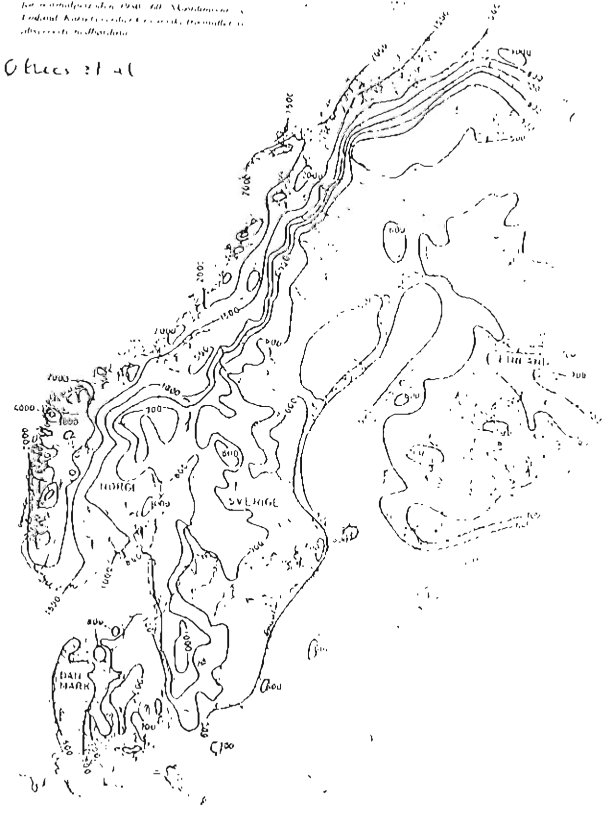


Fig. 23. Annual P, E, R in the Nordic countries

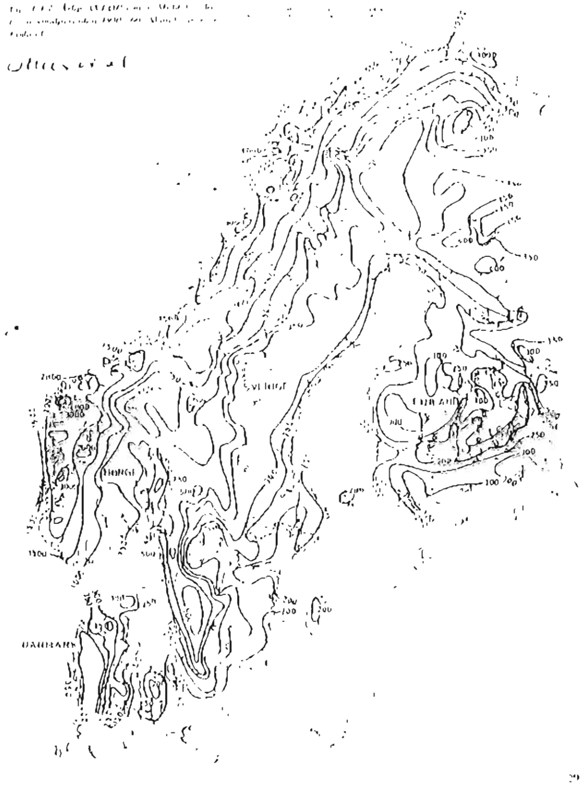
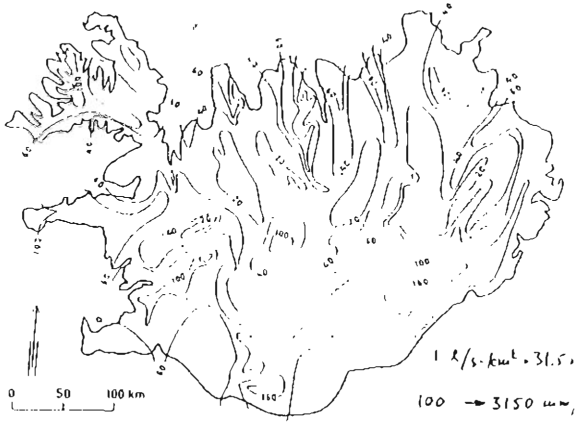


Fig. 24. Annual P, E, R in the Nordic countries

Melin, 1970

Avrinning l/s·km²



1 l/s·km² ≈ 31.5

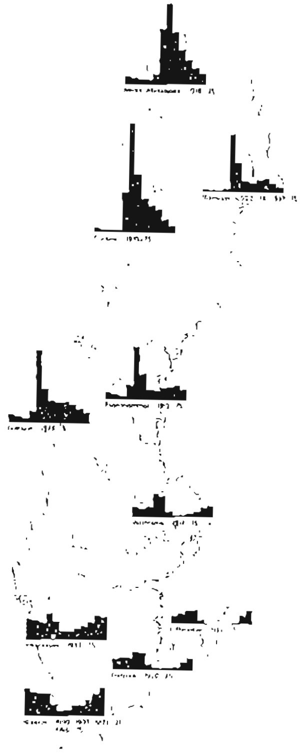
100 → 3150 mm/

Fig. 57 Medelavrinningen (R) på Island (l/s km²) OBS! Avd ≈ 200 mm/ ∴ P ≈ R + 200 mm/.

Island har många men relativt små flodområden. De största områdena avvattnas av följande floder.

		Nederbördsområde	Medelvattenföring	
Jokulsá á Fjollum	i norr	7950 km²	190 m³	24.0 l/s km²
Thjórsá	i söder	7530 "	394 "	52.0 "
Hvítá	"	6100 "	390 "	64.0 "

Fig. 25. Annual P, E, R in the Nordic countries



Bergström, 1993
Sveriges Hydrologi

Figur 10.2 Avrinningens variation under året vid några vattenföringsstationer (Källa SMHI [illegible])

Fig. 26 River regimes in Sweden

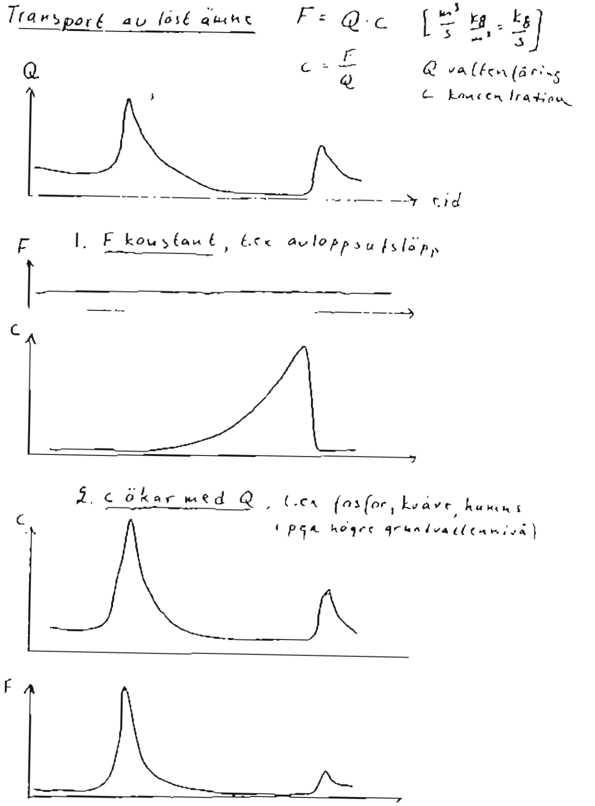


Fig. 27 Relationship between solute concentration and discharge

Litterature: Grip, H. and Rodhe A.(1991): Vattnets väg från regn till bäck. Hallgren och Fallgren, Uppsala, Sweden. (In Swedish, English title: Water flowpaths from rainfall to streamflow.)

An introduction to groundwater recharge

David N. Lerner
Department of Civil Engineering, University of Bradford,
Bradford, West Yorkshire. BD7 1DP.

WHAT IS RECHARGE?

Recharge is the water which enters the (saturated) groundwater reservoir. It can flow downwards, upwards or sideways to get there, and may only indirectly originate from precipitation.

There are alternative definitions, and terms with similar related meanings. For the purpose of this text, Figure 1 illustrates some of these related terms:

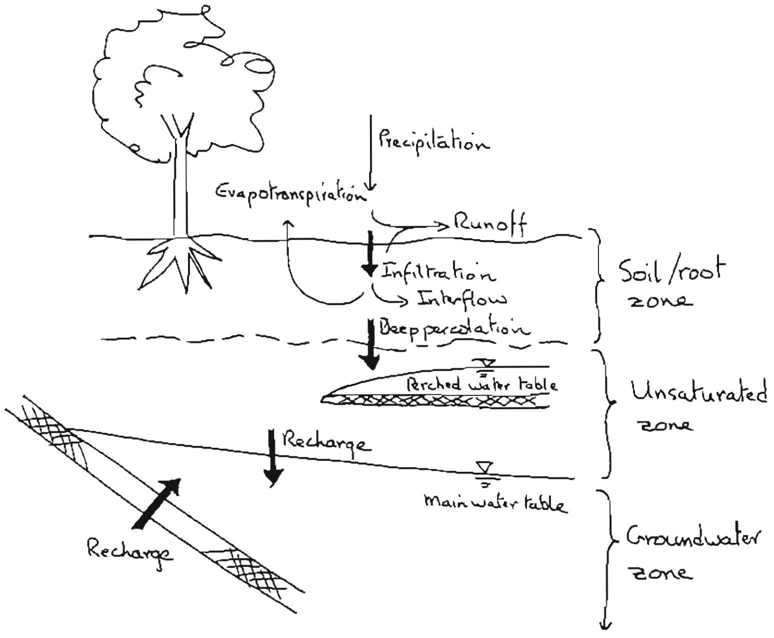


Figure 1. Infiltration, deep percolation and recharge.

- Infiltration is the water that enters the ground. It may come from precipitation or, for example, from leakage through a river bed. Infiltration is usually subject to soil moisture processes, such as evapotranspiration, and so only a proportion is available for downward migration.

- Deep percolation is the water that escapes downwards from the soil zone and gets below the influence of roots. Some may be lost on its way to the groundwater reservoir, due to perched water tables or capillary rise. There may be substantial delays in transit through the unsaturated zone.

A distinction can sometimes be usefully made between actual recharge (as defined above) and potential recharge (Rushton, 1988). The amount of recharge might alter if conditions in the aquifer change even though nothing had changed at the surface. A hypothetical example is illustrated in Figure 2, for the cases of river infiltration with a changing water table elevation. Assuming a simple, homogeneous, one dimensional system, the maximum infiltration rate equals K, the bed hydraulic conductivity, for a vertical hydraulic gradient of unity for gravitational flow. The water available for recharge is the river flow per unit bed area (Q/A).

Potential recharge (PR) is the maximum water available to become recharge, and in this example is the minimum of the two quantities defined above, i.e.

$$PR = \min(K, Q/A).$$

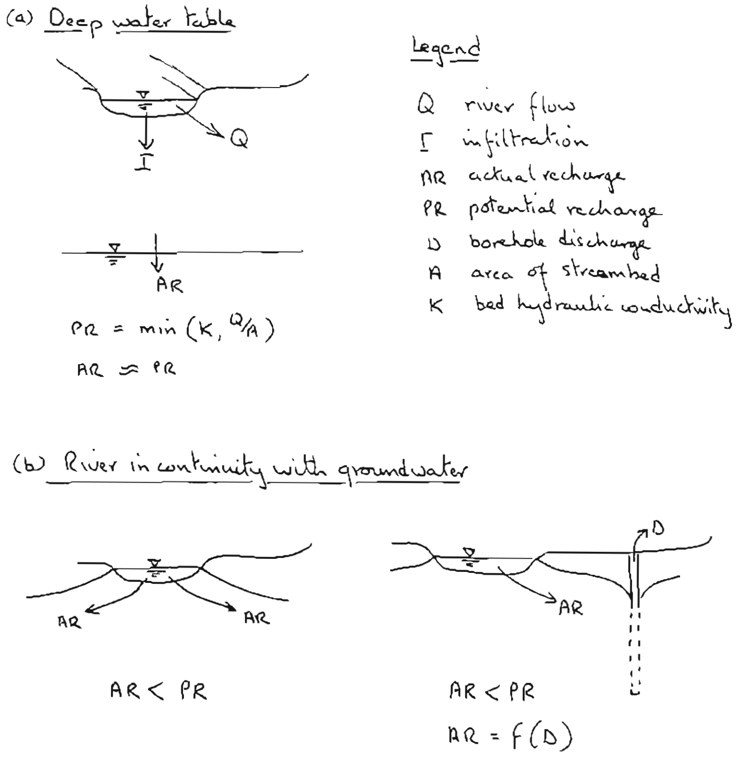


Figure 2. Differences between actual and potential recharge.

While the water table is deep and the river is hydraulically separate from groundwater, actual recharge (AR) will approximately equal PR. If the water is high, perhaps rising in response to recharge, and the river and groundwater are in continuity, then AR < PR. Other circumstances, such as a pumping well near the river, can change the relationship between AR and PR. A field example is given in Figure 27.

Further distinctions are conceptually useful for recharge which results from precipitation, depending on the route taken by the water. The routes are schematically illustrated in Figure 3, and define three categories:

- Direct recharge, which is a diffuse process, occurring beneath the point of impact of the precipitation.
- Localised recharge, resulting from the horizontal movement and concentration of water into joints, rivulets and depressions. These focused points are too frequent to be mappable and measurable, but clearly involve a different mechanism from direct recharge.
- Indirect recharge, also involves the concentration of water, usually into rivers; whether ephemeral or perennial. The intended distinction from localised recharge is that the water courses are sufficiently large to be mapped, counted and possibly gauged.

These distinctions are made on the bases of processes and of practicality, and cannot be rigidly adhered to. For example, preferential pathways for recharge such as fissures and root holes have been documented for situations which might have been described as leading to direct recharge (e.g. Sharma and Hughes, 1985). An example of transmission through fissures is given in Figure 4. The differences between localised and indirect recharge will depend on data availability and the methods of quantification being used; rapid consultancy may treat larger concentrations as localised that detailed research work.

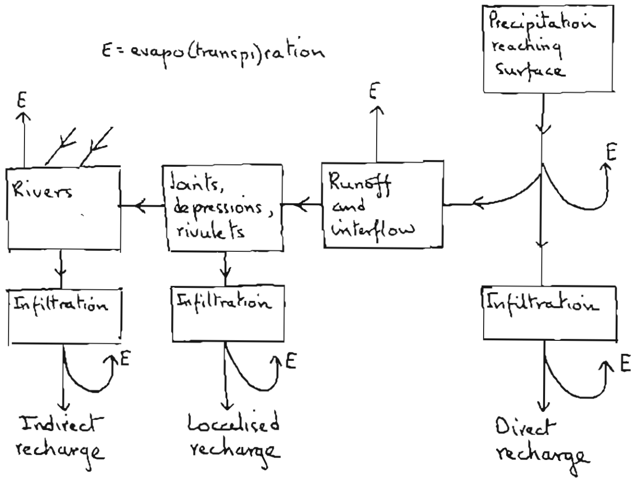


Figure 3. Subdivisions of precipitation derived recharge (adapted from Lloyd, 1986).

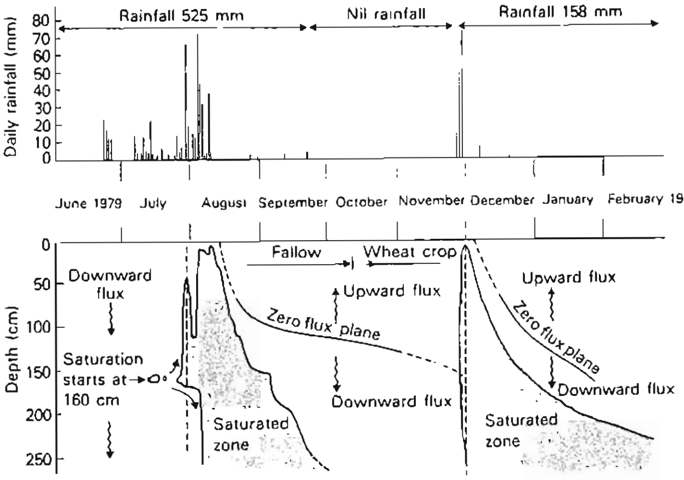


Figure 4. Annual cycle of water movement in a swelling clay soil in India, showing the development of saturated conditions starting at 1.6 m depth due to infiltration down shrinkage cracks. (From an original diagram by Hodnett and Bell, 1986).

RECHARGE IN THE HYDROLOGICAL CYCLE

Figure 5 shows the classical, simplified, hydrological cycle, as understood by my six year old. The purpose of this section is to point out that groundwater-surface water interactions are of course more complicated than shown. In particular, there are other sources of recharge than precipitation (Figure 6).

The section above indicated that natural recharge could occur in the subsurface, either across the boundaries of a study area when these do not coincide with groundwater system boundaries, or from deeper aquifers (Figure 1). Rivers and other surface waters are frequently a source of recharge (Figure 2). Artificial recharge is widely practised and is a research and engineering industry in its own right, and will not be discussed here.

In many environments, the hydrological cycle has been modified by human activity, and there changes are likely to influence recharge. For example, in lowland Britain almost all rivers of any significance flow in maintained channels with regular weirs to control levels and velocities. All aquifers are exploited with consequent lowering of water tables and opportunities for increased recharge (Figure 6).

Two other human interventions deserve special mention, irrigation and urbanisation. Both modify the hydrological cycle by importing water to an area, introducing new pathways for water (canals, pipelines), and altering recharge.

Figure 6 can be summarised to include four types of water that can become recharge: precipitation, surface water, groundwater and urban water (potable, waste). There are six situations to consider in which these waters may become recharge: natural precipitation,

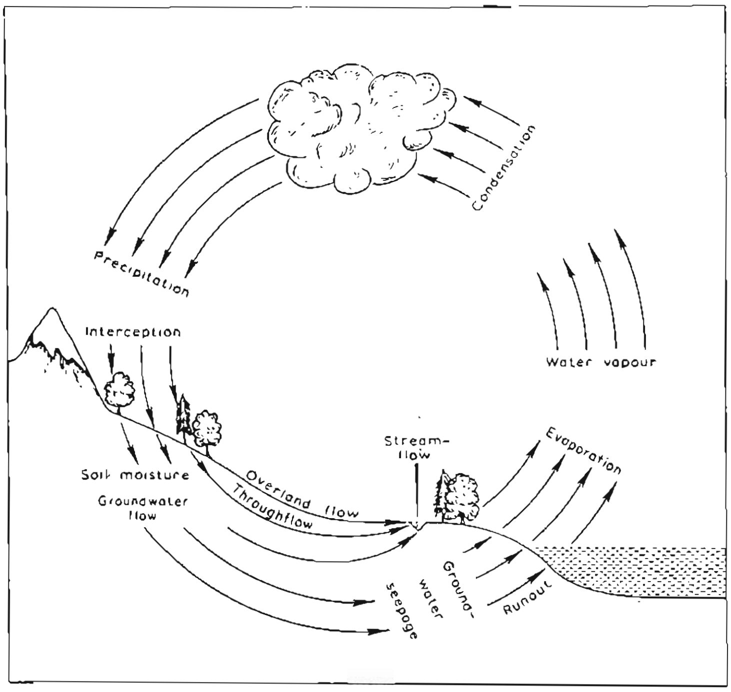


Figure 5. Schematic diagram of the hydrological cycle (Ward and Robinson, 1990).

Type of water	Common human interventions	Routes to groundwater
Precipitation	Agriculture Impervious surfaces (roads, etc.)	Direct infiltration Localised infiltration
Surface water	Flow regulation and abstraction Abstractions Flood control Channelisation Groundwater abstraction	Bed infiltration Overbank flooding Induced recharge (subsurface flow)
Groundwater	Abstraction	Subsurface flow (interaquifer)
Potable water Surface water Waste water	Artificial recharge: - lagoons - river spreading - injection boreholes	Bed infiltration Subsurface flow
Surface water Groundwater	Irrigation: - spray - flooding - canals	Direct infiltration Ponded infiltration Bed infiltration
Potable water Waste water	Urbanisation: - leaking mains and sewers - septic tanks and soakaways	Subsurface infiltration

Figure 6. Recharge in the hydrological cycle.

natural surface waters, interaquifer flows, artificial recharge, irrigation, urbanisation. The routes to groundwater may start at the surface, in the unsaturated zone, or in the groundwater zone. The routes may be diffuse (areal), linear or multi-linear, point or multipoint. Different geochemical processes will occur on the different routes; recharge estimation techniques will vary across the routes and sources of water.

HYDROGEOLOGICAL ENVIRONMENTS

It may well be possible to take a mechanistic approach to estimating recharge, that is going into the field, applying standard methods, and calculating results. Accuracy is more likely to be achieved when one has empathy with the environment, has a conceptual model of the processes operating, and has a prior mental estimate of the magnitude of the results. This understanding will lead to better design of measurements and enable one to spot the unusual results. The latter may be errors, requiring remeasurement, or may alert one that actual processes operation in the field are different from those expected.

This need to know the answers in advance can cause problems for the beginner! However, comparisons with similar hydrogeological environments can be very helpful, where a hydrogeological environment is defined by the combination of geology, climate, topography and human influence which gives an area its hydrogeological characteristics. Descriptions of, and case studies from, several environments are given by Lerner *et al.* (1990). They are geologically classified, and comprise alluvial provinces (riverbed and mountain front), sand and sandstone, limestone and dolostone, chalk, volcanic rock, and crystalline plutonic provision. The current text does not attempt to emulate that review, and will just give some examples to illustrate basic aspects.

Permo-Triassic sandstone of the UK

The second most important aquifer in the UK is the Permo-Triassic sandstone (Figure 7). It contains regional groundwater systems, and generally has a rolling topography. Soils are well developed and bedrock outcrops are rare. The climate is temperate, with year round rain, little snow, precipitation comfortably exceeding potential evapotranspiration, and few extreme events. There are very few areas of natural vegetation, expect for some managed forests. Agriculture is predominately arable in the drier east, with increasing livestock towards the west.

Rivers can be expected to be regional discharge points for groundwater, are often 10 km apart, and will only be perched above the water table if large boreholes are located nearby. Precipitation will lead to diffuse, direct recharge, and there is little to cause localised recharge (except as discussed below). Runoff is likely to occur by the variable source area concepts rather than Hortonian overland flow, and so flood conditions may cause local groundwater mounds along rivers.

Irrigation is confined to drier areas, sandy soils, and more valuable crops, and is always sprayed. It may lead to local increases in recharge because of the sandy soils.

Figure 7. Outcrop of the Permo Triassic sandstone aquifers in England and Wales.

The major complication to this picture is widespread presence of glacial drift cover. There is patchy superficial cover over most of England. It is very variable and unpredictable in character, from gravels to clays, and often with multiple units in any section (Figure 8). The effects of drift on recharge can be large and varied, including:

- different soil types,
- low permeability cover for the aquifer,
- confined aquifer conditions,
- perched water tables,
- increased surface runoff and interflow,

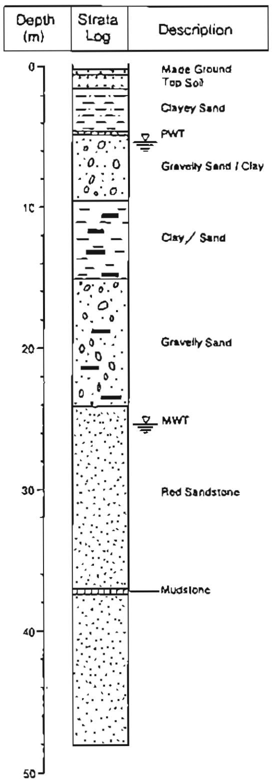


Figure 8. Typical mixed glacial sequence overlying the Triassic sandstone aquifer in Birmingham, UK.

Scandinavian conditions

My simple view of southern Scandinavian groundwater is that of contrast to the UK sandstone above (please correct me if wrong!). Aquifers are usually shallow glacial/alluvial sediments, relatively thin, of local rather than regional extent. Bedrock is crystalline, small outcrops are common, and experience localised recharge into joints. However bedrock is usually only of local importance as an aquifer, and the outcrops can cause localised recharge around the edges of the superficial aquifers. Large areas are naturally vegetated, but summer irrigation is common on arable lands. Snow and snow melt are important.

Mountain-front alluvium

A third and final example provides yet more contrasts of hydrogeological environment. Lima (Peru) is built on the alluvial fans of the Rimac and Chillon rivers where they emerge from the Andes into the Pacific Ocean (Lerner *et al*, 1982). There is no local rainfall, so all recharge ultimately comes from the rivers. At the apexes of the fans the rivers are well above the water table, so recharge the aquifer. Extensive canal systems distribute surface water across the aquifer surface for flood irrigation and subtract from river flows. High losses from canals and fields cause recharge.

Surface and groundwater are used for public water supply. The distribution system is in poor condition, losses are high, and plenty of urban recharge occurs (Figure 9). Sewers, where they exist, are likely to leak. Waste water, deliberately recharged in some areas, is reused for irrigation of vegetables in others. Overall, a simple system of only river recharge has been converted by human influence to a complex pattern of recharge sources and routes.

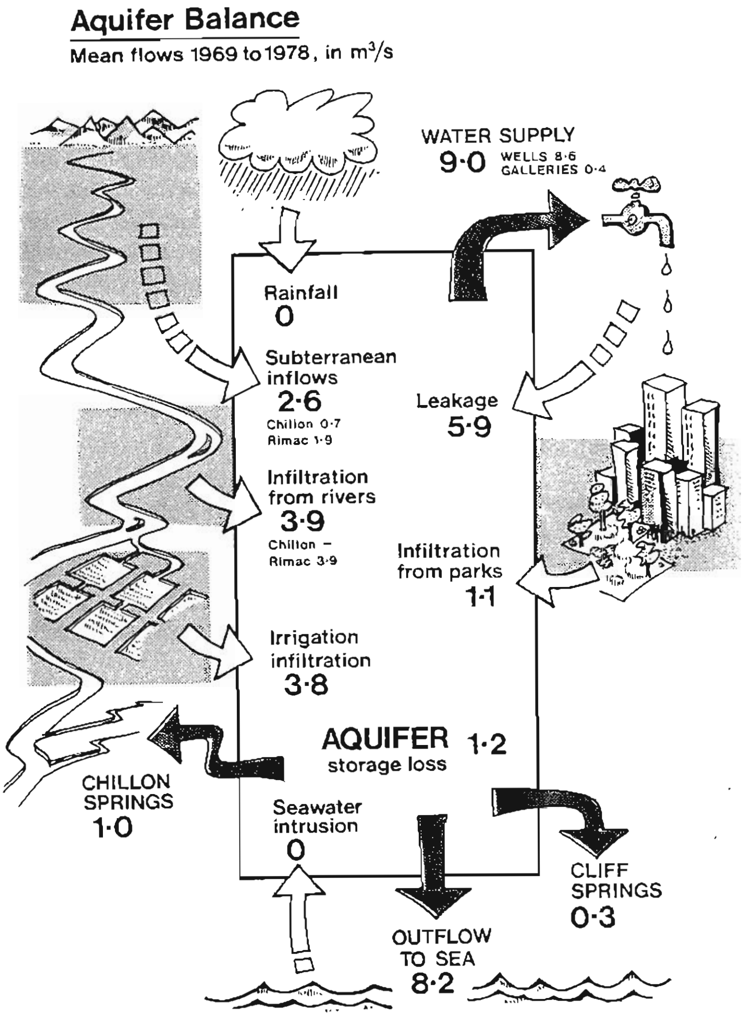


Figure 9. Overall water balance of the Lima (Peru) alluvial aquifer (Lerner et al., 1982).

PRECIPITATION RECHARGE

Introduction

A simplified view of the precipitation-recharge process is shown in Figure 1. In essence, some of the precipitation returns to the atmosphere by various evaporation processes, some runs off laterally, and the remainder becomes direct recharge. This chapter is concerned with estimating this direct recharge, that is recharge below the point of impact of the precipitation. Including moisture storage above the water table:

$$\text{recharge} = \text{precipitation} - \text{runoff} - \begin{array}{c}\text{actual}\\ \text{evapo-}\\ \text{transpiration}\end{array} \pm \begin{array}{c}\text{storage}\\ \text{change}\end{array} \qquad 1$$

An immediate conceptual difficulty is "How far can water move sideways and then infiltrate before counting as indirect recharge?". Any such movement implies variability in spatial properties which can invalidate many of the estimation methods for direct recharge. On the other hand, small and numerous movements cannot be counted or treated individually when building a regional groundwater model. Examples include:

(i) weathered, bare, hardrock or limestone terrain where recharge is into distinct fissures,

(ii) surface depressions, with sizes and spacings from centimetres upwards, where local runoff gathers and infiltrates,

(iii) bare rock outcrops in permeable terrain (e.g. sand or alluvium), where runoff infiltrates at the edge of the outcrop,

(iv) the many minor, ephemeral drainage channels with permeable beds which may only contribute to main channel flows on rare occasions; at other times all flow infiltrates into the bed.

These examples reinforce the need for a pragmatic division into three, (i) direct recharge, (ii) indirect recharge, and (iii) an intermediate category *localised recharge*. The last is the least well researched of the three and is discussed in the end section of this chapter.

Methods for estimating direct recharge can be classified as follows:

(i) Direct measurement over areas up to 100 m^2.

(ii) Totally empirical methods, which usually simplify eqn. 1 to:

$$\text{recharge} = f\,(\text{precipitation}) \qquad 2$$

where the function may be linear or non-linear and may involve easily measured variables like altitude or catchment area.

(iii) Water budget methods, (a) at a point or (b) at a catchment scale. The former are usually soil moisture budgeting methods.

(iv) Darcian approaches, that is making use of the equation of flow in the zone above the water table. These may be based on field measurements of moisture and head, or on numerical models.

(v) Environmental or applied tracers which track the movement of parcels of water in the unsaturated zone.

The differences between potential and actual recharge have already been discussed. Actual recharge is needed when simulating historical conditions of a groundwater system. However, when modelling possible future conditions, for example when new boreholes have lowered the water table, it may be more appropriate to use potential recharge. Many methods of estimating recharge estimate potential, rather than actual.

Lysimeters - direct measurement

The only practicable method of measuring recharge flux is with a lysimeter. This is a block of soil instrumented so that flows through it can be measured. The block is isolated from the surrounding soil but is representative because it has the same vegetation and climatic exposure. In order to minimise edge effects and average out local variations in soil and vegetation lysimeters need to be large (up to 10 metres in each of the three dimensions), and so are expensive to construct. Examples of their construction and use for recharge measurement are given by Kitching *et al.* (1977), Kitching *et al.* (1980) and Kitching and Shearer (1982).

A lysimeter design for estimating recharge should fulfil the following requirements:

(i) Contain undisturbed soil. Vertical flows through repacked soil will not be representative.

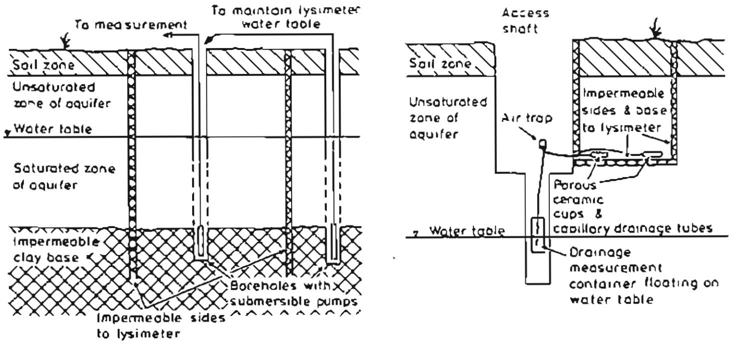
(ii) Be large enough to minimize edge effects and average small scale heterogeneities. A minimum plan area would be 1 m^2, the ideal would be 100 m^2.

(iii) Be large and deep enough to enclose complete root systems. This sometimes makes naturally vegetated lysimeters impractical, especially in arid and semi-arid areas where roots can be 50 m deep.

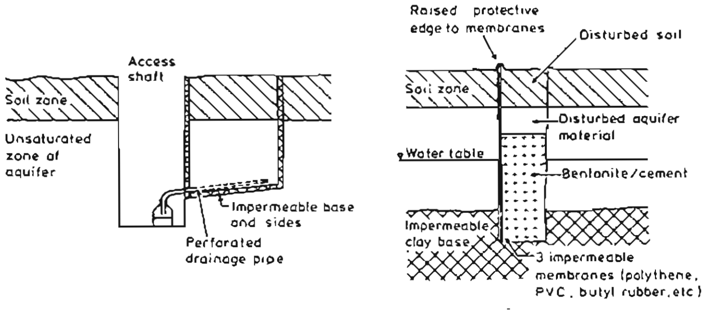
(iv) Be surrounded by similar vegetation to avoid oasis effects.

(v) Have the same hydraulic condition at the base as found at the same depth in the surrounding soil (Figure 10a-c).

(vi) Be watertight, except for the drainage to be measured.



(a) Lysimeter containing a water table (b) Suction drained lysimeter



(c) Gravity drained lysimeter (d) Trench construction for lysimeter sides

Figure 10. Design and construction of lysimeters. (a) requires pumping to equalise conditions inside and out. (c) is suitable for coarse materials where capillary use is insignificant.

Lysimeters, as with most other methods of estimating direct recharge, are better suited to humid than semi-arid climates. Careful construction, regular maintenance and frequent observation are needed to ensure that leaks, bypassing flows, faulty drainage or evaporation do not affect the amount of water collected.

Lysimeters are expensive to construct and only give point measurements of recharge. Kitching *et al.* (1977), set up two lysimeters within 50 metres of each other in a humid zone. Recharge was 159 mm/y in one and 114 mm/y in the other. The fact that the latter recharge value is only 72% of the former was not explained: does it represent true variability in recharge or errors in measurement? In either case it shows that lysimeter methods are as variable as other point methods.

The act of constructing a lysimeter will disturb the soil and moisture and time is needed for flows to settle back to a natural condition. At the extreme, this could be as long as the time for recharge to flush the lysimeter through; for a 2 m deep unit with a water content of 20% and recharge of 100 mm/y, this would be 4 years. For example, a lysimeter in Cyprus only recorded 5 mm of recharge in its first year of operation, although chloride and tritium profiles suggested average recharge was 50 mm/y (Kitching *et al.*, 1980; Edmunds and Walton, 1980).

Empirical methods

Many attempts have been made to find simple relationships between precipitation and recharge. Once derived by careful study, these are commonly used as 'black boxes', making recharge estimates without further consideration of hydrogeology or whether the results are feasible. The simplest empirical formula takes recharge as a proportion of precipitation:

$$r = f\,p \quad \text{(notation below eqn. 6)} \qquad 3$$

where f will probably vary with terrain and climate. The second level of formula includes a threshold. For example Mandel and Shiftan (1981) gives a formula for recharge in Mediterranean climates of:

$$r = 0.9\,(p - 360) \qquad 450 < p < 650 \text{ mm/y} \qquad 4$$

More complex formulae often do not preserve dimensionality, for example the Cheeturvedi formula (Sinha and Sharma, 1988) for recharge in India:

$$r = 50.8\,(p/25.4 - 15)^{0.4} \qquad p > 380 \text{ mm/y} \qquad 5$$

or Turc's (1954) formula which includes mean annual temperature:

$$r = p\,[1 - (0.9 + p^2/L^2)^{-0.5}] \qquad 6a$$

$$L = 300 + 25T + 0.05\,T^2 \qquad 6b$$

where r : annual average recharge rate (mm/y) f : empirical constant
p : annual precipitation (mm/y) T : mean annual temperature (^{o}C)

The origin of many formulae are lost in the darkness of history. In general they will have been obtained for a particular basin by correlation of precipitation with estimates of recharge obtained by other methods (water table rise, basin discharge, etc.).

There are two issues that must be resolved before an empirical formula can be used; 'How reliable is its derivation?' and 'Can it be transposed to either another catchment or another period in time?'. For example, how accurate were the recharge estimates used to derive the formula; were they checked, say by calibration of a groundwater model? When transposing, are conditions in the new catchment (or time period) the same as those in the original - depth to water table, unsaturated zone processes, land use, topographical characteristics, climate and type of rainfall.

An illustration of the low accuracy of simple precipitation-recharge relations can be found in a study of 63 desert catchments in Nevada, USA, carried out by Watson *et al.* (1976). They correlated recharge (estimated from measurements of groundwater discharge) with precipitation, obtaining results as follows:

Precipitation zone (mm)	Percentage recharge	95% confidence interval (%)
< 510	24	±15
380 - 510	19	±16
300 - 380	-0.1	± 6
200 - 300	4	± 2
< 200	0	± 1

The wide confidence intervals make the coefficients unusable for prediction, despite being derived from a large, carefully assembled database.

Given the difficulties outlined above, the occasions when empirical formulae will be most useful are:

(i) for reconnaissance studies, when high margins of error are acceptable;

(ii) for under exploited catchments, or those where groundwater conditions have little effect on recharge;

(iii) for extrapolation backwards in time within the catchment where the formula was derived;

(iv) when the resulting estimates can be checked and amended if necessary, for example when a groundwater model is to be constructed.

Many empirical formulae have been derived and used, especially in project studies by consultants and others, although few have found their way into the scientific literature. It is not possible to give enough background in this manual on each formula to allow its use, and so none can be recommended.

The use of empirical formulae has similarities to the use of *representative basins*. These well instrumented catchments have been set up in many parts of the world to provide accurate estimates of the components of the water balance, including recharge. The same difficulties of transposition will arise with representative basins as with empirical formulae, but at least the basic data will be of high quality.

Soil moisture budgeting methods

Soil moisture budgeting methods are those that measure or estimate the items on the right hand side of eqn. 1. At its simplest this leads to a conceptual model like that in Figure 11. Water is held in a soil moisture store; precipitation adds to the store, evapotranspiration depletes it. When full, excess precipitation is routed to groundwater as recharge. The most difficult item to measure is actual evapotranspiration, and in general a conceptual quantity called 'potential evapotranspiration' is defined. A budgeting procedure involving soil moisture is used to convert potential to actual evapotranspiration. There are many variations and

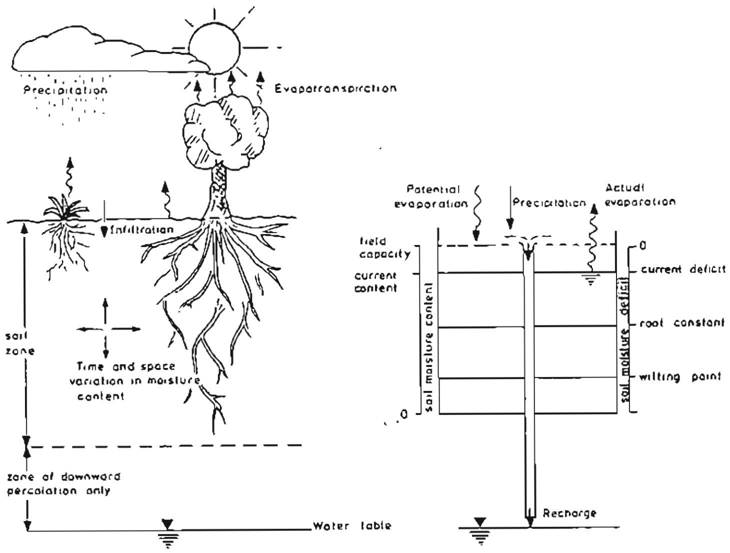


Figure 11. Soil moisture processes and a conceptual soil moisture budgeting procedure.

refinements on Figure 11 some of which are discussed below, but at this stage it is important to note these points:

- such models are only simple conceptual models of the precipitation-recharge process and may not be correct for your situation;
- the essence of the model is the relation between potential and actual evapotranspiration;
- estimates are for a uniform zone (see below).

Soil moisture budgeting models were developed for humid climates and have less validity in arid and semi-arid zones. They work best for seasonal patterns of recharge, well developed soils which do not dry completely, when potential and actual evapotranspiration are of similar sizes, and with precipitation that is widespread and relatively uniform. These conditions do not apply in arid and semi-arid zones where these models normally underestimate recharge, often giving zero values.

The following rule of thumb may help in deciding whether soil moisture budgeting models will be applicable to a given terrain:

Whole year: $p + i > 500$ 7a

Wet season: $et_p < 1.5\,(p + i)$ 7b

Dry season: $et_p < 3\,(p + i)$ 7c

et_p, et_a and $(p + i)$ similar in area where et_p data derived and area where model is applied 7d

where et_p : potential evapotranspiration (mm/y)
et_a : actual evapotranspiration (mm/y)
p : average precipitation (mm/y)
i : irrigation water applied (mm/y)

It may be possible to apply the methods for a wet season only on the basis that little recharge occurs in the dry season.

Figure 11 does not describe how soil moisture behaves; there can be vertical water flows (either up or down) when a soil moisture deficit exists. Nor does it describe the recharge process, which may be dominated by fissures, root channels or topographic depressions. Many models used in real situations need empirical adjustments to make them match field conditions.

Good data on actual evapotranspiration is equally important as good precipitation data. Unfortunately actual evapotranspiration is rarely measured except in research projects and so must be estimated from standard meteorological measurements. As mentioned above, this is usually done through a conceptual quantity called *potential evapotranspiration* or *reference crop evapotranspiration*. This is intended to be a measure of the energy available for evaporating and transpiring water.

The Penman-Grindley model (Penman 1950; Grindley 1967, 1969) is the simplest and most widely used soil moisture budgeting model. It was originally developed to estimate soil moisture deficit and actual evaporation, recharge estimates being a by-product. The soil moisture accounting is common to many models:

$$\begin{aligned} psmd_{i+1} &= smd_i + ae_i - p_i \\ r_i &= -\ psmd_{i+1} \qquad \text{when } psmd_{i+1} < 0 \\ smd_{i+1} &= psmd_{i+1} - r_i \end{aligned} \tag{8a}$$

Actual evapotranspiration is derived from potential as follows:

$$\begin{aligned} ae_i &= pe_i && \text{when } smd_i < C \text{ or } p_i \geq pe_i \\ ae_i &= p_i + F\,(pe_i - p_i) && \text{when } D > smd_i \geq C \text{ and } p_i < pe_i \\ ae_i &= p_i && \text{when } smd_i = D \text{ and } p_i < pe_i \end{aligned}$$

where smd_i : soil moisture deficit at the start of day or time period i
ae_i : actual evapotranspiration during day i (L)
pe_i : potential evapotranspiration during day i (L)
p_i : precipitation during day i (L)
r_i : recharge during day i (L)
$psmd_i$: intermediate variable (L)
C : root constant (L)
D : wilting point (L)
F : empirical constant

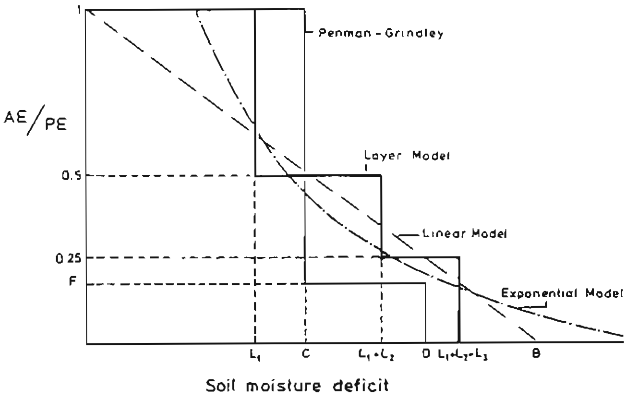


Figure 12. The form of actual - vs - potential evapotranspiration relationships used in various soil moisture budgeting models. (Based on Calder et al., 1983)

The shape of this function is shown in Figure 12, which shows the three parameters (C,D,F) of the model which must be calibrated or estimated. F is an empirical constant relating actual evapotranspiration to potential when deficits are greater than the root constant. All three are related to the vegetation cover and, as a second order effect, to soil characteristics. In the UK, monthly values of C and D, varying with crop type, are used (Lerner *et al.*, 1990, Table 11.2).

A number of other models have been used. They mainly differ from the Penman-Grindley model in the shape of the actual-potential evapotranspiration relationship, as shown in Figure 12. They include:

- layer models, such as those described by Calder *et al.* (1983);
- a linear model (Calder *et al.*, 1983);
- exponential models (Johannson, 1987).

It will usually be found that the difference between models will be less important than the accuracy of precipitation, potential evapotranspiration, irrigation and cropping data.

On less permeable soils a significant proportion of precipitation may become runoff. This is usually taken as an empirically derived proportion

$$q = f\,p \qquad \text{9c}$$

or a threshold may be included

$$\begin{aligned} q &= f\,(p - p_t) && \text{if } p > p_t \\ q &= 0 && \text{if } p < p_t \end{aligned} \qquad \text{9b}$$

or the runoff proportion may be related to the soil moisture deficit

$$q = f\,p\,(1 - smd/C) \qquad \text{9c}$$

where q : runoff in time period (L)
f : empirical factor (<1)
p : precipitation (L)
p_t : threshold precipitation below which no runoff occurs (L)
smd : soil moisture deficit (L)
C : empirical constant

The factors f, p_t and C are found by calibration against measured runoff, perhaps from an experimental catchment.

Recharge is seen to occur in some aquifers even when there is a soil moisture deficit. Rushton and Ward (1979) explored a number of ways of allowing such recharge in a soil moisture budgeting model. For a limestone aquifer in Eastern England, they finally chose a constant proportion of precipitation when it exceeded a threshold. The remaining precipitation entered a conventional Penman-Grindley model and could give rise to additional recharge. Various other empirical estimates of such rapid or bypassing recharge have been used; a method must be chosen and calibrated to suit local conditions.

Numerous authors (e.g. Howard and Lloyd, 1979) have pointed out that the time-step used in soil moisture models is critical. Longer time-steps, with the same parameters, lead to lower or zero recharge estimates. All recent work recommends a daily time step for humid zones. Intervals of less than a day, e.g. ones for both day and night, or storm based intervals, may be needed in arid and semi-arid areas, that is if the methods can be made to work at all.

There is no universally correct soil moisture budgeting model. Numerous authors have shown how different models, and different parameters, can significantly alter the recharge estimate (Alley, 1984). For any situation, a model should be chosen based on a conceptual model of the local recharge processes. This model should then be calibrated by one of the following methods, for point estimates:

(i) against a lysimeter,

(ii) against soil moisture measured by a neutron tube or tensiometer;

and for areal estimates:

(iii) against other estimates of recharge, e.g. from groundwater flow modelling,

(iv) against a catchment water balance.

In the UK, the Meteorological Office now provides useful data for hydrogeologists under the MORECS service. This is primarily intended for farmers, and issues 10 day, monthly and annual values of precipitation, actual and potential evapotranspiration, and effective precipitation. They are derived by the Penman-Montlith equation and are averages over 10 x 10 km squares of the National Grid. I recommend their use except in special cases, such as wetlands or where topography is very variable. Effort can then focus on hydrogeology instead of hydrology.

Darcian approaches

The flow of water in the unsaturated zone is governed by Darcy's Law, with the difference from fully saturated flow that hydraulic conductivity varies with moisture content. Moisture content and hydraulic conductivity also vary with pressure in the unsaturated zone. This section presents a brief review of the theory, and then describes how knowledge of these three parameters can be used to estimate the vertical flow of recharge, either by numerical modelling or by interpretation of field measurements.

Water, whether in the fully or partially saturated zones, has a total potential, h, which is the sum of two components; the gravitational component, z, and the pressure component, p:

$$h = z + p \tag{10}$$

It is convenient to work in terms of head of water, in which case z is the elevation relative to datum. For unsaturated flow, the ground surface is often taken as the datum. The variation of z with depth is always the same and is shown in Figure 13a as the gravitational component.

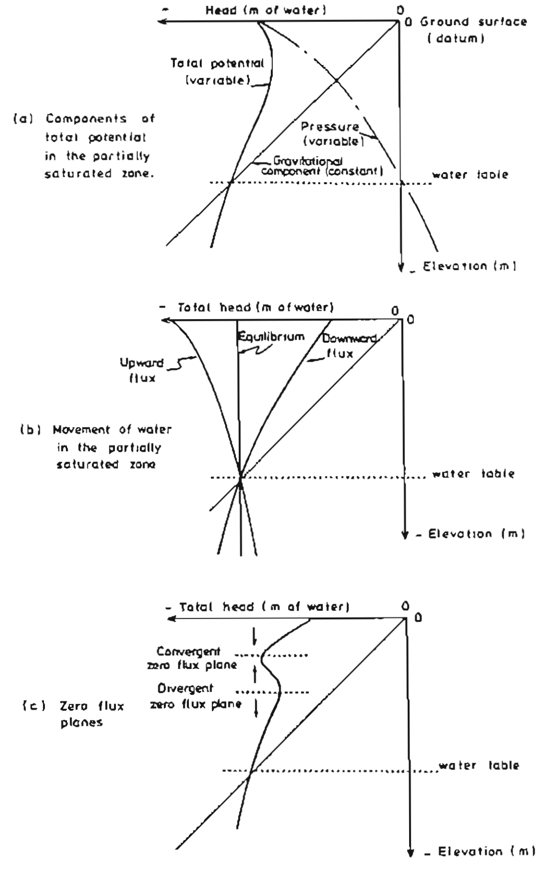


Figure 13. Head distributions and water movement in the unsaturated zone.

In the unsaturated zone, the pressure arises from the capillary forces that hold water in the smaller pores; it is always negative, that is below atmospheric pressure. Below the water table, pressure is a combination of hydrostatic forces (the weight of water above) and flow forces; and is always positive. An idealised curve is shown in Figure 13a.

Darcy's Law states that the seepage velocity is related to hydraulic conductivity and gradient of total head by:

$$q = -k \frac{\delta h}{\delta x} \tag{11}$$

(Notation for this section below eqn. 13)

Flow is usually vertical in the unsaturated zone, and hydraulic conductivity depends on moisture content, so Darcy's Law is rewritten:

$$q = -k_\theta \frac{\delta h}{\delta z}, \text{ where } k_\theta = f\,(k,\theta) \tag{12}$$

Thus flows depend on the profile of total potential, which is obtained by summing the gravitational and pressure components as shown in Figure 13a. When total potential is constant, the profile is in equilibrium and no flow is occurring (Figure 13b). When the gradient is upwards, flow is upwards, perhaps to satisfy evaporative demand in the root zone. Conversely, when the gradient of total potential is down, the flow is downwards. Both upward and downward flow can occur in the same profile, in response to time variant inputs (Figure 13c). The zones of upward and downward flow are separated by a "zero flux plane", so called because there is no flow across it.

In transient situations, water is taken up or released from storage by changing the moisture content or degree of saturation. A differential equation to describe this process can be obtained by combining Darcy's Law with a mass conservation equation:

$$\frac{\delta \theta}{\delta t} = \frac{\delta}{\delta z}\left[k_\theta \frac{\delta h}{\delta z}\right] - s \tag{13}$$

where q : seepage velocity (L^3/T)
k : hydraulic conductivity (L/T)
k_θ : unsaturated hydraulic conductivity (L/T)
θ : moisture content
h : hydraulic head (L)
s : rate of outflow of moisture (l/T)

A major difficulty with flow in the unsaturated zone is the nature of the relationships of hydraulic conductivity and moisture content to pressure. Both relationships have hysteresis, that is are different for wetting and drying (Figure 14).

Numerical models are widely used to solve eqn. 13 (e.g. Jansson and Halldin, 1979, 1980).

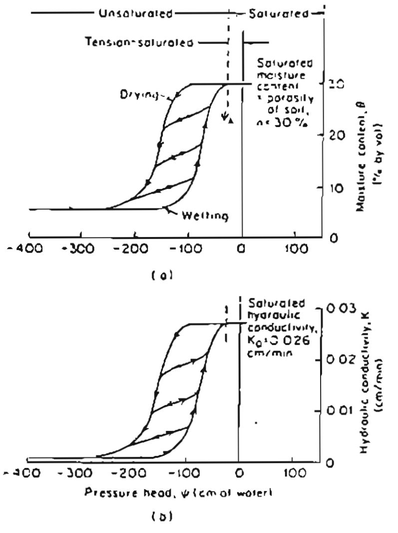


Figure 14. Characteristic curves relating hydraulic conductivity, moisture content and pressure head for a naturally occurring sandy soil (after Freeze and Cherry, 1979, Figure 2.13).

This is not straightforward, requiring good data on k-θ-p properties and their variation, and good field observations to use for calibration. With the 3-d variability of soil properties, and the presence of often unexpected features such as macropores, models often have limited value. They can be good for point estimates, developing understanding, and research; they can be poor for water resources and areal estimates.

There are three situations where measurements of pressure and moisture content in the unsaturated zone can potentially be used to estimate recharge:

(i) When there is no input to the soil profile at the surface, and the profile is draining to the water table or to evapotranspiration,

(ii) When there is input into a thick unsaturated zone, but it is insufficient to saturate the soil,

(iii) When there is sufficient surface input to saturate the profile.

No input. When there is no precipitation input to a soil profile, evaporation rapidly lowers the moisture content and pressure near the surface. A zero flux plane develops and can be detected by tension measurements. Below the zero flux plane, water drains downward to become recharge (Figure 15).

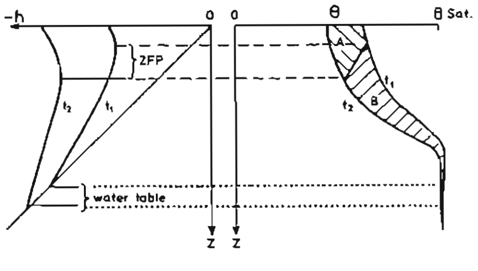


Figure 15. Evaporative (A) and drainage (B) fluxes from a profile with no surface input

The zero flux plane and water table will move over the time interval (t_1, t_2). The area A represents the volume evaporated, the area B represents the volume draining to the water table. Measurements of pressure and soil moisture have been used to estimate recharge in this manner. For example, Wellings (1984) gives examples for chalk in southern England where there is continuous drainage (recharge) throughout the summer despite the potential evapotranspiration greatly exceeding precipitation in these months. Cooper (1980) gives an example in eastern England where 43% of annual drainage occurred when there was a zero flux plane.

Small input. The inputs of water to a soil profile usually vary in time, and so there are transient changes in the soil moisture and pressure. Steady state conditions will never apply near the surface nor in the root zone. However, a thick unsaturated zone will tend to dampen and coalesce seasonal (or storm) pulses of recharge and so may reach an approximate steady state at depth. In this condition the unsaturated form of Darcy's Law (Eqn. 13) can be applied.

It is generally assumed that pressure is constant with depth under these steady state conditions, so that the hydraulic gradient is 1 and entirely due to the gravitational component of head. In this case:

$$r = k_\theta \qquad 14$$

where r : recharge rate (L/T)
k_θ : unsaturated hydraulic conductivity (L/T)

and both r and k_θ should be the same at all depths. This implies that layers which have different saturated conductivities will be at different saturations to make their unsaturated conductivities equal.

Sammis *et al.* (1982) describe briefly how the method was applied to a site in Arizona where the water table was 42 m below the ground. Push tube samples were taken every 3 m during drilling of a borehole. Pressure was found by inserting a tensiometer into the samples; moisture contents were measured by the gravimetric method. They give few details of their method for unsaturated conductivity. Their results are not impressive with k_θ estimates of 12, 790, 0.005, 226, 0.2 and 23 mm/y for various depths; under the method's assumptions, all these values should be the same.

Large input. If the infiltration rate is high enough, the ground will become saturated. Water will move downwards under gravitational forces only, that is the hydraulic gradient dh/dz = 1, provided that ponding on the surface is only slight so that pressure does not rise significantly above atmospheric. In this restricted circumstance:

$$r = k_{sat} \qquad 15$$

where r : recharge rate (L/T)
k_{sat} : saturated hydraulic conductivity (L/T)

that is infiltration rates can be estimated from saturated hydraulic conductivity. Although appealingly simple, this method has limited usefulness except when continuous applications of water are made, such as in irrigation schemes or river beds. Recharge will be controlled by the lowest permeability layer, which may not be at the surface.

Tracer techniques

Tracers (both isotopic and chemical) have been widely used in estimating recharge. They can be grouped into:

- environmental tracers, i.e. those that are already present in the geosphere,
- tracers applied by the researcher.

Environmental tracers can be assumed to be applied evenly over the whole land surface. Their movement in the unsaturated zone therefore is only in one dimension, with no lateral diffusion or dispersion. Applied tracers on the other hand are applied at a point or over a small area and so can disperse laterally as well as move vertically. This means that checks on the mass balance of an applied tracer are much more difficult to achieve.

There are several ways that tracers can be used:

(i) signature methods, in which particular parcels of water are labelled and traced. These methods have been used in three different situations:

 (a) environmental tracers in the unsaturated zone,

 (b)applied tracers in the unsaturated zone,

 (c) environmental tracers below the water table;

(ii) throughput methods, when fluxes of tracer and water are calculated in the unsaturated zone, usually for environmental tracers;

(iii) turnover or transit time calculations, which are used for whole aquifers, usually with environmental tracers.

The signature and throughput methods generally assume a single flow system and piston displacement in the unsaturated zone. If there are two flow systems, recharge can bypass the matrix of the unsaturated zone, for example by flowing in fractures, down rootways or other macropores. All of these phenomena have been observed in arid and semi-arid areas and may invalidate the methods. Some workers have interpreted their field data in terms of two flow systems; an example is discussed below.

Signature methods

General principles. Infiltrating water may be labelled (in time) by a particularly noticeable input of tracer. This may either be applied artificially, or be an environmental tracer, for example from an atmospheric nuclear explosion. This labelled water may be identified at depth in the soil profile at future dates, so providing information on water movement in the unsaturated zone.

Once the water at a particular depth has been dated, two methods are available to calculate average recharge. At a time of maximum soil moisture deficit, the total water content above the dated depth is the total recharge over the intervening years, i.e.:

$$r = 1/m \sum_{i=1}^{m} \theta_i \, l_i \qquad 16$$

(Notation below eqn. 20)

The second method is a mass balance of tracer, and is only applied to environmental tracers, or those applied over a large enough area to prevent lateral dispersion. The amount above the dated layer is measured:

$$T_p = \sum_{i=1}^{m} T_i \, \theta_i \, l_i \qquad 17$$

and equated with the input to recharge after allowing for radioactive decay:

$$T_p = \sum_{j=1}^{n} r_j \, T_j \exp(d_j) \qquad 18$$

Some workers estimate the relative sizes of r_i using rainfall and potential evapotranspiration data to define annual weightings:

$$r_j = w_j \, r, \sum w_j = n \qquad 19$$

Combining eqns. 17 - 19 gives average recharge as:

$$r = T_p \Big/ \left[\sum_{j=1}^{n} w_j \; T_j \exp(d_j) \right] \qquad 20$$

where n : number of years between signature and sampling (T)
r : mean annual recharge (L/T)
r_j : recharge in year j (L)
d : depth of dated water (L)
θ_i : moisture content in sampling interval i
l_i : length of sampling interval number i (L)
T_i : tracer concentration in sampling interval i (M/L^3)
T_p : tracer mass in profile (M)
T_j : tracer concentration in rainfall in year j (M/L^3)
$\exp(d_j)$: decay of tracer input since year j
n : number of years

m : number of sampling intervals
W_J : ratio of recharge in year j to mean recharge

Other workers have found that assuming equal recharge every year (i.e. $r = r_j$) is sufficiently accurate (Allison & Hughes, 1978).

For a conservative tracer, average recharge is estimated from eqn. 16. A balance of tracer (eqn. 20) can then be used to calculate the fraction of precipitation that becomes recharge.

Environmental tracers. The most commonly used environmental tracer for this method has been tritium which, as well as being produced naturally in the atmosphere, was introduced into the atmosphere in large amounts in 1952 by nuclear tests. There was a significant peak, particularly in the northern hemisphere, in 1963/4 (Figure 16). Atmospheric concentrations were always much less in the southern hemisphere and, as atmospheric testing stopped in 1963, tritium levels are now falling everywhere. More sensitive analytical techniques have extended the usefulness of environmental tritium for some time, but it is now rarely of use for signature methods.

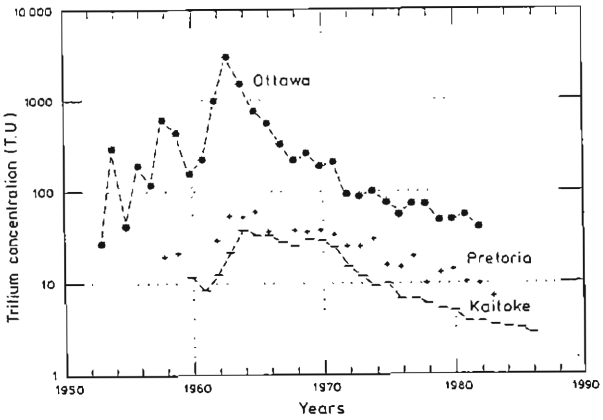


Figure 16. Tritium concentrations in rainfall (data courtesy of IAEA)

Oxygen-18 $(^{18}0)$ and Deuterium (^{2}H) are fractionated during precipitation and evaporation processes, a property which has so far prevented their use in any tracer balance method of estimating recharge. They have shown useful signatures in temperate, high recharge areas. For example Thoma *et al.* (1979) found seasonal, markers of deuterium in sand dunes; Bath *èt al.* (1982) and Saxena and Dressie (1984) sometimes found cyclic profiles of $^{18}0$ and ^{2}H corresponding to seasonal rainfall and recharge; snowmelt often contains a distinctive $^{18}0/^{2}H$ signature. However Allison (1988) reports that such seasonal peaks are not seen in arid and semi-arid areas because diffusion will redistribute the peaks over the very short interval between them.

Applied tracers. The signature method is mainly used with applied tracers. Radio-isotopes are usually used as they can be detected at low levels, thus they can be introduced in small quantities without large disturbances to soil or its moisture content. The tracer is introduced below the lowest depth of the zero flux plane, that is below the region where upward flow or evapotranspiration losses can occur (Figure 13). The method is widely used in India, and Figure 17 shows a typical site layout for tritium and Cobalt-60 (^{60}Co) injection.

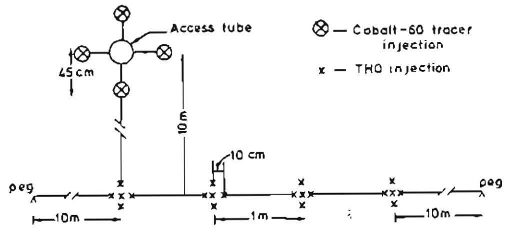


Figure 17. Layout of injection points for Tritium and Cobalt-60 (after Chandrasekharan et al., 1988).

The tracer position is determined at a future date, for example at the end of the wet season or one or more years later. The tracer may be detected by drilling a cored borehole and analysing the porewaters (^{3}H), or by non-destructive measurements of radiation from an adjacent open borehole (^{60}Co). Inevitably the injected tracer will have dispersed (Figure 18).

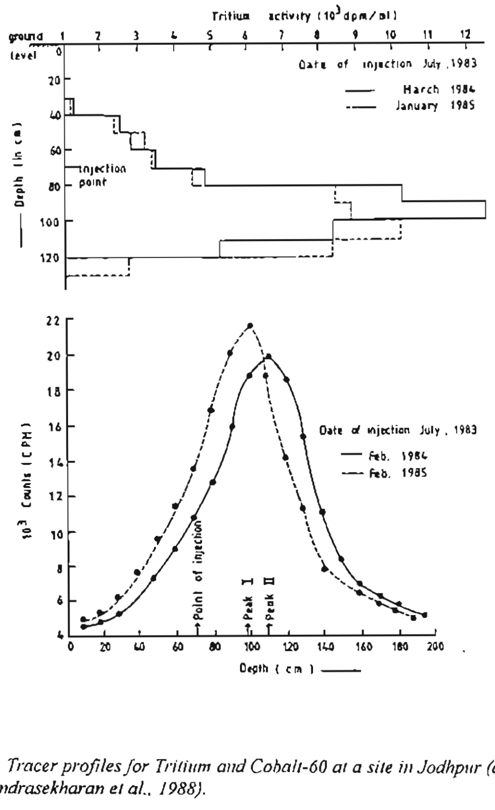


Figure 18. Tracer profiles for Tritium and Cobalt-60 at a site in Jodhpur (after Chandrasekharan et al., 1988).

Conventionally the centre of gravity of the tracer profile is taken to show how much displacement there has been; the moisture contained in the soil between injection point and centre of gravity is equivalent to recharge over the time period. Examples of the use of applied tracers are given by Sharma *et al.* (1985), Chandrasekharan *et al.* (1988) and Athavale and Rangarajan (1988).

Environmental tracers below the water table. Because the peak of bomb tritium has moved out of the unsaturated zone in most parts of the world, there have been attempts to use its distribution below the water table to estimate recharge. The method relies on obtaining a profile in an area where groundwater only moves vertically. Profiles have been taken on groundwater divides and in the centre of recharge mounds. The formulae are as above. In three published case studies of alluvial aquifers (Knott and Olimpio, 1986; Larsen *et al.*, 1987; Delcore and Larson, 1987), less tritium has been found in the profile than would be expected from recorded inputs, even after allowance for decay, seasonally varying inputs and dispersion. This suggests that there is an inconsistency in the conceptual model of the method of aquifer, for example that there are multiple flow pathways in granular materials, or that it has not been possible to estimate inputs correctly.

Mass balances of tracer are very useful because, as discussed above, they can reveal if some of the tracer (and recharge) have moved by a different pathway; for example along fissures or root channels. Repetitive sampling of the same profile is also useful in this context. For example, Smith *et al.* (1970) showed by repetitive profiling in the English Chalk that some tritium was lost from the unsaturated zone, and implied that 15% of recharge bypassed the matrix where piston flow was occurring.

Throughput method (chloride)

If a tracer does not have a signature that dates the profile, recharge can only be estimated if the concentration of the tracer is affected by evaporation processes that reduce the precipitation to recharge. Conservative, i.e. non-evaporated, tracers are of course concentrated in recharge as evaporation proceeds, so that the flux of tracer input at the surface equals the flux of tracer reaching the water table. As the age of the tracer reaching the water table is unknown, it is necessary to assume a steady input of tracer at the surface. Only environmental tracers have been used, as a long enough input is needed to reach steady state; chloride is the most commonly used ion.

Assuming no input of tracer from minerals and that water and tracer are transported at the same rate, the flux balance of tracer between surface and water table is:

$$r\,T_r = p\,T_p + f_d \tag{21}$$

(Notation below eqn. 22.)

The natural sources of chloride are from the ocean and terrestrial processes. Another major source (and sink) of chloride is agricultural, with deposition from fertilisers and animals, and removal in crops, soil erosion and animals. Agricultural fluxes of chloride can dominate the balance, and the method should be used with care in such areas.

Omitting dry deposition:

$$r = p\, T_p / T_r \qquad 22$$

where r : mean recharge rate (L/T)
T_r : mean tracer concentration in recharge (M/L^3)
p : mean precipitation (L/T)
T_p : mean tracer concentration in precipitation (M/L^3)
f_d : dry deposition flux ($M/L^2\,T$)

The method therefore consists of measuring precipitation, tracer concentration in precipitation, and tracer concentration profiles with depth from cored boreholes. Provided recharge and tracer inputs are in steady state and there is no secondary recharge mechanism, tracer concentration should increase with depth until a steady concentration is reached (Figure 19a). This indicates that no evaporation takes place from below this depth and eqn. 22 can be applied.

Profiles such as Figure 19b-d must be interpreted with care. They indicate that one of the assumptions of steady state, no secondary pathways and no input from minerals or agriculture do not apply.

Allison and Hughes (1983) interpret profiles like Figure 19b as a change in recharge rate at some time in the past, in their case due a change in land use. Allison *et al.* (1985) regard another profile like Figure 19b as showing a change in recharge rate some 16,000 years ago. In both cases, the recharge rates are so low as to be of little interest for groundwater resource modelling. Sharma and Hughes (1985) use profiles like Figure 19c by assuming that concentrations at the water table can be used in eqn. 22.

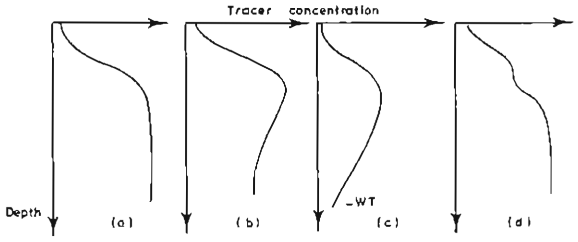


Figure 19. Schematic profiles of a conservative tracer. (a) Ideal case, recharge can be estimated. (b) Probable change in recharge history. (c) Secondary pathways for recharge. (d) Addition of tracer from soil.

Variability of recharge across catchments

Recharge varies across catchments because the controlling factors vary, both in their nature and size. These factors include:

- precipitation and other water supplies
- geology and soil
- topography and landform
- groundwater condition.

Because recharge is a non-linear process, it is not possible to use average values of each controlling factor to derive an average recharge. Recharge should be estimated separately for each homogeneous zone; the spatially varying values are of course essential for groundwater modelling studies.

In general, the more detailed the subdivision into zones, the more accurate will be the recharge values, but the more expensive and time consuming will be their estimation. The amount of data available (maps, climate stations, Landsat imagery, etc.) will partly determine the detail that can sensibly be achieved. Sensitivity analyses will help decide how important are variations in the controlling factors. Even where data on precipitation and climate are scarce, zoning on the basis of existing data will still be valuable. It will enhance understanding of processes in the catchment, and allow subjective adjustments to recharge estimates.

Table 1 (given on p. 52) suggests a set of factors that should be used to zone a catchment for recharge calculations. Initially, each factor should be mapped on a separate overlay, then the overlays combined to produce a master map of homogeneous zones. Recharge is then estimated for each zone by one of the methods discussed above.

Some authors have observed important changes in recharge controlling factors over very short distances (<1 to 100 m) in apparently homogeneous terrain (e.g. Berndtsson and Larson, 1987; Nielsen *et al.*, 1973). Others have reported significantly varying recharge over similarly small distances (e.g. Sharma and Hughes, 1985; Kitching *et al.*, 1977). The importance of these variations has not yet been established, and there is certainly no practical way to take account of them for engineering studies. Three to ten repeat samples within each zone (as defined by the factors in Table 1) would certainly be worthwhile if they can be afforded, but first priority should go to obtaining one value from each zone.

Localised recharge

The introduction to this chapter pointed out that there is a category of recharge intermediate between direct (at the spot where the precipitation falls) and indirect (along main river channels). This intermediate category is called localised, implying some horizontal movement of water before recharging groundwater. This movement is on a scale too detailed to map for engineering studies, and therefore causes great difficulty in estimating recharge. Unfortunately this type of recharge is often the largest in arid and semi-arid areas, and can be important in humid zones such as UK (drift cover) and Scandinavia (bed rock outcrops).

A number of situations where localised recharge will occur are identified in the following sections. Suggestions are made about the most appropriate methods for estimating recharge in each.

In weathered, bare hardrock or limestone terrain, recharge is into distinct fissures. There are few satisfactory ways of estimating the localised recharge component in such terrain. Tracers, Darcy's law below the water table, and water balances o^c project or representative basins offer the best possibilities.

Topographical depressions can range in size from centimetres to enclosed catchments of several square kilometres. Studies in a wide range of terrains have shown that recharge is focused in such depressions, where they exist (e.g. Freeze and Banner, 1970); Rehm *et al.*, 1982). This is true even for highly permeable materials.

At the smallest and intermediate sizes, there is no alternative but to ignore the depressions.

The largest sizes are typified by Qatar, where 7000 km^2 or 61% of the total land area drains to 850 depressions; individual catchment areas range from 0.25 to 45 km^2. Estimating recharge in such cases is similar to estimating it from minor wadis below, but is simpler. In Qatar, recharge was estimated by water balance methods for the period of detailed observation (Lloyd *et al.*, 1987). Rainfall data and post-storm surveys of individual depressions were used to estimate rainfall-runoff characteristics; 15-25% of rainfalls over 10 mm/day became runoff. Water level measurements showed how long depressions took to empty by evaporation and infiltration. Allowances were made for soil moisture storage in depressions, which would later be evaporated. Once a relationship between rainfall and recharge had been developed for the period of observation, it was used to extrapolate backwards in time when only precipitation data were available.

In arid and semi-arid areas there are many small channels which carry ephemeral flows only in response to significant storms. These flows will often not reach the main channel, where flow gauging stations are usually located, but will infiltrate into the channel alluvium. These channels may drain small bare rock outcrops, in which case they will hardly be visible, or may drain catchments of several thousand km^2 in areas where little flow gauging is carried out.

The most common approach to estimating recharge in these cases in practice is a mixture of water balance and empirical formula. Rainfall-runoff relationships are estimated or transposed from other areas, and the proportions of runoff going to each destination (evaporation, soil moisture storage, irrigation, etc.) is measured, estimated or guessed; the residual is recharge.

Many alluvial aquifers can be described as mountain front systems, including alluvial fans, piedmont plains and subsidence basins. The major sources of recharge to these systems are often along the mountain boundary, consisting of two components:

(i) subsurface inflow from the mountain mass to the basin infill,

(ii) infiltration of runoff in defined channels, minor unmapped channels, and from mountain slopes.

These are illustrated in Figure 20.

The two components of mountain front recharge are difficult to separate in practice. The runoff component can be estimated by water balance methods, as discussed above, but the

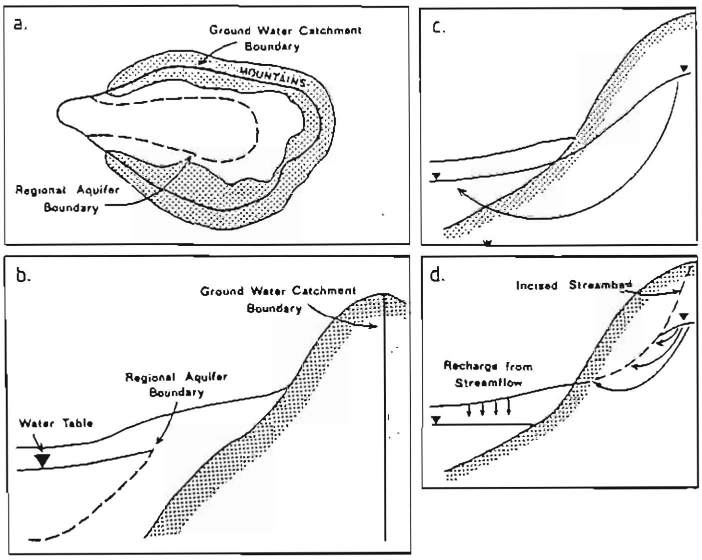


Figure 20. Principles of groundwater recharge in mountain front basins. (a) Plan view of boundaries of groundwater catchment and regional aquifer. (b) Cross-section of boundaries. (c) Mountain front recharge by subsurface inflow. (d) Mountain front recharge through streamflow infiltration (Wilson et al., 1980).

subsurface flow is difficult to isolate. The best solution is probably to estimate the total by a Darcy throughflow calculation within the main aquifer. The section over which flow is estimated should be well within the aquifer, away from the mountain boundary, for two reasons. All of the mountain front recharge will occur upslope of the section as all the wadis flows will have infiltrated. Secondly, it is probable that more data on groundwater gradients, hydraulic conductivities and aquifer cross-section will be available in the main body of the aquifer, so giving greater accuracy and confidence in the results. (This section is based upon an excellent review of mountain front recharge by Wilson in WRRC, 1980).

RECHARGE FROM RIVERS

Recharge from rivers is probably the most difficult type of natural recharge to estimate. Temporal variability of flows is high, and measurement is difficult. The most important part of any technique is to understand the hydrogeological context before making any calculations.

River types

Rivers can be classified in several ways, for example by flow characteristics (Figure 21):

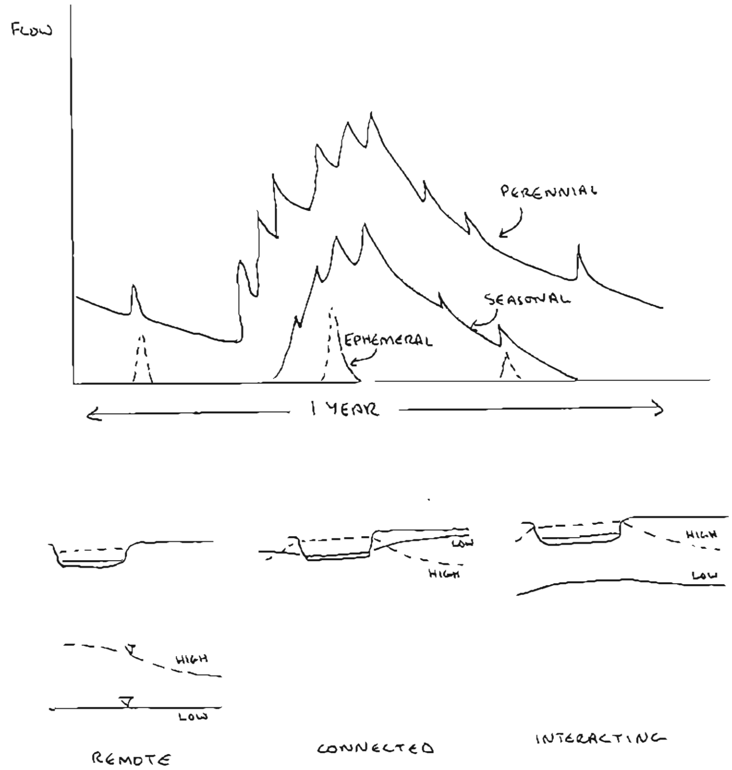


Figure 21. Classification of rivers for recharge estimation.

(i) Perennial, that is flowing all year. This implies a river source in a higher rainfall catchment, or a river fed by groundwater.

(ii) Seasonal, that is flowing for part of the year.

(iii) Ephemeral rivers flow only in response to storms and are the classical type found in arid areas.

An alternative grouping is by their connection to the water table. This will affect the ability of the aquifer to accept water and hence may control recharge rates. Three broad types can be defined:

(i) Remote from (high above) the regional water table. In general, perennial and seasonal rivers are perched in a low conductivity material which reduces infiltration rates and keeps the river flowing, while ephemeral rivers need not be perched.

(ii) Connected to groundwater.

(iii) Above the water table, but close enough for the water table to rise to the river in response to recharge. For the purposes of this chapter, this intermediate category has been included with (ii) connected rivers.

Different methods of recharge estimation apply to perennial/seasonal rivers on one hand and ephemeral rivers on the other. Remote rivers require differences of approach from those connected to the water table.

Controls on river recharge

The amount of recharge is controlled by the river flow, the river bed, and the aquifer. Controlling factors may include:

> river flow: flow rate, depth, flow volume, peak flow, velocity, duration of flow, frequency of flows, temperature (affects hydraulic conductivity), silt content;
>
> river bed: antecedent conditions, width, hydraulic conductivity.
>
> aquifer: boundaries, depth to water table, hydraulic conductivity (especially variations in conductivity), moisture retention in unsaturated zone.

As an example of such controls, Figure 22 shows the effect of antecedent conditions on transmission losses in a perched ephemeral river. The first flow (Figure 22a) occurred when the channel was dry and almost all flow is lost. In the second event (Figure 22b) there are tributary inflows between the two gauging stations before the main flood event arrives at the downstream end. These wet the channel bed and the transmission losses are much lower.

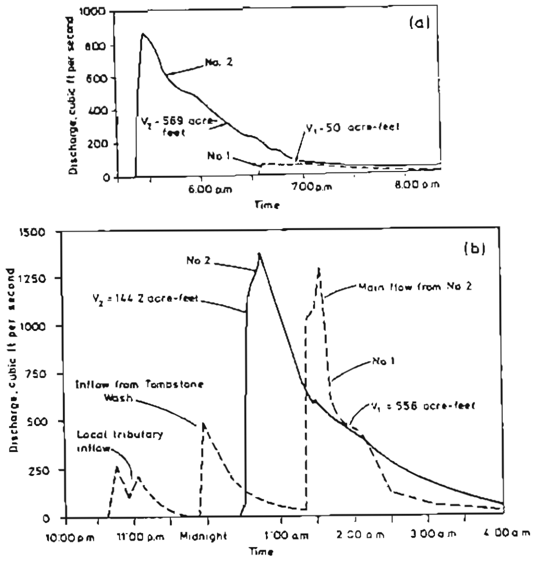


Figure 22. The effect of antecedent conditions on transmission losses. (a) High losses on 10th August 1959. (b) Low losses on 3-4th August 1959. No. 2 is the upstream gauging station, No. 1 is downstream (Keppel and Renard, 1959).

Figure 23 shows the importance of low conductivity layers in the unsaturated zone beneath a remote river. The water table responds directly to flood events (Figure 23b) but water is also stored in the unsaturated zone above two low conductivity layers; Figure 23b shows the variation in the elevation of the top of one perched mound, while Figure 23c shows profiles of moisture content with time from ground to water table. The perched mounds do not finally disappear until several months after the flood event.

Not all the water that leaves a river to move downwards becomes recharge. The distinction between transmission losses, deep percolation and recharge is useful. Transmission losses are the river flows that don't arrive at the downstream end of a river, deep percolation is the water that enters the aquifer, and recharge is that which crosses the water table. Transmission losses can become bank storage and evaporation from the surface and temporary pools, with the remainder becoming deep percolation. This in turn can be reduced by evaporation, perched water tables and underflow before becoming recharge.

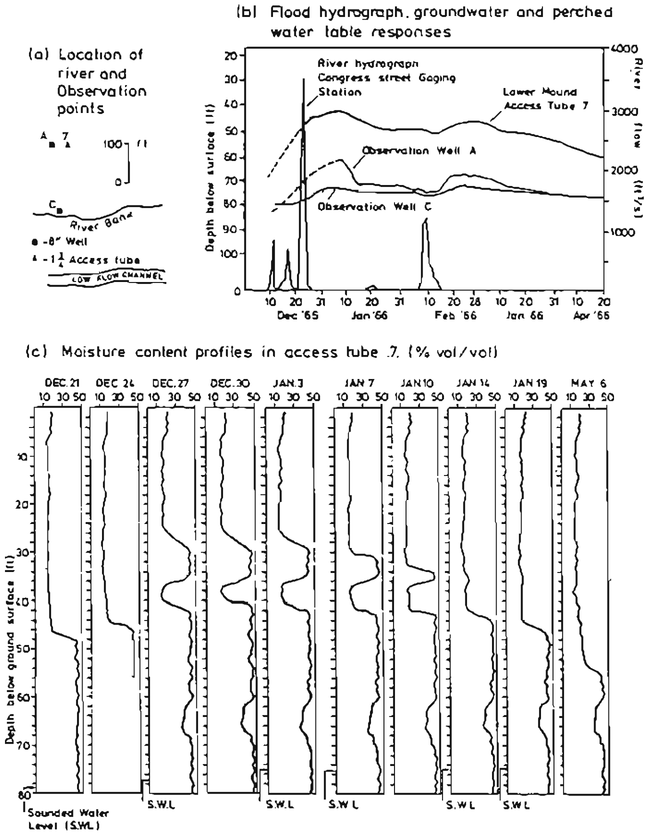


Figure 23. Effect of low conductivity layers on river recharge, with perched water tables forming at two levels (Wilson and De Cook, 1968).

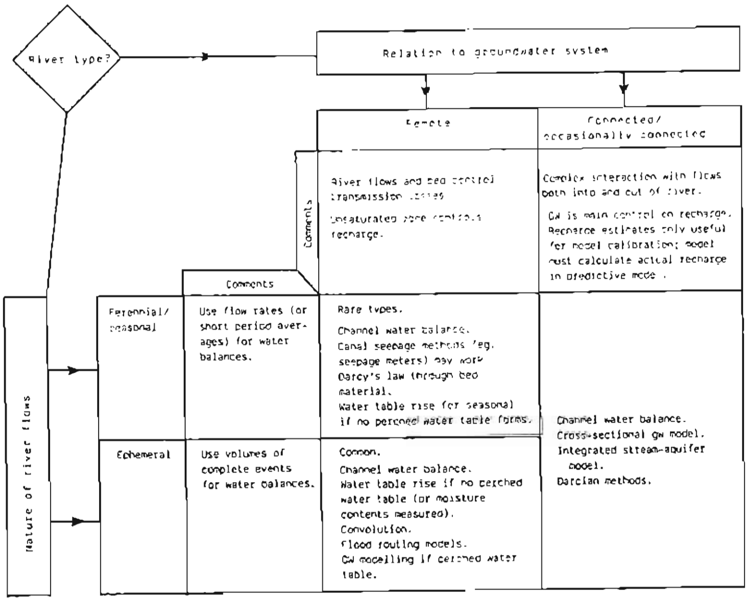
River recharge estimation methods

Direct measurement of river recharge is not possible. A number of empirical methods have been devised and are discussed below; they include a number of methods based on groundwater response. Water balance methods are also possible for all three items, transmission loss, deep percolation and recharge. Darcian approaches, that is based on the groundwater flow equation, are possible in theory but are difficult to use in practice without important simplifications. Tracer techniques have very limited use for quantifying river recharge. Models of river hydraulics and of the complete hydrological cycle in catchments have been used for river recharge estimation. Figure 24 outlines the choices for the various river conditions.

Empirical methods

In an empirical method, recharge is related to one or more of the controlling variables, for example river flow rate. Some examples are shown in Figure 25.

There are several general comments about the usefulness of these methods. Firstly they are likely to work best for transmission losses in perched rivers, for which there is less influence of the unsaturated zone and aquifer. Good estimates of recharge are needed to develop such relationships; once data collection networks have been installed, is it not better to continue to use them than to rely on a derived relationship? Finally, the relationships can only be transferred to other rivers which have the same hydrogeological conditions.



	Comments	Remote	Connected/ occasionally connected
Comments		River flows and bed control transmission losses. Unsaturated zone controls recharge.	Complex interaction with flows both into and out of river. GW is main control on recharge. Recharge estimates only useful for model calibration; model must calculate actual recharge in predictive mode.
Perennial/ seasonal	Use flow rates (or short period averages) for water balances.	Rare types. Channel water balance. Canal seepage methods (eg. seepage meters) may work. Darcy's law through bed material. Water table rise for seasonal if no perched water table forms.	Channel water balance. Cross-sectional gw model. Integrated stream-aquifer model. Darcian methods.
Ephemeral	Use volumes of complete events for water balances.	Common. Channel water balance. Water table rise if no perched water table (or moisture contents measured). Convolution. Flood routing models. GW modelling if perched water table.	

Figure 24. Relation of type of river to recharge estimation methods.

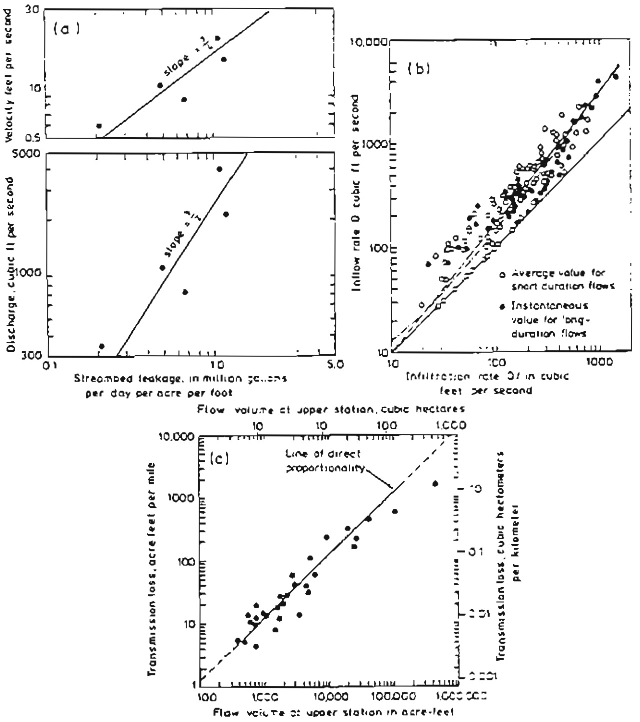


Figure 25. Examples of empirical relationships for transmission losses from remote rivers. (a) Norris, 1970, (b) Burkham, 1970, (c) Jordan, 1977.

Other examples of empirical relationships for transmission losses are given by the papers from which Figures 22 and 25 were taken, and by Lane *et al.* (1971). WRRC (1980) reviewed a large number of American studies of transmission losses in remote ephemeral rivers. They suggested some general rules for transmission losses.

Groundwater response under ephemeral rivers

There are a number of empirical techniques based upon analysing groundwater response. They are typified by the <u>convolution</u> approach, and include a spectral analysis method (Gelhar *et al.*, 1979) and a linear transformation of the annual distribution of rainfall (Flug *et al.*, 1980). The latter two require more simplifying assumptions but are more mathematically complex than convolution. The convolution method has been described by Besbes *et al.* (1978) and others.

All of the groundwater response methods require a good independent estimate of recharge for at least one event. Without this, the parameters of the method cannot be found. Therefore any results for other events will be less accurate than the one independent estimate.

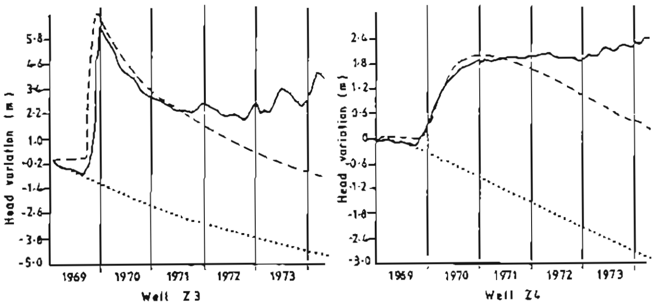


Figure 26. Groundwater responses to a single large recharge event, Kairowan, Tunisia (Besbes et al., 1978). Solid lines: observed heads. Dotted lines: heads computed by groundwater model assuming no recharge, Dashed lines: heads computed with estimated recharge.

Figure 26 shows some examples from Besbes *et al.* who give a thorough comparative study of several methods in a semi-arid area of Tunisia. They observed the initial and peak response and constructed the remainder using a model. The volume of the recharge event was initially estimated by a water table rise calculation using a guessed specific yield, then refined by calibrating the model. Other authors have used transmission losses to estimate recharge volume.

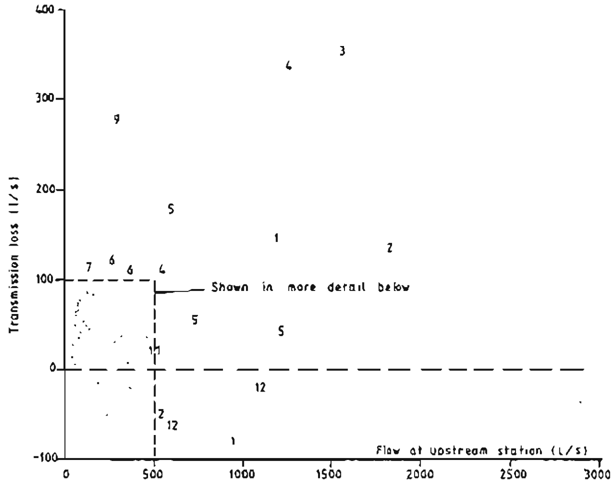
Rivers in contact with the water table

Recharge from rivers that contact the water table is affected by conditions below that water table. In these cases, the empirical methods outlined above rarely work. Figure 27 shows data for a perennial river crossing a major alluvial aquifer in N. Africa. There is no clear relation between transmission loss and river flow because of the effect of the groundwater system. The data can be interpreted to show an upper envelope to transmission loss as 100% of flow, with many months having lower losses. The high loss months are in late summer when groundwater levels are low, which suggests that the state of groundwater is the second control on recharge. No general relationship can be devised for these rivers and recharge for each time period must be explicitly calculated from measurements, or assessed through modelling.

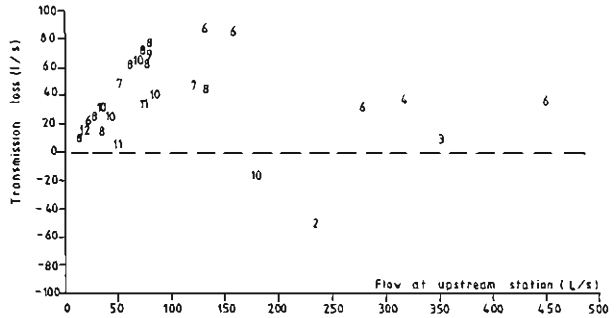
Channel water balance

A water balance of the flows along a river reach is the most straightforward way to estimate transmission losses. By careful consideration of processes, this may be extended to deep percolation or recharge. Using river flow rates,

$$R = Q_{up} - Q_{down} + \sum Q_{in} - \sum Q_{out} - E_a - \delta S / \delta t$$



(a) Complete flow range



(b) Low flows

Figure 27. Transmission loss - vs - flow for a river in contact with the water table showing little correlation. Each point is one month, identified by its month number (e.g. 1 = January).

or in volume terms,

$$R\,\delta t = V_{up} - V_{down} + \sum V_{in} - \sum V_{out} - E_a\ \delta t - \delta S$$

where Q : flow rate (L^3/T), V : flow volume (L^3)
Q_{up}, V_{up} : the flow at the upstream end of the reach
Q_{down}, V_{down} : the flow at the downstream end of the reach
Q_{in}, V_{in} : inflows from tributaries, urban effluents and irrigation returns
Q_{out}, V_{out} : outflows for water supply, irrigation
E : evaporation from water surface or stream bed (L^3/T)
δS : the change in channel or unsaturated zone storage (L^3)

Total flow volumes should be used in preference to instantaneous flow rates for ephemeral rivers, as lag between the two gauging stations will be difficult to estimate accurately.

Any time period can be used with flows for seasonal and perennial rivers, but there are advantages in using a short period. More data points are generated, and conditions are relatively constant during each period. This may make the process controlling recharge more obvious. More importantly, the higher errors associated with the higher flows will not dominate the whole period as would occur if averages were taken over long periods, including low flow periods. Travel time of flood waves between gauging stations should be allowed for if short time periods are used.

Channel water balances are probably the most accurate way to estimate recharge, and often provide the data on which other methods are based. They are essential when modelling rivers which may receive as well as donate water. The data collected may have many other uses, including rainfall-runoff modelling for transposition to ungauged catchments.

The disadvantages include cost and accuracy. Setting up, maintaining and running river gauging stations is expensive. They must be able to measure large floods accurately. In arid and semi-arid areas where access may be difficult, floods are infrequent and river beds shift, it may be too expensive.

The method is prone to inaccuracy as recharge is calculated as the difference between large numbers. Measurement errors on high river flows are commonly ±25%, and for flash floods in arid and semi-arid areas can easily be a factor of 2. Thus the possible errors in recharge estimates can be be ±100%.

Water table rise

An alternative to a surface water balance for transmission loss is a groundwater balance for river recharge. Observation wells perpendicular to the axis of the river will show the profile of a recharge mound. Under ideal conditions the volume of water in the mound is the amount of recharge. Several points should be noted:

(i) The groundwater flows away from the mound may be significant in relation to the change in storage.

(ii) Storage in the unsaturated zone is common below ephemeral rivers on alluvial aquifers. Wilson and De Cook (1968) found that 33% of recharge showed up immediately as a water table rise; the remainder took several months to arrive (Figure 23).

(iii) Specific yield is a critical parameter but difficult to estimate.

Darcian approaches

There are three ways to use Darcy's law in estimating recharge from rivers as follows:

(i) using infiltration equations and flow nets, that is desk studies of homogeneous and isotropic aquifers with assumed boundary conditions,

(ii) gathering field data on aquifer properties, moisture content and pressure in the unsaturated zone, and piezometric heads in the saturated zone, and using these data in the equations of flow to analyse real events,

(iii) using numerical models.

The infiltration approach assumes no preferential pathways and no lateral flow (Figure 28), so would not be valid in the case of Figure 23 where low conductivity layers will cause perching and lateral flow. The equation ceases to be valid once the wetting front has reached the water table, when horizontal flow occurs.

Field data on moisture content and pressure in the unsaturated zone beneath river beds have been collected for research projects on recharge estimation (for example Figure 23). I know of no instances where such data have been collected for resources studies, presumably because of the expense and difficulty of collecting enough data to describe a transient process with three dimensional variability.

Head and hydraulic conductivity data from connected rivers could be used to estimate flow away from the river; velocity measurements using tracers could also be used. The difficulty in using them for rivers is the transient, three dimensional nature of the flow pattern.

Numerical modelling techniques for river recharge

Dillon (1982) gives a thorough review of stream-aquifer models, quoting over 100 references. He classifies models into three broad types:

(i) Surface water models, including flood wave models considered below, and linear response models, typified by the convolution approach discussed above.

(ii) Groundwater models, discussed below;

(iii) Integrated stream-aquifer models.

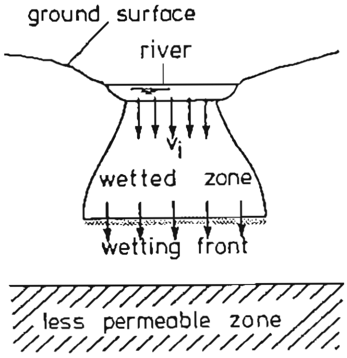


Figure 28. Idealised infiltration from a perched river into a homogeneous material.

Dillon (1982) lists and comments upon the assumptions commonly used to make models tractable. He reports that practical studies and field observations have shown that the following assumptions, if used, can have significant effects on estimated recharge:

- homogeneous, isotropic and infinite aquifer,
- Dupuit-Forchheimer approximation,
- fully penetrating river,
- constant recharge rate,
- isothermal,
- initially dry soil,
- no resistance to air flow.

Most of these assumptions can be dispensed with by using a more complex computer code. A better description of the aquifer requires more (expensive) field data. In practice this has meant that most of these models have been used for theoretical or parametric studies of hypothetical rivers, rather than real resource studies.

Integrated river-aquifer models. A number of authors have written models which link river and groundwater flows and heads. They are for hydraulically connected rivers, and model river flow in some detail. Almost none have been used in groundwater resource studies, and all require a large data collection, programming and calibration effort. In general, this type of model is too sophisticated for recharge estimation. They are concerned with the immediate locality of the river and with short time steps. Water may well enter the aquifer on this scale that never recharges the regional groundwater system. Such interactions only confuse a regional groundwater modelling effort. A pragmatic policy for connected rivers is to draw the model boundary outside this zone of fluctuating interaction rather than at the river bed.

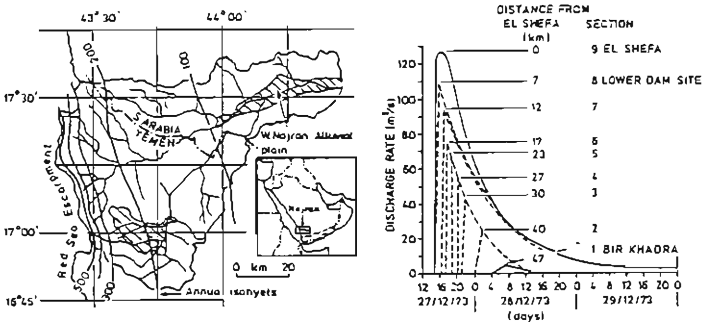
Tracer techniques for river recharge

Environmental tracers in groundwater seem to have little role in quantifying river recharge. Almost all of the groundwater in the vicinity of a recharging river will have come from that river. Therefore the local groundwater will carry the same environmental tracers as the recharging water, and the two cannot be separated.

Of course isotopes and other tracers are very valuable in identifying recharge sources and in delineating zones of river recharged groundwater. Some discussion of methods is given in the Guidebook on Nuclear Techniques in Hydrology (IAEA, 1983).

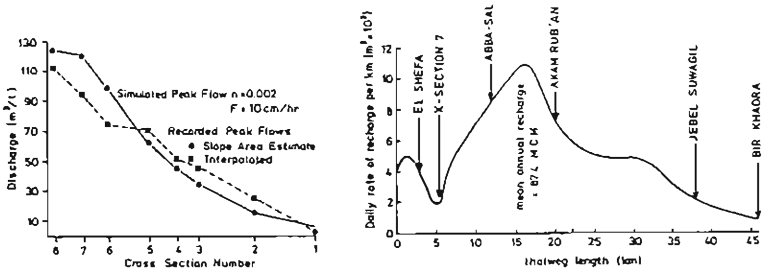
Surface water modelling

There are often more data available on surface water than groundwater and so many practical recharge studies have had to rely on surface data to predict recharge. Flood routing models of river flows have been used to estimate transmission losses by Cornish, Smith and Binnie and Partners (see FAO, 1981, Section 18.3.3.4). River level, or preferably flow, hydrographs are required for several stations along the river. Transmission losses and their distribution are estimated by calibrating a model of flood travel along the channel. Figure 29 illustrates the



a) Wadi Najran catchment and alluvial aquifer.

b) Flood hydrograph at various points along alluvial plain



c) Calibration of flood routing model.

d) Distribution of average recharge.

Figure 29. Recharge from the Wadi Najran, Saudi Arabia, estimated by a flood routing model (FAO, 1981).

method for the Wadi Najran in Saudi Arabia. In this case a single well monitored flood wave was used to calibrate transmission loss rate and Manning's n in order to match recorded flood volumes. Both the historical 7 year river flow hydrograph and two 30 year synthetic sequences were routed through the model to estimate mean annual transmission loss as 86% of inflow at the upstream end of the aquifer.

Numerical modelling of groundwater systems with connected rivers

The ideas of potential and actual recharge are relevant to river recharge, especially for rivers which are connected to groundwater (Rushton, 1988). Groundwater conditions partly control recharge from connected rivers. As groundwater conditions change, for example with pumping from new boreholes, recharge will change. Estimates of historical average recharges would not be valid under such changed conditions.

Estimates of historical, i.e. actual, recharge are needed for calibrating groundwater models. On the other hand, for predictive modelling it is essential to know what the recharge would be if groundwater conditions changed; this quantity is clearly different from actual historical recharge, but may not be the same as potential recharge.

The best way to proceed is:

(i) estimate historical recharges by one of the methods discussed in this chapter;

(ii) calibrate the groundwater model treating the river as a specified flux boundary, that is using these recharges as inflows to the model;

(iii) once the aquifer section of the model is calibrated, alter the model to treat the river as a leaky boundary, that is where the model calculates recharge from both river and groundwater levels. Rushton and Tomlinson (1979) discuss possible relationships to use. The model must keep an account of the water left to flow downstream in the river to ensure that recharge does not exceed river flow;

(iv) calibrate the new river part of the model to reproduce the historical recharges;

(v) use the model in predictive mode. It will now estimate recharge for future groundwater conditions.

IRRIGATION

Importance of irrigation recharge

Irrigation schemes are frequently a major source of recharge to aquifers, whether they use groundwater or surface water. For example, Lerner *et al.* (1982) estimated that the alluvial fan aquifer under Lima, Peru received 20% of its recharge from irrigation losses. There have been many studies of groundwater in the Indus Basin where groundwater levels rose at least 20 m in the 30 years after irrigation began (Figure 30).

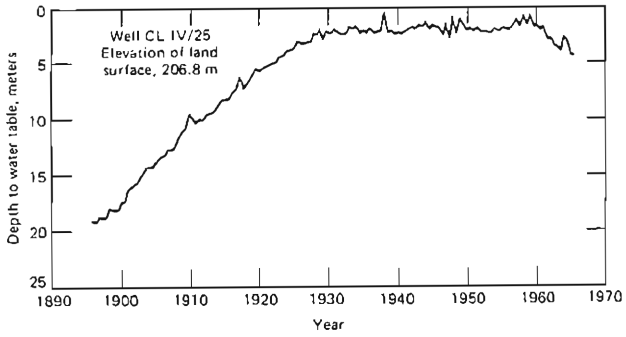


Figure 30. Depth to water for a well in the irrigated portion of the Indus river plain, Pakistan (Greenman, 1967).

Losses from irrigation systems are potential recharge, as defined earlier. They occur in two overlapping areas (a) canals and other delivery systems, (b) on fields. Canal losses have been studied widely, particularly in colonial situations and, more recently, in aid-funded projects.

Field losses will vary with the type of irrigation: flooded paddy, intermittently flooded field, furrow or spray irrigation. Drip irrigation is not considered here; if there are any losses they will probably be measured by the operators.

Recharge from irrigation schemes is sometimes underestimated because it is neglected. On the other hand, once irrigation has been recognised as a recharge source, the quantities are often overestimated because all losses are assumed to become recharge. The difference between potential and actual recharge has been discussed. WRRC (1980) consider the issue in more detail for irrigation, and Rushton (1990) gives a good example of the difference in the Mehsana aquifer in Gujarat, India, where the water table is 20 m below ground. Substantial losses of water were observed from the canal irrigation scheme which has recently spread over the aquifer, and groundwater levels were expected to rise in response. This has not occurred. Instead waterlogging is now seen along many canals; the losses have formed a perched water table on top of low conductivity layers in the aquifer, and are evaporating or being taken away by drains.

Water losses from irrigation canals

The loss of water from canals can amount to a significant percentage of the total irrigation supply. This often results in the waterlogging and consequent abandonment of surrounding arable land, as well as representing a substantial depletion of the water resource. Case studies reveal canal losses which when expressed as a loss over the wetted area of the bed, vary from 0.05 m/d to 1.5 m/d. Kraatz (1977) lists the estimated water losses from 15 canal systems around the world; seepage losses ranged from 3-50% of the water carried.

Total water loss from a canal is the sum of evaporative and seepage losses. Losses due to evaporation will probably be in the range of 1-10 mm/d. Seepage losses may either flow to shallow water tables and thence to drains and evaporation, or recharge deep groundwater systems.

As with most recharge calculations, there is not a perfect method for all, or even individual cases. Direct measurement has been used for seepage rates through canal beds. Empirical formulae are perhaps more widely used for this type of recharge than any other, but most formulae have been developed to estimate canal losses for design and operation purposes. Water balances are commonly and successfully used to estimate canal losses and there are a number of tracer techniques but these have not been widely used.

Seepage meters have been developed to measure canal losses. They involve a seepage bell or cylinder which is pushed into the sediment at the base of the canal and the infiltration rate measured by constant or falling head techniques (Figure 31).

254

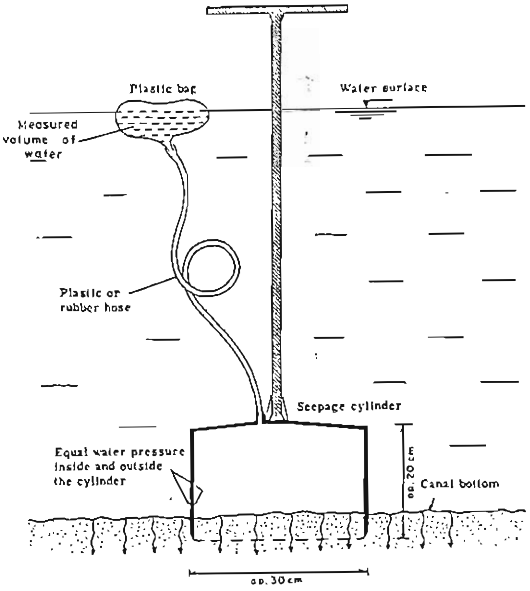


Figure 31. Constant head seepage meter for canal and stream beds (Kraatz, 1977).

There are many empirical formulae used to estimate canal losses. Few, if any, formulae are designed to estimate actual recharge. They are usually based on prolonged observation of canals within a region, and so are specific to that area.

Water balances on canals can be of two types, inflow-outflow or ponded. The former is similar to a river channel water balance, and is undertaken while the canal is flowing. If a section is closed off and ponded, a water balance can be more easily obtained from level measurements, but suffers from the effects of siltation.

Recharge from irrigated fields

Recharge by deep percolation from irrigated fields has many similarities with recharge by precipitation, and the same methods can often be used to estimate it. It is an areal process, water is applied at intervals, some may run off, the remainder becomes soil moisture where evaporation forces are applied before the excess becomes recharge. However there are differences between irrigation and recharge which may be important in individual cases:

(i) field canals may be present, carrying water at all times;

(ii) in flooded (paddy) irrigation, substantial depths of water are held on the surface leading to saturated conditions in at least the upper part of the soil profile;

(iii) soils are commonly puddled in flooded irrigation to reduce infiltration rates, however the bunds between fields may be a highly permeable pathway for deep percolation (Walker and Rushton, 1984).

Provided account is taken of such special circumstances, irrigated field recharge can be estimated by the same methods as used for precipitation recharge which were outlined earlier.

NET RECHARGE OVER A REGION

Introduction

The previous chapters have dealt with the various components of recharge to an aquifer, based on the origin of the water (precipitation, irrigation, etc.) and its route to groundwater. Alternative approaches are to estimate the total recharge. Such approaches seek to avoid describing any of the complex processes controlling recharge and groundwater movement, such as evaporative losses from the unsaturated zone, rivers with both influent and effluent reaches, and phreatophyte evapotranspiration.

There are four types of methods described in this chapter:

(i) Storage change or water table rise, i.e.

$$\text{total recharge} = \text{change in saturated volume} \times \text{specific yield} + \text{outflows}$$

(ii) Discharge from the aquifer, that is the sum of the measured discharge of groundwater across the boundary of a study area plus internal discharges by wells, etc. is equal to the net recharge inside the study area.

(iii) Inverse techniques, that is inverting the groundwater flow equation to solve for recharge instead of groundwater heads.

(iv) Tracer techniques, either estimating residence time of water in an aquifer, or modelling the distribution of several tracers throughout an aquifer.

All the methods have advantages and disadvantages as discussed below. Numbers (i), (ii) and (iv) are useful as checks on a water balance for the aquifer calculated by considering all the components of recharge, by the methods outlined above. The methods are only accurate for groundwater systems which are in hydraulic equilibrium.

Water table rise

The total recharge to an aquifer is often estimated from the volume of water stored as the water table rises during the wet season. The method clearly only applies to aquifers with a well defined recharge season. Allowances are made for pumping wells and other discharges as required, and recharge is estimated as:

$$r = (\delta s + \sum Q_A \, \delta t + V_D) \; / \; A \, \delta t$$

where r : total recharge per unit area (L)
δs : volume of water stored between lowest and highest water table positions (L^3)
Q_A : abstraction rate from wells pumping during the water table rise (L^3/T)
δt : the time interval between high and low water table positions (T)
A : area of the aquifer (L^2)
V_D : the volume of water discharged to springs, river bed seeps, etc. (L^3)

The method is appealing because of its simplicity and because it appears to be a direct measurement of total recharge.

There are plenty of reasons why it may not give a good answer, which need to be identified because the method is in such widespread use. Minor problems include (Figure 32):

- Seasonal abstractions inducing recharge from perennial water courses at some times of year,
- The common association of observation wells with abstraction zones, overestimating the degree of regional drawdown,
- Observation wells in deeper parts of the system showing confined responses but being interpreted as phreatic,
- Subjective contouring,
- Use of specific yield from pumping tests. Most experienced modellers recognise that short term pumping tests give gross underestimates of regional Sy values.

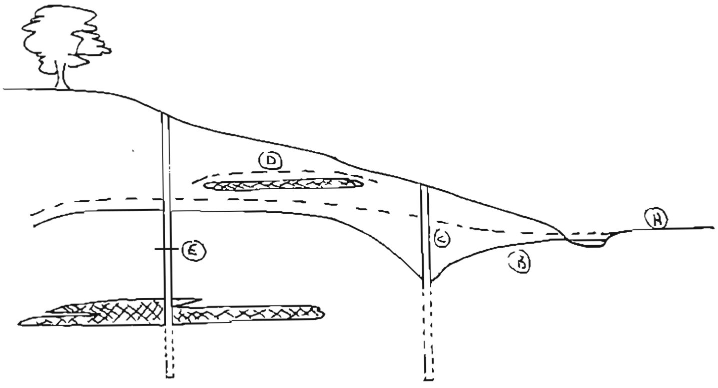


Figure 32. Total recharge by water table rise. A - riparian saturation, B - induced recharge, C - seasonal pumping, D - perched water tables, E - deep confined responses.

The most important problem is the assumption that recharge travels rapidly through the unsaturated zone. The water table peak then coincides with the end of recharge, which is a necessary but often invalid assumption. Recent regional studies of recharge in the Chalk aquifer in the Thames area of the UK suggest that up to 50% of recharge is still in transit in the unsaturated zone when the water table peaks (M. Lewis, British Geological Survey, pers. comm.). This water drains to the saturated zone over the summer at a slower rate than groundwater discharges to rivers. Similar effects are seen in Figure 23 where a thick unsaturated zone and perching layers hold up recharge. For this case involving river recharge, Wilson and De Cook (1968) found that only 33% of recharge arrived by the peak of the water table rise.

Hydrograph analysis for groundwater discharge

One method to estimate total recharge over a whole catchment relies on analysis of river hydrographs. The concept can be expressed as:

$$\text{recharge} = \text{baseflow} + \begin{array}{c}\text{withdrawals}\\ \text{and discharges}\end{array} + \begin{array}{c}\text{rate of storage}\\ \text{depletion}\end{array}$$

where all terms are averaged over a substantial time. Baseflow is taken to mean groundwater discharges to springs and rivers, and is estimated by hydrograph separation, as discussed below. Withdrawals by wells, discharges to lakes, seas and sabkahs, and evapotranspiration from the water table must be estimated and allowed for. Significant storage depletion will only occur in over exploited aquifers or during long droughts, and is more likely to be significant in arid lands.

Runoff processes, and stream-aquifer interactions, are not simple processes. A lengthy discussion is possible on the definition of baseflow and its relation to groundwater discharge - but it will not be conducted here. The following assumes that baseflow separation is a useful exercise, and discusses how it might be done.

It is straightforward to identify baseflow during the recession season, when soil moisture deficits are high enough to prevent further recharge, surface runoff is negligible, and all river flow is groundwater derived. The difficult part is during the winter recharge season, when surface runoff must be separate, and multiple events follow each other, masking each others' tails. There are three ways to approach the problem:

(1) An empirical approach, sketching in a baseflow separation subjectively or according to some simple rules. The Institute of Hydrology (1980) developed an automatic procedure based on five day minima which has the advantage of being routine and consistent, and the disadvantage of having no underlying conceptual model.

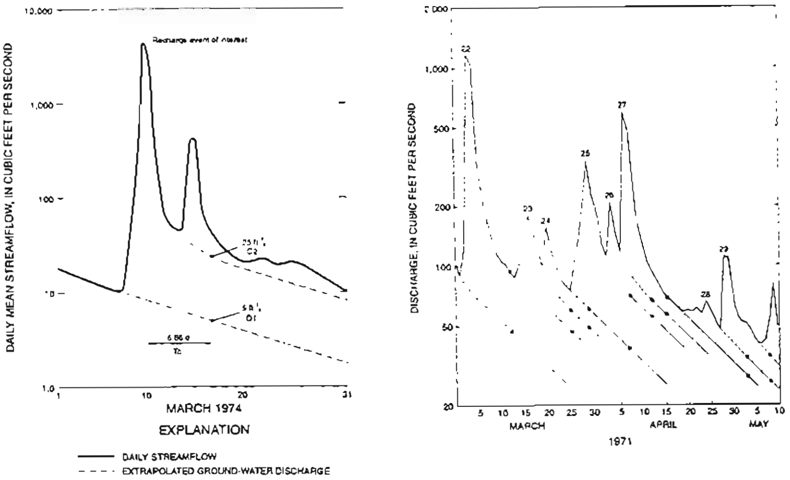
(2) Developing a groundwater rating curve, that is a graphical relationship between a representative groundwater level and baseflow. The rating curve is derived for recession periods when all flow is groundwater, and then used to estimate the groundwater component at other times of the year.

(3) A recession-curve-displacement method developed by Rutledge and Daniel (1994) from earlier analytical solutions by Roraborough (1964). In essence, the upward displacement of the recession by a recharge event is calculated and used to estimate the increase in groundwater volume (Figure 33). The method requires extrapolation of the recession and assumes, in common with most approaches, that delays in the unsaturated zone are unimportant.

Inverse techniques

The flow of groundwater can be represented by an equation that includes (i) groundwater heads, (ii) aquifer properties, (iii) recharges and boundary fluxes. Knowing the boundary conditions and two of these sets of data allows solution of the equation for the third. Conventional groundwater models solve for groundwater heads, inverse models solve for aquifer properties or, more unusually, recharge.

Inverse techniques to find aquifer properties were a popular research field in the late 1970's. Three problem occur with most methods; great sensitivity to errors in water level and recharge data, low sensitivity in unstressed areas of an aquifer, and non-uniqueness of solutions. These problems have meant that inverse techniques have had little practical application.



PROCEDURE
1. Compute recession index, K (32 d/log cycle)
2. Compute critical time, T_c (0.2144K, or 6.86 d)
3. Locate time that is 6.86 days after peak
4. Extrapolate pre-event recession to Q1 (5 ft³/s)
5. Extrapolate post-event recession to Q2 (23 ft³/s)
6. Compute total recharge by equation 6,

$$\frac{2\times(18\ ft^3/s)\times 32\ d}{2.3026}\times\frac{86400\ s}{1\ d}=4.32\times10^7\ ft^3$$

1971 water year storm-event number	Q2 (ft³/s)	Q1 (ft³/s)	Ground-water recharge ((ft³/s)×d)
22	93.0	47.0	1,800
23	59.0	47.0	470
24	60.0	49.0	430
25	72.0	38.5	1,311
26	66.0	56.0	470
27	70.0	57.0	509
28	34.5	28.0	254
29	35.0	26.0	352

Figure 33. Recession-curve-displacement method for estimating groundwater discharge. (a) Procedures for a single event, (b) example application to Big Alamance Creek, North Carolina (Rutledge and Daniel, 1994).

Few authors have attempted inverse techniques to estimate recharge. In general they can be expected to suffer from the same problems identified for aquifer property solutions above. Inverse models are probably best left to experienced modellers; anyone insisting on trying should use the MODFLOWP package (Hill, 1992).

Aquifer wide tracers

There are a number of ways that the concentrations of environmental tracers throughout an aquifer can be used to estimate net recharge. Although the methods are inter-related, it is useful to subdivide them into groups as follows:

(i) Analytical transit time models,

(ii) Distributed numerical models of single or multiple tracers.

Comprehensive reviews of mathematical transit time models are given by Maloszewski and Zuber (1982) and Zuber (1986). The models relate the mean age of water leaving a groundwater system to the mean transit time of a tracer through alternative mathematical models of the way the tracer moves through the system.

In order to use data resulting from a continuous, and possibly varying, input concentration, we need to derive the weighting function or impulse response of the system and use the convolution integral. The impulse response could theoretically be obtained by a tracer experiment, but it is not normally possible to apply an impulse of tracer and wait long enough to observe the complete output response. Therefore a number of conceptual models have been developed, and their impulse responses theoretically determined. Details can be obtained in the references cited above.

The method consists of the following:

(i) measuring input and output concentrations of tracer over a sufficiently long period, preferably at least as long as the expected turnover time,

(ii) choosing an appropriate groundwater flow and tracer sampling model, and so the form of impulse response,

(iii) solving the convolution equation for the numerical values of the parameters of the impulse response, including a, usually by trial and error comparison of measured and predicted output values,

(iv) estimating the volume of mobile groundwater (V) and hence the recharge rate (R) on the basis that mean age = V/R.

The limitations of the mathematical transit time models can sometimes be overcome by using distributed, numerical models of tracer movement. The aquifer is divided into cells and the movement of tracer and water between inputs, cells and outputs is modelled using mass conservation and assuming each cell is fully mixed. The various fluxes in the model are calibrated until the model predicts the same distribution of tracer as observed in the aquifer. All models that I know of use a mixing cell approach instead of a numerical solution to the solute transport equation. Examples include Campana and Simpson (1984) and Campana and Mahin (1985), who looked at a single tracer, and Adar and Neuman (1988) and Adar *et al.* (1988) who use a quadratic programming algorithm to find a weighted least squares solution for recharge to a semi-arid alluvial aquifer with multiple tracers.

Table 1. Factors for classifying recharge zones

Factor	**Example values or classification scheme**
Precipitation type	monsoon/thunderstorm/winter/summer/etc.
Precipitation amount	20% increase in pptn between zones.
Irrigation type	sprinkler/furrow/paddy/flood/canal/etc.
Irrigation amount	as precipitation.
Evaporation potential	20% increase in Penman potential evapotranspiration between zones. Solar or net radiation can be used in arid and semi-arid areas.
General lithology	alluvium/chalk/limestone/sand/etc.
Soil classification	Important factors are infiltration capacity, moisture storage and depth. Natural vegetation or drainage density may be adequate substitutes if no soil data available.
Land cover	grass/arable (crop type)/natural forest/plantation/phreatophyte/ urban/etc.
Landform	floodplain/rolling hills/plateau/etc.
Depth to groundwater or capillary fringe	> root depth/ within reach of plants/ within soil zone/ direct evaporation possible.

Notes

This classification is not intended to be exhaustive. On the other hand, it will often be unnecessary to use all classifying factors.

Data sources: topographic maps
geological and soils maps
Landsat and other satellite imagery
aerial photography
climate maps and meteorological stations
precipitation records
irrigation scheme maps and records
field visits!

REFERENCES

ADAR, E.M., NEUMAN, S.P. & WOOLHISER, D.A., 1988. Estimation of spatial recharge distribution using environmental isotopes and hydrochemical data. I. Mathematical model and application to synthetic data. J. Hydrol. 97, 251-277.

ADAR, E.M. & NEUMAN, S.P., 1988. Estimation of spatial recharge distribution using environmental isotopes and hydrochemical data. II. Application to Aravaipa Valley in southern Arizona, USA. J. Hydrol. 97, 251-277.

ALLEY, W.M., 1984. On the treatment of evapotranspiration, soil moisture accounting and aquifer recharge in monthly water balance models. Water Resour. Res. 20, 1137-1149.

ALLISON, G.B., 1988. A review of some of the physical, chemical and isotopic techniques available for estimating groundwater recharge. In: I. Simmers (ed), 49-72.

ALLISON, G.B. & HUGHES, M.W., 1983. The use of natural tracers as indicators of soil water movement in a temperate semi-arid region. J. Hydrol. 60, 157-173.

ATHAVALE, R.N. & RANGARAJAN, R., 1988. Natural recharge measurements in the hard-rock regions of semi-arid India using tritium injection - a review. In: I. Simmers (ed), 175-194.

BATH, A.H., DARLING, W.G. & BRUNSDEN, A.P., 1982. The stable isotopic compositon of infiltration moisture in the unsaturated zone of the English Chalk. Stable Isotopes (H.L.Schmidt *et al* (ed)), Elsevier, Amsterdam, 161-166.

BERNDTSSON, R. & LARSON, M., 1987. Spatial variability of infiltration in a semi-arid environment. J. Hydrol. 90, 117-133.

BESBES, M., DELHOMME, J.P. & DE MARSILY, G., 1978. Estimating recharge from ephemeral streams in arid regions: a case study of Kairouan, Tunisia. Water Resour. Res. 14, 281-290.

BURKHAM, D.E., 1970. A method for relating infiltration rates to streamflow rates in perched streams. USGS Prof. Paper 700-D, D266-D271.

CALDER, I.R., HARDING, R.H. & ROSIER, P.T.W., 1983. An objective assessment of soil-moisture deficit models. J. Hydrol. 60, 329-355.

CAMPANA, M.E. & MAHIN, D.A., 1985. Model-derived estimates of groundwater mean ages, recharge rates, effective porosities and storage in a limestone aquifer. J. Hydrol. 76, 247-264.

CAMPANA, M.E. & SIMPSON, E.S., 1984. Groundwater residence times and recharge rates using a discrete-state compartment model and 14C data. J. Hydrol. 72, 171-185.

CHANDRASEKHARAN, H., NEVADA, S.V., JAIN, S.K., RAO, S.M. & SINGH, Y.P., 1988. Studies on natural recharge to groundwater by isotope techniques in arid Western Rajasthan, India. In: I. Simmers (ed), 205-220.

COOPER, J.D., 1980. Measurement of moisture fluxes in unsaturated soil in Thetford Forest. Report No. 6, Inst. Hydrol., Wallingford.

DELCORE, M.R. & LARSON, G.J., 1987. Application of the tritium interface method for determining recharge rates to unconfined drift aquifers, II. Non-homogeneous case. J. Hydrol. 91, 73-81.

DILLON, P.J., 1982. Stream-aquifer interaction models: a review. Proc. of the Hydrology and Water Resources Symposium, Melbourne, May 1979, 132-138.

EDMUNDS, W.M. & WALTON, N.R.G., 1980. A geological and isotopic approach to recharge evaluation in semi-arid zones - past and present. In: Arid Zone Hydrology: Investigations with Isotope Techniques, 47-68. IAEA, Vienna.

FAO, 1981. Arid zone hydrology. Irrigation and Drainage paper 37, FAO, Rome, 271 pages.

FLUG, N., ABI-GHANEM, G.V. & DUCKSTEIN, L., 1980. An event-based model of recharge from an ephemeral stream. Water Resour. Res. 16, 685-690.

FREEZE, R.A. & BANNER, J., 1970. The mechanism of natural groundwater recharge and discharge, 2. Laboratory column experiments and field measurements. Water Resour. Res. 6, 138-155.

FREEZE, R.A. & CHERRY, J.A., 1979. Groundwater. Prentice-Hall Inc., New Jersey, 595 pages.

GELHAR, L.W., GROSS, G.W. & DUFFY, C.J., 1979. Stochastic methods of analysing groundwater recharge. In: Hydrology of areas of low precipitation. Proc. of the Canberra Symp., December 1979. IAHS-AISH Publ. No. 128, 313-321.

GREENMAN, D.W., SWARZENSKI, W.V. & BENNETT, C.D., 1967. Groundwater hydrology of the Punjab, West Pakistan, with emphasis on problems caused by canal irrigation. USGS Water Supply Paper 1608-H, 66 pages.

GRINDLEY, J., 1967. The estimation of soil moisture deficits. Meteorol. Mag. 96, 97-108.

GRINDLEY, J., 1969. The calculation of actual evaporation and soil moisture deficits over specified catchment areas. Meteorol. Off. Bracknell, Hydrol. Mem. no. 38, 3 pages.

HILL, M.C., 1992. A computer program (MODFLOWP) for estimating parameters of a transient, three dimensional, groundwater flow model using non-linear regression. USGS Open-File Report, 91-484.

HODNETT, M.G., & BELL, J.P., 1986. Soil moisture investigations of groundwater recharge through black cotton soils in Madhya Pradesh, India. Hydrological Sciences Journal, 31, 361-381.

HOWARD, K.W.F., & LLOYD, J.W., 1979. The sensitivity of parameters in the Penman evaporation equations and direct recharge balance. J. Hydrol. 41, 329-344.

IAEA (INTERNATIONAL ATOMIC ENERGY AGENCY), 1983. Guidebook on nuclear techniques in Hydrology. Technical Report Ser. No. 91, IAEA, Vienna. 439 pages.

JANSSON, P-E. & HALLDIN, S., 1980. Soil water and heat model. Technical description. Swedish Coniferous Forest Project, Uppsala. Tech. Rep. 26, 81 pages.

JOHANSSON, P-O., 1987. Estimation of groundwater recharge in sandy till with two different methods using groundwater level fluctuations. J. Hydrol. 90, 183-198.

JORDAN, P.R., 1977. Stream flow transmission losses in western Kansas. J. Hyd. Div. ASCE 103, HY8, 905-919.

KITCHING, R., EDMUNDS, W.M., SHEARER, T.R., WALTON, N.R.G. & JACOVIDES, J., 1980. Assessment of recharge to aquifers. Hydrol. Sci. Bull. 25, 217-235.

KITCHING, R. & SHEARER, T.R., 1982. Construction and operation of a large undisturbed lysimeter to measure recharge to the chalk aquifer, England. J. Hydrol. 58, 267-277.

KITCHING, R., SHEARER, T.R. & SHEDLOCK, S.L.., 1977. Recharge to the Bunter Sandstone determined from lysimeters. J. Hydrol. 33, 217-232.

KNOTT, J.F. & OLIMPIO, J.C., 1986. Estimation of recharge rates to the sand and gravel aquifer using environmental tritium, Nantucket Island, Massachusetts. USGS Water Supply Paper 2297, USGS, Federal Center, Box 25425, Denver, Co 80225, USA.

KRAATZ, D.B., 1977. Irrigation canal lining. FAO Land and Water Development Ser. No. 1, 199 pages.

LANE, L.J., DISKIN, M.H. & RENARD, K.G., 1971. Input-output relationships for an ephemeral stream channel system. J. Hydrol. 13, 22-40.

LARSON, G.J., DELCORE, M.R. & OFFER, S., 1987. Application of the tritium interface method for determining recharge rates to unconfined drift aquifers, I. Homogeneous case. J. Hydrol. 91, 59-72.

LERNER, D.N., ISSAR, S.A. & SIMMERS, I., 1990. Groundwater recharge - a guide to understanding and estimating natural recharge. Heise, Hannover, 345 pages.

LERNER, D.N., MANSELL-MOULLIN, M., DELLOW, D.J. & LLOYD, J.W., 1982. Groundwater studies in Lima, Peru. From: Optimal application of water resources, Proc. Exeter Symp. July 1982. IAHS Publ. No. 135, 17-30.

LLOYD, J.W., 1986. A review of aridity and groundwater. Hydrological Processes 1, 63-78.

LLOYD, J.W, PIKE, J.G., ECCLESTON, B.L. & CHIDLEY, T.R.E., 1987. The hydrogeoelogy of complex lens conditions in Qatar. J. Hydrol. 89, 239-258.

MALOSZEWSKI, P. & ZUBER, A., 1982. Determining the turnover time of groundwater systems with the aid of environmental tracers. 1. Models and their applicability. J. Hydrol. 57, 207-231.

NIELSEN, D.R., BIGGAR, J.W. & ERH, K.T., 1973. Spatial variability of field measured soil-water properties. Hilgardia 42, 215-259.

NORRIS, S.E., 1970. The effect of stream discharge on streambed leakage to a glacial outwash aquifer. USGS Prof. Paper 700-D, D262-D265.

PENMAN, H.L., 1950. The water balance of the Stour catchment area. J. Inst. Water Eng. 4, 457-469.

REHM, B.W., MORAN, S.R. & GROENEWOLD, G.H., 1982. Natural groundwater recharge in an upland area of central North Dakota, USA. J.Hydrol. 59, 293-314.

RORABAUGH, M.I., 1964. Estimating changes in bank storage and groundwater contribution to streamflows. IAHS Publ. 63, 432-441.

RUSHTON, K.R., 1988. Numerical and conceptual models for recharge estimation in arid and semi-arid zones. In: I. Simmers (ed), 223-238.

RUSHTON, K.R. 1990. Recharge in the Mehsana alluvial aquifer, India. Ch. 23 in Lerner *et al.* (1990).

RUSHTON, K.R. & TOMLINSON, L.M., 1979. Possible mechanisms for leakage between aquifers and rivers. J. Hydrol. 40, 49-65.

RUSHTON, K.R. & WARD, C.J., 1979. The estimation of groundwater recharge. J. Hydrol. 41, 345-361.

RUTLEDGE, A..T. & DANIEL, C.C., 1994. Testing an automated method to estimate groundwater recharge from streamflow records. Ground Water, 32, 180-189.

SAMMIS, T.W., EVANS, D.D. & WARWICK, A.W., 1982. Comparison of methods to estimate deep percolation rates. Water Resour. Bull. 18, 465-470.

SAXENA, R.K. & DRESSIE, Z., 1984. Estimation of groundwater recharge and moisture movement in sandy formations by tracing natural oxygen-18 and injected tritium profiles in the unsaturated zone. Isotope Hydrology 1983, Proc. Symp. Vienna, IAEA, 139-150.

SHARMA, M.H., CRESSWELL, I.D. & WATSON, J.D., 1985. Estimates of natural groundwater recharge from the depth distribution of an applied tracer. Proc. 21st Int. Assoc. Hydraulic Res., Melbourne, 65-70.

SHARMA, M.L. & HUGHES, M.W., 1985. Groundwater recharge estimation using chloride, deuterium and oxygen-18 profiles in the deep coastal sands of Western Australia. J. Hydrol. 81, 93-109.

SIMMERS, I. (ed), 1988. Estimation of natural groundwater recharge. NATO ASI Series C, Vol. 222 (Proc. of the NATO advanced research workshop, Antalya, Turkey, March 1987). D. Reidel Publ. Co., Dordrecht.

SINHA, B..P.C. & SHARMA, S.K., 1988. Natural groundwater recharge estimation methodologies in India. In: I. Simmers (ed), 301-311.

SMITH, D.B., WEARN, P.L., RICHARDS, H.J. & ROWE, P.C., 1970. Water movement in the unsaturated zone of high and low permeability strata by measuring natural tritium. In: Isotope Hydrology, IAEA, Vienna, 73-87.

THOMA, G. ESSER, N., SONNTAG, C., WEISS, W., RUDOLPH, J. & LEVEQUE, P., 1979. New technique of in-situ soil-moisture sampling for environmental isotope analysis applied at Pilat sand dune near Bordeaux. Isotope Hydrology 1978. Proc. Symp. Neuherberg, IAEA, Vol. 2, 753-766.

WALKER, S.H. & RUSHTON, K.R., 1984. Verification of lateral percolation losses from irrigated rice fields by a numerical model. J. Hydrol. 71, 335-351.

WARD, R.C. & ROBINSON, M., 1990. Principles of Hydrology (3rd ed), McGraw-Hill.

WATSON, Ph, SINCLAIR, P. & WAGGONER, R. 1976. Quantitative evaluation of a method for estimating recharge to the desert basins of Nevada. J. Hydrol. 31, 335-357.

WELLINGS, S.R., 1984. Recharge of the upper chalk aquifer at a site in Hampshire, England. 1. Water balance and unsaturated flow. J. Hydrol. 69, 259-273.

WILSON, L.G. & DE COOK, K.J., 1968. Field observations on changes in the subsurface water regime during influent seepage in the Santa Cruz River. Water Resour. Res. 4, 1219-1234.

WRRC (WATER RESOURCES RESEARCH CENTER), 1980. Regional recharge research for southwest alluvial basins. Final report on USGS contract 14-08-0001-18257. Dept. of Hydrology and Water Resources, Univ. of Arizona, Tuscon. 417 pages.

ZUBER, A., 1986. Mathematical models for the interpretation of environmental radio-isotopes in groundwater systems. In: Fritz, P. & Fontes, J-Ch (eds), Handbook of environmental geochemistry. Vol. 2. The terrestrial environment, B.

Atmospheric Environment Vol. 24B, No. 1, pp. 29–33, 1990.
Printed in Great Britain.

0957-1272/90 $3.00 + 0.00

GROUNDWATER RECHARGE IN URBAN AREAS

DAVID N. LERNER
Hydrogeology Research Group, School of Earth Sciences, University of Birmingham, P.O. Box 363, Birmingham B15 2TT, U.K.

(*First received 1 August 1988 and in final form 1 April 1989*)

Abstract—The two interlinked networks of hydrological pathways in urban areas are described with particular reference to the links with groundwater. As well as reducing direct recharge, urbanization creates new pathways and sources of water for recharge, including leaking water mains, sewers, septic tanks and soakaways. The net effect is often to increase recharge to pre-urbanization rates, or higher in dry climates and cities with high densities and large imported water supplies.

Key word index: Groundwater, recharge, urban, leaking pipes, sewers, irrigation, septic tanks, soakaways, water resources.

1. INTRODUCTION

Much of the surface of urban areas is rendered impermeable by buildings, roads and surface coverings. Because of this covering, the classical view of the effect of urbanization on groundwater is that recharge is reduced. For example,

> "... groundwater outflow ... decreases with urbanization, with direct runoff ... increasing." (Douglas, 1983),

and similarly from Lindh (1983)

> "Infiltration to the groundwater is markedly reduced.... with less water reaching the aquifer, wells may have to be deepened."

In fact urbanization alters all parts of the hydrological cycle so much that no simple analysis of the effects on groundwater is possible. Many of the alterations will increase recharge where there is permeable ground below the city. This has been recognized in principle by some writers, for example Gray and Foster (1972) and Monition (1977), but the increases have rarely been quantified. It has not been realized that urban recharge is often as high, or higher, than pre-urbanization rates (Lerner, 1986). This is due in part to the lack of appropriate data. Those concerned with water in urban areas (i.e. urban hydrologists, water supply engineers, sewage engineers, etc.) usually have little interest in urban groundwater except as a nuisance, and do not quantify the interactions of their systems with it.

This paper has two purposes. Firstly it is a review of the few data available on urban recharge, in order to show that it is not a single, simple process. Secondly it is a plea for urban recharge to be recognized as an area for research, and especially for the collection of suitable data.

2. URBAN HYDROLOGICAL PATHWAYS

Water does not often follow a cycle in the urban environment because it enters and leaves across the urban boundary. Rather it follows a network of pathways. There are two such networks of pathways in urban areas which are interlinked at many points. These are the (heavily modified) natural pathways and the water supply–sewage pathways.

The principal pathways of the natural network are precipitation, evapotranspiration, runoff, infiltration, recharge and groundwater flow. The principal paths of the water supply network are shown in Fig. 1, which particularly emphasizes the interconnections with groundwater. Boreholes bring groundwater into the network; a variety of paths including leakage from pipes, over-irrigation and septic tanks carry water down to the groundwater system.

The water supply network carries large flows. Table 1 compares, for a number of major cities, precipitation—usually the main input to the natural network—with imports of water and local groundwater abstraction, which are the inputs to the water supply network. In highly urbanized areas, e.g. Hong Kong, and in arid or semi-arid climates, e.g. Doha and Lima, flows in the water supply network exceed those in the natural system. Even in temperate, moderate density cities (Vancouver, Birmingham) flows in the two systems are comparable in size.

3. RECHARGE IN URBAN AREAS

3.1. *Recharge from the natural network*

Large urban areas often have a meso-climate which may alter rainfall and evapotranspiration rates (e.g. Shaw, 1988). These are probably second order effects on recharge when compared to the changes caused by

58

268

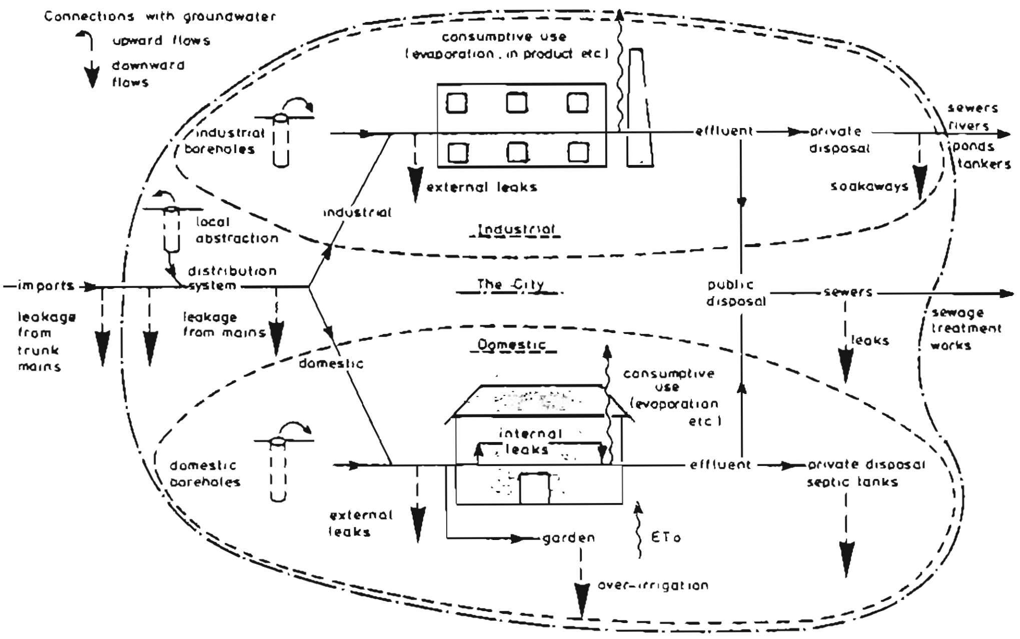


Fig. 1. Flows of water and sewage in a city.

Table 1. Relative sizes of inputs to the urban hydrological networks

City	Area[1] (km^2)	Date	Precipitation	Imports	Local g'water	Units
Urban Sweden[2]	4024	c 1970	701	235	nd[3]	mm
Mexico City[2]	nd[3]	1980	86	14	nd[3]	%
Hong Kong[2,4]	1046	1971	1912	1310	64	mm
Hong Kong[5,6]	0.61–0.35	1980	2000	650–7500[7]	0	mm
Sydney[2]	1035	1962–1971	1150	333	16	mm
Vancouver[2]	0.21	1982	1215	576	0	mm
Lima[8]	400	1978	10	1650[9]	950[9]	mm
Doha, Qatar[10]	294	1981–1982	167	175	27	mm
Birmingham[11]	500	1985	730	675	30[12]	mm

[1] Many of the areas are for a supply zone and include rural and semi-rural land.
[2] Grimmond and Oke, 1986.
[3] nd—No data given in source document.
[4] Whole city.
[5] Lerner, 1986; Geotechnical Control Office, 1982.
[6] Thirteen districts on Hong Kong Island.
[7] Assuming minimum night flows are 40% of supply.
[8] Lerner *et al.*, 1982; Wild and Ruiz, 1987.
[9] Includes agricultural use.
[10] Pencol, John Taylor and Sons, 1985.
[11] Written communications, Severn Trent Water Authority.
[12] Local groundwater only used by industry.

surface coverings which generally reduce infiltration, and increase and accelerate runoff. This runoff is often carried in storm sewers, drains, or other artificial waterways. Thus it is probable that direct recharge is reduced in urban areas. It should be noted in passing that permeable pavements are sometimes used to reduce runoff, and these will increase recharge by allowing infiltration while reducing vegetation cover and so reducing evapotranspiration. This was demonstrated by van Dam and van der Ven (1984). There is potential for recharge from storm sewers and drains, even when these are designed to carry water out of the city. Sewers are well known to leak and Lerner (1986) presents evidence that such leakage recharges groundwater in Hong Kong. In a water resources study of Liverpool, the University of Birmingham (1984) concluded that significant amounts of storm water leaked from sewers into the groundwater system, more than replacing the reduction in direct recharge due to impermeable cover. Storm water is often deliberately recharged. In the U.K. for example, soakaways are commonly used to dispose of runoff from domestic roofs, and for road runoff from some motorways and in some cities. As no evapotranspiration takes place from such water, the amount of recharge is probably higher than the direct recharge that would have occurred without urbanization. Recharge basins for storm water are used, for example on Long Island where they are thought to bring recharge up to pre-urbanization rates (Seaburn and Aronson, 1974). In arid climates there is often no provision for storm runoff, and the (rare) increased runoff from impermeable surfaces will infiltrate into the permeable surroundings.

3.2. *Recharge from the water supply network*

All water supply networks leak. Few authorities claim to be able to reduce leakage below 10% of supply, and rates of 50% have been reported. This source alone can generate a potential recharge of up to 3000 mm a^{-1}, and Lerner (1986) showed that some of this leakage did in fact recharge groundwater. He gave examples from Hong Kong and Lima, Peru, with several lines of evidence. Groundwater balances and flow modelling showed that additional sources of recharge were needed to sustain groundwater levels. There was hydrochemical evidence of mains water in groundwater, and in the Hong Kong study there was piezometric evidence under completely paved areas. More recently Lerner (1989a) presents informal evidence of the downward percolation of leaking mains water from a large underground Byzantine cistern. Rivett *et al.* (1989) report localized contamination of groundwater in Birmingham by trihalomethanes, and attributed it to leaking mains. Lerner (1986) concludes that recharge rates of 100–300 mm a^{-1} are common.

Sewers leak, as shown by the many examples of groundwater pollution below cities by sewage. Many urban areas do not have sewers, relying on septic tanks and soakaways to dispose of effluents—this water must recharge groundwater. For example in Bermuda, with an average rainfall of 1460 mm a^{-1}, water supply is from roof catchments and sewage is disposed of to septic tanks. These, and soakaways for storm drainage, increase recharge from 365 mm a^{-1} in rural areas to 575 mm a^{-1} in the urban area (Thomson and Foster, 1986).

The third main source of recharge from the water supply network is over-irrigation of parks and gardens. These are irrigated for aesthetic rather than commercial reasons. Application is often by unskilled labour. The amount of water applied rarely depends on plant water needs, but on the affluence of consumers, pricing policies for water supplies and, in the case of municipal parks, on bureaucratic procedures. For these reasons over-irrigation is normal with many multiples of the potential evapotranspiration being applied. Excess water percolates deep to recharge groundwater. An example is Doha in Qatar where over-irrigation has raised groundwater levels to the surface in many low lying areas (La Dell, 1986).

3.3. *Net effect of urbanization*

Figure 2 summarizes the changes in recharge that urbanization can cause. With so many changes, and with different conditions in every city, it is difficult to generalize about the net effect. However it is clear that there will always be man-made sources of recharge. In drier climates, or with large imported supplies, or with

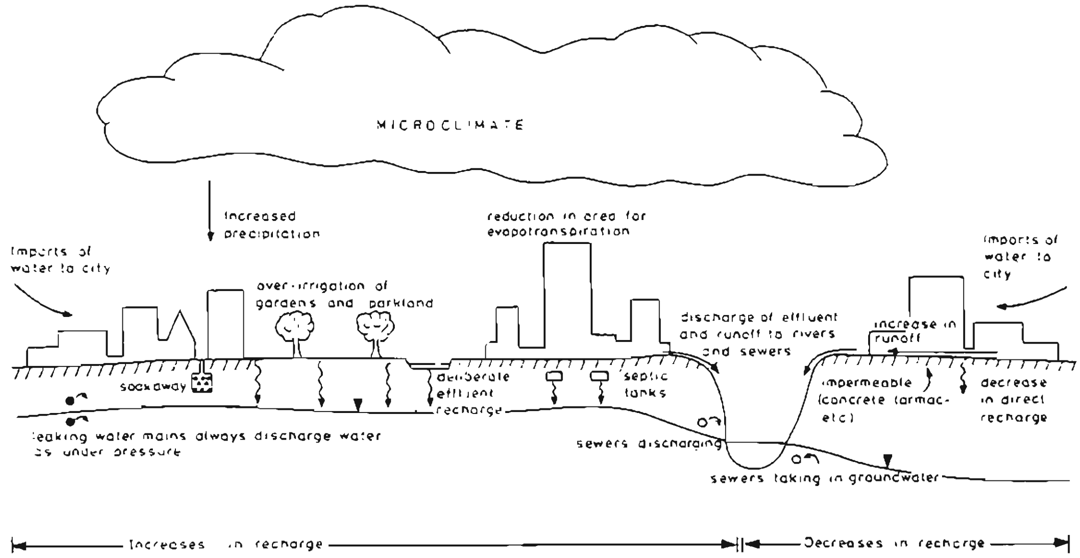


Fig. 2. Urban effects on groundwater recharge.

60

poor maintainance of piped systems, recharge in urban areas is likely to exceed that in rural areas. Table 2 gives some examples of water balances for urbanized aquifers.

3.4. *Urban groundwater quality*

There are many potential threats to groundwater quality in urban areas, including:

- —waste disposal activities,
- —industrial premises,
- —chemical stockpiles and storage tanks,
- —spillages during transport,
- —sewers, and
- —surface drainage, including roads.

Many of these threats have been realized, and there are numerous cases of contaminated groundwater beneath cities, of which just a few need be cited here. The Triassic sandstone aquifer beneath Birmingham is widely polluted by chlorinated hydrocarbon solvents (Rivett *et al.*, 1989) as is the alluvial aquifer beneath Milan, Italy (Cavallaro *et al.*, 1986). Examples of microbiologically polluted groundwater are given by Galbraith *et al.* (1987), while Eisen and Anderson (1980) showed that Milwaukee's groundwater is contaminated by Cl^- and SO_4^{2-}. Bermuda's limestone aquifer is contaminated with NO_3^- and some microbiological organisms (Thomson and Foster, 1986).

4. ESTIMATING RECHARGE IN URBAN AREAS

4.1. *Scientific studies*

This review has shown that there are few accurate data on the effects of urbanization on recharge, either for individual components or overall, but that such data are needed because there are significant effects to quantify. Those familiar with urban environments will realize that collecting accurate data on such episodic (in time and space) phenomena will be difficult. As just one example, consider the problem of leaking storm sewers. Water balances will be of especially low accuracy during storms. Methods based on measuring infiltration, soil moisture or pressure, or tracer concentrations require knowledge of the leak location; more importantly they need a method to move from point measurement to a more regional scale.

Table 2. Example water balances of urbanized aquifers (10^3 m^3 day^{-1})

City	Lima*	Doha†	Bermuda‡
Area (km^2)	400	294	6.3
Recharge from:			
precipitation	0	11.5	4.83
rivers	280	0	0
agricultural irrigation	390	0	0
park irrigation		37.6	0
leaking mains	340	25.3	0
sewers and septic tanks	0?	17.6	3.13
soakaways	0	-	-

* Lerner *et al.*, 1982; Wild and Ruiz, 1987.
† La Dell, 1986.
‡ Thomson and Foster, 1986.

4.2. *Water resource and other engineering studies*

The multiplicity of recharge sources and complexity of landuse and surface cover, make it extremely difficult to estimate recharge in urban areas for engineering studies. For example, consider the possible sources of recharge from the water supply network:

recharge = leakage from water mains
+ external losses from consumers' properties
+ leakage from sewers
+ flow to septic tanks
+ deep percolation from domestic irrigation
+ deep percolation from municipal irrigation. (1)

Considering all these components individually, with their associated errors, can lead to a large accumulated error in the recharge estimate. To reduce errors, it is preferable to consider the overall balance of the network and estimate net recharge as follows:

net recharge = imports of water
+ local abstractions of groundwater
− consumptive use
− effluent leaving area. (2)

Methods of using both equations are discussed by Lerner (1989b).

Estimating recharge from the natural network presents even greater problems, the most difficult of which is losses from storm sewers. All studies of leaking sewers have been of those which collect water, not lose it, because this is an important factor in their design. Except in research studies, water balance over individual storms are unlikely to be accurate enough because of the uncertainties in both measurement of precipitation and runoff, and in estimates of detention and evaporation. Field tests in combined sewers at times of low flow, or injecting and tracking known flows or tracers, or analysis of groundwater responses may provide more accurate estimates for small parts of a city, but are impractical for the whole system.

The best approach will be to set upper and lower bounds on recharge by water balance methods. These estimates can then be refined by calibration of a groundwater flow model against groundwater responses and outflows.

5. CONCLUSIONS

This review argues that there are more sources and pathways of groundwater recharge in urban than rural areas. The large imports of water to cities give rise to large recharges from leaking pipes, and storm water is often partly diverted to groundwater, whether intentionally or otherwise. Despite the increased impermeable area, the net result is that total urban recharge is often at least as large as would occur without the city. Groundwater quality is often poor beneath cities where there is no geological (low permeability) protection. Few reliable data are available on urban recharge, either for the individual components or the total, and further scientific and engineering studies are needed.

Acknowledgements—An earlier version of this paper was presented to UNESCO's International Symposium on Hydrological Processes and Water Management in Urban Areas held in Duisburg, F.R.G., during April 1988.

REFERENCES

Cavallaro A., Corradi C., De Felice G. and Grassi P. (1986) Underground water pollution in Milan and the Province by industrial chlorinated organic compounds. In *Effects of Land Use on Fresh Waters* (edited by J. F. de L. G. Solbe), pp. 68–84. Ellis Horwood, Chicester.

Douglas I. (1983) *The Urban Environment*. Edward Arnold, London.

Eisen C. and Anderson M. P. (1980) The effects of urbanization on groundwater quality, Milwaukee, Wisconsin, U.S.A. In *Aquifer Contamination and Protection* (edited by R. E. Jackson), pp. 378–390. SRH no. 30, UNESCO Press, Paris.

Galbraith N. S., Barrett N. J. and Stanwell-Smith R. (1987) Water and disease after Croydon: a review of water borne and water associated disease in the U.K. 1937–1986. *J. Inst. Wat. Envir. Manag.* 1(1), 7–21.

Geotechnical Control Office (1982) Mid-levels study: report on geology, hydrology and soil properties. Government Printer, Hong Kong.

Grimmond C. S. B. and Oke T. R. (1986) Urban water balance—2. Results from a suburb of Vancouver, British Colombia. *Water Resources Res.* **22**, 1404–1412.

Gray D. A. and Foster S. S. D. (1972) Urban influences on groundwater conditions in the Thames flood plain deposits of central London. *Phil. Trans. Royal Soc. London* **A272**, 135–168.

Hogland W. and Niemczynowicz J. (1986) The unit superstructure—a new construction to prevent groundwater depletion. *IAHS Publ.* **156**, 513–522.

La. Dell T. (1986) Rising groundwater in Doha, Qatar. *Geolog. Soc. Lond. News.* **15**, (Abstract).

Lerner D. N. (1986) Leaking pipes recharge groundwater. *Ground Water* **24(5)**, 654–662.

Lerner D. N. (1989a) Large Byzantine water storage cisterns. *Q. J. Eng. Geol.* **22**.

Lerner D. N. (1989b) Recharge due to urbanization, chapter 15. In *Groundwater Recharge Estimation, a Guidebook to Practice*. Heise, Hannover, F.R.G.

Lerner D. N., Mansell-Moullin M., Dellow D. J. and Lloyd J. W. (1982) Groundwater studies for Lima, Peru. *IAHS Publ.* **135**, 17–30.

Lindh G. (1983) *Water and the City*. UNESCO Press, Paris.

Monition L. (1977) Effects de l'urbanization sur leas eaux souterrraines. Proc. Amsterdam Symp. IAHS/UNESCO, pp. 162–166.

Pencol J. *et al.* (1985) Doha stormwater and groundwater management action report. Ministry of Public Works, State of Qatar.

Rivett M. O., Lerner D. N., Lloyd J. W. and Clark L. (1989) Organic contamination of the Birmingham aquifer. Report PRS 2064-M, Water Research Centre, Marlow, U.K.

Seaburn G. E., Aronson D. A. (1974) Influence of recharge basins on the hydrology of Nassau and Suffolk Counties, Long Island, New York. USGS Water Supply paper 2031.

Shaw E. M. (1988) *Hydrology in Practice* (2nd edition). Van Nostrand Reinhold (International), London.

Thomson J. A. M., Foster S. S. D. (1986) Effect of urbanization on groundwater of limestone islands: an analysis of the Bermuda case. *J. Inst. Wat. Eng. Sci.* **40(6)**, 527–540.

University of Birmingham (1984) Saline groundwater investigation. Phase 2—North Merseyside. Final report to the North West Water Authority (unpublished).

Van Dam C. H. and Van de Ven F. H. M. (1984) Infiltration in the pavement. *Proc. 3rd Int. Conf. on Urban Storm Drainage*, Gothenburg, Sweden **3**, 1221–1230.

Wild J. and Ruiz J.-C. (1987) Lima groundwater modelling revisited. Presented to National Hydrology Symposium, University of Hull, U.K., Sept 1987.

Inorganic aspects of catchment mass balance

James I. Drever
University of Wyoming

THE CONCEPT OF MASS BALANCE

The concept of mass balance is deceptively simple. It is simply a budget that describes fluxes into and out from a catchment, and assigns the solutes to specific sources and sinks such as atmospheric deposition, biomass change, or bedrock weathering. Although the concept is simple, identifying and quantifying to the individual fluxes is not easy, and the final budget often has ambiguities and uncertainties.

Budgets can be constructed incorporating different levels of detail. As an example, a geochemist would generally view "biomass uptake" as a single process, whereas he/she would probably try to resolve "chemical weathering" into a detailed set of mineral reactions. A biologist, on the other hand, would tend to regard "chemical weathering" as a single process but subdivide the biomass into different compartments.

A simple view of some processes occurring in a soil (or a catchment) is shown in Figure 1.

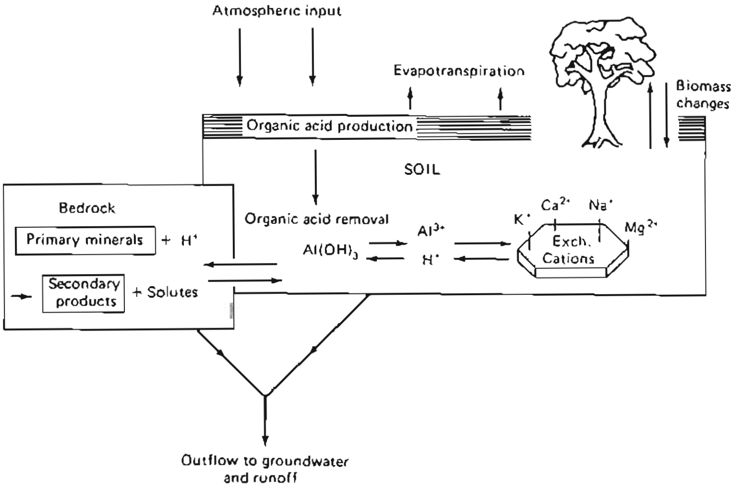


Figure 1: Some processes affecting solute outputs from a catchment (after Drever, 1988)

For this system, we can write a mass balance equation

$$\text{[solutes in outflow]} = \text{[solutes from atmosphere]} + \text{[solutes from weathering]} \pm \text{[solutes from change in biomass]} \pm \text{[change in exchange pool]} \qquad (1)$$

In writing this equation we have already made the assumption that our time-scale is long enough that changes in water storage can be ignored. The question of time-scale is very important in mass-balance studies: it is rarely possible to construct a meaningful catchment budget for a time-scale of less than a year. Let us consider the individual terms in the above equation.

Solutes in outflow

Questions of gauging and analysis are discussed elsewhere in this book. The information that is needed is the volume of water leaving the catchment and its chemical composition. The main assumption that is made is that the gauged outflow represents all the water leaving the catchment and that discharge via groundwater is negligible. It is difficult to test this assumption. One approach is test the input-output budget for chloride balance, assuming that there are no sources or sinks for chloride in the catchment. The problems with this approach may include uncertainties as to whether precipitation gauges accurately measure the amount of incoming precipitation (blowing snow can be a problem), a highly variable chloride concentration in precipitation (making it difficult to quantify precisely the annual input), and unknown dry deposition of chloride (Probst et al., 1990). The assumption of no sources or sinks for chloride on a relatively short time-scale may also be questionable.

There may also be practical problems with measuring outflow in cold environments. The greatest water flux and solute flux occur at the time of snowmelt, a time when access may be difficult and instrumentation may be affected by intermittent freezing.

Solutes from the atmosphere

Again, this topic is discussed elsewhere in this book. Atmospheric deposition occurs in the form of rain and snowfall, condensation from mist or fog, and dry deposition. Dry deposition in turn consists of solid particles and gases such as SO_2 and NO_x that are taken up directly by vegetation or moist foliage. Dry deposition is difficult to quantify; however, the main elements involved are nitrogen and sulfur, which are generally not directly involved in weathering budgets.

Changes in the exchange pool

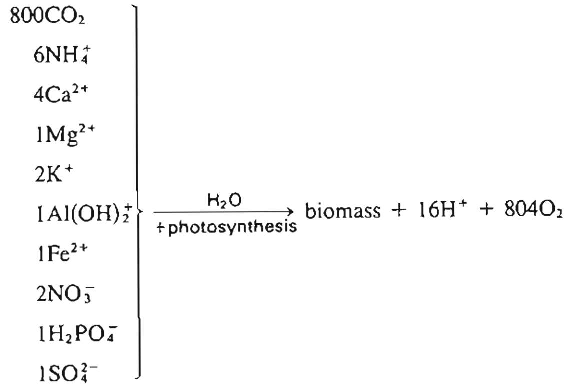
Soils contain exchangeable cations and anions that are in equilibrium with the soil solution. As the composition of the soil solution changes, ions will be exchanged between the solid phase and solution. If the soil solution composition does not change with time, adsorbed ions will not change either and ion exchange will make no contribution to the solute budget. In the short term, changes in solution composition occur as a result of precipitation events, evaporation, and growth cycles of plants. It is often assumed that these changes average out over an annual cycle so that, on a time scale longer than a year they cause no net change to the exchange pool and hence no net contribution to solute flux. This assumes that year-to year variations in solution chemistry (caused, for example by alternating wet and drought years) are insignificant.

If there is a permanent change in the chemistry of atmospheric input, this will cause a unidirectional change in the exchange pool. If, for example, the concentration of H^+ in precipitation increased, the H^+ (or Al^{3+} generated from dissolution of $Al(OH)_3$ in response to the input of H^+) would displace other cations (commonly mostly Ca^{2+}) from exchange sites, and these cations would appear in the outflow from the catchment (e.g. Reuss and Johnson, 1986). If the composition of precipitation remained constant, the composition of the exchange pool would gradually adjust to the new input and a new steady state, in which there was no net change in the exchange pool over time, would be established. The time taken by the exchange pool to adjust to a change in atmospheric input is generally quite long—on a time-scale of decades—because the reservoir of exchangeable ions in the soil is generally large compared to the annual input from the atmosphere.

It is very difficult to determine whether cations in the output from a catchment are derived from exchange sites or from mineral weathering. If the catchment does not receive a large input of anthropogenic acidity from the atmosphere, it is generally assumed that the exchange pool is in steady state, and is not a net contributor of solutes. In catchments that are influenced by acid deposition, one of the major tasks of modeling is to distinguish between cations from weathering and cations displaced from exchange sites.

Changes in biomass

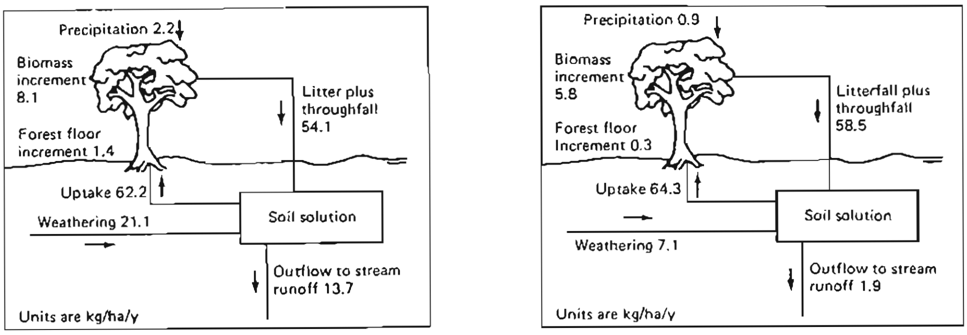
As plants grow, they extract inorganic nutrients from the soil solution and incorporate them into plant tissue. The stoichiometry is approximately (Schnoor and Stumm, 1985)



$$\left.\begin{array}{l} 800CO_2 \\ 6NH_4^+ \\ 4Ca^{2+} \\ 1Mg^{2+} \\ 2K^+ \\ 1Al(OH)_2^+ \\ 1Fe^{2+} \\ 2NO_3^- \\ 1H_2PO_4^- \\ 1SO_4^{2-} \end{array}\right\} \xrightarrow[+\text{photosynthesis}]{H_2O} \text{biomass} + 16H^+ + 804O_2$$

Plant growth affects the budgets of the major cations and protons and will affect both the exchange pool and the net output from a catchment. When plants die and decompose, the process is reversed and the elements are returned to the soil. This topic is covered in more detail in another chapter.

If a forest is in steady state, that is to say the growth of new vegetation is exactly balanced by the death and decay of old vegetation, the biomass will be neither a source nor a sink in the mass balance equation. In forested catchments, however, the biomass is rarely in a steady state. Even without human intervention such as tree cutting or planting, forests typically go through cycles of gradual biomass increase interrupted by catastrophic events, such as fire or disease, that result in rapid biomass loss. It is only on a very long time scale, a scale of centuries, that the biomass of a forest can be considered in steady state. Hubbard Brook, New Hampshire, is an example of a catchment where the biomass term is large compared to the weathering term (Figure 2). According to the data of Likens et al. (1977), the net uptake of Ca in the form of biomass increment and forest floor (organic) increment was 45% of the



(a) (b)

Figure 2: Calcium (a) and Potassium (b) fluxes at Hubbard Brook Experimental Forest, New Hampshire, USA. Data from Likens et al., 1977; after Drever, 1988).

amount released by weathering. For potassium, the biomass increment was 86% of the amount released by weathering: the net outflow of K from the catchment (runoff - precipitation) was only 14% of the K calculated to be released by chemical weathering. More recent measurements have decreased the biomass uptake term relative to the weathering term, but the conclusion remains: in forested catchments with silicate bedrocks, the biomass uptake term is likely to be important.

The biomass term in Eq. (1) can be measured directly by measuring the rate of growth of trees and the chemical composition of different tissues. In principle, it can be determined by difference if the other terms in Eq. (1) are known (Velbel, 1986). This approach is rarely useful because the other terms can rarely be measured independently with sufficient accuracy.

Chemical weathering

There is no simple way of measuring the contribution of chemical weathering to the solutes measured at the output of a catchment. The most common approach is to assign values to (or simply ignore) the biomass change and ion exchange terms in Eq. (1) and determine the weathering term by difference. The utility of this approach depends very much on the catchment. In catchments where the bedrock weathers rapidly, the weathering term is much larger than the other two, so there is not a problem. Where atmospheric deposition has not been significantly affected by pollution, neglecting the ion exchange term should not introduce much error, and in catchments that are largely unvegetated, the biomass term should be negligible. Unfortunately, the catchments that are of greatest interest, particularly in the Nordic countries, are those that are characterized by low weathering rates, forests (often complicated by changes in land use in historic times), and significant anthropogenic inputs of acid (Wright, 1988). In these situations, the weathering term may be small compared to the other terms in Eq. (1) and it cannot be determined with much confidence.

An alternative approach is look at the budget of an element that is not significantly affected by ion exchange or biomass uptake, such as sodium (Stauffer and Wittchen, 1991). If the flux of sodium from weathering is known (output – atmospheric input), and the stoichiometry of weathering is known, then the fluxes of the other cations, notably Ca^{2+} and Mg^{2+}, can be calculated by multiplying the sodium flux by the appropriate stoichiometric ratio. There are two problems with this approach: in coastal areas with low weathering rates, the flux of Na from weathering may be small compared to the flux from the atmosphere, making it difficult to measure accurately. The other problem is that there is no easy way of measuring the stoichiometry of weathering. Silica can be used instead of sodium. The advantage is that the silica budget can be established with more certainty than that of sodium (usually the atmospheric input is negligible and the flux from weathering relatively large). The disadvantage is that silica is affected by adsorption-desorption reactions and plant uptake.

In some catchments, the isotopic composition of strontium on exchange sites is different from that of bedrock minerals (Miller et al., 1993). This can be used to determine the proportion of the Sr in catchment outflow that is derived from each of the two sources. If it is assumed that Ca^{2+} behaves exactly as Sr^{2+}, then the Ca^{2+} in the outflow can be assigned to the two sources in the same ratio as for Sr. Since Ca^{2+} is the most important cation released by weathering, establishing a budget for Ca^{2+} is almost equivalent to establishing a budget for weathering as a whole.

MASS BALANCE AND MINERAL WEATHERING

The Sierra Nevada of California, USA

The classic example of the use of mass balance to characterize weathering reactions was by Garrels and Mackenzie (1967). They interpreted the mean compositions of perennial and ephemeral springs on the east side of the Sierra Nevada of California in terms of atmospheric input and bedrock weathering. Ion exchange and biomass uptake were ignored. The compositions of the springs is shown in Table 1 and Figure 3.

Table 1: Mean values for compositions of springs of the Sierra Nevada (after Garrels and Mackenzie, 1967)

	Ephemeral Springs		*Perennial Springs*	
	ppm	*molality × 10^4*	*ppm*	*molality × 10^4*
SiO_2	16.4	2.73	24.6	4.1
Al	0.03	—	0.018	—
Fe	0.03	—	0.031	—
Ca	3.11	0.78	10.4	2.6
Mg	0.70	0.29	1.70	0.71
Na	3.03	1.34	5.95	2.59
K	1.09	0.28	1.57	0.40
HCO_3	20.0	3.28	54.6	8.95
SO_4	1.00	0.10	2.38	0.25
Cl	0.50	0.14	1.06	0.30
F	0.07	—	0.09	—
NO_3	0.02	—	0.28	—
Dissolved solids	36.0		75.0	
pH	6.2[a]		6.8[a]	

[a] Median value.

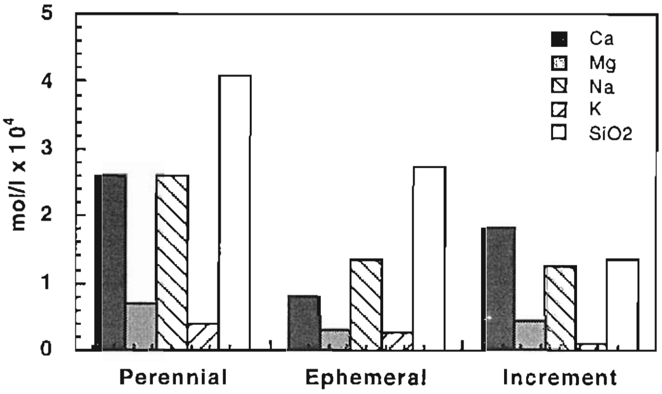


Figure 3: Major solutes in springs from the Sierra Nevada (data from Garrels and Mackenzie, 1967)

The two sets of springs were considered separately: the calculation procedure for the ephemeral (flowing part of the year only) springs was as follows (Table 2):

1. The atmospheric input (which is minor) was subtracted.
2. All the remaining Na and Ca were ascribed to the weathering of plagioclase feldspar. The Na/Ca ratio of the feldspar chosen was equal to the Na/Ca ratio in the water, which was reasonably consistent with the composition of the feldspar in the bedrock. The corresponding amounts of silica and bicarbonate were calculated.
3. All the Mg was ascribed to the weathering of biotite (actually end-member phlogopite) to kaolinite. The corresponding amounts of K, silica and bicarbonate were calculated.
4. The remaining K (which was a small amount) was ascribed to the weathering of K-feldspar to kaolinite. The corresponding amounts of silica and bicarbonate were calculated.

These three reactions could "explain" all the major solutes in the water (the minor amount of silica remaining was within the precision of the analyses). Bicarbonate balance is inevitable from charge balance, but there is no such constraint on silica. From the excellent balance obtained, several conclusions were drawn:

1. The reactions chosen were a reasonable description of what was occurring in nature.
2. Because the three reactions accounted for all the observed silica in solution, silica was derived from the breakdown of silicate minerals and not from dissolution of quartz.
3. About 80% of the rock-derived solutes came from the dissolution of plagioclase alone. Even though K-feldspar was abundant in the rocks, little dissolution occurred. This illustrates the selectivity of weathering discussed in Chapter **: more reactive minerals contribute disproportionately to the solute flux.

Table 2: Mass balance calculation for the ephemeral springs of the Sierra Nevada (after Garrels and Mackenzie, 1967).

Reaction (coefficients × 10^4)	Na^+	Ca^{2+}	Mg^{2+}	K^+	HCO_3^-	SO_4^{2-}	Cl^-	SiO_2	*Products (mol/l × 10^4)*
Initial concentrations in spring water	1.34	0.78	0.29	0.28	3.28	0.10	0.14	2.73	
Minus concentrations in snow water	1.10	0.68	0.22	0.20	3.10	—	—	2.70	Derived from rock
Change kaolinite back into plagioclase Kaolinite				Minus plagioclase					
1.23 $Al_2Si_2O_5(OH)_4$ + 1.10 Na^+ + 0.68 Ca^{2+} + 2.44 HCO_3^- + 2.20 SiO_2 = Plagioclase 1.77 $Na_{0.62}Ca_{0.38}Al_{1.38}Si_{2.62}O_8$ + 2.44 CO_2 + 3.67 H_2O	0.00	0.00	0.22	0.20	0.64	0.00	0.00	0.50	1.77 $Na_{0.62}Ca_{0.38}$ feldspar
Change kaolinite back into biotite Kaolinite				Minus biotite					
0.037 $Al_2Si_2O_5(OH)_4$ + 0.073 K^+ + 0.22 Mg^{2+} + 0.15 SiO_2 + 0.51 HCO_3^- = 0.073 Biotite $KMg_3AlSi_3O_{10}(OH)_2$ + 0.51 CO_2 + 0.26 H_2O	0.00	0.00	0.00	0.13	0.13	0.00	0.00	0.35	0.073 biotite
Change kaolinite back into K-feldspar				Minus K-feldspar					
0.065 $Al_2Si_2O_5(OH)_4$ + 0.13 K^+ + 0.13 HCO_3^- + 0.26 SiO_2 = K-feldspar 0.13 $KAlSi_3O_8$ + 0.13 CO_2 + 0.195 H_2O	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.12	0.13 K-feldspar

[a] Concentrations are in mol/l × 10^4.

Perennial springs

The perennial (flowing year-round) springs (Table 1, Figure 2) represent water that has circulated deeper through the bedrock and had more time to react with minerals. Garrels and Mackenzie made the assumption that when the water first infiltrated, it acquired the composition represented by the ephemeral springs, and then the deeper circulation provided additional solutes. The "increment", the difference between compositions of the ephemeral and perennial springs represents the solutes acquired by deeper circulation. Chemically, the increment is not the same as the ephemeral springs (Figure 3), which means that different weathering reactions must determine its composition. The ratio of Ca to Na is higher, and the ratio of silica to Na lower. In calculating a budget, Garrels and Mackenzie assumed the same feldspar composition as that used to interpret the ephemeral springs. [An alternative approach might have been to assume that the deeper waters interacted with a more calcic plagioclase, but there was no geological reason to suggest this.] The lower silica to sodium ratio implied that more silica was retained in the solid phase than was the case for the ephemeral springs. This was interpreted to mean that a secondary phase with a higher silica/Al ratio than kaolinite was being formed as a secondary phase. The phase hypothesized was a beidellite (montmorillonite in the original terminology of Garrels and Mackenzie), which was consistent with the idea (Chapter **, Fig. 3) that more concentrated solutions would be in equilibrium with a smectite rather than kaolinite. A reaction was then written (Table 3) for plagioclase altering to a mixture of kaolinite and beidellite, and the relative proportion of the two phases was adjusted to give exact balance for silica. This left the problem of excess calcium, which was attributed to "minor amounts of carbonate encountered en route." The balance for the perennial springs is less satisfying than that for the ephemeral springs. Is a smectite really forming, and what is its chemical composition? Is calcite really present, or does the excess

Table 3: Mass balance calculation for the perennial springs of the Sierra Nevada (after Garrels and Mackenzie, 1967).

Reaction (coefficients × 10^4)	*Na^+*	*Ca^{2+}*	*Mg^{2+}*	*K^+*	*HCO_3^-*	*SO_4^{2-}*	*Cl^-*	*SiO_2*	*Mineral altered and product (mol/l × 10^4)*
Initial concentrations (perennial minus ephemeral)	1.25	1.82	0.42	0.12	5.67	0.15	0.16	1.37	
Remove $Ca^{2+} = SO_4^{2-}$, and $Na^+ = Cl^-$	1.09	1.67	0.42	0.12	5.67	0.00	0.00	1.37	
Adjust HCO_3^- = total electrical charge of cations	1.09	1.67	0.42	0.12	5.39	0.00	0.00	1.37	
Change kaolinite back into biotite									
$0.07\ Al_2Si_2O_5(OH)_4 + 0.42\ Mg^{2+} + 0.14\ K^+ + 0.28\ SiO_2 + 0.98\ HCO_3^- = 0.14\ KMg_3AlSi_3O_{10}(OH)_2 + 0.98CO_2 + 0.49H_2O$	1.09	1.67	0.00	0.02	4.41	0.00	0.00	1.09	0.14 biotite 0.07 kaolinite
Change kaolinite back into plagioclase									
$0.26\ Al_2Si_2O_5(OH)_4 + 0.235\ Na^+ + 0.144\ Ca^{2+} + 0.47\ SiO_2 + 0.52\ HCO_3^- = 0.38\ Na_{0.62}Ca_{0.38}Al_{1.38}Si_{2.62}O_8 + 0.52\ CO_2 + 0.78\ H_2O$	0.85	1.53	0.00	0.00	3.89	0.00	0.00	0.62	0.38 plagioclase 0.26 kaolinite
Change smectite back into plagioclase									
$0.81\ Ca_{0.17}Al_{2.33}Si_{3.67}O_{10}(OH)_2 + 0.85\ Na^+ + 0.38\ Ca^{2+} + 0.61\ SiO_2 + 1.62\ HCO_3^- = 1.37\ Na_{0.62}Ca_{0.38}Al_{1.38}Si_{2.62}O_8 + 1.62\ CO_2 + 1.62\ H_2O$	0.00	1.15	0.00	0.00	2.27	0.00	0.00	0.01	1.37 plagioclase 0.81 smectite
Precipitate $CaCO_3$									
$1.15\ Ca^{2+} + 2.30\ HCO_3^- = 1.15\ CaCO_3 + 1.15\ CO_2 + 1.15\ H_2O$	0.00	0.00	0.00	0.00	−0.03	0.00	0.00	0.01	1.15 calcite

[a] Concentrations are in mol/l × 10^4.

calcium come from some other source? We shall return to these questions later in the chapter.

The importance of the Garrels and Mackenzie (1967) paper here is not so much the conclusions as the concept of explaining quantitatively the composition of a water in terms of mineral weathering reactions. The same approach has been used subsequently in numerous quantitative studies. The type of calculation can be generalized and solved by matrix methods (Velbel, 1986; Bowser and Jones, 1993). Mathematically speaking, if the number of solid phases available as reactants or products is equal to the number of independent solutes used in the mass balance calculation, a balanced weathering reaction can generally be written. The calculated reaction may or may not be an accurate description of what is happening in nature. The computer programs BALANCE (Parkhurst et al., 1982) and NETPATH (Plummer et al., 1991.

Absaroka Mountains, Wyoming, USA

Miller and Drever (1971) applied the same approach as Garrels and Mackenzie to a river draining andesitic volcanic rocks in Wyoming. They also analyzed soils and bedrock to see if the behavior of the solid phases was consistent with that predicted from the runoff chemistry. The Na, Ca, and silica concentrations in the water could be "explained" by the equation

$$\underset{\text{plagioclase}}{Na_{0.6}Ca_{0.4}Al_{1.4}Si_{2.6}O_8} + 2.69H_2O + 1.25CO_2$$

$$= 0.18\underset{\text{kaolinite}}{Al_2Si_2O_5(OH)_4} + 0.22\underset{\text{smectite (beidellite)}}{Ca_{0.33}Al_{4.67}Si_{7.33}O_{20}(OH)_4}$$

$$+0.6Na^+ + 0.32\ Ca^{2+} + 1.25HCO_3^- + 0.63H_4SiO_4$$

The water chemistry thus predicts that smectite and kaolinite should be forming in approximately equal proportions. However, kaolinite was a rare mineral in the soils; the only common secondary mineral was smectite. Furthermore, the smectite present was a montmorillonite, which can be represented approximately by the formula

$$Ca_{0.33}(Mg, Fe)_{0.67}Al_{3.33}Si_8O_{20}(OH)_4$$

and not a beidellite. In fact the composition of the minerals in the soil was such that a reasonable mass balance reaction for

Bedrock minerals + atmospheric input = soil minerals + solutes in runoff

could not be written. Miller and Drever concluded that slight alteration of large volumes of rock below the soil zone, rather than the soil zone itself, was the major source of solutes. This conclusion may apply only to the Absaroka mountains, where the andesine bedrock is porous and permeable. It illustrates the point, however, that even if an apparently satisfactory mass balance is achieved with the calculation procedure of Garrels and Mackenzie (1967), the calculated reactions may not be an accurate representation of reality.

Adirondack Mountains, New York, USA

The Adirondack Mountains have been the site of many studies related to acid deposition. The question of mineral weathering was addressed in detail by April et al. (1988). The focused particularly on two apparently similar catchments, Woods Lake and Panther Lake. Despite the close proximity and apparent similarity of the catchments, Woods lake is "acidic", with a typical outlet pH between 4.5 and 5, whereas Panther Lake was "neutral", with a typical outlet pH near 7. It was suggested that the difference was related to thicker tills in the Panther Lake basin, which allowed for more chemical weathering. April et al used three different mass-balance approaches to estimate weathering rates:

1. The present-day (1978-1980) flux of solutes leaving the basins, which reflects the present-day weathering rate ± ion exchange reactions ± the effect of biomass change.

2. Depletion of weatherable minerals in soil profiles in the basins. The abundance of weatherable minerals such as hornblende and plagioclase decreases towards the surface of the soil, presumably as a result of chemical weathering. If it is assumed that there has been no physical erosion or lateral transport, the decrease towards the surface reflects the integrated effect of weathering since the soil was first exposed at the surface, presumably at the end of glaciation 14,000 y ago. The amount of minerals weathered can be corrected for changes in soil volume by normalization to a mineral that is not weathered significantly, such as ilmenite. By this procedure, a mean flux of cations over the last 14,000 y can be calculated. This will be a "true" weathering rate in the sense that it is unaffected by plant uptake or cation exchange.

3. The change in the bulk chemistry of the soil with depth. This method is strictly analogous to the mineral depletion method above, but is based on bulk chemistry (normalized to an element such as titanium) rather than mineralogy.

A comparison of the cation fluxes calculated from methods 1 and 3 is shown in Table 4, and a comparison of mineral weathering rates for the Panther Lake catchment calculated from methods 1 and 2 is shown in Table 5. The results from the different methods clearly disagree. According to the soil chemistry profiles, the rates of weathering in the two catchments are indistinguishable, whereas the present-day cation fluxes differ by a factor of 8. The rate of hornblende weathering at Panther Lake calculated from the stream flux is about 40 times the rate calculated from the mineral depletion curves.

Table 4: Cation fluxes (eq/ha/y) from Woods Lake and Panther Lake basins in the Adirondack Mountains (after April et al, 1986; Drever, 1988).

	Net present-day stream flux		Calculated from bulk composition of soil	
	Woods	Panther	Woods	Panther
Ca	167	1149	178	100
Mg	28	294	122	84
K	−21*	32	129	150
Na	24	204	189	108
Total	198	1679	618	449

*K input from atmosphere was greater than stream output.

Table 5: Mineral weathering rates (kg/ha/y) for Panther Lake basin calculated from net present-day stream flux and from mineral depletion curves in soil (after April et al., 1986; Drever, 1988).

	From net present-day stream flux	From mineral depletion curves
Hornblende	312	8
Plagioclase	7	32
K-feldspar	−22	31

The differences between the present and long-term rates at Woods Lake could be explained by a gradual decrease over time in the amount of weatherable minerals exposed to incoming precipitation. As reactive minerals were dissolved, the rate of weathering would decrease. There is no easy or unique explanation for the discrepancies at Panther Lake. One possible explanation is that weathering is taking place deeper in the hydrologic system that the 1 m of soil analyzed by April et al. The discrepancy between the present-day and long-term apparent weathering rates for hornblende suggest that ion exchange may be a significant present-day source of calcium. An alternative explanation would be that the weathering rate at Panther Lake has increase in response to the recent increase in acid deposition. There is no obvious reason, however, why mineral weathering would increase in one catchment and not the other.

This study illustrates the importance of using solid phases as a constraint on a mass balance calculation. It also illustrates the difficulty in distinguishing between ion exchange and weathering as a source of cations.

Sogndal, Norway

Sogndal was the site of an experiment in which two small catchments on gabbroic gneiss were artificially acidified by application of acid (H_2SO_4) on one, a mixture of H_2SO_4 and HNO_3 on the other over a four-year period (Frogner, 1990). The experiment was part of the 8-year RAIN Project (Wright et al., 1986) directed towards understanding the effects of acid deposition on surface waters. Catchments adjacent to the acidified catchments served as controls. The area was characterized by treeless alpine vegetation and thin (about 30 cm) sandy soils. In addition to the field studies, laboratory studies were conducted on the dissolution of minerals from the soil. The discussion here is based on Frogner (1990).

Acidification caused an increase of about 45% in the annual fluxes of both base cations (dominantly Ca^{2+}) and silica, which suggests an increase in the weathering rate of plagioclase (plagioclase was the dominant source of solutes). Sodium, however, showed no significant increase, which suggests no increase in the plagioclase weathering rate. Immediate increases in Ca and Mg concentrations associated with applications of acid were clearly related to ion exchange. The question is whether the change in annual flux represents increased weathering or ion exchange. The exchange in the exchange pool, even over the four years of the experiment, would have been too small to measure directly. One could argue that the increased silica flux clearly indicates increased weathering, and the sodium data are inconclusive, or that the silica flux represents pH-dependent desorption and the sodium data are to believed. The important point, in my opinion, is that even in a small, intensively studied catchment with several years of data, the question cannot be unambiguously resolved.

THE PROBLEM OF EXCESS CALCIUM

In many catchments, particularly high-elevation catchments where there has been little weathering and soil development since the Pleistocene glaciation, Ca makes up a larger proportion of the cations in runoff than can be explained by stoichiometric weathering of feldspar. The Sierra Nevada perennial springs of Garrels and Mackenzie discussed above are an example of the phenomenon. Garrels and Mackenzie attributed the excess to calcite without really discussing the issue.

South Cascade Glacier

A fairly extreme example is the South Cascade Glacier area of Washington State in the northwest United States (Drever and Hurcomb, 1986). The compositions of waters draining different lithologies in the area are shown in Fig. 4. The relatively high abundance of Ca and K, and low abundances of Na and silica are striking. Plagioclase in the bedrock ranged in composition from about An_{28} ($Na_{0.72}Ca_{0.28}Al_{1.28}Si_{2.72}O_8$) to An_{55} ($Na_{0.45}Ca_{0.55}Al_{1.55}Si_{2.45}O_8$). Weathering of plagioclase should produce waters with Na/Ca ratios of about 1 or greater, so stoichiometric weathering of plagioclase could not be the source of the calcium in the water. Drever and Hurcomb concluded that the only reasonable source of Ca, without large amounts of Mg or silica, was dissolution of calcite. They proposed a mass balance equation

$$\begin{aligned}
&\underset{\text{calcite}}{28.4\,CaCO_3} + \underset{\text{biotite}}{5.9KMg_{1.75}Fe_{1.25}AlSi_3O_{10}(OH)_2} + \underset{\text{actinolite}}{1.5Ca_2Fe_{2.5}Mg_{2.5}Si_8O_{22}(OH)_2}\\
&+ 0.9FeS_2 + 81.2H_2O + 48.9CO_2 + 2.9O_2 = \underset{\text{vermiculite}}{4.5Ca_{0.28}Mg_{1.5}Fe_{1.25}Al_{1.3}Si_{2.7}O_{10}(OH)_2}\\
&+ 0.6Fe_2O_3 + 30.2Ca^{2+} + 7.3Mg^{2+} + 5.9K^{+} + 77.3HCO_3^{-} + 1.8SO_4^{2-} + 17.5H_4SiO_4
\end{aligned}$$

In fact, calcite was present in the catchment as veins, as a deposit in joint plains, in occasional marble bands, and as a superficial subglacial deposit associated with regelation processes under glacial ice. The high (relatively) K concentration they attributed to alteration of biotite to vermiculite, which is often an important reaction under cold conditions.

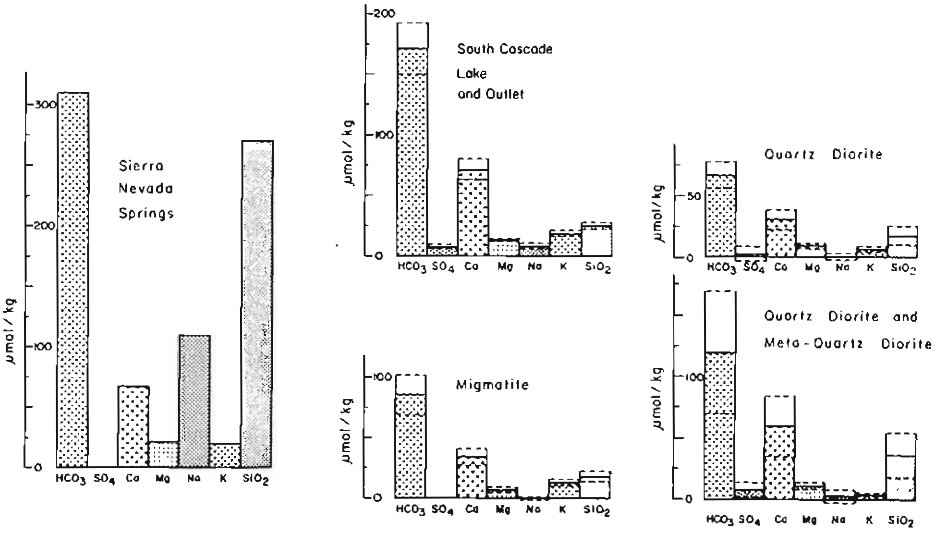


Figure 4: Compositions of surface waters from South Cascade Lake basin compared to Sierra Nevada ephemeral springs, all corrected for atmospheric input. Error bars represent one standard deviation of mean surface water composition (after Drever and Hurcomb, 1986).

Loch Vale catchment, Colorado

The Loch Vale catchment in Colorado, USA, also shows an excess of calcium over what would be predicted from feldspar weathering (Mast et al., 1990). On the basis of the compositions of the minerals present in the catchment, Mast et al. came up with a mass-balance equation for weathering

$$\begin{array}{rcll} \text{125 plagioclase} & & \text{74 kaolinite} & \\ \text{29 biotite} & & \text{30 smectite-illite} & \\ & + \text{ atmos input } = & & + \text{ solutes} \\ \text{8 chlorite} & & 12\ Al(OH)_3 & \\ \text{113 Calcite} & & 52\ FeO(OH) & \end{array}$$

Mast et al. were unable to find any macroscopically visible calcite in the catchment, but demonstrated, using cathodoluminescence microscopy, that very fine-grained calcite was present in the bedrock in hydrothermally altered zones and in grain boundaries adjacent to plagioclase. They argued that deglaciation had occurred relatively recently, and rates of physical erosion were high, so that there was enough of this calcite in contact with meteoric waters to supply the observed excess of calcium. Over time, accessible calcite would be depleted and it would not serve as a significant source of calcium. One could use this argument, speculatively, to explain the excess calcium in the Sierra Nevada springs. The shallow, near-surface rocks (ephemeral springs) had been extensively flushed and no calcite remained. The deeper rocks (perennial springs) were less flushed and some calcite still remained.

Velbel (manuscript in preparation), on the other hand, explained the excess calcium in the runoff in Loch Vale on the basis of selective weathering of isolated regions of mafic amphibolite (amphibole and calcic plagioclase) in the bedrock. This alternative explanation cannot be ruled out on the basis of presently-available information.

Discussion

The question of excess calcium is important because it is almost universal in high-elevation lakes in North America (Turk and Spahr, 1990; Stauffer, 1990) and also in waters of Scandinavia. If we are to model surface water chemistry successfully, we need to understand the underlying reactions. Several mechanisms could give rise to it:

1. Weathering of calcite in the bedrock. I would argue that, on the scale of a catchment, calcite will always be present initially in a granitic rock. It may be present in veins, in zones of hydrothermal alteration, or along grain boundaries as a result of deuteric alteration or simply closed-system alteration in the presence of groundwater. [Petrologic descriptions of rocks are often misleading; petrologists systematically avoid sampling altered rocks, which are the ones most likely to contain calcite.] Although calcite may be present initially in a granitic rock, it will be rapidly depleted as the rock is exposed to meteoric water. Whether enough calcite remains at the present time to influence water chemistry will vary from catchment to catchment. It is most likely to occur in catchments that have recently been glaciated and that are undergoing rapid physical erosion.

2. Selective weathering of more calcic plagioclase in the bedrock. Clayton (1988) showed that zoned plagioclases in the bedrock of the Idaho Batholith did not weather uniformly: the calcic cores dissolved while the more sodic rims remained intact. He used this observation to explain why the Ca/Na ratio in the water was higher than that of the average plagioclase in the rock. This is a plausible mechanism *provided the Ca/Na ratio in the water is not higher than the highest ratio present in the feldspar.* Some authors have claimed, on the basis of Clayton's results, that the albite and anorthite components of a feldspar could weather independently: the anorthite component could dissolve, leaving the albite component behind. I believe this is wrong: feldspars alter congruently, but different compositions alter at different rates.

3. Ca comes from alteration of minor amounts of reactive Ca-containing silicates such as pyroxenes, amphiboles, or epidote. This is certainly the case in some catchments. Such reactions are often difficult to constrain by mass-balance equations because the secondary products are often smectites or vermiculites, whose compositions are poorly known. The overall plausibility of this mechanism must assessed on a catchment by catchment analysis.

4. Ca comes from cation exchange sites. Calcium is greatly favored over sodium on the exchange sites of clays in dilute solutions. Thus the Ca/Na ratio of ions on exchange sites in a soil will be much higher than the Ca/Na ratio of plagioclase in the underlying bedrock. If exchangeable ions are displaced by, for example, an increased input of acidity from the atmosphere, the result will be an addition of Ca to runoff water. This is a plausible explanation for catchments affected by anthropogenic acidity (perhaps Scandinavia). It is not a plausible explanation for the Rocky Mountain region of North America, where the input of anthropogenic acidity is minor. Excess Ca could also be derived from a net decrease in biomass, but catchments showing a net decrease in biomass are relatively rare.

Documenting the source of calcium is an essential part of catchment mass balance studies, particularly for modeling future responses to acid deposition. The predictions of mechanistic model will be quite different depending of which source for calcium is assumed.

REFERENCES

APRIL R., NEWTON R., AND COLES L.T. (1986) Chemical weathering in two Adirondack watersheds: Past and present-day rates. *Geol. Soc. Amer. Bull.* **97**, 1232-1238.

BOWSER C.W AND JONES B.F. (1993) Mass balances of natural waters: Silicate dissolution, clays, and the calcium problem. *Biogeomon* Symposium on Ecosystem Behaviour: Evaluation of Integrated Monitoring in Small Catchments, Prague, Czech Republic, Abstracts P. 30-31.

CLAYTON J.L. (1988) Some observations of the stoichiometry of feldspar hydrolysis in granitic soils. *J. Environmental Quality* **17**, 153-157.

DREVER J.I. (1988) *The Geochemistry of Natural Waters*, 2nd Edition. Englewood Cliffs, Prentice-Hall, 437 p.

DREVER J.I. AND HURCOMB D.R. (1986) Neutralization of atmospheric acidity by chemical weathering in an alpine drainage basin in the North Cascade Mountains. *Geology* **14**, 221-224.

FROGNER T. (1990) The effect of acid deposition on cation fluxes in artificially acidified catchments in western Norway. *Geochim. Cosmochim. Acta* **54**, 769-780.

GARRELS R.M. AND MACKENZIE F.T. (1967) Origin of the chemical compositions of some springs and lakes. *Equilibrium Concepts in Natural Water Systems.* Am. Chem. Soc. Adv. Chem. Ser. 67, 222-242.

LIKENS G.E., BORMANN F.H., PIERCE R.S., EATON J.S., AND JOHNSON N.M. (1977) *Biogeochemistry of a Forested Ecosystem.* Springer-Verlag, 135 p.

MAST M.A., DREVER, J.I., AND BARON, J. (1990) Chemical weathering in the Loch Vale watershed, Rocky Mountain National Park, Colorado. *Water Resources Research* **26**, 2971-2978.

MILLER E.K., BLUM, J.D., AND FRIEDLAND A.J. (1993) Determination of soil exchangeable-cation loss and weathering rates using Sr isotopes. *Nature* **362**, 438-441.

MILLER W.R. AND DREVER J.I. (1977) Chemical weathering and related controls on surface water chemistry in the Absaroka Mountains, Wyoming. Geochim. *Cosmochim. Acta* **41**, 1693-1702.

PARKHURST D.L., PLUMMER L.N., AND THORSTENSON D.C. (1982) BALANCE--a computer program for calculating mass transfer for geochemical reactions in ground water. U.S. Geological Survey Water-Resources Investigations Report 82-14, 29 p.

PLUMMER L.N., PRESTEMON E.C., AND PARKHURST D.L. (1991) An interactive code (NETPATH) for modeling net geochemical reactions along a flow path. U.S. Geological Survey Water-Resources Investigations Report 91-4078, 149 p.

PROBST A., DAMBRINE E., VIVILLE D., AND FRITZ B. (1990) Influence of acid atmospheric inputs on surface water chemistry and mineral fluxes in a declining spruce stand within a small granitic catchment (Vosges Massif, France). *J. Hydrology* **116**, 101-124.

REUSS J.O. AND JOHNSON D.W. (1986) *Acid Deposition and the Acidification of Soils and Waters.* Ecological Studies 59, Springer-Verlag, 119 p.

SCHNOOR J.L AND STUMM W. (1985) Acidification of aquatic and terrestrial systems. *Chemical Processes in Lakes* , W. Stumm, ed., Wiley-Interscience, New York, 311-338.

STAUFFER R.E. (1990) On granite weathering and the sensitivity of alpine lakes to acid deposition. *Limnology and Oceanography* **50**, 1109-1130.

STAUFFER R.E. AND WITTCHEN B.D. (1991) Effect of silicate weathering on water chemistry in forested, upland, felsic terrane of the USA. *Geochim. Cosmochim. Acta* **55**, 3253-3271.

TURK J.T. AND SPAHR N.E. (1990) Rocky Mountains: Controls on lake chemistry. *Acid Deposition and Aquatic Ecosystems: Regional Case Studies*, D.F. Charles, ed., Springer-Verlag.

VELBEL M.A. (1986) The mathematical basis for determining rates of geochemical and geomorphic processes in small forested watersheds by mass balance: Examples and implications. *Rates of Chemical Weathering of Rocks and Minerals*, S.M. Colman and D.P. Dethier, eds., Academic Press, 439-451.

WRIGHT R.F. (1988) Influence of acid rain on weathering rates. *Physical and Chemical Weathering in Geochemical Cycles*, A. Lerman and M. Meybeck, eds., NATO ASI Series 251, Kluwer, 181-196.

WRIGHT R.F., GJESSING E., CHRISTOPHERSEN N., LOTSE E., SEIP H.M., SEMB A., SLETAUNE B., STORHAUG R., AND WEDUM K. (1986) Project RAIN: changing acid deposition to whole catchments. The first year of treatment. *Water, Air, Soil Pollution* **30**, 47-63.

EXAMPLE OF AMBIGUITIES IN A MASS BALANCE CALCULATION-- WEST GLACIER LAKE BASIN, WYOMING

J.I. Drever

West Glacier Lake catchment is a high-elevation (3270 m) catchment located in the Medicine Bow Mountains of Wyoming, which has been the subject of ongoing studies for several years. The bedrock is 85% quartzite (mostly quartz, also alkali feldspars and muscovite) cut by mafic dikes (44% actinolite, 30% epidote, 10% albite). Biotite, minor amounts of pyrite and very minor amounts of calcite have also been identified in the bedrock. Soils are thin and rocky; vegetation is mostly bare rock and tundra, with minor forest (*Abies lasiocarpa*) at the lower elevations.

A significant fraction of the soil is aeolian in origin. It is an open question whether the aeolian material is a significant source of solutes.

Long Creek, a tributary to West Glacier Lake, has the following base-flow composition, corrected for atmospheric input (µmol/L):

Ca 21.55; Mg 12.85; Na 11.6; K 3.12; H_4SiO_4 62.46

Can we write a balanced weathering reaction to "explain" this composition?

Plausible primary minerals are actinolite, epidote, albite, K-feldspar, biotite, quartz, and calcite (we shall omit muscovite for the sake of simplicity). Identified minerals in the soil include vermiculite, smectite, and kaolinite: gibbsite (or an amorphous equivalent) is also a plausible secondary phase. Only smectite is definitely forming as a weathering product. The other minerals could be secondary or could be aeolian. The molar proportions in these phases are shown in Table 1. In this example, the compositions of epidote, actinolite, and smectite are analyses from the field area; the others are from a similar area (vermiculite) or idealized formulae. I shall ignore Fe: including it adds very little additional information.

Table 1: Compositions of phases involved in weathering reactions, West Glacier Lake basin.

	actin	albite	K-spar	biotite	epid.	calcite	quartz	vermic	smect.	kaolin.	gibb.	stream
Ca	1.85	0	0	0	1.95	1	0	0.28	0.09	0	0	21.55
Mg	3.00	0	0	1.5	0	0	0	1.5	0.21	0	0	12.85
K	0.00	0	1	1	0	0	0	0	0.02	0	0	3.12
Na	0.15	1	0	0	0	0	0	0	0.01	0	0	11.60
Si	7.50	3	3	3	3.02	0	1	2.7	3.7	2	0	62.46
Al	0.80	1	1	1	2.4	0	0	1.3	1.93	2	1	0.00

Mathematically, we have 6 "components", Ca, Mg, Na, K, Si, Al (anions, specifically bicarbonate, have been ignored because of the requirement of charge balance: if the cations balance, bicarbonate inevitably balances also). We can therefore, in principle, write a balanced reaction involving 6 minerals, which may be either reactants or products. This produces 6 equations (one for the sum of each component) in 6 unknowns (the amount of each mineral). The set of equations can readily be solved on a microcomputer spreadsheet program by matrix methods (the one I use was written by J.B. Finley). With more than 6 minerals the solution is indeterminate, with fewer than 6, there is generally no solution to the matrix of equations. I have listed 6 possible primary minerals and 4 possible secondary minerals for a total of 10. We therefore need to choose subsets of the 10, and see which subsets yield a reasonable mass balance. Because the mathematical solution is indeterminate does not, of course, mean that only 6 minerals are involved in nature. There are some immediate constraints: albite is the only primary phase containing much Na, so it must be included. K-feldspar and biotite are the only primary phases containing K, so 1 (at least) of them must be included (the amount of K in the water, which is small, determines the amount of K-containing primary phase; since the amount is small, it does not make much difference whether K-feldspar or biotite is used), for Mg, either actinolite or biotite must be included, but because of the high Mg/K ratio in the water, only solutions involving actinolite are reasonable; for Ca, actinolite or epidote or calcite must be included; the relatively high Ca/Mg ratio in the water excludes actinolite as the sole source of Ca. The requirement that primary minerals dissolve and not precipitate, and that secondary minerals precipitate and not dissolve also constrains the number of possible reactions. Let us look at the set of mineral reactions that satisfy all the listed constraints.

1. Reactions not involving quartz. The assumption has generally been made since the original Garrels & Mackenzie (1967) paper that quartz neither dissolves nor precipitates in the weathering environment. Whether or not this is a reasonable in a basin where 80% of the bedrock is quartz is an open question.

The set of reactions that meet the above criteria is shown in Table 2:

Table 2: Weathering reactions not involving quartz[1].

	µmoles dissolved or precipitated							
React. No.	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Actinolite	4.68	9.34	8.24	5.84	7.05	8.39	9.56	5.97
Albite	10.90	10.30	10.42	10.72	10.54	10.37	10.24	10.71
K-feldspar	3.12	3.32	3.23	3.12	3.12	3.18	3.27	3.12
Epidote	6.72	3.90					4.16	6.32
Calcite			8.81	11.62	10.93	9.38		
Vermic.	0.80	8.72	7.11	3.11	8.66	10.96	12.78	6.48
Smectite		9.91	5.70			3.12	8.53	
Kaolinite	16.43			7.23	4.02			12.67

Table 2 (cont.)

	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Actinolite	5.01	4.28	3.22	2.72	4.55	4.28	2.74	2.72
Albite	10.95	10.96	11.20	11.19	10.96	10.96	11.19	11.19
K-feldspar	3.33	3.12			3.12	3.12		
Biotite			3.28	3.12			3.13	3.12
Epidote	6.78	6.99	8.38	8.47				
Calcite					13.48	13.63	16.51	16.51
Smectite	10.38		8.22		3.82		0.28	
Kaolinite		16.50		13.24		5.95		0.45
Gibbsite	14.51	1.27	21.30	10.34	10.42	5.61	15.96	15.59

[1] the combination vermiculite-gibbsite gives no solutions.

Epidote and calcite are essentially interchangeable. Both supply only Ca as a cation and the additional Si does not have a major effect on the calculation. On the other hand, the calculation is rather sensitive to the choice of secondary phase. The ratio of actinolite:epidote changes from 2.72:8.47 (Reaction 12) to

9.34:3.90 (Reaction 2) when one of the secondary minerals is changed from vermiculite to gibbsite. In general, the compositions of secondary phases, on a catchment scale, are rarely well constrained; as a result detailed mineral budgets are rarely well constrained.

2. Reactions involving quartz. If we allow quartz to act as both a primary phase and a secondary phase, the number of possible reactions is large. If we restrict the set to reactions in which quartz is a reactant and not a product, the pattern of reactions is relatively simple. 1) If we add quartz as a primary phase, we need to delete one other phase from our matrix. The only reasonable choice is to delete one of the secondary clay minerals, leaving only one secondary phase. After that is done: 2) K-feldspar and biotite continue to be interchangeable, likewise epidote and calcite. 3) No solutions involving vermiculite as the secondary phase exist. 4) All permutations of primary minerals altering to smectite or kaolinite "work". The difference between the smectite and kaolinite is essentially Si/Al ratio. Solutions involving kaolinite differ from those involving smectite largely in the amount of quartz dissolved. 5) No solutions involving gibbsite "work", because silica must precipitate rather than dissolve to provide balance.

The set of reactions that provide balance is shown in Table 3.

Table 3: Weathering reactions involving a single secondary phase and quartz as a primary phase.

	µmoles dissolved or precipitated							
React. No.	(17)	(18)	(19)	(20)	(21)	(22)	(23)	(24)
Actinolite	2.72	2.72	4.28	4.28	3.89	3.25	5.54	4.94
Albite	11.19	11.19	10.96	10.96	11.21	11.20	10.95	10.95
K-feldspar			3.12	3.12			3.48	3.31
Biotite	3.12	3.12			3.51	3.30		
Epidote	8.47		6.99		8.26		6.62	
Calcite		16.51		13.63		16.33		13.25
Quartz	10.34	15.59	1.27	5.61	36.34	27.34	24.23	17.52
Kaolinite	18.47	8.25	17.14	8.75				
Smectite					19.51	8.86	18.01	9.44

Summary: This exercise has shown that, even with a relatively well-constrained system, the mathematical solution to the mass-balance equation in terms of mineral weathering reactions is non-unique. In reality, the compositions of the minerals involved, particularly of the secondary phases, are not well constrained. The solution to the mass balance equation is rather sensitive to the assumed composition of the secondary phases. Also, K-feldspar/biotite and epidote/calcite are presented as either/or options. In reality, all four phases could be contributing solutes (as could muscovite).

Despite these uncertainties, we can draw some important qualitative conclusions:

1. Albite must be a significant contributor to the solutes, even though the amount of albite in the bedrock is not large, and albite is not a particularly reactive mineral.

2. Actinolite must likewise be a major contributor.

3. Although the Ca/Na ratio in the water is relatively high, there is no "need" to invoke calcite as a source of Ca: epidote is a plausible alternative. On the other hand, calcite is not ruled out as a source of calcium.

4. There is no way of identifying the source of K (a minor solute) by this type of mass-balance calculation.

In this example, at least, the composition of the water does very little to constrain the composition of the secondary clay minerals formed during weathering. This is a somewhat surprising result; in several of the early mass balance calculations of this type (Garrels and Mackenzie, 1967; Tardy, 1971) the main outcome of the calculation was the identity of the secondary mineral(s) being formed.

The role of natural organic matter (NOM) in geochemical processes in natural and polluted aquatic systems

Donald L. Macalady and James F. Ranville
Department of Chemistry and Geochemistry
Colorado School of Mines
Golden, Colorado 80401 USA

TOPICAL OUTLINE

I. Introduction
II. The origin and nature of natural organic matter
  A. Terrestial Detritus and NOM
  B. Humus and humic materials
    1. Traditional definitions
    2. Modern definitions and isolation/separation methods
    3. NOM fractions vs. unfractionated samples
    4. Chemical characteristics of NOM
  C. Aquatic organic matter
    1.Dissolved and colloidal forms
    2. "Dissolved" organic matter (DNOM)
    3. Colloidal organic matter (CNOM)
    4. Pedogenic and aquagenic materials
III. Geochemical reactions of NOM
  A. Weathering and NOM
  B. NOM profiles in soils and associated aqueous systems.
  C. Hydrologic controls on the transport of NOM
  D. Oxidation/reduction chemistry of metal-organic complexes
  E. NOM and metal ion transport
IV. Interactions between NOM and anthropogenic chemicals
  A. Transport of pollutants as DNOM complexes
  B. Effects of sorption/partitioning to NOM on the transport of contaminant organic chemicals
  C. Effects of NOM on hydrolytic reactions
  D. Oxidation/reduction reactions facilitated by NOM
  E. Colloidal NOM and facilitated transport
V. NOM....sink for pollutants or facilitator of transport? Mysteries and research questions.

VI. Bibliography, literature cited, and notes

I. INTRODUCTION

Natural organic matter (NOM), in its broadest sense, refers to the complex and chemically and physically diverse substances which result directly or indirectly from the photosynthetic activity of plants. In the geosciences we typically limit the use of this term, however, to substances resulting from the partial decay of detrital materials from terrestrial and aqueous plants. This definition includes everything along the diagenic pathway from living plants to fossil fuels such as coal and petroleum.

For the purposes of this paper, a further limitation is required. In this discussion, nothing further along the diagenic pathway than materials roughly resembling peat will be considered. In other words, the NOM discussed here has retained a considerable portion of its original oxygen and hydrogen, and may have picked up a bit of additional sulfur or metal content along the diagenic pathway or in other geologic processes not accurately described as diagenic. It's the kind of stuff you're likely to find as you dig in your garden, walk through a bog or fen, or examine the colored substances found naturally in lakes, rivers and ground waters.

This chapter is <u>not</u> primarily about the mass balances for the transport of organic materials in catchments. Rather, it concerns the origin and nature of NOM as defined above, and the role of NOM in the transport and transformations of substances, natural and anthropogenic, in aquatic systems. Thus, the goal of this discussion is to provide the background and technical understanding necessary for the evaluation of the role of NOM in the many mass balances which are important to catchment studies. Hopefully, it will also stimulate the reader to additional study into the nature of the involvements of NOM in geochemical processes in aquatic systems.

The text of the paper is followed by a general bibliography that should facilitate further study. Specific references and notes from the text are listed separately.

II. THE ORIGIN AND NATURE OF NOM

Natural organic matter has been classified according to a wide variety of schemes, one of which involves the origin of the plants which serve as a starting material for the NOM (Figure II-1). *Aquogenic* refers to NOM originating from the decomposition of aquatic organisms (macrophytes, algae, bacteria); *pedogenic* refers to NOM from terrestrial plants and microorganisms, including pedogenic material leached from soils into an aquatic system. For most catchments which do not include a large lake, river or reservoir, and for all soils, pedogenic materials can be expected to dominate. For large lakes, and, of course, the open oceans, aquogenic materials are the overwhelming majority. Another pair of terms is often used to refer to a similar distinction among types of organic materials is *autochthonous and allochthonous*. Some authors are careful to distinguish between these pairs of terms (Buffle, 1984), particularly with reference to dissolved or colloidal organic materials.

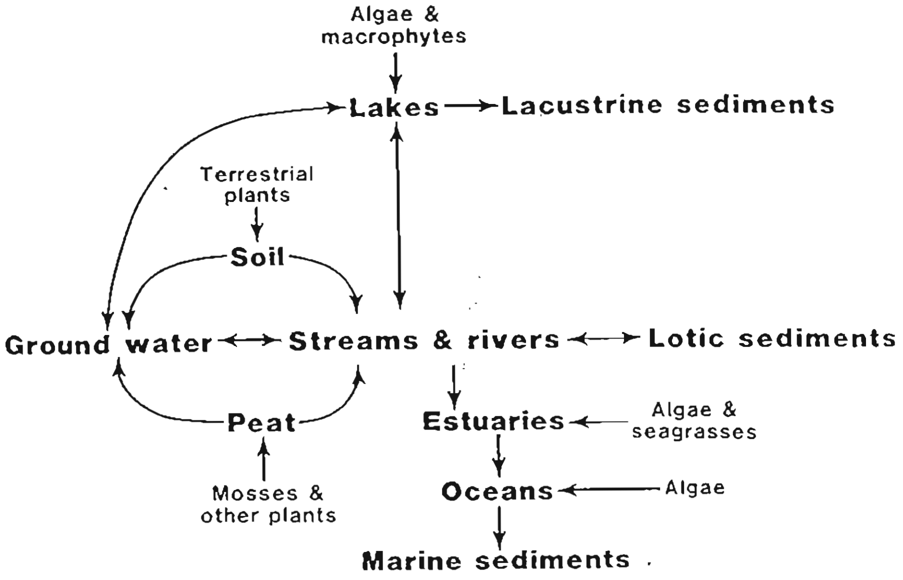


Figure II-1. Some of the many possible flowpaths for NOM (from Aiken, et al., 1985)

A. Terrestrial Detritus and NOM

Detritus is a general term which refers to dead plant parts which are in various stages of decomposition. In a global sense, the production of plant litterfall has well-understood geographical patterns which roughly follow global patterns in net primary production (Figure II-2). Generally speaking, production of terrestrial detritus varies inversely with latitude from tropical to boreal forests (Lonsdale, 1988).

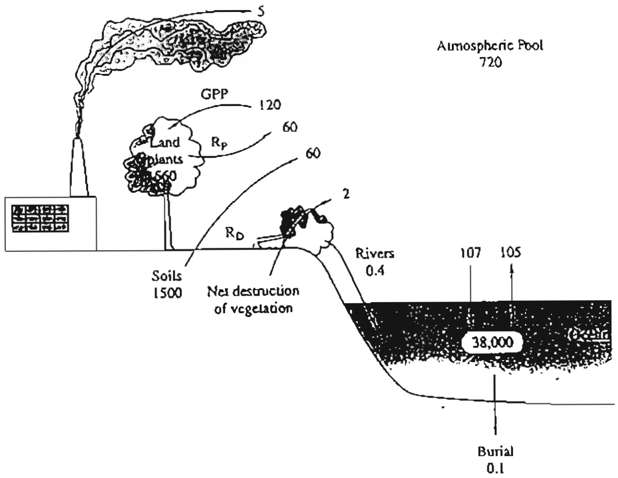


Figure II-2. The present-day global carbon cycle. All pools are expressed in units of 10^{15} gC and all fluxes in units of 10^{15}gC/yr. (From Schlesinger, 1991. p . 309)

Most terrestrial detritus results from either the above-ground portions or the root systems of plants, and is thus located in the upper layers of the soil. Here it is subject to a variety of decomposition processes, mostly through heterotrophic metabolism involving bacteria, fungi and other microfauna (Swift, et al., 1979). One result of such processes is the release of CO_2, H_2O and nutrients such as nitrogen and phosphorus compounds. Another, more important for the purposes of this paper, is the production of a group of substances which are highly resistant to further aerobic or anaerobic decomposition, collectively known as *humus*. These early diagenic processes can be viewed in terms of the rapid turnover of the majority of the detrital material near the soil surface and the much slower production, accumulation and turnover of humus in the deeper soil layers (Schlesinger, 1977).

A simple, first-order model for the rate of decomposition of forest litterfall provides a basis for comparing a variety of ecosystems, allowing the estimation of a mean residence time for plant debris as 1/k, where k is the first order decay rate constant (yr^{-1}). Values for many tropical rain forest ecosystems, for example are greater that 1.0, indicating little surface accumulation and rapid turnover of organic matter (Cuevas and Medina, 1988). In contrast, some peat lands have k values of ca. 0.001 (Olsen, 1963). Esser, et al.(1982) estimated a global mean value of $k = 0.33$ for soil surface litter. In various geographical and ecological regions, the rate of decomposition may be limited by a variety of environmental factors, including temperature, moisture and the origin of the litter material. The role of litter material in soil genesis and soil profiles is discussed in some detail in a later section.

B. Humus and Humic Materials

The discussion above indicates that incomplete turnover of plant litter in most of the world's ecosystems results in the accumulation of soil organic matter. Accumulation rates vary from about 0.2 g $C/m^2/yr.$ in polar desert environments up to 15 in cool, wet locations such as boreal forests (Bockheim, 1979; Ugolini, 1968). Much of this accumulated material is soil humus, which is extremely resistant to further degradation. Radiocarbon dating techniques indicate ages of several thousand years for soil organic matter in lower soil profiles (Obrien and Stout,1978; Schlesinger, 1977).

Chemical characterization of humus is often based on a solubility scheme originally developed by soil scientists (Figure II-3, see also Stevenson, 1986). Soil or sediment is first extracted with alkali. The insoluble organic matter remaining after alkaline extraction is termed *humin*. The portion of the soil solubilized by the alkali is subsequently treated with acid to near-neutral pH. The material which precipitates in this step is termed *humic acid*, while the organic material remaining in solution is called *fulvic acid*. The precipitated humic acid is sometimes further differentiated into *hymatomelanic acid* (alcohol soluble portion of humic acid), gray humic acid (precipitated from an alkaline humic acid solution with added electrolyte), and brown humic acid (soluble in alkaline solution with added electrolyte), but these latter categories are not widely discussed in geochemistry.

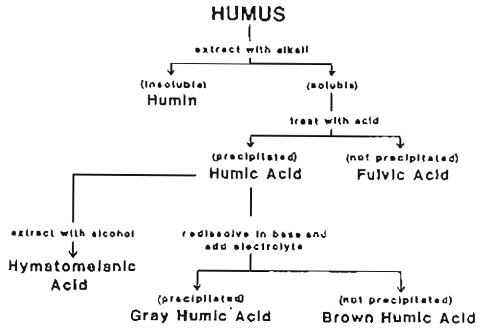


Figure II-3. Scheme for the extraction of humic substances (from Stevenson, 1982, p.43)

More recently, NOM fractions in aquatic systems have been redefined in terms of a separation scheme based on retention characteristics on macro-reticular non-ionic resin (Figure II-4, Thurman and Malcolm, 1981). In the most commonly applied method, the hydrophobic characteristics of different humic fractions are exploited to enable separation after adsorption (at pH = 2) on a column of a non-ionic acrylic resin. Application of a pH gradient to the material adsorbed onto the columns provides a separation of the aquatic NOM into fractions identifiable as fulvic and humic acids (MacCarthy et al., 1979), which typically accounts for about 50% of the total DOC (Figure II-5).

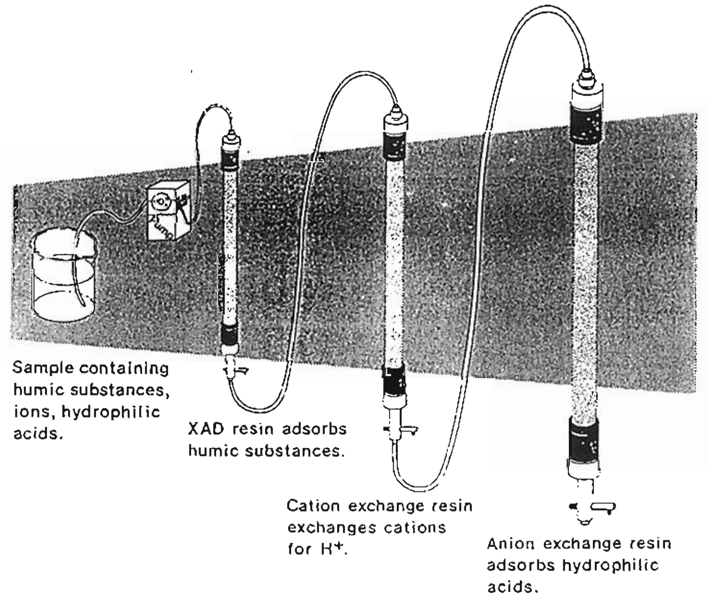


Figure II-4. *Separation of humic substances and hydrophilic acids by adsorption and ion exchange chromatography. (from Thurman, 1985, p. 109)*

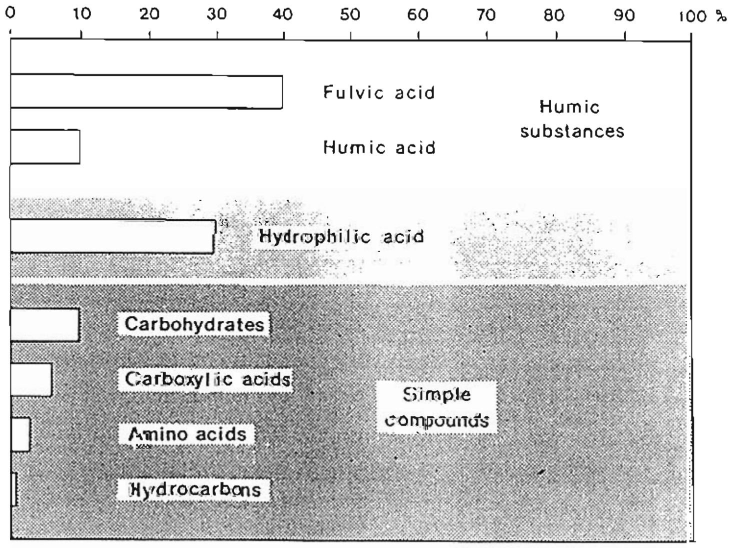


Figure II-5. *Dissolved organic carbon histogram for an average river water with a DOC of 5 mg/L. (from Thurman, 1985, p. 105)*

Regardless of the isolation/definition scheme, humic and fulvic acids constitute a major fraction of any NOM sample. Soil NOM typically has a larger humic/fulvic ratio, and smaller contents of identifiable non-humic materials. Humic and fulvic acids account for widely varying proportions of aquatic NOM, from around 25% for sea water to almost 90% for some wetlands and highly colored rivers. The remaining NOM in aquatic samples consists of more hydrophilic organic acids and a variety of identifiable compounds including amino acids, some simple sugars and several small organic acids (Thurman, 1985).

Much of the effort to chemically and physically characterize NOM has focused on the humic and fulvic acid fractions. In general, aquatic NOM can be described as a polymolecular array of organic molecules which contain about 50-60% carbon, 4-6% hydrogen, 23-40% oxygen, 0.3-6% nitrogen (by weight), with traces of phosphorus, sulfur and "ash" (residue after combustion). Soil NOM is similar in elemental composition, with naturally larger proportions of highly insoluble and/or hydrophobic constituents such as lipids, proteins and polysaccharides. Isolates of soil NOM generally have considerably higher ash contents (Figure II-6).

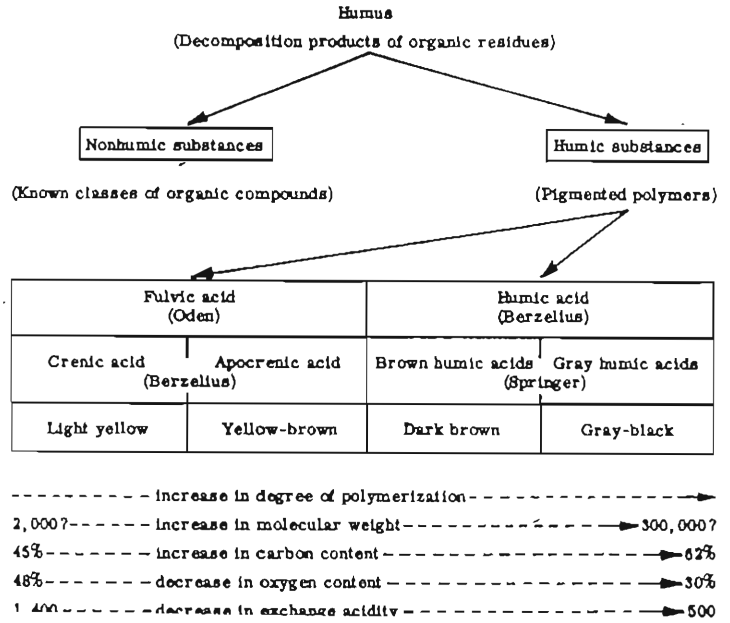


Figure II-6. Classification and chemical properties of humic substances. (from Stevenson, 1982, p. 45)

In addition to elemental composition, NOM fractions have been characterized according to molecular weight distributions, the presence of certain functional groups, acid-base and oxidation/reduction characteristics, aromatic/aliphatic character, chelating/complexation abilities, and trace element composition. An excellent summary of many characterizations of aquatic NOM fractions is given by Thurman (1985). A similarly useful source for soil NOM is given by Stevenson (1982).

The utility of such analytical efforts, for the purposes of this paper, is limited by the fact that the NOM has been fractionated into fulvic acid, humic acid, etc. prior to analysis. It is a given that, for the purposes of studies into processes which occur in nature, one is interested in the behavior of NOM as it exists *in situ*, not in the behavior of the humic or fulvic acid fraction of that NOM. Thus, the large amount of scientific endeavor which has been expended to isolate and characterize NOM fractions, though extremely useful in providing a basis for the understanding of the physical and chemical characteristic of NOM, is not directly applicable to the considerations of this paper. The brief look at some of these characteristics which follows is intended to provide general picture of the types of interactions we might expect between NOM and the surrounding geochemical matrices.

The fractions of NOM from soil or aquatic environments called fulvic and humic acids differ primarily in their hydrophobicity and solubility characteristics, with concomitant differences in average molecular size and mass. In general, fulvic acids are more soluble, smaller in average molecular weight, less aromatic, and more highly charged than humic acids. Fulvic acids also typically have higher oxygen content, with higher carboxylic acid (COOH) and lower phenolic (ArOH) content than humic acids (Hayes et al., 1989).

Fulvic acids are therefore probably more representative of aquatic NOM, especially the "dissolved" fraction (Malcolm 1985). A typical "molecule" of fulvic acid might contain one COOH per 6 carbon atoms, with about 65% aliphatic (cf. aromatic) carbons (Thurman, 1985, p. 356). Fulvic acid and NOM in general are known from ESR evidence to produce free radicals in aqueous solutions (Senesi and Steelink, 1989). The most likely groups to contribute this free radical character are quinones, and quinone functionalities have recently been shown by ^{13}C NMR evidence to be present (Thorn et al.,1992). The emerging picture of aquatic NOM is one of polyfunctional organic acids containing significant metal-complexing and oxidation-reduction capabilities. Further, various portions of individual fulvic and humic acid "molecules" are expected to exhibit significant hydrophobic character, more so for humic acids and particulate NOM (PNOM) than for the "dissolved" fraction of aquatic NOM (DNOM). Finally, the dominance of carboxyilic acid and phenolic functional groups in NOM means that the net "surface" charge of aqueous NOM and the overwhelming majority of soil or sediment particles, which contain NOM-dominated surfaces, can be expected to be negative at essentially all environmentally-relavent pH values (Tipping and Cooke, 1982)

C. Aquatic Organic Matter:

Dissolved and colloidal forms

Organic matter present in raw natural waters may be considered to physically occur as an essentially continuous series which we will refer to as the dissolved/colloidal/particulate organic matter continuum (Figure II-7). While the pure end member modes are easily conceptualized as: a) dissolved (solvated) organic molecules; and b) macroscopic solid particles which may occur with or without sorbed ions or molecules, the intermediate colloidal mode requires further explanation. The word `colloid' as used here collectively identifies the intermediate continuum of modes. Organic matter species which may be referred to as 'colloidal' include:

a) discreet chemical species with sufficient size or mass to behave as colloids (macromolecules);

b) aggregates of smaller organic molecules, and

c) organic matter associated with (usually by sorption) compositionally distinct, colloid-size particles, such as clays or oxides.

Traditionally, filtration through a 0.45 μm pore-size membrane, although quite arbitrary, has been widely used as a means of separating particulate from "dissolved" species. Such an operational definition ignores the continuum of suspended particle sizes and the issue of colloids entirely. For environmental biogeochemical investigations this separation may be insufficient to fully describe the processes involved in both organic matter and contaminant transport. Unfortunately, it has only been recently that studies have attempted to rigorously differentiate between dissolved, colloidal, and particulate organic matter in natural waters.

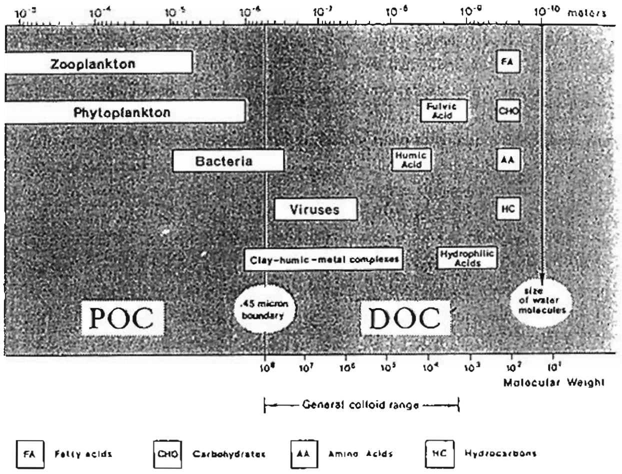


Figure II-7. *Continuum of particulate and dissolved organic carbon in natural waters. (from Thurman, 1985, p. 3)*

Dissolved Organic Matter (DNOM)

Dissolved NOM, in its most rigorous sense, refers to those compounds which are fully solvated by water. This consists of a wide range of compounds which generally share the properties of relatively low molecular weight (<1000 daltons) and significant numbers of polar and/or ionizable functional groups. Most significant in this group are the complex mixtures of polyelectrolytes called humic substances which are described above. Humic and fulvic acid are the major classes of humic substances and are defined by their pH-dependent solubility and their adsorption onto non-ionic acrylic resins. Another class of complex mixtures are known as hydrophilic acids. Hydrophilic acids are similar to fulvic acid but are more soluble, higher in oxygen content, and less readily isolated on acrylic resins. Other major components of DNOM are simpler organic compounds, the major classes being carbohydrates, carboxylic acids, amino acids, and hydrocarbons (Thurman, 1985, Ch. 10).

Experimentally, DNOM is often determined by measurements of dissolved organic carbon (DOC). In general DOC can be converted to DNOM by multiplying by a factor of two. Most modern DOC determinations rely on either chemical or thermal oxidation of the organic matter to CO_2 followed by detection using spectrophotometric or coulometric techniques (See, for example, Lee and Macalady, 1989). DOC measurements are complicated by the low values often encountered, differing oxidation efficiencies for components of DOC, and the presence of interfering species such as chloride (Hedges and Lee, 1992). Distinction between particulate organic carbon (POC) and DOC is made by analyzing raw and filtered samples. As previously pointed out the result of this-two phase classification scheme is that most DOC and POC values will contain variable amounts of colloidal organic carbon (COC). The amount of COC will vary depending on the source of NOM and the aqueous environment.

Colloidal Organic Matter (CNOM)

The major components classified as colloidal organic matter are: high-molecular weight compounds, aggregates of smaller molecules, organic coatings on other colloidal-sized particles (i.e. clays and oxides),and viral and bacterial cells (both viable and senescent). These colloidal materials can be classified as either hydrophilic or hydrophobic colloids. Macromolecular organics such as humic and fulvic acids, polysaccharides, proteins, peptides, and amino acids tend to be polar and maintain their colloidal stability via interactions between charged functional groups and water molecule dipoles. They are best classified as hydrophilic colloids. For hydrophobic solids, permanent dispersion or suspension is maintained by the random thermal activity of water molecules (Brownian movement). In order for this mechanism to be effective, the particle must be sufficiently small to allow spatially uneven bombardment by water molecules and must have at least some electrostatic charge. If not for electrostatic repulsions, the hydrophobic nature of these particles would tend to force them together during particle collisions and promote aggregation which would lead to larger and larger particle sizes and eventual destruction of the colloidal system. These electrostatic repulsions promote coulombic repulsion between particles and also allow interactions with water molecules, thus making the colloids effectively more hydrophilic. The importance of organic coatings to establish a substantial negative charge on mineral colloids and thereby impart significant stability has been widely demonstrated (Tiller and O'Melia, 1993).

The "molecular weight" ranges of aquatic fulvic and humic acids are generally considered to be approximately 500-1000 and 1000-5000 daltons respectively (Thurman, 1985, Ch. 10). These values, along with the significant number of ionizable functional groups, suggest that fulvic acid may best be considered a dissolved species, while humic acid molecules probably span the dissolved-colloidal interface. It should be noted that these molecular weights are generally determined on de-salted, isolated fractions of NOM. Fulvic and humic acids may occur in natural waters as aggregates with much higher apparent molecular weights and exist in colloidal forms. Aggregation is enhanced by the presence of multi-valent cations which reduce the electrical double layer. Ultrafiltration studies of NOM have suggested that a significant portion of NOM exists as colloids with "molecular weights" exceeding 10,000 daltons . In general it is not known whether this CNOM exists as discrete, high-molecular weight molecules or represents some of the other forms of CNOM mentioned above such as aggregates or NOM coatings. Ultrafiltration results must be interpreted carefully as artifacts can result from processes such as charge repulsion by the ultrafilter membrane, clogging of the membrane, and aggregation at the membrane surface (Buffle et al., 1992). Some of the experimental difficulties of using ultrafiltration, particularly membrane fowling have been obviated by the introduction of tangential-flow filtration methods (Figure II-8 and 9, Ranville et al., 1991).

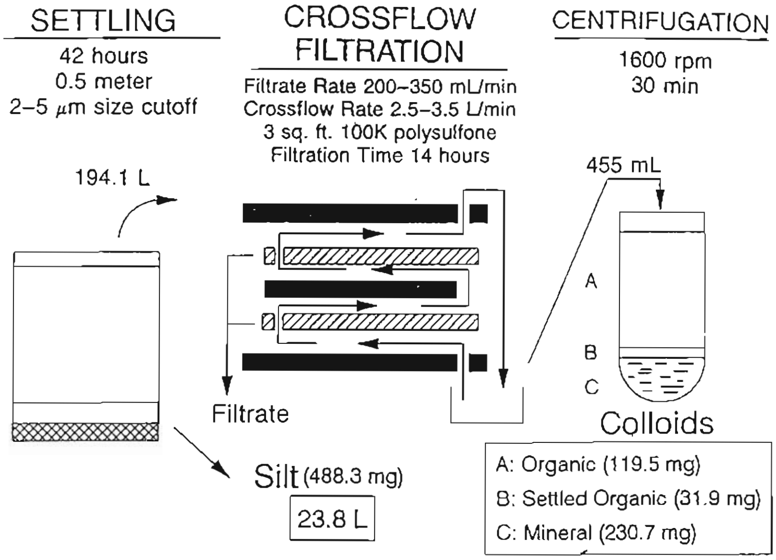
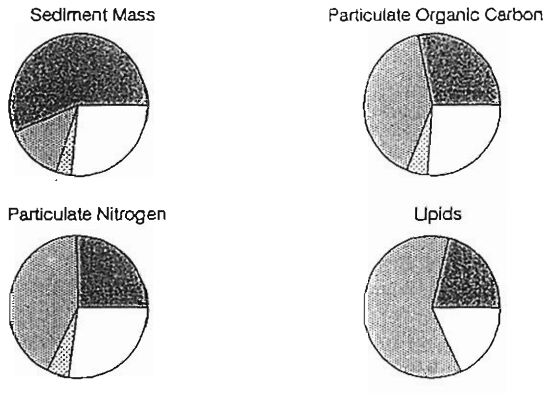


Figure II-8. Diagram showing the operation of cross flow ultrafiltration for collection of a colloid sample.(from Ranville, et al., 1991)



% Distribution	Sediment Mass	Particulate Organic Carbon	Nitrogen	Lipids
▩ Silt	56.1	28.4	24.3	21.0
▦ Organic Colloid	13.7	40.6	40.7	61.1
▨ Settled Organic Colloid	3.7	4.9	6.0	0
☐ Mineral Colloid	26.5	26.1	29.0	17.9

Figure II-9. *Relative mass distribution (in weight percent) of sediment mass, particulate organic carbon, nitrogen, and lipids in the silt, organic colloid, settled organic colloid, and mineral fraction from Pueblo Reservoir. (from Ranville, et al., 1991)*

Some of the CNOM is present as adsorbed coatings on mineral particles. Particles with very similar negative surface charge, regardless of underlying mineralogy, dominate in most aquatic environments, a reflection of the universal presence of organic coatings (Hunter, 1980). These NOM coatings significantly modify the ability of mineral colloids to partition metal and organic contaminants in natural waters. The negative charge imparted by these coatings also facilitates the transport of colloids and associated contaminants in surface and ground waters by preventing aggregation in the surface waters and capture by the aquifer matrix in ground waters. In addition to altering the properties of the mineral colloid, the adsorption process can also significantly alter the distributions of NOM fractions within an aquatic system. McKnight et al. (1992) documented preferential sorption of certain components of NOM to iron and aluminum oxides in an acidic stream. The NOM remaining in solution was shown to be less aromatic than that adsorbed to the oxides. Changes also occurred in the elemental composition and molecular weight.

Pedogenic and aquogenic materials

As outlined above, dissolved and colloidal organic matter in natural waters may be composed of an exceedingly wide variety of natural and synthetic materials, including simple compounds, macromolecules (molecular weight > 1000), biological particles, NOM-coated oxides and phyllosilicates (clays). The relative importance of these various types of organic colloids will vary with the nature of the aqueous environment. For example, it is likely that in productive surface waters, living microorganisms and detrital organic materials will dominate the colloid population. In waters affected by mining wastes or agricultural runoff, iron and aluminum oxides and clays with adsorbed organic matter may be the most abundant colloids. Aquatic colloids originate in at least three ways (Figure II-10):

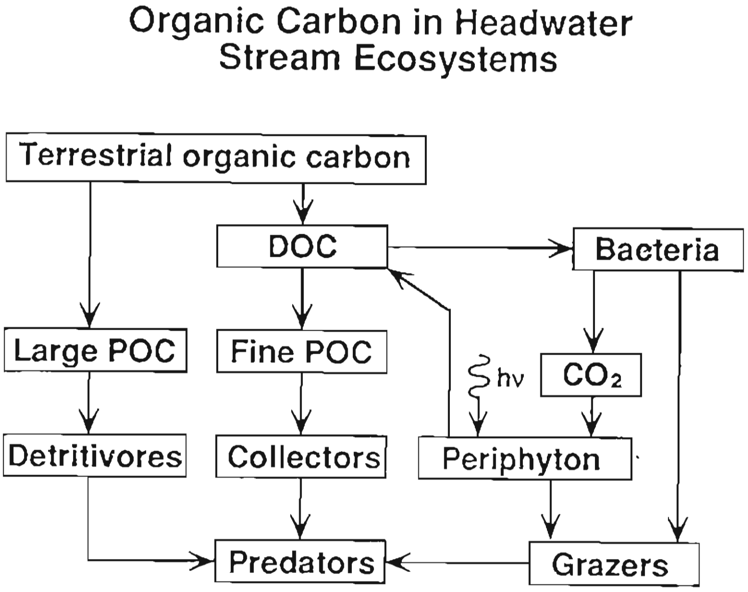


Figure II-10 Schematic diagram representing carbon cycling in a headwater stream.(from Ranville, et al., 1991)

1) Pedogenic NOM includes both DNOM leached out of the soil of the drainage basin (Leppard et al., 1986) and particulate NOM carried in runoff. This latter form, at least in streams and rivers, is comprised principally of plant fragments and woody debris, and may be chemically degraded during transport to the aquatic system (Ranville, et al., 1991)

2) Aquogenic NOM, on the other hand, forms from two distinct types of processes:

a. as a result of biological activity (eg., growth and death of microbes, production of fibrils, organic skeletons and protien-rich cell fragments (Leppard, 1992)); and

b. as a result of in-situ aqueous phase inorganic precipitation due to changing chemical or physical conditions followed by NOM adsorption (eg., precipitation of iron oxides and/or sulfides due to redox changes, formation of complex phosphate mineral suspensions at chemical interfaces in lakes (Buffle, et al. 1989).

III. GEOCHEMICAL REACTIONS OF NOM

The geochemistry of NOM is as complex and varied as the diverse origins and nature of the material suggest. In this section, some of the more important geochemical interactions of NOM with unaltered geological systems are briefly reviewed. For a more complete discussion of any of the topics in this section, refer to the bibliography.

A. Weathering and NOM

The occurrence of specific interactions involving NOM in the chemical weathering of soils and minerals is difficult to determine because such interactions are universally accompanied by a complex array of biological and biochemical reactions as well as simple hydrolysis (Figure III-1). The role of microbial processes in weathering has been discussed in detail by Robert and Berthelin (1986). It is the opinion of these and other (e.g. Loughman, 1969) authors that the direct involvement of NOM in chemical weathering is extremely limited; carbonic and other mineral acids produced by biological activity being the primary agents in chemical weathering.

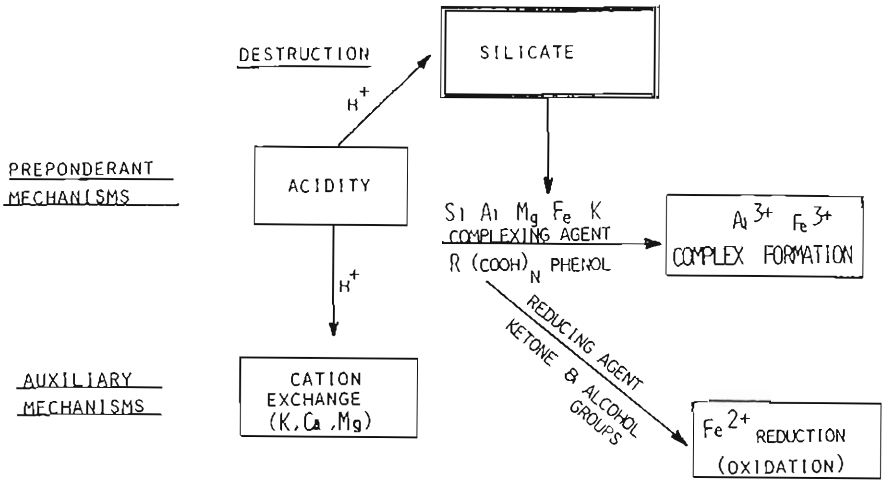


Figure III-1. Mechanisms in the alteration of silicates by organic acids. (from Robert and Berthelin, 1986, p 479)

Nevertheless, both low-molecular-weight biochemical compounds and humic and fulvic acids have been implicated in the degradation of mineral matter in nature. The ability of many microorganisms, particularly lichen fungi, to bring about the weathering of rocks and minerals due to the synthesis of biochemical chelating agents is well-known (Stevenson, 1982). Specific weathering effects of low-molecular-weight organic acids of natural origin are well-documented (Bennett and Siegel, 1987, Mast and Drever, 1987, Schalascha, et al., 1967). with both the acidic and chelating characteristics of the acids implicated in weathering processes. Antweiler and Drever (1983) correlated the weathering of volcanic ash with the DOC of percolating waters. The DOC consisted predominately of humic materials but contained traces of oxalate, acetate and formate with pH values between 4.3 and 5.2.

Data obtained by Baker (1973) show that humic acids exhibit an activity of the same order of as that of several simple organic acid chelating agents in the weathering of a number of minerals and metals. The ability of NOM, more specifically humic substances, to decompose common soil minerals such as biotite, muscovite, illite and koalinite has been demonstrated by a number of authors, including Huang and Keller (1971), Schnitzer and Kodama (1976) and Tan (1980).

Because of their low molecular weights and abundance in soil solutions, fulvic acids may be particularly effective in dissolving silicate minerals (Stevenson, 1985). Zunino and Martin (1977) have advanced a detailed concept of the involvement of humic substances in weathering and subsequent translocation of trace elements to biological systems.

An additional role ascribed to humic substances in weathering processes depends upon their oxidation/reduction (redox)properties. The reduction of iron minerals, releasing ferrous iron into solution is an obvious possibility. However, little direct evidence has been reported for direct involvement of NOM in redox transformations of minerals. The role of NOM in redox transformations of metals, however, is well-established (see below).

B. NOM profiles in soils and associated aqueous systems

The general pattern of NOM content in a soil profile includes a gradation from the highly organic layers containing litterfall from the surface vegetation to the lower highly mineral layers containing relatively little organic content. The precise pattern for a particular catchment depends upon a variety of geographical, ecological and mineralogical factors, as well as the extent of agricultural and other anthropogenic disturbances (Figure III-2). For example, in typical tropical forest soils, the organic matter is largely recycled in the surface soil zones with little or no transport of soluble NOM fractions to underlying layers. In contrast, in regions between the arctic and cooler temperate climates, much of the forested area is dominated by coniferous forests, which produce litterfall rich in phenolic compounds and organic acids (Cronan and Aiken, 1985). In these ecosystems, decomposition is also slower and correspondingly less complete, resulting in a soil solution rich in NOM percolating into the lower soil horizons (Dethier et al., 1988). A detailed discussion of the variations in soil development and concomitant NOM profiles is beyond the scope of this review, but a bief description of soil NOM profiles in a typical Northern forest ecosystem will be presented as an example. For further discussion the reader is referred to Schlesinger's treatise on Biogeochemistry (1988). Much of the discussion below is based on Schlesinger's treatment of temperate forest soils.

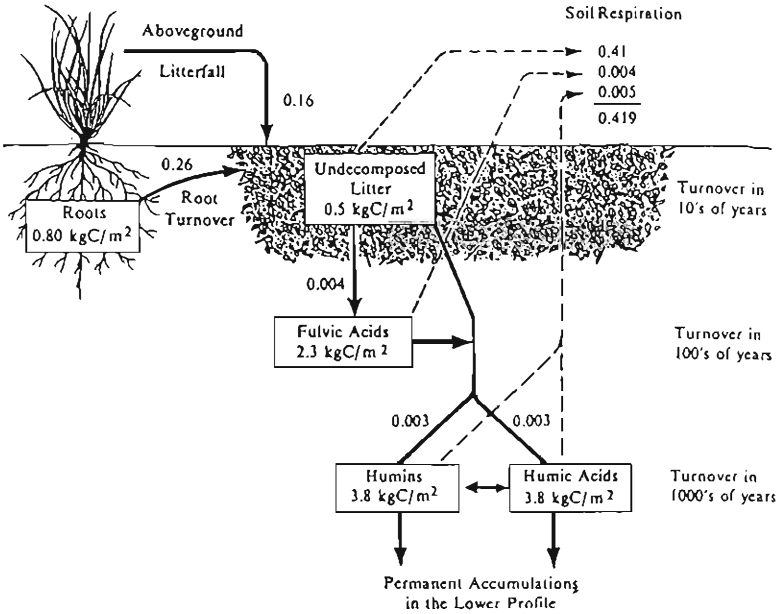


Figure III-2. Turnover of litter and soil organic fractions in a grassland soil. Flux estimates are in kg C/m² yr. (from Schlesinger, 1991, p. 136)

Soil profiles are typically separated into layers, or horizons, which define the characteristics of the soil as a function of depth. Forest soils generally contain an organic layer which is clearly separated from the underlying mineral layers. The organic, or O, layer can be further divided into zones of increasingly-decomposed organic materials on the forest floor, from the largely cellular surface material to a humus layer consisting of amorphous, degradation-resistant organic substances, and often containing significant quantities of minerals. Differentiation of sublayers within this surface zone is often difficult, with large regional and seasonal variations. In temperate and boreal coniferous forests, this organic layer often accumulates into a thick organic mat, known as a mor, that is sharply differentiated from the underlying soil (Romell, 1935). Many arctic soils are waterlogged and contain primarily organic material in the entire rooting zone (peatland soils, or Histosols).

Beneath this organic layer lies the upper mineral soil, designated the A horizon, which has a significant NOM fraction, and varies from several cm to 1 m in thickness. This is a zone of eluvial processes, characterized by significant weathering by organic acids contained in the water percolating through the forest floor. Iron and aluminum are typical removed by chelation with NOM. Such downward movement of Fe and Al in conjunction with NOM is known as *podzolization* (Antweiler and Drever, 1983, Chesworth and Macias-Vasquez, 1985). Known throughout the world, podzolization is particularly intense in temperate to boreal coniferous forest ecosystems (Figure III-3). In such soils, the pH of the soil solution is often as low as 4.0 (Dethier et al., 1988). When podzolization is particularly pronounced, a whitish layer of nearly pure quartz can form at the base of the A horizon (the E horizon, Pedro et al., 1978).

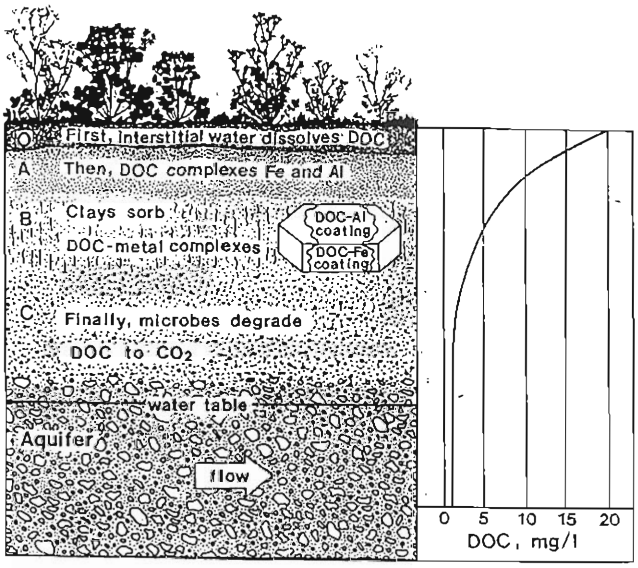


Figure III-3. Podzolization and decrease of organic carbon in interstitial water of soils. (from Thurman, 1985, p. 17)

Substances leached from the A and E horizons are deposited during soil development in the underlying, or B horizon, defined as a zone of deposition or illuvial processes, where secondary clay minerals accumulate. These clay minerals retard the downward movement of soluble NOM fractions that are carrying Fe and Al (Greenland, 1971). Soils of varying degrees of podzolization are characterized by B horizons which contain varying amounts of clay and organic matter. Spodosols are highly podzolized soils which have an upper B horizon (B_S) which is dark and rich in Fe and NOM. Less highly podzolized soils may have B horizons which vary from nearly pure clays to zones which are orange-red to yellow from varying iron contents. In forests in New England (USA), the accumulation of NOM in the B_S horizon appears to control the loss of solution-phase NOM to streams (McDowell and Wood, 1984).

Beneath the B horizon, the C horizon is characterized by soil material with little NOM content. Depending upon whether or not the soil has developed from local materials, the C horizon may or may not resemble the underlying bedrock. In any case, carbonation weathering is generally the dominant process in the C horizon (Ugolini et al., 1977). In many temperate regions, anthropogenic activities such as agriculture and subsequent erosion substantially alter these forest profiles. For example, in the Piedmont region of the southeastern US, the forest floor often resides directly atop the B horizon due to such erosion. Acid rain can also substantially alter this analysis through the introduction of strong acids such as sulfuric into the soil solution, particularly with respect to the weathering of Al.

The NOM content of aqueous systems associated with a given catchment is largely controlled by the soil processes outlined above. The export of NOM from a catchment is largely through DNOM and CNOM in surface and ground waters. The NOM content of ground water is generally a direct result of the nature and extent of the soil processes, though some NOM is imported into a given catchment through rainfall, which typically contains about 1 mg C/L above the forest canopy and 2-3 mg C/L below the canopy . (Rainfall that drips from leaf and plant matter may contain 5-25 mg C/L (Thurman, 1985, p 21)). Because of variations in soil development processes and hydrological considerations, the measured NOM content of ground water and streams draining various catchments shows dramatic regional variations. For example, streams and ponds in very sandy regions, which show little retention of NOM in soils, may have up to several hundred mg/l of organic carbon in the DNOM and CNOM forms. Shallow ground waters in peat lands or bogs may have similarly high NOM contents (see below).

However, NOM concentrations in both surface and ground waters are generally much lower than these extremes (Figure III-4). In a broad study of ground water NOM concentrations, Leenheer et al. (1974) analyzed 100 ground water samples from a variety of aquifer types in the United States, finding a median DOC concentration or 0.7 mg C/L, with a range of 0.2 to 15 mg C/L. Their study and others show that the majority of ground waters have a DOC concentration of 2 mg C/l or less, providing areas dominated by kerogen or petroleum and oil-field brines are excluded. Surface waters in rivers and streams show generally higher NOM levels, with significant fractions often present as DNOM, CNOM and larger particulate forms. Thurman (1985, Chapter 1) lists DNOM concentrations, expressed as DOC, ranging from 2 mg C/L in arctic and alpine streams to above 25 mg C/L in streams draining swamps or wetlands. Inclusion of particulate NOM forms can increase these organic stream loadings considerably. Particulate organic carbon loadings have been estimated to show a range of 1 to 30 mg C/L for 99% of all rivers (Meybeck, 1981, 1982).

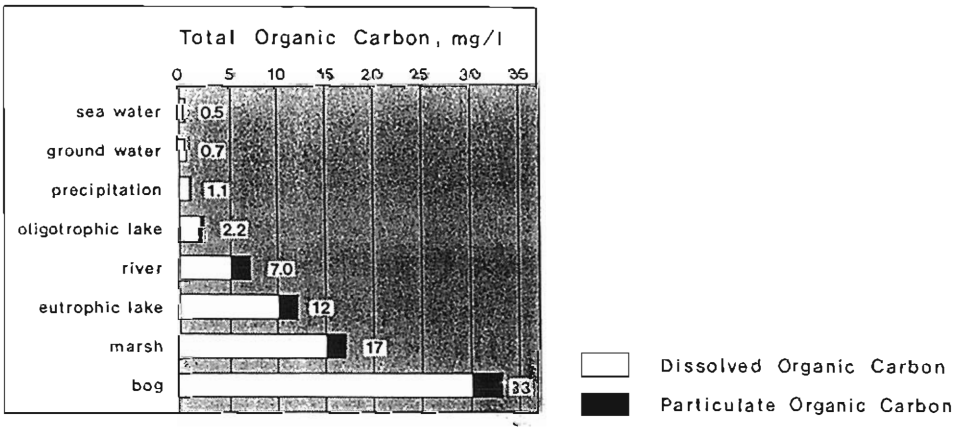


Figure III-4 Approximate concentrations of dissolved and particulate organic carbon in natural waters. (from Thurman, 1985, p. 8)

Thus, the export of NOM by ground and surface water flow is typically a relatively minor process in terms of the overall effect on catchment mass balance. The fraction of pedogenic NOM generated in catchments which is exported via ground and surface water flow is difficult to estimate in a general sense, but these fractions can be estimated for a given catchment given adequate data on NOM production, ground and surface water compositions and hydrology. The problems associated with such estimates are discussed briefly in the following section of this review. Concomitant roles of such exported NOM on the mass balances of other natural and anthropogenic materials may, as discussed in subsequent sections of this review, be a more significant consideration than the fractional export of NOM may indicate.

C. Hydrologic controls on the transport of NOM

Transport of pedogenic NOM to streams is controled by the amount and pathways of water movement through the catchment. Both these variables show seasonal variations which are translated to changes in the NOM concentration and possibly the NOM composition of streams. Flow entering a stream is the sum of the overland and subsurface inputs:

$$\text{Flow}_{\text{stream}} = \text{Flow}_{\text{overland}} + \text{Flow}_{\text{subsurface}}$$

The subsurface component may be more accuratly described as two components (Hornberger et al., 1994)

$$\text{Flow}_{\text{subsurface}} = \text{Flow}_{\text{upper soil}} + \text{Flow}_{\text{lower soil}}$$

The mass balance of chemical species is therefore

$$[C]_{\text{stream}}\text{Flow}_{\text{stream}} = [C]_{\text{overland}}\text{Flow}_{\text{overland}} + [C]_{\text{upper soil}}\text{Flow}_{\text{upper soil}} + [C]_{\text{lower soil}}\text{Flow}_{\text{lower soil}}$$

These equations were incorporated into the hydrological model TOPMODEL (Beven and Kirkby, 1979) by Hornberger et al. (1994).

A seasonality in the relative importance of NOM transport by overland flow versus subsurface flow was found by Chittlborough et al. (1992). They found that most NOM transported in autumn and winter was through subsurface flow wheras in spring and summer overland flow was dominant. During the dry Australian summer the surface organic matter became more hydrophobic which prevented infiltration and therefore overland flow was favored. During the wetter months a well developed system of biopores and channels allowed subsurface flow to predominate. Preferential flow through macropores has been observed by a number of authors (Jardine et al., 1990; Wilson et al., 1990) and is evidenced in part by rapid stream response to rainfall input. These processes have a number of potential effects on the amount and nature of NOM. Overland flow is likely to be enriched in PNOM and possibly CNOM. Flow through macropores reduces the amount of contact between the NOM and the soil matrix and may limit, but not eliminate, the interaction and imobilization of NOM by soil surfaces (Jardine et al., 1989). Preferential flow may also allow enhanced PNOM and CNOM transport and has been shown to allow transport of significant quantities of soil minerals (Chittleborough, in prep) which may carry adsorbed NOM.

Because of the variability of NOM in various soil horizons, it is clear that the relative amounts of flow in the two subsurface compartments defined in the TOPSMODEL will have major influence on NOM transport (Hornberger et al. 1994). During low baseflow conditions (Figure III-5), water travels primarily through the lower soil compartment where NOM conentrations are lower. This NOM is likely composed of less reactive components as interaction with soil minerals will have occured in the upper horizons. During high flow conditions (Figure III-6), a water table ridge may form at the stream-aquifer boundary thus allowing water movement through the upper soil compartment. Not only will the NOM concentration likely be higher, but this NOM may represent more reactive compounds formed from leaf litter leachates, which have had minimal interaction with the soil matrix. This 'flushing' of high NOM concentrations from upper soil horizons is particularly noticable in systems where the stream hydrology is dominated by snowmelt. NOM builds up from microbial activity in the soil during the winter and is rapidly released during snowmelt. NOM concentrations are highest during the rising limb of the spring hydrograph (Hornberger et al., 1994).

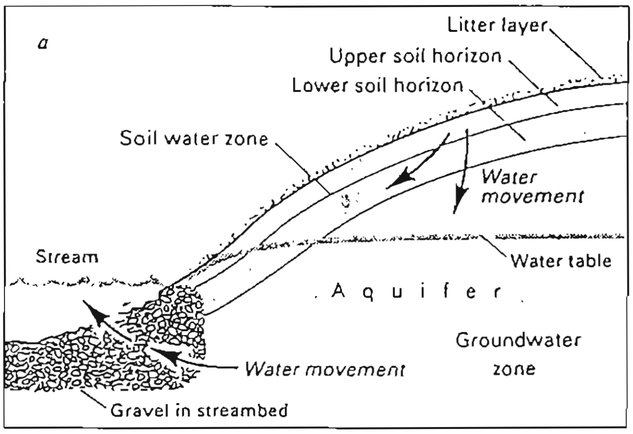


Figure III-5. Schematic diagram of soil-water flushing under low baseflow conditions. (from Hornberger, et al., 1994).

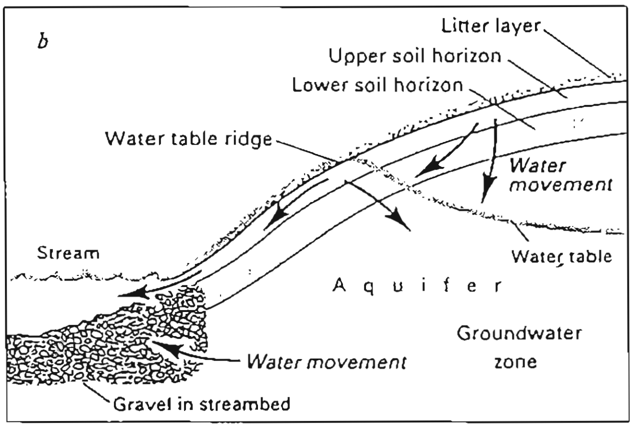


Figure III-6 Schematic diagram of soil-water flushing under high-flow conditions. (from Hornberger, et al., 1994)

Aquogenic NOM may also be influenced by hydrology but is in general more directly related to biogeochemical processes occuring within the stream. Major processes occuring are the utilization of coarse PNOM by heterotrophic organanisms to produce finer PNOM and DNOM and the mineralization of NOM to carbon dioxide. In small streams net primary production is often small so that most of the NOM available for respirtation is allochthonous (Fisher and Likens, 1973). A mass balance of various NOM fractions in a small stream in New Hampshire was obtained by Meyer et al.(1981) and shown in Table 1.

Table 1 *Yearly flux on NOM components (grams/m^2)in Bear Brook, New Hampshire.*

Stream Inputs	
DNOM	260
Fine PNOM	12
Coarse PNOM	340
Gases	1
Total Input	613
Stream Outputs	
DNOM	260
Fine PNOM	25
Coarse PNOM	100
Gases	230
Total Outputs	615

In this case no net production of DNOM occured but significant changes in the nature of the PNOM and the amount of carbon dioxide did occur.

In larger streams/rivers primary productivity by rooted plants and phytoplankton is significant (Lewis, 1988). This atutochthonous NOM is more readily utilized and recyled by the aquatic organicms than is allochthonous NOM. Hydrology plays a role in productivity in that turbidity is often a limiting factor to plant growth (Edwards and Meyer, 1987). In large rivers a large amount of NOM is obtained from the floodplain during seasonal flooding (Cuffney, 1988). Finally it has been shown that total organic carbon load is strongly correlated to discharge with $R^2 = 0.987$ (Schlesinger, 1991).

D. Oxidation/reduction chemistry of metal-organic complexes

One of the most important considerations in the geochemistry of NOM is its role in the transport and transformations of metals. The interactions of NOM with base and soil minerals has already been discussed briefly in the preceding sections on weathering ans soil profile development, The transport of NOM, and necessarily, associated organic and inorganic materials has also been discussed. This and the following section will more fully explore the role of NOM in determining the ultimate fate of the metals released in weathering and soil development processes and associated with NOM as chelated ions.

NOM plays an important role in the cycling of iron in surface waters. Reduction of Fe(III) to Fe(II) in ozygenated waters by NOM can occur by both photolytic and non-photolytic reactions. Non-photolytic reactions involving organic compounds such as oxalate can promote the dissolution of iron oxides by forming soluble iron complexes. These may in turn be reduced in solution by species such as sulfide. Catalysed reductive dissolution of Fe(III) oxides by Fe(II) oxalate complexes is another important mechanism. Organic compounds such as ascorbate can form inner-sphere surface complexes with Fe(III) oxides and diretly transfer electrons. Photolytic reduction of Fe(III) oxides is greatly enhanced by formation of surface complexs with electron donating ligands such as oxalate. The cyling of iron in surface waters is shown in Figure III-7, taken from Stumm (1992).

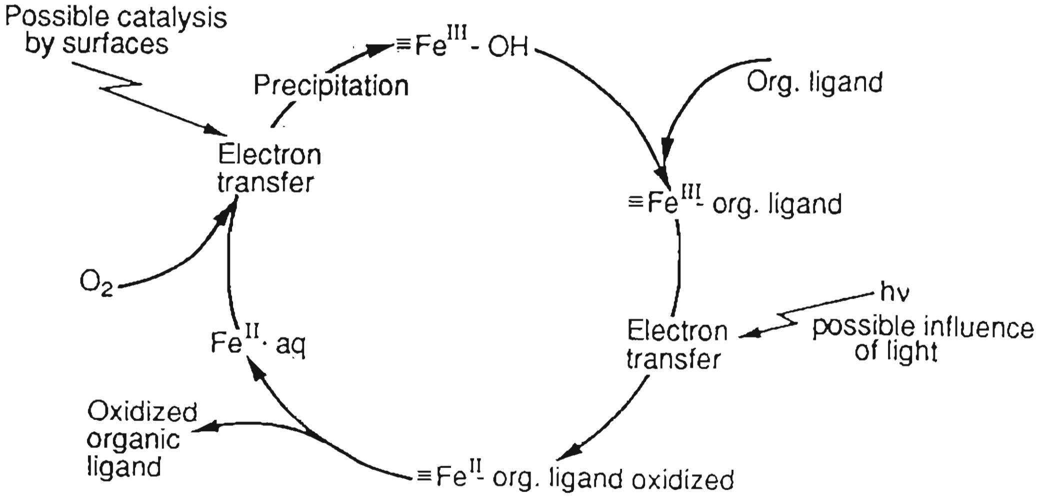


Figure III-7. Schematic representation of the aquatic redox cycling of iron. Triple lines represent the lattice surface of an iron(III)(hydr)oxide. (from Stumm, 1992, p. 362)

In groundwaters, NOM (2 mgC/L) has been found to increase the rate of Fe(II) oxidation at low partial pressures of oxygen (P_{O_2} = 0.005 atm.) by a factor of five (Figure III-8, Liang et al., 1993). In contrast NOM showed little effect on oxidation rates of Fe(II) at high partial pressures of oxygen (P_{O_2} = 0.2 atm.). In addition, total concentrations of Fe(III) species in solution were enhanced by Fe(III) NOM complex formation and stabilization of colloidal Fe(III) hydroxides by adsorption of NOM.

Finally recent observations in our laboratories (Peiffer and Macalady, unpublished data) indicate that NOM may play an additional role in metal redox reactions in the presence of varying amounts of sulfide and oxygen. Reduction of ferric iron, present as non-filterable(0.2 micrometer) ferric-NOM complex, in an iron and NOM rich ground water from a shallow aquifer beneath a wetland, proceeded rapidly upon the introduction of millimolar quantities of sulfide. Upon removal of the sulfide (by N_2 sparging), and the introduction of air, the iron was quickly re-oxidized. However, the NOM was apparently re-oxidized much more slowly. After exposure to air for periods of up to several days, the NOM retained the ability to reduce the iron back to the ferrous state after the solution was sparged with N_2 to remove oxygen. This suggests that NOM reduced by sulfide is reoxidised very slowly by molecular oxygen, providing a kind of redox-buffering of kinetic origin. The extent to which these observations are related to sulfide-NOM interactions as a unique form of redox reactions remains to be demonstrated.

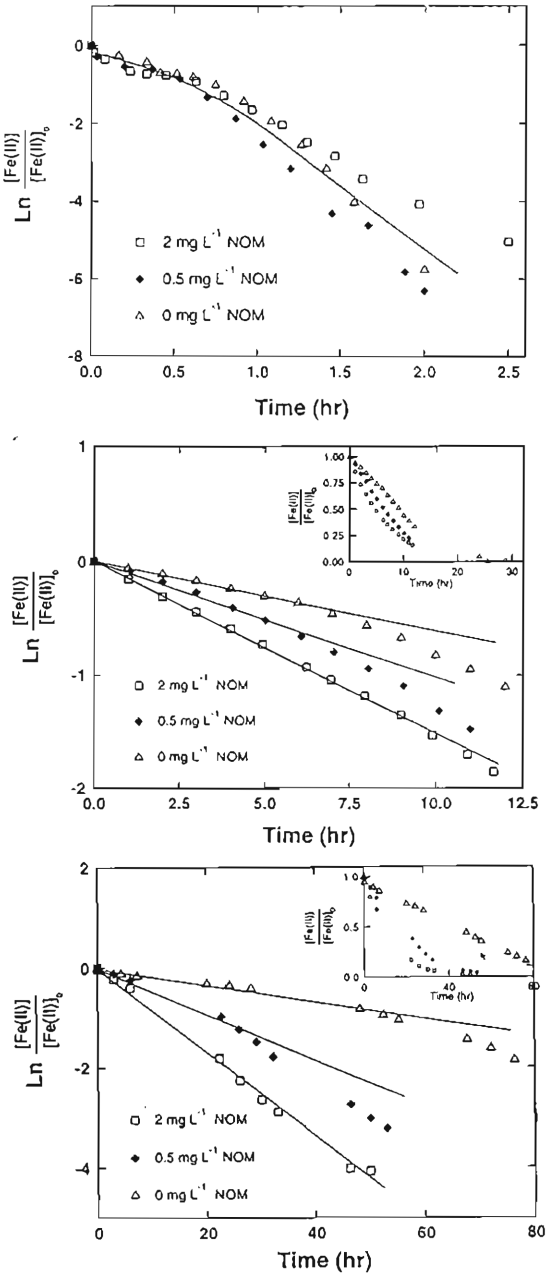


Figure III-8. The effect of NOM on the rate of oxidation of Fe(II) at ca. 20° C. under high (upper frame, $P_{O_2} = 0.2$ atm.), intermediate (middle frame, $P_{O_2} = 0.02$ atm), and low (lower frame, $P_{O_2} = 0.005$ atm) dissolved oxygen concentrations. (from Liang, et al., 1993)

E. NOM and metal ion transport

The binding of a metal ion to an organic complexing agent is generally expressed in terms of a *binding constant.* Binding constants can take a variety of forms depending upon the protonation state of the complexing agent used in the expressions, but they generally represent (for monodentate complexes) the equilibrium ratio of bound metal (metal-organic complex, ML) concentration (or activity) to the product of the concentrations (activities) of unbound ligand (L) and free aqueous metal ion (M), e.g..:

$$HL + M^{n+} \text{<---->} ML^{(n-1)+} + H^+$$

$$K_b = [ML^{(n-1)+}] / [M^{n+}] [HL],$$

where K_b is the binding constant and the square brackets represent concentrations, or more accurately, activities.

Because of the multiplicity of sites for metal complex formation on NOM "molecules," precisely-defined binding constants for NOM or its defined fractions are conceptually difficult. Nevertheless, the binding constant concept can be used in a general way to discuss the relative strengths of metal-binding interactions of NOM in aqueous systems. Binding constants can be used as a surrogate for binding strength.

As previously discussed, the presence of functional groups which may form bonds with metals, suggests NOM is likely to play a major role in metal transport. The strength of the binding of most metals to these groups generally follows the series:

$$-O^- > -NH_2 > -N=N- > -COO^- > -O- > C=O$$

The order of metal binding affinity is then: enolate followed by amines, azo compounds, ring nitrogen, carboxyl, ether, and ketones (Charbeck and Martell, 1959). In some cases the adjacent placement of two groups on the NOM "molecule" is favorable for bidentate complexes, (M_2L, Figure III-9), creating sites with very large binding constants. Some examples of the more important chelation sites, according to Thurman (1985), are: salicylic acid type: aromatic carboxyl and a phenol; phthalic acid type: two aromatic carboxyls; picolinic acid type: aromatic carboxyl and a ring nitrogen; and malonic acid type: two aliphatic carboxyls (especially with a nearby ether group).

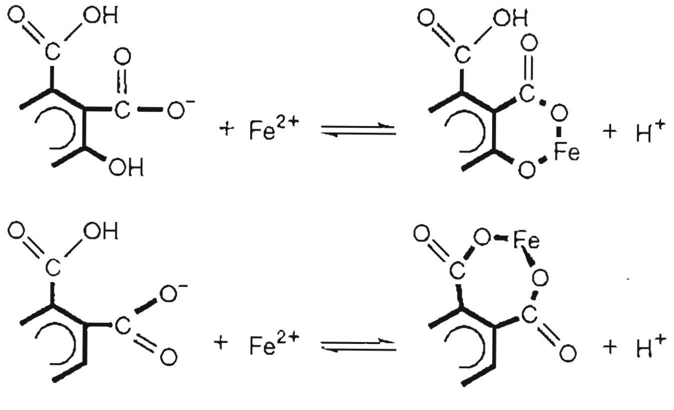


Figure III-9. An example of the formation of a bidentate complex with a metal ion by portions of an NOM molecule. (from Liang, et al., 1993)

The components of NOM which are likely to be most responsible for metal mobility are: humic and fulvic acids, hydrophilic acids, pigments, and amino acids. However, humic and fulvic acids account for roughly 50-90 percent of DNOM, so they are often considered to be the major complexers of metals in most natural waters. Certain other components of NOM, however, such as pigments (e.g.. chlorophyll), other porphyrins, and low-molecular-weight polyfunctional organic acids such as oxalate and citrate are known to have very strong binding constants, perhaps several orders of magnitude greater than humic substances. Though these components are generally present at very low concentrations in natural waters, their extremely high binding constants may impart a measurable impact on metal mobility in certain systems. For example, the upper leaf litter zone of forest floor soils, where pigments may influence metal transport. Amino acids constitute about 3-5 percent of DNOM. Their binding constants are similar to that of humic substances (Tuschall and Brezonik, 1980). Simple acids such as acetate, formate, etc. have been identified in water and their metal binding constants are generally well known (Martel and Smith,1977,1982). Little is known however, about the more complex components of the hydrophilic acid class, which can account for up to 50 percent of DNOM. McKnight et al. (1983) determined copper binding constants for hydrophilic acids and found them similar to those of fulvic acid.

Schnitzer and Khan (1978) reviewed the early work on metal binding by pedogenic humic substances. Most workers found that in general the order of metal binding constants follow the Irving-Williams series (1948, as quoted by Thurman, 1985 page 415). Part of the series found for pedogenic humic substances is:

$$Pb^{2+} > Cu^{2+} > Ni^{2+} > Co^{2+} > Zn^{2+} > Cd^{2+} > Fe^{2+} > Mn^{2+} > Mg^{2+}$$

and at pH = 3

$$Fe^{3+} > Al^{3+} > Cu^{2+} > Ni^{2+} > Co^{2+} > Pb^{2+} > Ca^{2+} > Zn^{2+} > Mn^{2+} > Mg^{2+}$$

For aquogenic humic substances Mantoura et al. (1978) found:

$$Hg^{2+} > Cu^{2+} > Ni^{2+} > Co^{2+} > Mn^{2+} > Cd^{2+} > Ca^{2+} > Mg^{2+}$$

In addition to the magnitude of the binding constants, the number of sites present on DNOM is important to the mobility of metals. The estimates of the number and strength of sites varies considerably based on the fraction of NOM studied, the environment of isolation, and the metal of interest. As an example, McKnight and Wershaw (1989) studied Cu^{2+} binding by aquatic fulvic acid. They found that the Cu^{2+} binding could be modeled using two types of sites: a strong binding site with a $\log K_b = 8.1$, present at a level of 2.7×10^{-7} moles per mg carbon; and a weak site with $\log K_b = 5.9$, present at 1.6×10^{-6} moles per mg carbon. This result suggests that a few micromoles of copper will be complexed by fulvic acids in most freshwaters. This will also be influenced by competition with major divalent cations and other trace metals. The exact role of DNOM in metal transport is still poorly understood.

313

IV. INTERACTIONS BETWEEN NOM AND ANTHROPOGENIC CHEMICALS

A. Transport of pollutant metals as NOM complexes

NOM in an aquifer exists as DNOM or CNOM in the aqueous phase and as adsorbed organic coatings on the aquifer matrix. Consequently the role of NOM as either a sink for metal contaminants or a facilitator of transport is complex and difficult to ascertain. It is likely that NOM may act in both roles with different NOM components contributing in different ways. Consider for example the role of NOM in radionuclide transport. Righetto et. al. (1991) studied the sorption of actinides to mineral oxides in the presence of humic acid (Figure IV-1). The humic acid was strongly sorbed to the oxides and significantly enhanced the actinide sorption. Thus the humic acid component of NOM can act as a sink for actinides through adsorption. Conversely, McKnight et al. (in prep) observed that 70-85 % of the Pu present in surface and ground water at a nuclear weapons facility was associated with fulvic acid. Their results further suggested that fulvic acid contains very strong binding sites for Pu. Marley et al.(1994) performed a forced-gradient tracer injection using surface water NOM and monitored the breakthrough of NOM and actinides. Their NOM consisted of 81-98 % fulvic acid and only 2-19 % humic acid. Their results indicated that americium was transported by lower molecular weight components which were less sorbed than the more hydrophobic components. These studies and others (Figure IV-2)suggest that the fulvic acid component of NOM may provide a means of facilitated transport of actinides.

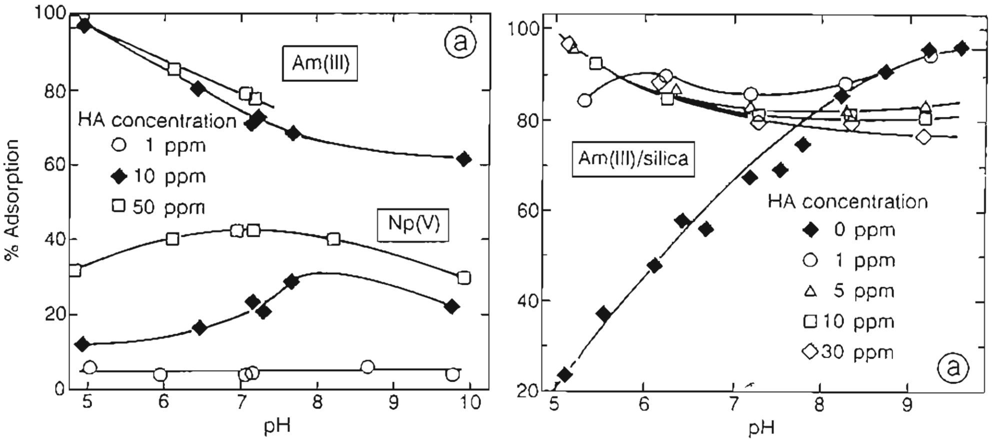


Figure IV-1. Left : pH dependence of Am(III) and Np(V) association with particulate humic acids.
Right : Adsorption behavior of Am(III) in the amorphous silica (1200 ppm)-humic acid system as a function of pH and humic acid concentration.
(from Righetto, et al., 1991)

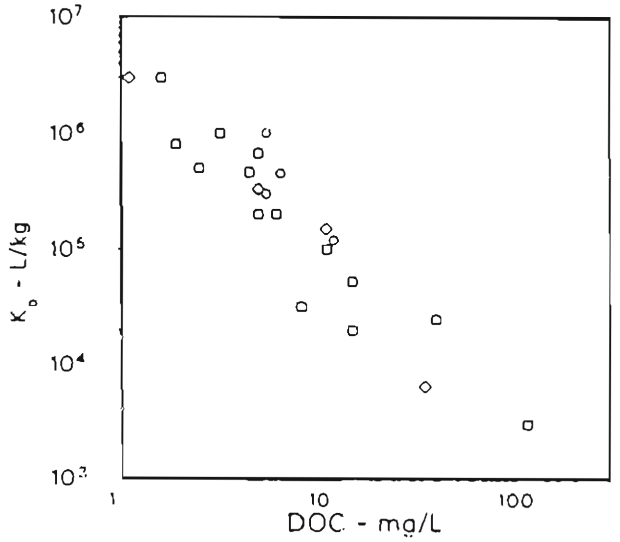


Figure IV-2. *Variation of K_d of reduced plutonium as a function of total DOC (from Nelson, et al., 1985)*

Another example of metal transport by NOM complexes was the work of Dunnivant et al. (1992). They performed column experiments using aquifer sediments and NOM collected from a stream near a peat deposit to study cadmium transport. At an NOM concentration of 5.2 mg C/L the breakthourgh volume was reduced by roughly 25 % over that when NOM was absent. Increasing the NOM concentration further enhanced the breakthrough although the relative effect was not as great per mg C/L (Figure IV-3).

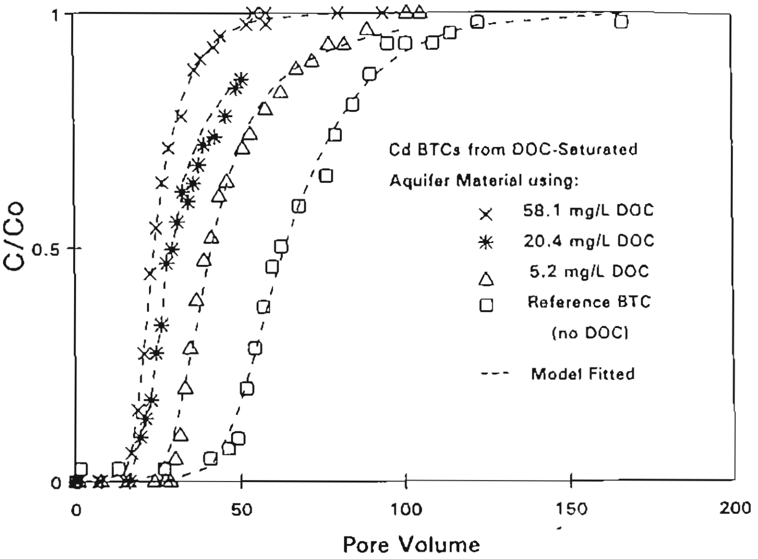


Figure IV-3. *Breakthrough of cadmium from DOC-saturated aquifer material (0.058% by weight) in the presence of four different DOC concentrations. Dashed lines represent model-fitted curves. (from Dunnivant, et al., 1992)*

These two examples, along with numerous other literature sources and the large body of evidence on metal NOM complexes in surface waters, suggests that metal NOM complexes are a very important if not predominant transport mechanism.

B. Effects of sorption/partitioning to NOM on the transport of organic contaminants

Hydrophobic interactions between soil or sediment organic matter and sparingly-soluble, non-ionic organic chemicals has been characterized in terms of a partitioning (or sorption) of the organic chemical into the NOM fraction of the soil or sediment. In a series of papers, (Karickhoff, et al., 1979; Karickhoff, 1980, 1981, 1984; Chiou et al., 1983; Voice and Weber, 1983), this partitioning was shown to be adequately described in terms of an equilibrium between the dissolved and particulate phases. The partition coefficient, K_p, defines the ratio of the concentration of the organic contaminant in the solid soil or sediment phase, expressed on a weight or mole of contaminant per kilogram of solid basis to the concentration in the aqueous phase, also expressed on a weight or mole per kilogram basis, or:

$$K_p = C_s/C_w$$

(In some literature, K_p called K_d, the solid-water distribution coefficient). This relationship is valid over a wide range of conditions for dilute (C_s less than 10^{-5} M or 1/2 the aqueous solubility) solutions of hydrophobic contaminants. The value of K_p for a particular soil or sediment and a particulate contaminant is of course related to both the characteristics of the solid phase and the nature of the contaminant organic chemical. For most soils or sediment samples (f_{oc} greater than about 0.002, Schwarzenbach et al., 1993, Abdul et al., 1987), the only relevant soil characteristic is its NOM fraction, and, for a given contaminant and a series of soils or sediments,

$$K_p = K_{oc} (f_{oc}),$$

where f_{oc} represents the NOM fraction of the soil or sediment expressed as a fraction of the soil mass which is organic carbon (Figure IV-4). K_{oc} is a characteristic of the contaminant chemical and represents its hypothetical partitioning to the organic carbon fraction of the soil or sediment. For soils with a dominant sand size fraction, the NOM content of the finer size fractions (<50 micrometers) has been shown to be most useful for predictions of solid/water partitioning. For most soils and sediments, whole soil f_{oc} values can be used without substantial error (Karickhoff, 1981).

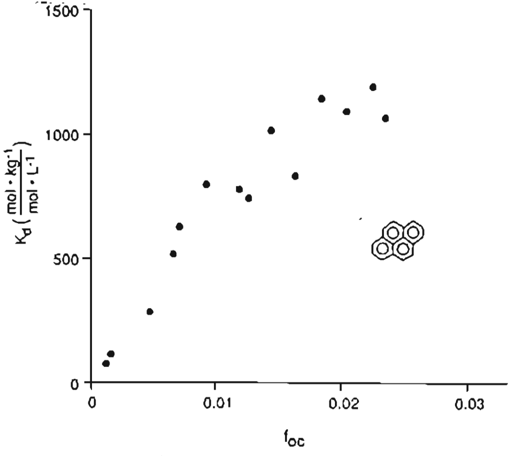


Figure IV-4. Observed increase in solid-water distribution ratios for a hydrophobic compound, pyrene, as a function of the NOM content of the solid (measured as organic carbon) in a variety of soils and sediments. (from Schwarzenbach, et al., 1993. p. 268)

Values of K_{oc} for contaminant organic chemicals were in turn shown to be related to a variety of molecular parameters which attempt to quantify the hydrophobicity of the organic substance. The most widely discussed of such parameters is the octanol/water partitioning coefficient, K_{ow}, which has been measured for a large number of chemicals of interest in studies of contaminant behavior (Pomona College, 1984). Measured values of K_{ow} for pollutant organic chemicals range over at least seven orders of magnitude. Many empirical and semi-empirical relationships between K_{ow} values and other properties such as water solubility and molecular size (Karickhoff, 1981 and others), and molecular size and topology indices such as molecular connectivity (Bahnick and Doucette, 1988) have been developed. The most widely-used relationships fro the estimation of K_{oc} values involve empirical relationships of the form:

$$\log K_{oc} = a \log K_{ow} + d$$

A large number of such relationships have been reported in the literature (for example, Figure IV-5 and Brown and Flagg, 1981; Schwarzenbach and Westall, 1981; Karickhoff, 1981; Hassett et al., 1983,; Kenaga and Goring, 1978; Chiou et. al., 1983). Baker and Mihelcic (1994) recently developed a relationship valid for log K_{ow}'s between 1.59 to 7.32 (spanning a wide variety of chemical classes) which gives a value of 1.015 for "a" and -0.404 for "d" which accounts for 90% of the variability in the data. Such correlations can provide an adequate approximation for the prediction of the equilibrium partitioning of a wide variety of hydrophobic organic chemicals to soil or sediment NOM, but more precise predictions of K_{oc} are not possible using existing data and models.

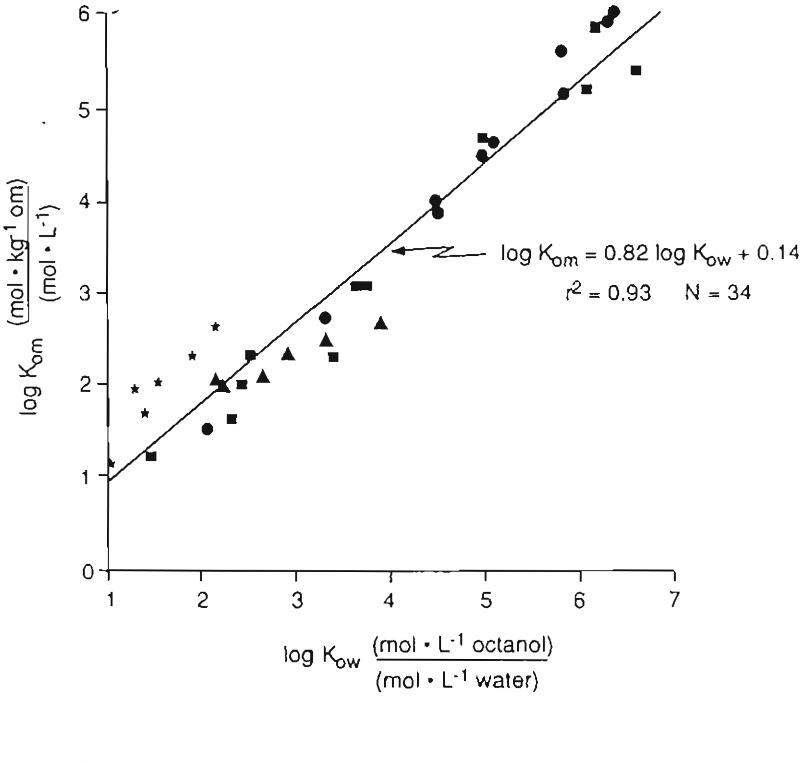


Figure IV-5. Relationship of log K_{om} (K_{oc}) and log K_{ow} for a series of neutral organic compounds: (●) aromatic hydrocarbons, (•) chlorinated hydrocarbons, (△) chloro-S-triazines, and (*) phenyl ureas. (from Schwarzenbach, 1993, p. 275. Data compiled by Karickhoff, 1981)

Nevertheless, the role of NOM in solid phases in the transport of hydrophobic organic chemicals in soil and aqueous systems is well-established and widely accepted as a primary means of the retardation of pollutant migration in soils, streams and ground waters. Knowledge of the organic carbon content of a soil or sediment can be used to predict the aqueous phase concentration of a given hydrophobic chemical in an associated aqueous phase and provide an accurate indication of the retardation of pollutant migration in a given system.

Numerous problems with such predictions have been discussed in the literature, including non-equilibrium between solid and aqueous phases caused by slow kinetics in the partitioning process (Karickhoff, 1984), the role of mobile colloidal phases in considerations of pollutant transport (see below), and the difficulties associated with predictions for marginally hydrophilic substances and ionizable organic compounds (Grundl and Small, 1993) and for solid phases of very low organic carbon contents (Southworth and Keller, 1985). All of these illustrate that the application of such partitioning estimates to a given system must be made with attention to the limitations of the formalism used. In the best of situations, the partition coefficient can be predicted for a given system to within 10-20% based on f_{oc} and K_{ow} values. Values accurate to within a factor of two are possible for most chemicals in most soils or sediments (Karickhoff, 1981).

C. NOM effects on hydrolytic reactions

The degradation chemistry of many anthropogenic organic compounds is dominated by abiotic or biologically-mediated hydrolysis reactions. Hydrolysis refers to nucleophilic attack on a molecule by water, resulting in the breakdown of the molecule into smaller, more hydrophilic products. Examples of hydrolysis reactions relevant to considerations of ground and surface water pollution are shown in Figure IV-6.

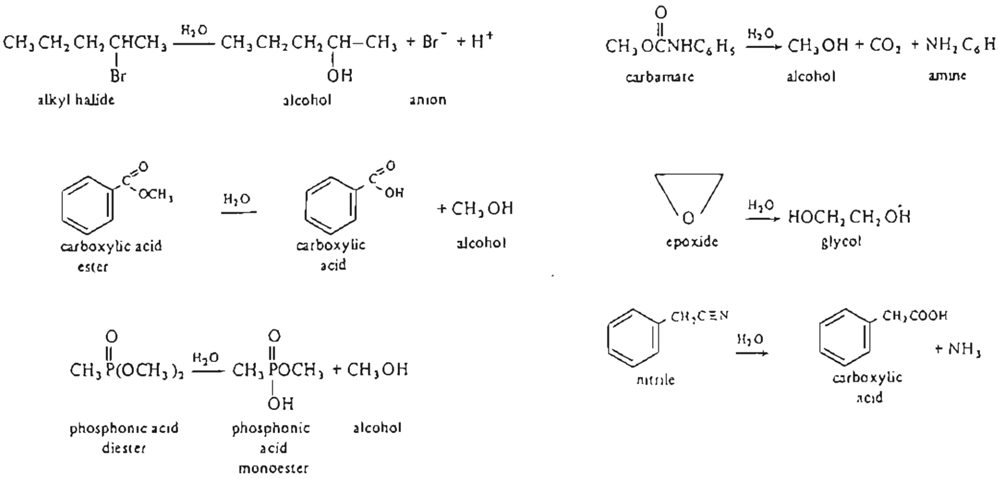


Figure IV-6. Some examples of hydrolysis reactions and reaction products. (from Harris, 1982)

Abiotic hydrolysis reactions, which are commonly fast compared to biologically-mediated reactions for many contaminant molecules in most aquatic systems (Macalady, et al., 1989) are formally second order reactions, and involve pH independent and base or acid-catalyzed components. At constant pH, hydrolysis reactions assume a pseudo-first order character, with the observed rate constant expressed formally as:

$$k_{obs}\,(\text{time}^{-1}) = k_A[H^+] + k_W + k_B[OH^-],$$

where k_A, k_W, and k_B are the acid-catalyzed, pH-independent, and base-catalyzed rate constants respectively, and $[H^+]$ and $[OH^-]$ represent acid and base concentrations. Observation of hydrolysis kinetics at constant temperature and a series of fixed pH values enables determination of all three rate constants. Studies of the organophosphorothioate insecticide chlorpyrifos, for example, reveal measurable pH-independent and base-catalyzed contributions only, with values in distilled water at 298 K of (6.2+/-0.9)x 10^{-6} min^{-1} and 0.5+/-0.2 $M^{-1}\text{min}^{-1}$ respectively (Macalady and Wolfe, 1983). Figure IV-7 shows acid and base-catalyzed hydrolysis data for the herbicide atrazine.

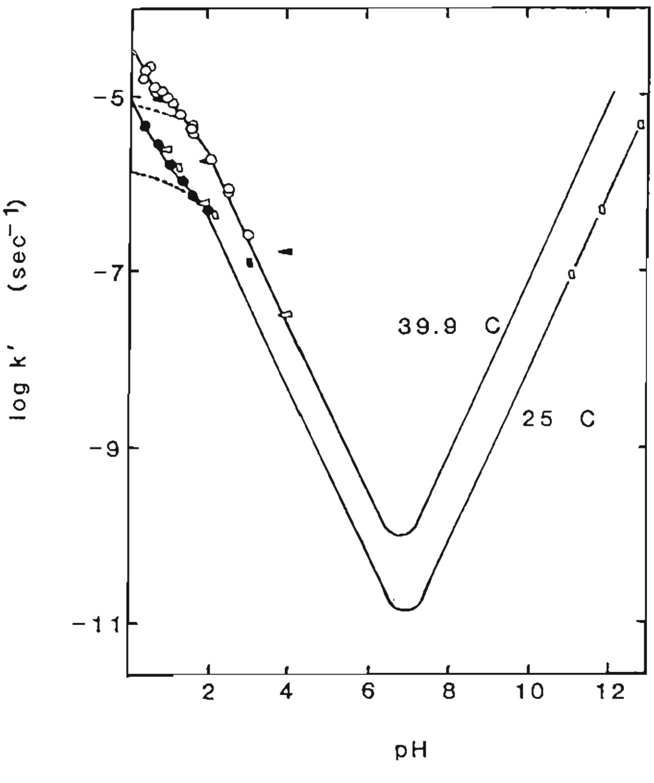


Figure IV-7. The pH-rate profile for the hydrolysis of the aromatic chlorine of atrazine. (from Plust, et al., 1981, as reproduced in Macalady, et al., 1989)

Since most pesticides and many other organic contaminants are quite hydrophobic and would therefore be expected to partition to NOM fractions of soils, sediments, and colloids, it is important to assess the effect of NOM-pollutant interactions on hydrolysis reactions. Beginning with the work of Macalady and Wolfe (1984, 1985), a substantial body of evidence has established the effects of NOM in both solid and aqueous phases on hydrolysis reactions. Partitioning to sediment or colloidal aqueous phase NOM has no measured effect on values of pH-independent hydrolysis rate constants (cf. distilled water). Base-catalyzed hydrolysis reactions, on the other hand, are generally considerably retarded (by factors of 10-1000) for the fraction of a given substance associated with particulate, colloidal, or dissolved NOM as compared to the rate for the molecules in the associated dissolved phase (Figure IV-8). The effects of NOM on acid-catalyzed hydrolysis reactions are less well characterized, partly due to the fact that observations of acid-catalyzed hydrolysis are relatively uncommon for organic contaminants, especially at pH values at which NOM retains its net negative surface charge. In addition, other pH effects, such as inactivation of soil enzymes, may be more important in the acid regime (Macalady, et al., 1989)

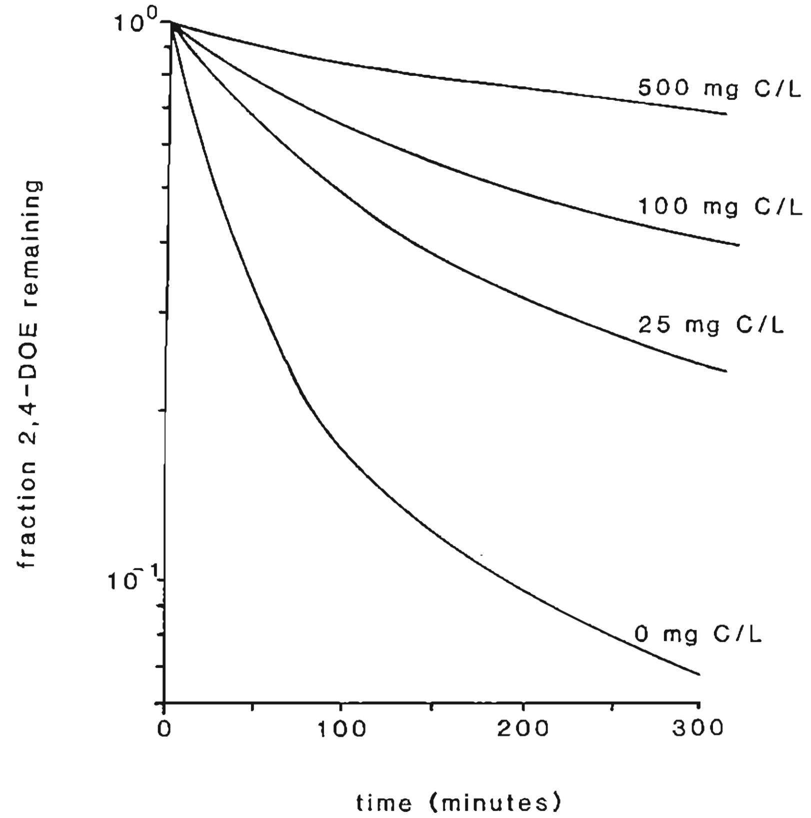


Figure IV-8. Effect of partitioning to DNOM and CNOM on the hydrolysis of the n-octyl ester of 2,4-dichlorophenoxyacetic acid at pH 10 with varying amounts of NOM. (from Perdue and Wolfe, 1982)

Thus, the presence of NOM may have a negligible or substantial effect on the rates of hydrolyses of organic contaminants depending upon the relative importance of acid, base, and pH-independent contributions to the hydrolysis reactions at the ambient pH of the aqueous system and, of course, on the hydrophobicity of the contaminant molecule.

D. Oxidation/reduction reactions facilitated by NOM

Oxidation/reduction (redox) reactions represent another important mode of transformation of anthropogenic organic chemicals in aqueous systems. These reactions can be defined as electron-transfer reactions, in which electrons are exchanged between components of an aqueous system and the contaminant molecule. As a consequence of a redox reaction, both dramatic and relatively subtle changes in the characteristics of organic contaminants can be realized. For example, biologically-mediated oxidations of a wide variety of organic substrates, both natural and anthropogenic, result in the conversion of the carbon/hydrogen framework of organic molecules to carbon dioxide and water, a process often referred to as "mineralization."

Less complete oxidations also occur, as in conversions of carbohydrates to simple organic alcohols or acids, thioethers to sulfoxides and sulfones, or amines to the corresponding nitroso or nitrate derivatives. Reduction reactions generally produce less dramatic changes, including the important class of reductive dehalogenation reactions and the inverse reactions of the partial oxidation processes listed above. Comprehensive reviews of redox reactions of anthropogenic organic chemicals in aquatic systems have been provided by Macalady, et al. (1986) and Wolfe and Macalady (1992).

A clear definition of redox reactions is complicated by the wide variety of redox reactions which have been characterized for organic molecules. Simple definitions of redox reactions as involving the addition or removal of hydrogen or oxygen to an organic substrate are helpful but not sufficiently comprehensive or unambiguous. A more useful definition involves the separation if the reaction into oxidation and reduction *half reactions*. In aqueous systems, this procedure involves separating related pairs of reactants and products and adding water and/or protons to provide an elemental balance for the reactant/product pair. Electrons are then added to provide charge balance. Half reactions with electrons as reactants are reductions, those with electrons as products are oxidations. For example, the oxidation of ethanol to acetic acid by molecular oxygen is written:

$$CH_3CH_2OH + O_2 \text{---->} CH_3COOH + H_2O$$

The relevant half-reactions are:

$$CH_3CH_2OH + H_2O \text{--->} CH_3COOH + 2H^+ + 2e^- \text{ (oxidation)}$$

$$O_2 + 2H^+ + 2e^- \text{---->} 2H_2O \text{ (reduction)}$$

Another example is the reduction of 1,2-dichloroethane to ethene by iodide ions:

$$CH_2Cl\text{-}CH_2Cl + 2I^- \text{---->} H_2C{=}CH_2 + I_2 + 2Cl^-$$

Which has the following half-reactions:

$$CH_2Cl\text{-}CH_2Cl + 2e^- \text{---->} H_2C{=}CH_2 + 2Cl^- \text{ (reduction)}$$

$$2I^- \text{---->} I_2 + 2e^- \text{ (oxidation)}$$

Because the nature of the oxidation or reducing agents for redox transformations of anthropogenic chemicals in environmental aqueous systems are often unknown, it is often necessary to express the pollutant transformation half reaction without being able to write the corresponding half reactions for the environmental oxidant or reductant.

321

Because oxidation of organic molecules is the primary activity of a diverse and widespread group of biological systems, and because many organic molecules of concern in the pollution of aqueous systems are intentionally designed to be resistant to oxidation, reductive processes are of considerably more interest in considerations of the redox transformations of anthropogenic organic chemicals. A wide variety of reductive transformations of anthropogenic organic chemicals have been observed in natural systems. A few examples are illustrated in Figure IV-9.

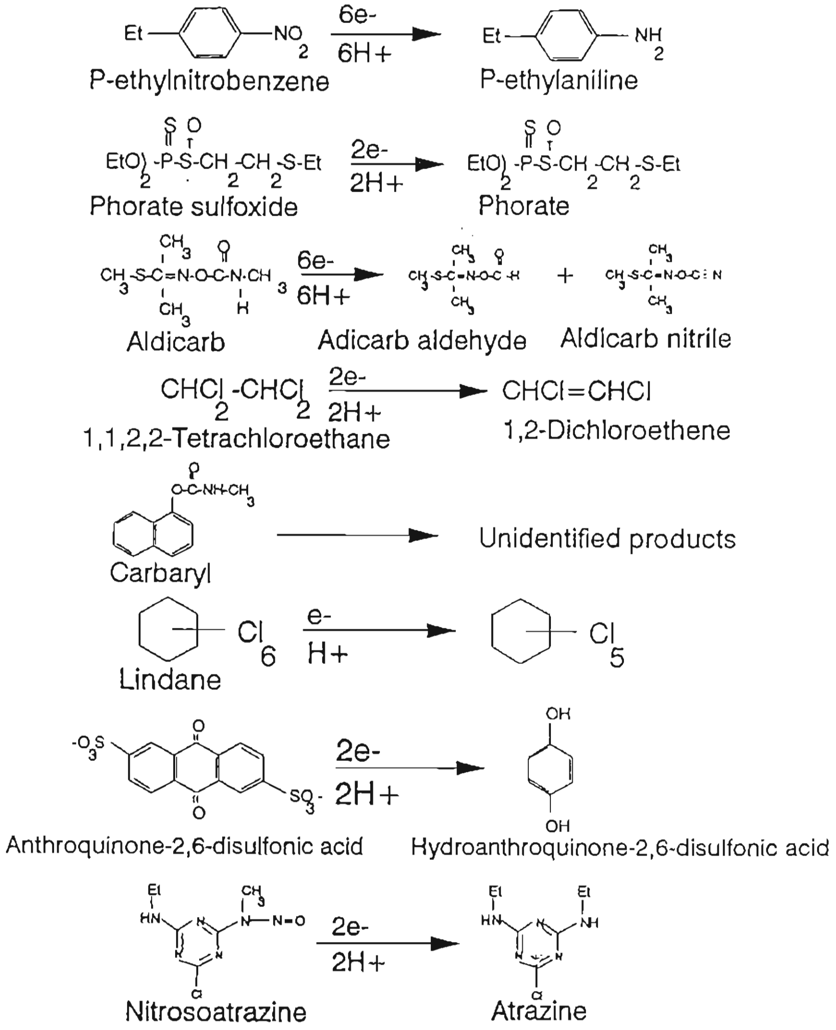


Figure IV-9. Examples of reductions which have been shown to occur in environmental samples under laboratory conditions. (from Wolfe and Macalady, 1992)

Considerations of the role of NOM in reductive transformations have only been studied in detail relatively recently. In some senses the present understanding of the role of NOM in reductive transformations began with the observation of the extremely rapid reduction of the nitro-aromatic group on the pesticide methyl parathion to the corresponding aromatic amine. Experiments in our laboratories at the Colorado School of Mines with sediment/water slurries sampled in a variety of local ponds and reservoirs demonstrated first-order reductive transformations with half-lives from several hours to as fast as 15 seconds. Reactions proceeded in open beakers of unmodified sediment slurries, with no precautions to exclude atmospheric oxygen. Related studies at the US. Environmental Protection Agency (USEPA) laboratories subsequently demonstrated rapid reduction of methyl parathion in additional sediment/water samples. Efforts to correlate reduction rates with sediment parameters revealed a direct relationship of the rate of these reductions to the NOM content of the sediments, with reaction rates also directly proportional to the sediment to water ratio in the slurries (Wolfe, et al., 1987)

Similar results were obtained for reductions of azo compounds (Weber and Wolfe, 1987) and halogenated organic compounds (Peijnenburg et al., 1992). Direct involvement of microbial systems in these reactions seemed improbable due to the extremely fast reaction rates observed in some systems. The emerging hypothesis was that NOM played key role in the observed transformations.

Additional efforts were clearly necessary in order to more precisely define the nature of the role of NOM in the reduction of nitroaromatic compounds. Nitro-reduction can be viewed as a series of 2 electron, 2 proton reductions to intermediate reduction products:

$$--NO_2 \xrightarrow{2e^-,\ 2H^+} --NO\ (+\ H_2O) \xrightarrow{2e^-,\ 2H^+} --NHOH \xrightarrow{2e^-,\ 2H^+} --NH_2$$

though there is little evidence that these are the actual mechanistic steps in the reaction.

One series of efforts to characterize the role of NOM in these reductive processes began with the hypothesis that functional groups in NOM similar to quinone/hydroquinone redox pairs (Figure IV-10) are responsible for the observed NOM-related redox activity. Reduction of methyl parathion and a series of substituted nitrobenzenes and nitrophenols in homogeneous aqueous systems containing reduced model quinones such as anthroquione disulfonate and the common redox indicator dye indigo carmine (Tratnyek and Macalady, 1989) and natural napthoquinone compounds such as juglone and lawsone (Schwarzenbach et al., 1990), demonstrated the ability of quinones to reduce nitroaromatic compounds.

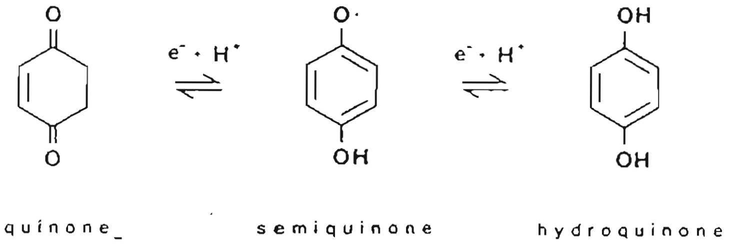


Figure IV-10. Redox half-reactions for the conversions among quinones, semiquinones and hydroquinones. (from Macalady, et al., 1986)

In the studies using model, highly water soluble quinones, an important aspect of the reduction of nitroaromatics by hydroquinones was revealed. Studies of the reaction rate as a function of pH demonstrated increasing reaction rates as the pH increased from 5 to 9. This is precisely the opposite pH trend expected based on the involvement of 6 protons in the overall reduction sequence. This suggests two important features of quinone-mediated reductions. First, the rate-limiting step in the reduction reaction probably does not involve the addition of a proton to the molecule. Second, the reductive reactivity of hydroquinones must increase dramatically as the pH is increased. Since hydroquinones are diphenols, this further suggests that the anionic forms of hydroquinones are more effective reductants than the neutral molecule. These possibilities were supported in the first study quoted above by the fact that the rate of the reduction of methyl parathion by the model quinone compounds was shown to be directly proportional over the pH range 5-9 to the concentration of the hydroquinone anion.

In experiments using natural quinones, an important additional feature of this NOM-mediated reduction was elucidated. Since juglone, lawsone, and many other natural quinones have limited water solubility, and thus very limited capacity to reduce nitroaromatic compounds, a more abundant bulk reducing agent must be involved if such compounds and similar functional groups in the larger array of NOM molecules are to be postulated as responsible for the reduction of nitrobenzenes in homogeneous systems. Addition of a large excess of sulfide, which itself reduces nitrobenzenes only very slowly, to solutions of juglone or lawsone facilitated rapid reduction of nitrobenzenes to the corresponding anilines. A model of the role of quinones in these reductive processes emerged, illustrated for a more general case in Figure IV-11, in which a bulk electron donor, in this case sulfide, rapidly reacts with the mediator, which in this case would be represented as quinones, to reduce them (to the corresponding hydroquinones), which in turn react more slowly to reduce the nitroaromatic compounds. Sulfide reacts much more slowly with the nitroaromatic compounds, producing little or no observed reductions in the time scale of these experiments (several days). Furthermore, pH variations in these reaction rates followed a pattern similar to that observed in the sulfide-free solutions of the more water-soluble model quinones. Studies with a series of substituted nitrobenzenes further revealed a clear relationship between the reduction rate constants in a solution of fixed pH and sulfide and napthoquinone concentrations and the one-electron redox potential of the nitrobenzenes. This confirms the notion that the rate-limiting step for these reactions is the addition of a single electron to the nitro group.

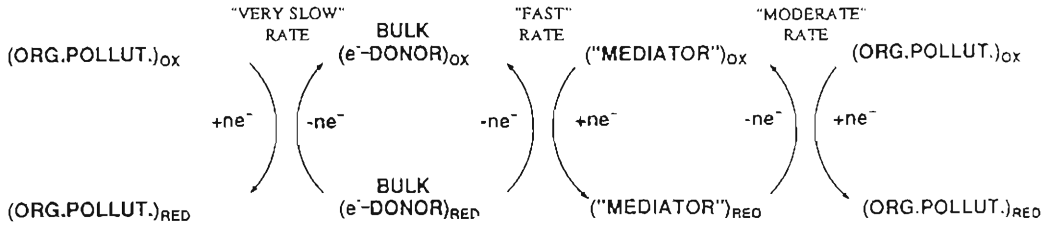


Figure IV-11. Model for the mediation of redox reactions by NOM in natural waters. (from Dunnivant, et al., 1992)

Subsequent studies by Macalady and Dunnivant, conducted at the Swiss Federal Institute for Water Resources and Water Research (EAWAG) with homogeneous (0.2 micrometer filtered) solutions of natural organic matter demonstrated the ability of NOM to act as an effective reductant for aromatic compounds in the presence of a bulk reducing agent such as sulfide (Figure IV-12). Furthermore, a wide variety of NOM sources, including simple water leachates of leaf-litter and forest by-products, ground waters with high NOM concentrations (10-30 mg C/L), surface waters from NOM-rich rivers and ponds, and landfill leachates, all exhibited carbon-normalized reduction rate constants of very similar magnitude. A fixed pH and sulfide concentration, the second-order reduction rate constants agreed within a factor of about 3, suggesting the ubiquitous presence of similar reductive ability in NOM. The pH dependencies and structure-reactivitiy relationships (Figure IV-13) observed for the natural napthoquinones were paralleled for the NOM samples (Dunnivant, et al., 1992).

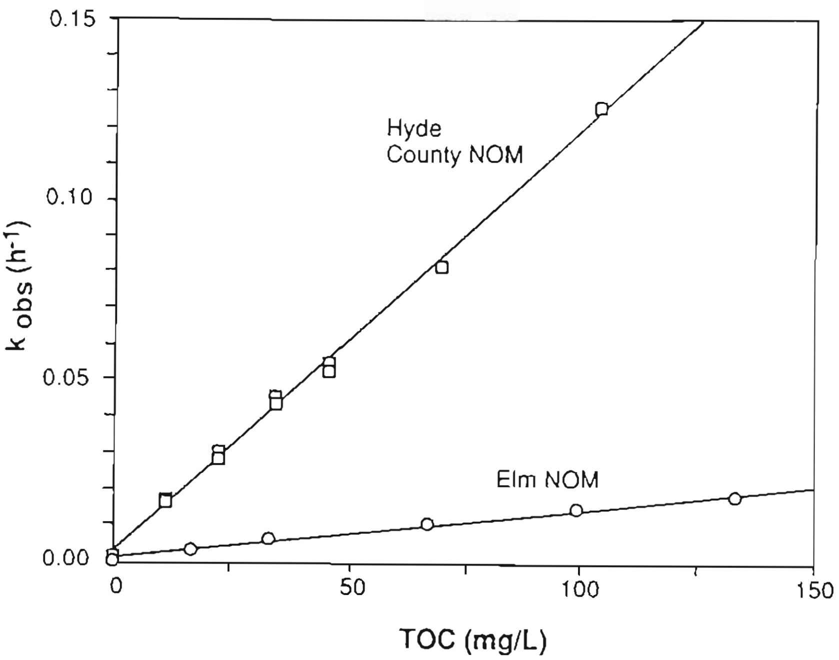


Figure IV-12. Data plot showing the pseudo-first-order rate constants for the reduction of 3-chloronitrobenzene at 25° C as a function of NOM concentration for two water sources.(from Dunnivant, et al., 1992)

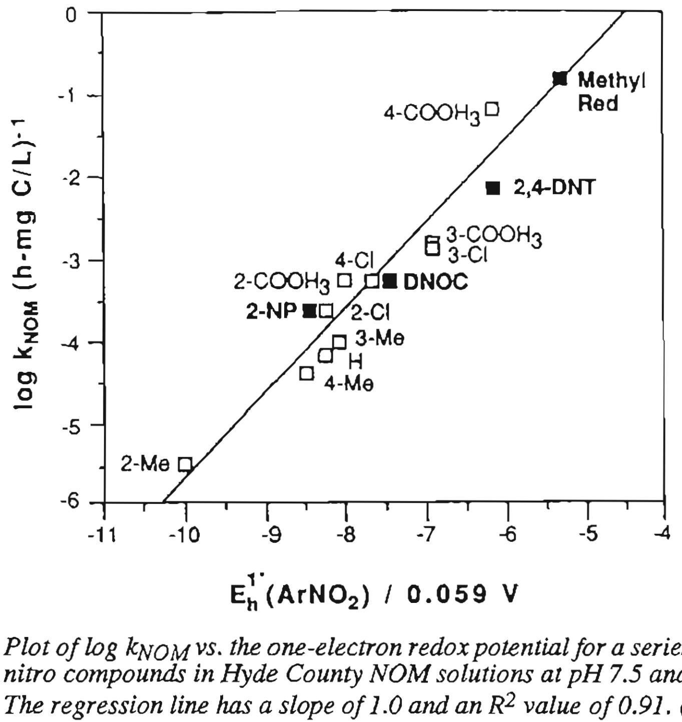


Figure IV-13. Plot of log k_{NOM} vs. the one-electron redox potential for a series of aromatic nitro compounds in Hyde County NOM solutions at pH 7.5 and 5 mM sulfide. The regression line has a slope of 1.0 and an R^2 value of 0.91. (from Dunnivant, et al., 1992)

Several additional features of the experiments with NOM are worth highlighting. Analyses of the metals contents of the NOM solutions revealed wide variations in the contents of redox-active metals such as iron and manganese in these solutions, with no discernible relationships between metal contents and reductive reactivity. This strongly suggests that, at least over the pH ranges investigated in these studies (6-9), metal-organic complexes do not play a primary role in the observed reduction reactions. Work by Schwarzenbach, et al. (1990), among others, has demonstrated that metal organic complexes such as iron porphyrins can also act as effective reductants for nitroaromatic compounds in the presence of a bulk reducing agent such as sulfide. Deviations from expected behavior of NOM-mediated reductions at the extremes of the pH range studied, however, suggest the possibility of the involvement of other, non-quinone-like reactants, especially at lower pH's.

Thus a considerable body of evidence has been accumulated supporting the hypothesis that quinone-like functionalities in NOM can act as effective mediators of reductive transformations in natural systems containing a bulk reducing agent such as sulfide. However, very recent (unpublished) research by Judith Perlinger at EAWAG which attempted to further verify this hypothesis strongly suggests that this interpretation may be an oversimplified view of the actual redox pathways in NOM-mediated reductions of anthropogenic organic compounds such as nitroaromatics and halogenated organics. Electron-spin-resonance studies, which can detect the free radical semiquinone intermediates known to be present in hydroquinone/quinone redox reactions, indicate that other types of free radicals may be involved, perhaps more characteristic of the bulk reductant used (sulfide) than the naturally-present quinones. Further research is clearly necessary in order to fully understand the mechanisms by which aqueous-phase NOM participates in the reductions of anthropogenic organic compounds.

Other efforts to characterize the nature of the involvement of NOM in reductions of nitroaromatic and halogenated organic compounds have focused directly on reactions mediated by PNOM in sediments. Continuing work in the laboratories of N. Lee Wolfe at the USEPA Laboratories in Athens, GA has resulted in isolation and characterization of a several robust sediment enzymes which facilitate reductions of anthropogenic chemicals. Immunospecific assay techniques have been developed which provide a method to identify the sources of biological redox proteins in soils, sediments and aquifer materials (Carreira and Wolfe, 1994). Proteins isolated from plant materials around soils and ponds have been shown to exhibit reductive activity which can account for a substantial fraction of the redox activities observed in sediment slurries. Thus, the reductive activity of sediment NOM may be largely due to biological components such as proteins.

Clearly, NOM can facilitate the rapid reduction of a variety of anthropogenic chemicals in sediment/water and homogeneous aqueous systems. It is equally clear that we have an incomplete understanding of the nature and scope of such redox activities.

E. Colloidal NOM and facilitated transport

The importance of colloids, particularly colloidal NOM, in the transport of contaminants in the environment has recently received a great deal of attention. Part of this interest arose from "anomalous" observations obtained in laboratory investigations of hydrophobic organic contaminant partitioning. It was observed that as the total solids content increased the observed partitioning coefficient, which should be invariant, decreased. This variation was termed the "solids effect" and has since been explained by the presence of colloids which were not efficiently separated from the aqueous phase by standard techniques (i.e. filtration, centrifugation) due to their small size and/or low density (Morel and Geschwend, 1987). These results highlighted the potential for colloidal influences on contaminant distributions.

Further evidence of the role of colloids in contaminant transport came investigations of the distribution of radionuclides in the environment. For example, Penrose et. al. (1991) found elevated levels of plutonium and americium in ground water more than 3 km from a radioactive liquid waste outfall. Laboratory partitioning studies with soils from the site indicated the plutonium should be highly retarded by sorption to the soil and should not have traveled greater than a few meters. Ultrafiltration studies showed that Pu was associated with material in the size range of 0.025-0.45 microns. This process of enhanced transport (over model predictions) of strongly sorbing contaminants has been termed "facilitated" transport.

In the traditional view of partitioning in ground water, a solute exists as partitioned to an immobile solid phase and dissolved in the mobile aqueous phase. It has become clear that a third phase must be considered. Facilitated transport occurs as the result of the presence of this third mobile sorbing phase. This phase shares the property of the immobile solid phase in that it binds the solute. However, due to its small particle size (generally sub-micron) and/or hydrophilic character, and it is mobile and transported in the aqueous phase. The transport of the contaminant in the environment thereby becomes linked to the transport of the colloidal phases. Colloid transport in surface waters is mainly influenced by aggregation followed by settling. In ground waters colloid transport is influenced by straining of larger particles by pores in the porous media and by interception and capture of smaller colloids by the surfaces of the aquifer particles. These processes are all dependent on colloid size and surface charge.

Colloidal NOM is particularly well suited to provide a means of facilitated transport. CNOM, although not well-characterized, has ionizable functional groups of the same type as DNOM (primarily carboxyl) which may bind metals and other positively charged species. Some of these groups are in favorable positions to allow bidentate chelate formation and therefore have large binding constants for metals (Schnitzer and Kahn ,1972). Other minor components of CNOM, nitrogen in particular, may provide additional sites for metal interactions. Electrostatic repulsion between the negatively charged CNOM and soil minerals, which are predominantly negatively charged aluminosilicates (clays) and quartz, promotes the transport of the CNOM. Adsorption of NOM to iron(III) oxides will inhibit transport, but under anoxic, reducing conditions common in contaminated ground waters, NOM has been shown to stabilize iron oxide colloids and thereby promote transport (Liang et al.,1993). Also present on CNOM are regions of relatively hydrophobic character (i.e. aromatic rings, aliphatic chains) which can bind hydrophobic contaminants. It is suspected that CNOM, in part due to its higher apparent molecular weight, should provide a greater amount of hydrophobic character than DNOM and should therefore exert a greater influence on the transport of sorbed hydrophobic contaminants.

Considerable evidence demonstrates the importance of the association of organic pollutants with CNOM . Less is known about the specific role of CNOM in enhancing the transport of contaminants in aqueous systems, especially ground water. A number of studies have demonstrated that CNOM may adsorb nonpolar organic pollutants in a manner similar to soil and sediment organic matter. Chiou et al. (1986) demonstrated an increase in the "apparent solubility" of DDT and PCBs in the presence of humic colloids (Figure IV-14). Estuarine colloids have been shown to be 10-35 times better than soil or sediment organic matter at adsorbing the herbicides atrazine and linuron (Means and Wijayaratne,1982) and 10 times greater for PAH's (Wijayaratne and Means, 1984). Accounting for the presence of CNOM improved the modeling of the partitioning of PCB's in Lake Superior (Baker et al., 1986). This study indicated a three-phase model, including non-filterable CNOM, best explained the data and also suggested that colloid-associated PCB's may be the dominant species in surface waters. In addition to hydrophobic partitioning, which is related to the organic carbon content of the colloid and the hydrophobicity of the pollutant (Karikhoff et al., 1979), specific interactions between certain types of pollutants are also important. Interaction of pollutants with NOM can occur by various mechanisms including hydrogen bonding, charge transfer complexes through aromatic pi-electrons, cation echange, and conjugate formation through biochemical processes (Leenheer, 1991). As an example, the sorption of benzidine and toluidine onto estuarine colloids was enhanced over that explained by hydrophobic partitioning (Means and Wijayaratne, 1989). Under pH conditions favoring the cationic form of the amine, sorption was enhanced, presumably by cation exchange.

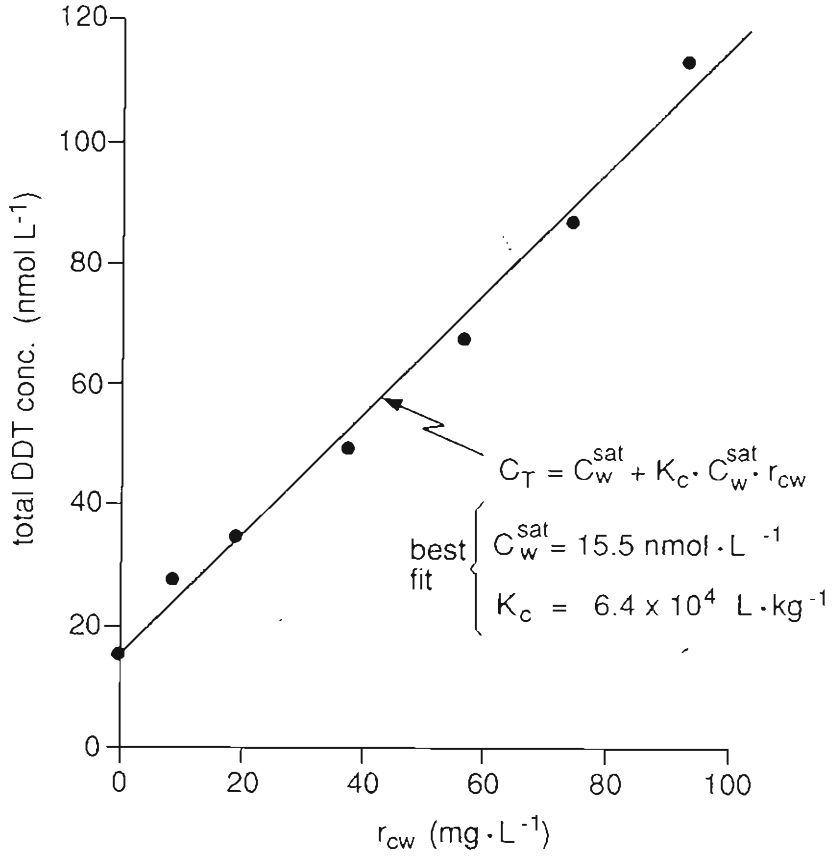


Figure IV-14. Plot of apparant DDT concentration in the aqueous phase as a function of the concentration of humic colloids, r_{cw}*. (from Chiou, et al. 1986)*

Despite overwhelming evidence showing interaction of NOM with organic and inorganic pollutants, the role of NOM in enhancing transport, especially in groundwaters, is poorly known.

V. NOM....SINK FOR POLLUTANTS OR FACILITATOR OF TRANSPORT? MYSTERIES AND RESEARCH QUESTIONS.

NOM is a complex mixture of components, with a wide range of molecular weight and functional group contents. As scientists, we seek to understand how this complex mixture of chemical and physical properties determines NOM's distribution among various environmental compartments (i.e. water, soil, sediment), its role in soil development and weathering processes, its ability to bind metals and moderate hydrolysis reaction rates and its role in inorganic and organic redox reactions. It should be clear from the discourse above that our present understanding of relationships between the properties of NOM, the time, source and spatial variations in these properties, and the role of NOM in geochemical processes is superficial and incomplete. Consequently research to deepen and broaden the role of NOM in these processes must proceed along a wide variety of pathways.

The research outlined above provides a confusing array of conclusions which indicate that, in various soil and aqueous environments, NOM can act as either as an agent for retardation of contaminant transport and transformtion or as facilitator of such processes. If we have as our goal the ability to predict the roles of NOM in geochemical processes, both natural and those induced by anthropogenic disturbances, a considerable effort must continue to define in a fundamental way the nature of the chemistry and physics of NOM. There is no lack of mysteries, and the attempt to obtain a coherent picture of the interplay of NOM with its surrounding chemical and geological environments will continue to provide exciting and important research callenges.

VI. BIBLIOGRAPHY AND LITERATURE CITED

General bibliography

Aquatic Humic Substances: Influence on Fate and Treatment of Pollutants. 1989. I. H. Suffet and P. MacCarthy, eds., Advances in Chemistry Series 219, American Chemical Society, Washington, D.C., 864p.

Biogeochemistry: An Analysis of Global Change. 1991. Schlesinger, W. H. Academic Press. San Diego, CA, 443p.

Environmental Organic Chemistry. 1993. Schwarzenbach, R. P., Gschwend, P. M., and Imboden, D. M., Wiley-Interscience. New York, NY, 681p.

Humic Substances in Soil, Sediment, and Water: Geochemistry, Isolation, and Characterization. 1985. G. A. Aiken, D. M. McKnight, and R. L. Wershaw, eds., Wiley-Interscience. New York, NY, 692p.

Humic Substances II: In Search of Structure. 1989. M. H. B. Hayes, P. MacCarthy, R. L. Malcolm, and R. S. Swift, eds., Wiley-Interscience. Chichester, UK, 764p.

Humus Chemistry: Genesis, Composition, Reactions. 1982. Stevenson, F. J., Wiley-Interscience. New York, NY, 443p.

Interactions of Soil Minerals with Natural Organics and Microbes. 1986. P. M. Haung and M. Schnitzer, eds., SSSA Special Publication Number 17, Soil Science Society of America, Inc., Madison, WI, 606p.

Organic Geochemistry of Natural Waters. 1985. Thurman, E. M., Martinus Nijhoff/Dr. W. Junk Publishers. Dordrecht, 497p.

Organic Substances and Sediments in Water. Volume 1: Humics and Soils. 1991. R. A. Baker, ed., Lewis Publishers Chelsea, MI, 392p.

The Geochemistry of Natural Waters. 1988. Drever, J. I., 2nd Ed., Prentice Hall. Englewood Cliffs, N.J., 437p.

Literature cited

Abdul, A. S., Gibson, T. L., and Rai, D. N. 1987. Statistical correlations for predicting the partition coefficents for non-polar organic contaminants between aquifer organic carbon and water. Hazardous Waste and Hazardous Materials. 4. pp. 211-

Aiken, G.R., McKnight, D.M., Wershaw, R.L., and MacCarthy, P. 1985. An introduction to humic substances in soil, sediment and water, in Aiken, G.R., McKnight, D.M., Wershaw, R.L., and MacCarthy, P. (eds), Humic Substances in Soil Sediment and Water: Geochemistry, Isolation, and Characterization. John Wiley and Sons, New York. p. 6

Antweiler, R. C. and Drever, J. I. 1983. The weathering of a late Tertiary volcanic ash: importance of organic solutes. Geochim. Cosmochim. Acta. 47. pp 623-629.

Bahnick, D. A., and Doucette, W. J. 1988. Use of molecular connectivity indices to estimate soil sorption coefficents for organic chemicals. Chemosphere. 17. pp. 1703.

Baker, J .E., Capel, P. D., and Eisenreich, S. J. 1986. Influence of colloids on sediment-water partition coefficents of polychlorobiphenyl congeners in natural waters. Environ. Sci. Technol. 20. pp 1136-1143.

Baker, J. R., and Mihelcic, J. R. 1994. Development and evaluation of models for predicting soil/water partition coefficents. (submitted for publication).

Baker, W. E. 1973. The role of humic acids from Tasmanian podzolic soils in mineral degradation and metal mobilization. Geochim. Cosmochim. Acta. 37. pp 269-281.

Bennet, P. and Siegel, D. I. 1987. Increases solubility of quartz in water due to complexing by organic compounds. Nature. 326. pp 684-686.

Beven, K. J. and Kirkby, M. J. 1979. A physically based, variable contributing area model of basin hydrology. Hydrological Sci. Bull. 24. pp 43-69.

Bockheim J. G. 1979. Properties and relative ages of soils of southwestern Cumberland Peninsula, Baffin Island, N.W. T., Canada. Arctic and Alpine Res. 11. pp 289-306.

Brown, D. S. and Flagg, E. W. 1981. Emprical prediction of pollutant sorption in natural sediments. J. Environ. Qual. 10. pp. 382-386.

Buffle, J. 1984, Natural Oraganic Matter and Metal-Organic Interactions in Aquatic Systems. in Metal Ions in Biological Systems. H. Siegel, Ed. Marcel Dekker. New York. pp 154-221.

Buffle, J., DeVitre, R. R., Perret, D., and Leppard, G. G., 1989, Physico-chemical characteristics of a colloidal iron phosphate species formed at the oxic-anoxic interface of a eutrophic lake: Geochim. Cosmochim. acta.,v. 53, pp. 399-408.

Buffle, J., Perret, D., and Newman, M. 1992. The use of filtration and ultrafiltration for size fractionation of aquatic particles, colloids, and macromolecules. in Environment Particles. Volume 1. Lewis Publishers. Boca Raton. 554p.

Carreria, L. H. and Wolfe, N. L. 1994. Use of field immunoassays to identify redox proteins in natural systems. (submitted for publication).

Charbeck, S. and Martell, A. E., 1959. Organic Sequestering Agents. John Wiley & Sons. New York. ???p

Chesworth, W. and Macias-Vasquez, F. 1985. Pe, pH, and podzolization. Am. J. of Sci. 285. pp. 128-146.

Chiou, C. T., Porter, L. J., and Schmedding, D. W. 1983. Partition equilibria of nonionic organic compounds between soil organic matter and water. Environ. Sci. Technol. 17. pp 227-231.

Chiou, C. T., Malcolm, R. L., Brinton, T. I., and Kile, D. E. 1986. Water solubility enhancement of some organic pollutants and pesticides by dissolved humic and fulvic acids. Environ. Sci. Technol. 20. pp 502-508.

Chittleborough, D. J., Smettem, K. R. J., Cotsaris, E., and Leaney, F. W. 1992. Seasonal changes in the pathways of dissolved organic carbon through a hillslope soil (Xeralf) with contrasting texture. Aust. J. Soil. Res. 30. pp 465-76.

Cronan, C. S. and Aiken, G. R. 1985. Chemistry and transport of soluble humic substances in foresested watersheds of the Adirondack Park, New York. Geochim. Cosmochim. Acta. 49. pp 1697-1705.

Cuevas, E. and Medina, E. 1988. Nutrient dynamics within Amazonia forests. II. Fine root growth, nutrient availablity and leaf litter decomposition. Oecologia. 76. pp 222-235.

Cuffney, T. F. 1988. Input, movement and exchange of organic matter within a subtropical coastal blackwater river-floodplain system. Freshwater Biol. 19. pp 305-320.

Dethier, D. P., Jones, S. B., Fiest, T. P., and Ricker, J. E. 1988. Relations among sulfate, aluminum, iron, dissolved organic carbon , and pH in upland forest soils of northwestern Massachusetts. Soil Sci. Soc. Amer. J., 52. pp. 506-512.

Dunnivant, F. M., Jardine, P. M., Taylor, D. L., and McCarthy, J. F. 1992. Cotransport of cadmium and hexachlorobiphenyl by dissolved organic carbon through columns containing aquifer material. Environ. Sci. Technol. 26. pp 360-368.

Dunnivant, F. M., Schwartzenbach, R. P., and Macalady, D. L. 1992. Reduction of substituted nitrobenzenes in aquous solutions containing natural organic matter. Environ. Sci. Technol. 26. 2133-2141.

Edwards, R. T. and Meyer, J. L. 1987. Metabolism of asub-tropical low gradient blackwater river. Freshwater Biol. 17. pp 251-263.

Esser, G., Aselmann, I., and Lieth, H. 1982. Modeling the carbon reservoir in the systemcompartment 'litter'. in. Mitteilungen aus dem Geologish-Palantologischen Institut der Universitat Hamburg. Vol. 52, University of Hamburg, Germany. pp 39-58.

Fisher, S. G. and Likens, G. E. 1973. Energy flow in Bear Brook, New Hampshire: An integrative approach to stream ecosystem metabolism. Ecological Monographs. 43. pp 421-439.

Greenland, D. J. 1971. Interactions between humic and fulvic acids and clays. Soil Sci. 111. pp. 34-41.

Grundl, T. J. and Small, G. 1993. Mineral contribution to atrizine and alachlor sorption in soil mixtures of variable organic carbon. J. Contam. Hydrol. 14. pp.

Harris, J.C.,1983. Rate of hydrolysis, in Lyman, W., Reehl, W. and Rosenblatt, D. (eds), Handbook of Chemical Property Estimation Methods. McGraw-Hill, New York. Chapter 7.

Hassett, J. J., Banwart, W. L., and Griffen, R. A. 1983. Correlation of compound properties with sorption characteristics of non-polar compounds by soils and sediments: Concepts and limitations. in: C. W. Francis and S. I. Aurebach, eds., Environmental and Solid Waste Characterization, Treatment and Disposal. Butterworth Publishers, MA.

Hayes, M. H. B., MacCarthy, P., Malcom, R. L., and Swift, R. S. 1989. Structures of humic substances: the emergance of 'forms'. in M. H. B. Hayes, P. Macarthy, R. L. Malcom, and R. S. Swift (eds.). Humic Substances II. In Search of Structure. John Wiley and Sons. New York. pp. 690-733.

Hedges, J. I., and Lee, C. 1992. Measurement of dissolved organic carbon in natural waters. Marine Chem. 41, pp 290.

Hornberger, G. M., Bencala, K. E., and McKnight, D. M. 1994. Hydrological controls on dissolved organic carbon during snowmelt in the Snake River near Montezuma, Colorado. 25. pp 147-165.

Huang, W. H. and Keller, W. D. 1971. Dissolution of clay minerals in dilute organic acids at room temperature. Am. Mineral. 56. pp. 1082-1095.

Hunter, K. A. 1980. Microelectrophoretic properties of natural surface-active organic matter in coastal seawater. Limno. Oceanogr. 25. pp 807-822.

Jardine, P. M., Wilson, G. V., Luxmoore, R. J., and McCarthy, J. F. 1989. Transport of inorganic and natural organic tracers through an isolated pedon in a forest watershed. Soil Sci. Soc. Am. J. 53. pp 317-323.

Jardine, P. M., Wilson, G. V., McCarthy, J. F., Luxmoore, R. J., Taylor, D. L., and Zelazny, L. W. 1990. Hydrogeochemical processes controlling the transport of dissolved organic carbon through a forested hillslope. J. Contam. Hydrol. 6. pp 3-19.

Karickhoff, S. W. 1980. Sorption kinetics of hydrophobic pollutants in natural sediments. in R. A. Baker, ed. Contaminants and Sediments. Vol. 2: Analysis, Chemistry, and Biology. Ann Arbor Science, Ann Arbor. pp. 193-205.

Karickhoff, S. W. 1981. Semi-empirical estimation of sorption of hydrophobic pollutnats on natural sediments and soils. Chemosphere. 10. pp 833-846.

Karickhoff, S. W. 1984. Organic pollutant sorption in aquatic systems. J. Hydraulic Eng. 110. pp 707-735.

Karickhoff, S. W., Brown, D. S., and Scott, T. A. 1979. Sorption of hydrophobic pollutants on natural sediments. Water Res. 13, pp 241-248.

Kenaga, E. E. and Goring, C. A. I. 1978. Relationship between water solubility, soil sorption, octanol-water partioning and bioconcentration of chemicals in biota. Proceedings of the American Society for Testing and Materials, 3rd Aquatic Toxicology Symposium. No. STP 707. pp. 78-115.

Lee, C. M. and Macalady, D. L. 1989. Towards a standard method for the measurement of organic carbon in sediments. Intern. J. Environ. Anal. Chem. 35. pp. 219-225.

Leenheer, J. A. 1991. Organic substance structures that facilitate contaminant transport and transformations in aquatic sediments. in R. A. Baker, ed. Organic Substances and Sediments in Water. Vol 1: Humics and Soils. Lewis Publishers. Chelsea. MI. pp. 3-22.

Leenheer, J. A., Malcom, R. L., McKinley, P. W., and Eccles, L. A. 1974. Occurance of dissolved organic carbon in selected groundwater samples in the Uninited States. U.S. Geol. Surv. J. Res. 2, pp 361-369.

Leppard, G. G., 1992, Evaluation of electron microscope techniques for the description of aquatic colloids, *in* Buffle, J., and van Leeuwen, H. P., ed., Environmental Particles, Vol. 1: Boca Raton, FL, Lewis Publishers, pp. 231-290.

Leppard, G. G., Buffle, J., and Baudat, R. 1986. description of the aggregation properties of aquatic pedogenic fulvic acids-Combining physico-chemical data and microscopical observations. Water Res. 20. pp. 185-196.

Lewis, W. M. 1988. Primary production in the Orinoco River. Ecology. 69. p 679-692.

Liang, L., McNabb, J. A., Paulk, J. M., Gu, B., and McCarthy, J. F. 1993. Kinetics of Fe(II) oxygenation at low partial pressure of oxygen in the presence of natural organic matter. Environ. Sci. Technol., 27. pp 1864-1870.

Lonsdale, W. M. 1988. Predicting the amount of litterfall in forests of the world. Annals of Botany. 61. pp 319-324.

Loughman, F. C. 1969. Chemical Weathering of Silicate Minerals. Elsevier, New York, 154 pp.

Macalady, D. L. and Wolfe, N. L. 1983. New perspectives on the hydrolytic degradation of the organophosphorothioate insecticide chlorpyrifos. J. Agric. Food Chem. 31. pp. 1139-1147.

Macalady, D. L. and Wolfe, N. L. 1984. Abiotic hydrolysis of sorbed pesticides. in R. F. Krueger and J. N. Seiber, eds. Treatment and disposal of pesticide wastes. ACS Symposium Series 259. American Chemical Society, Washington, D.C. pp. 221-244.

Macalady, D. L. and Wolfe, N. L. 1985. Effects of sediment sorption and abiotic hydrolyses. 1. Organophosphorothioate esters. J. Agric. Food Chem. 33. pp. 167-173.

Macalady, D. L., Tratnyek, P. G., and Grundl, T. J. 1986. Abiotic reduction reactions of anthropogenic chemicals in anaerobic systems: A critical review. J. Contam. Hydrol. 1. pp. 1-28.

Macalady, D. L., Tratnyek, P. G., and Wolfe, N. L. 1989. Influence of natural organic matter on the abiotic hydrolysis of organic contaminants in aqueous systems. in. I. H. Suffet and P. MacCarthy, eds. Aquatic Humic Substances: Influence on Fate and Treatment of Pollutants. Advances in Chemistry Series 219, American Chemical Society, Washington, D.C., pp 323-332.

MacCarthy, P., Peterson, M. J., Malcom, R. L., and Thurman, E. M. 1979. Separation of humic substances by pH gradient desorption from a hydrophobic resin. Anal. Chem. 51. pp 2041-2043.

Malcolm, R. L. 1985. Humic substances in rivers and streams. in. Humic Substances in Soil, Sediment, and Water: Geochemistry, Isolation, and Characterization.1985. G. A. Aiken, D. M. McKnight, and R. L. Wershaw, eds., Wiley-Interscience. New York, NY, 692p.

Mantoura, R. F. C., Dickson, A., and Riley, J. P. 1978. The complexation of metals with humic materials in natural waters. Estuarine and Coastal Marine Sci. 6. pp 387-408.

Marley, N. A., Gaffney, J. S., Orlandini, K. A., and Cunningham, M.M. 1993. Evidence for radionuclide transport and mobilization in a shallow, sandy aquifer. Environ. Sci. Technol. 27. pp. 2456-2461.

Martell, A. E., and Smith, R. M. Critical Stability Constants. Vol. 3; 1977, Vol.5; 1982, Plenum Press, New York.

Mast. M. A. and Drever, J. I. 1987. The effect of oxalte on the dissolution rates of oligoclase and tremolite. Geochim. Cosmochim. Acta. 51. pp 2559-2568.

McDowell, W. H., and Wood, T. 1984. Podzolization: Soil processes control dissolved organic carbon concentrations in stream water. Soil Sci. 137. pp 23-32.

McKnight, D. M., Feder, G. L., THurman, E. M., Wershaw, R. L., and Westall, J. C. 1983. Complexation of copper by aquatic humic substances from different environments. in. Biological Availability of Trace Metals. eds. R. E. Wildung and E. A. Jenne. Elsevier, Amsterdam. pp 65-76.

McKnight, D. M. and Wershaw, R. L. 1989. Complexation of copper by fulvic acid from the Suwannee River-Effect of counter-ion concentrations. in R. C. Averett, J. A. Leenheer, D. M. McKnight, and K. A. Thorn, eds. Humic Substances in the Suwannee River, Georgia: Interactions, Properties, and Proposed Structures. U. S. Geol. Survey, Open-File Report, 87-557. pp 59-80.

McKnight, D. M., Bencala, K. E., Zellweger, G. W., Aiken, G. R., Feder, G. L., and Thorn, K. A. 1992. Sorption of dissolved organic carbon by hydrous aluminum and iron oxides occuring at the confluence of Deer Creek with the Snake River, Summit County, Colorado. Environ. Sci. Technol. 26. pp. 1388-1396.

Means, J. C., and Wijayaratne, R. 1982. Role of natural colloids in the transport of hydrophobic pollutants: Science. 215. pp. 968-970.

Means, J. C., and Wijayaratne, R. 1989. Sorption of benzidine, toluidine, and azobenzene on colloidal organic matter. in. Aquatic Humic Substances: Influence on Fate and Treatment of Pollutants. 1989. I. H. Suffet and P. MacCarthy, eds., Advances in Chemistry Series 219, American Chemical Society, Washington, D.C., pp 209-222.

Meybeck, M. 1981. River transport of organic carbon to the ocean. in Flux of organic carbon by rivers to the ocean., G. E. Likens, ed. U. S. Department of Energy. NTIS report # CONF-8009140, UC-11, Springfield, VA. pp. 219-269.

Meybeck, M. 1982. Carbon, nitrogen, and phosphorous transports by world rivers. Amer. J. Sci. 282. pp. 401-450.

Meyer,J. L., Likens, G. E., Sloane, J. 1981. Phosphorous, nitrogen, and organic carbon flux in a headwater stream. Archiv. fuer. Hydrobiologie. 91. pp 28-44.

Morel, F. M. M. and Gschwend, P. M. 1991. The role of colloids in the partitioning of solutes in natural waters. in. W. Stumm, ed. Aquatic Surface Chemistry. John Wiley & Sons. New York. 508 p.

Nelson, D.M., Penrose, W.R., Karttunen, J.O., and Melhoff, P. 1985. Effects of dissolved organic carbon on the adsorption properties of plutonium in natural waters. Environ. Sci. Technol. 19, pp. 127-131.

Penrose, W. R., Polzer, W. L., Essington, E. H., Nelson, D. M., and Orlandini, K. A. 1990. Mobility of plutonium and amercium through a shallow aquifer in a semiarid region: Envron. Sci. Technol. 24. pp. 228-234.

O'Brien, B. J. and Stout J. D. 1978. Movement and turnover of soil organic matter as indicated by carbon isotope measurements. Soil Biology and Biochemistry. 10. pp 309-317.

Olson, J.S. 1963. Energy storage and the balance of producers and decomposers in ecological systems. Ecology. 44. pp 322-331.

O'Melia, C. R. and Tiller, C. L. 1993. Physicochemical aggregation and deposition in aquatic environments. in J. Buffle and H. P. van Leeuwen, eds. Environmental Particles. Volume 2. J. Buffle and H.P van Leeuwen (eds.). Lewis Pubs. Boca Raton. pp. 353-386.

Pedro, G., Jamagne, M., and Begon, J. C. 1978. Two routes in the genesis of strongly differentiated acid soils under humid, cool-temperate conditions. Geoderma. 20. pp. 173-189.

Peijnenburg, W. J. G. M., 't Hart, M. J., den Hollander, H. A., van de Meent, D., Verboom, H. H., Wolfe, N. L. 1992. Reductive transformations of halogenated aromatic hydrocarbons in anaerobic water-sediment systems: Kinetics, Mechanisms, and Products. Enron. Toxicol. and Chem. 11. pp 289-300.

Pomona College. 1984. Log P and Parameter Database. Pomona College, Claremont, CA. Seaver Chemistry Labortory, Medicinal Chemistry Project. Technical Database Services Inc. New York.

Ranville, J. F., Harnish, R. A., and McKnight, D. 1991. Particulate and colloidal organic material in Pueblo Reservoir, Colorado: Influence of autochthonous source on chemical composition. in R. A. Baker, ed. Organic Substances and Sediments in Water. Volume 1. Lewis Pub. Chelsea. pp.47-74.

Righetto, L., Bidoglio, G., Azimonti, G., and Bellobono, I. R. 1991. Competitive acitinde interactions in colloidal humic acid-mineral oxide systems. Environ. Sci. Technol. 25. pp 1913-1919.

Robert, M. and Berthelin, J. 1986. Role of biological and biochemical factors in soil mineral weathering. in. P. M. Huang and M. Schnitzer, eds., Interactions of soil minerals with natural organics and microbes. SSSA Special Publication Number 17. Madison. pp. 453-496.

Romell, L. G. 1935. Ecological problems of the humus layers in the forest. Cornell University, Agricultural Experiment Station Memoir 170, Ithaca, New York.

Schalscha, E. B., Appelt, H., and Schatz, A. 1967. Chelation as a weathering mechanism-I. Effect of complexing agents on the solubilization of iron from minerals and granodiorite. Geochim. Cosmochim. Acta., 31, pp 587-596.

Schlesinger, W. H. 1977. Carbon balance in terrestrial detritus. Annual Review of Ecology and Systematics. 8. pp 51-81.

Schlesinger, W. H. 1991. Biogeochemistry: An analysis of Global Change. Academic Press, San Diego, 443p.

Schnitzer, S., and Kahn, S. U. 1972. Humic Substances in the Environment. Marcel Dekker. New York. p ?

Schnitzer, M. and Kodama, H. 1976. The dissolution of micas by fulvic acid. Geoderma. 15. pp 381-391.

Schnitzer, S., and Kahn, S. U. 1978. Soil Organic Matter. Elsevier. Amsterdam. ???p

Schwartzenbach, R. P., and Westall, J. 1981. Transport of non-olar organic compounds from surface water to groundwater: Laboratory sorption studies. Environ. Sci. Technol. 15. pp 1360-1367.

Schwartzenbach, R. P., Stierli, R., Lanz, K., and Zeyer, J. 1990. Quinone and iron porphyrin mediated reduction of nitroaromatic components in homogeneous aqueous solution. Environ. Sci. Technol. 24. pp. 1566-1574.

Schwarzenbach, R. P., Gschwend, P. M., and Imboden, D. M. 1993. Environmental Organic Chemistry. Wiley-Interscience. New York, NY, 681p.

Senesi, N. and Steelink C. 1989. Application of ESR spectroscopy to the study of humic substances. in M. H. B. Hayes, P. Macarthy, R. L. Malcom, and R. S. Swift (eds.). Humic Substances II. In Search of Structure. John Wiley and Sons. New York. pp. 373-408.

Southworth, G. R. and Keller, J. L. 1986. Hydrophobic sorption of polar organics by low organic carbon soils. Water, Air and Soil Poll. 28. pp. 239-248.

Stevenson, F. J. 1982. Humus Chemistry. John Wiley and Sons. New York. 443p.

Stevenson, F. J. 1985. Geochemistry of soil humic substances. in: G. A. Aiken, D. M. McKnight, and R. L. Wershaw, eds. Humic Substances in Soil, Sediment, and Water: Geochemistry, Isolation, and Characterization., Wiley-Interscience. New York, NY, pp. 13-52.

Stevenson, F. J. 1986. Cycles of Soil. John Wileyand Sons. New York. ??

Stumm, W. 1992. Chemistry of the Solid-Water Interface. John Wiley & Sons.New York, 428p.

Swift, M. J., Heal, O.W., and Anderson, J.M. 1979. Decomposition in Terrestrial Ecosystems. University of California Press, Berkeley, p ??

Tan, K. H. 1980. The release of silicon, aluminum, and potassium during decomposition of soil minerals by humic acid. Soil Sci. 129. pp 5-11.

Thorn, K. A., Arterburn, J. B. and Mikita, M. A. 1992. ^{15}N and ^{13}C NMR invstigation of hydroxylamine-derivatized humic substances. Environ. Sci. Technol., 26, pp 107-116.

Thurman, E. M. 1985. Organic Geochemistry of Natural Waters. Martinus Nijhoff/Dr. W. Junk Pubs. Dordrecht. 497p.

Thurman, E. M. and Malcom, R. L. 1981. Preparative isolation of aquatic humic substances. Env. Sci. and Technol. 15. pp 463-466.

Tiller, C. L. and O'Melia, C. R. 1993. Natural organic matter and colloidal stability: models and measurements. Colloids and Surfaces A: Physiochem. and Eng. Aspects. 73. pp 89-102.

Tipping, E. and Cooke, D. 1982. The effects of adsorbed humic substances on the surface charge of goethite (a-FeOOH) in freshwaters. Geochim. Cosmochim. Acta. 46. pp 75-80.

Tratnyek, P. G. and Macalady, D. L. 1989. Abiotic reduction of nitroaromatic pesticides in anaerobic laboratory systems. J. Agric. Food. Chem. 37. pp. 248-254.

Tuschall, J. R. and Brezonik, P. L. 1980. Characterization of organic nitrogen in natural waters: Its molecular size, protein content, and interactions with heavy metals. Limnol. Oceanog. 25. pp 495-504.

Ugolini, F. C. 1968. Soil development and alder invasion in a recently degalaciated area of Glacier Bay, Alaska. in J. M. Trappe, J. F. Franklin, R. F. Tarrant and G.M. Hansen (eds). Biology of Alder. U.S. Forest Service, Pacific Northwest Forest and Range Experiment Station, Portland, Oregon. pp 115-140.

Ugolini, F. C., Minden, R., Dawson, H., and Zachara, J. 1977. An example of soil processes in the *Abies amabilis* zone of central Cascades. Washington. Soil Sci. 124. pp 291-302.

Voice, T. C. and Weber, W. J. 1983. Sorption of hydrophobic compounds by sediments, soils, and suspended solids. I. Theory and background. Water Res. 17. pp. 1433-1441.

Weber, E. J. and Wolfe, N. L. 1987. Kinetic studies of the reduction of aromatic azo compounds in anaerobic sediment/water systems. Environ. Toxicol. Chem. 6. pp. 911-919.

Wijayaratne, R. D. and Means, J. C. 1984. Sorption of polycylic aromatic hydrocarbons by natural estuarine colloids. Mar. Environ. Res. 11. pp 77-89.

Wilson, G. V., Jardine, P. M., Luxmoore, R. J., and Jones, J. R. 1990. Hydrology of a forested watershed during storm events. Geoderma. 46. 119-138.

Wolfe, N. L.,Macalady, D. L, Kitchens, B. E., and Grundl, T. J. 1987. Physical and chemical factors that influence the anearobic degredation of methyl parathion in sediment systems. Environ. Toxicol. Chem. 6. pp. 827-837.

Wolfe, N. L. and Macalady, D. L. 1992. New perspectives in aquatic redox chemistry: Abiotic transformations of polutants in groundwater and sediments. J. Contam. Hydrol. 9. pp. 17-34.

Zunino, H. and Martin, J. P. 1977. Metal-binding organic macromolecules in soils:I. Soil Sci. 123. pp. 65-76.

Relationship between rock, soil and groundwater compositions

John Mather
Lyell Professor of Geology, Royal Holloway, University of London, Egham, Surrey, UK.

THE SOURCE TERM

The ultimate source of most groundwater is atmospheric deposition, rain and snow melt, although deep groundwater may reflect contributions from additional sources such as fluid inclusions or waters which existed in the aquifer when the materials forming the aquifer were deposited (known as connate waters). The atmospheric or meteoric water infiltrates through the soil and the unsaturated zone before entering the saturated zone.

The composition of this source term is controlled by the dissolution of atmospheric gases particularly carbon dioxide, and by the wash out of components derived from the sea, the land and pollution sources. Thus rain is essentially a dilute solution of carbonic acid and a sea-salt aerosol plus a variable mixture of sulphuric, nitric and hydrochloric acids and usually some ammonium ion. In the United Kingdom the average pH of rainfall varies from 5.0 to 4.4 in the west to 4.3 to 4.1 in the east. This latter figure is comparable to that for rainfall in parts of Eastern Europe and Scandinavia which have been affected by industrial pollution.

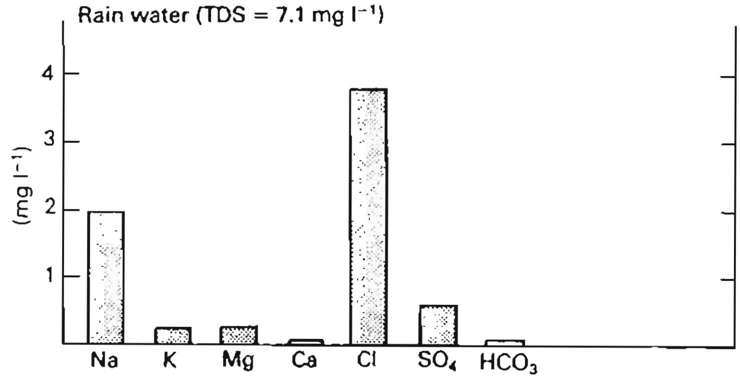


Fig. 1. Typical composition of rainwater with a total dissolved solids content (TDS) of 7.1 mg/l (Raiswell, Brimblecombe, Dent and Liss, 1980).

In areas near the sea, rainfall has a variable load of the major sea-salt components such as Cl, Na and Mg; further inland both the variability and mean concentration of these components decline. The concentrations of most minor elements are higher than can be attributed to a marine origin and probably result from the partial dissolution of particles of terrestrial and industrial origin by the acidic rain. However, the concentration of dissolved material is generally low and frequently below 20 mg/l. Figure 1 shows a typical rainwater with a total dissolved solids content of 7.1 mg/l and Figure 2 demonstrates how the chloride content of rainwater decreases with increasing distance inland from the coast (from, Raiswell, Brimblecombe, Dent and Liss, 1980).

Vegetation can have a significant effect on the composition of rainfall, depending on the type of vegetation, meteorological conditions and even the presence or absence of daylight. Studies with forest trees have shown that there is an increase in the concentration of most solutes in the throughfall and stemflow and an increase by a factor two or three is typical (Kinniburgh and Edmunds, 1984). Some of this may be due to concentration of solutes by evaporation from the surface of the vegetation but some at least is likely to be derived from the vegetation itself.

It is important to recognise the role of evaporation in concentrating solutes. Thus, to look at it very simplistically, if the chloride content of precipitation is 10 mg/l and 75% of this precipitation is lost by evaporation and transpiration the chloride concentration in the water interacting with the soil could be increased to 40 mg/l.

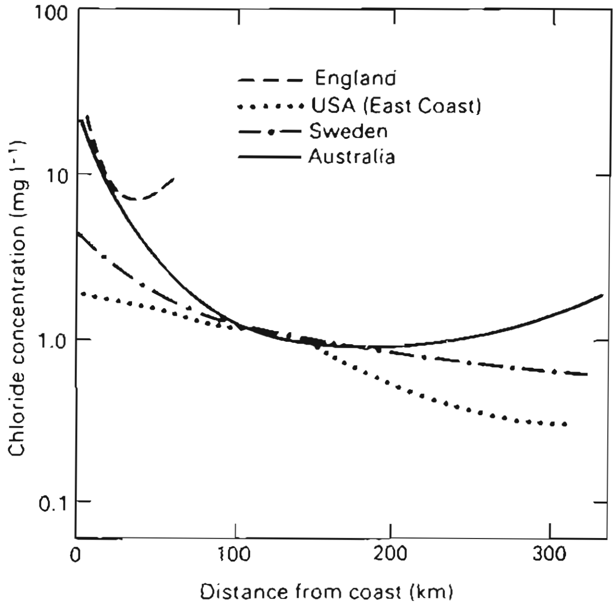


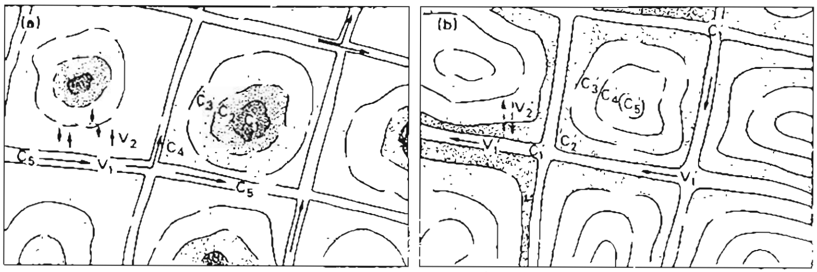
Fig. 2. Decrease in chloride content of rainwater with increasing distance inland. (Raiswell, Brimblecombe, Dent and Liss, 1980).

FACTORS AFFECTING GROUNDWATER CHEMISTRY

A number of natural processes can be identified by which groundwater gains its intrinsic properties. Most of the chemical properties of groundwater are attained rapidly, during movement through the soil and the unsaturated zone, and an understanding of the reactions which occur there are of particular importance in hydrogeochemical studies. Although there may be little change in total mineralisation within the saturated zone it is here that exchange reactions and other processes slowly modify groundwater chemistry. It is here also that there is an opportunity for the removal or modification of harmful substances by adsorption, precipitation or oxidation/reduction reactions.

However, before considering the geochemical reactions themselves in more detail it is necessary to consider the physical factors which are important with respect to the chemical composition of groundwater. These are residence time and the pathways or routes along which water moves through the system. A longer residence time provides an opportunity for reactions to occur between water and the materials with which it is in contact and the pathways along which water moves determine the materials water contacts during its passage. In general water which follows shallow pathways contacts more weathered and consequently less reactive materials than water which moves along deeper pathways.

The flow system is very important as it brings in new reactants and removes reaction products. Flow in fractures is quantitatively more important then intergranular flow in the majority of aquifers. However, fissure storage is commonly only a small fraction of the total storage in most sedimentary aquifers. Because, the rate of solute transport in the intergranular pore space may be several orders of magnitude lower than in the fissures there is a potential for compositional disequilibrium to exist between water in rapid transit via fissures and relatively immobile interstitial fluids. This is illustrated in Figure 3 (from Edmunds, 1983). Two situations are shown (a) displacement of old formation water (C_1) by recently induced recharge (C_5) and (b) dispersion of a pollutant (C_1) with the gradual contamination of the intergranular pore fluid with an initial composition C_5.



V_1 (fissure flow) >> V_2 (intergranular flow)
$C_1 > C_2 > C_3 > C_4 \geqslant C_5$

V_1' (fissure flow) >> V_2' (intergranular flow)
$C_1 \geqslant C_2' > C_3' > C_4' > C_5'$

Fig. 3. Schematic representation of the effects of heterogeneous groundwater flow on water quality distribution in the fissure and intergranular pore spaces of an aquifer (Edmunds, 1983). For definition of symbols see text.

The chemical composition of groundwater is dependent primarily upon the mineralogy and geochemistry of the parent rock. Groundwater chemistry strongly reflects the composition of the aquifer although not in the way which might be immediately supposed. Two concepts are important here - geochemical abundance and geochemical mobility. Those elements most abundant in the earth's crust (Si, Al and Fe) are not those most mobile in groundwaters (Na, Cl). Taking one rock type - average sandstone - and comparing elemental abundance with that in sea water gives an approximate summary of geochemical mobility for the most important elements (Figure 4, from Edmunds, 1986).

REACTIONS IN THE UNSATURATED ZONE

As meteoric waters infiltrate downwards through the soil and the unsaturated zone chemical evolution takes place as they react with the soil and aquifer minerals. The principal processes affecting the inorganic constituents are summarised in Figure 5 (from Domenico and Schwartz, 1990).

(1) Gas dissolution and redistribution.

The dissolution and redistribution of CO_2 (gas) are very important soil zone processes.

$$CO_2(g) + H_2O = H_2CO_3$$

$$H_2CO_3 = HCO_3^- + H^+$$

$$HCO_3^- = CO_3^{2-} + H^+$$

Rainwater is acidic, has low dissolved solids and rapidly dissolves CO_2 which occurs in the soil as a result of root and microbial respiration and the oxidation of organic matter. This dissolved CO_2 is further redistributed amongst the weak acids of the carbonate system. One direct result of dissolving CO_2 is a rapid increase in the total carbonate content of the water and a decrease in pH.

Another important soil zone process is the dissolution of O_2 (gas). The resulting concentrations of dissolved oxygen control the redox chemistry in shallow groundwaters.

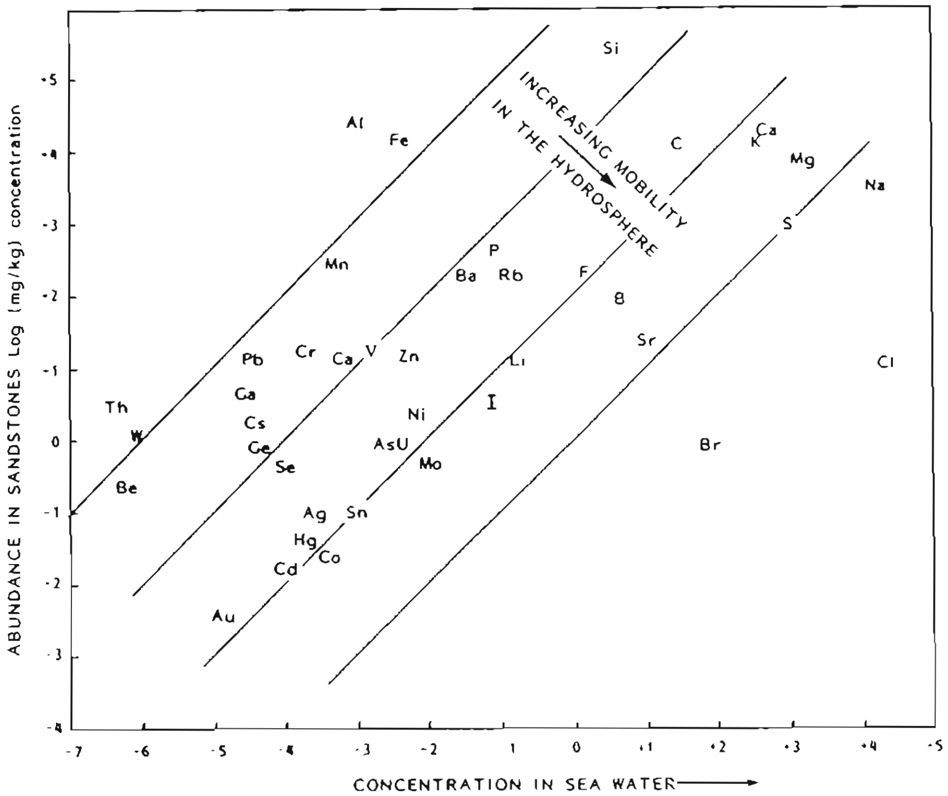


Fig. 4. Geochemical abundance of the elements, based on their abundance in sandstone compared with sea water (Edmunds, 1986).

(2) *Weak acid - strong base reactions.*

The CO_2-charged water is effective in dissolving minerals. The most common reactions involve the weak acids of the carbonate and silicate systems and strong bases from the dissolution of carbonate, silicate and aluminosilicate minerals (examples are given in Figure 5). The dissolution of calcite is probably the most important reaction in the unsaturated zone and a model involving the dissolution of CO_2 gas, accompanied by calcite dissolution, adequately describes the geochemistry of groundwater in the unsaturated zone in carbonate terrains.

(1) Gas dissolution and redistribution, for example,

$$CO_2(g) + H_2O = H_2CO_3^*$$
$$H_2CO_3^* = HCO_3^- + H^+$$
$$HCO_3^- = CO_3^{2-} + H^+$$

(2) Weak acid–strong base reactions, for example,

calcite: $CaCO_3(s) + H^+ = Ca^{2+} + HCO_3^-$
anorthite: $CaAl_2Si_3O_8(s) + 2H^+ + H_2O =$ kaolinite $+ Ca^{2+}$
albite: $2NaAlSi_3O_8(s) + 2H^+ + 5H_2O =$ kaolinite $+ 4H_2SiO_3 + 2Na^+$
enstatite: $MgSiO_3(s) + 2H^+ = Mg^{2+} + H_2SiO_3$

(3) Sulfide mineral oxidation, for example,

$$4FeS_2(s) + 15O_2 + 14H_2O = 4Fe(OH)_3 + 16H^+ + 8SO_4^{2-}$$

(4) Precipitation-dissolution of gypsum,

$$CaSO_4 \cdot 2H_2O(s) = Ca^{2+} + SO_4^{2-} + 2H_2O$$

(5) Cation exchange, for example,

$$Ca^{2+} + 2Na\text{-}X = 2Na^+ + Ca\text{-}X$$
$$Ca^{2+} + Mg\text{-}X = Mg^{2+} + Ca\text{-}X$$
$$Mg^{2+} + 2Na\text{-}X = 2Na^+ + Mg\text{-}X$$

where Na-X is Na adsorbed onto a clay mineral.

Fig. 5. Summary of the important processes affecting inorganic constituents in the unsaturated zone (after Domenico and Schwartz, 1990).

When there are no carbonates present the ion chemistry will be controlled by the dissolution of silicate and aluminosilicate minerals. The relatively low solubility of these minerals, however, means that the concentration of most ions is low. Figure 6 (from Bricker, 1993) shows theoretical groundwater compositions resulting from carbonic acid weathering of some common primary silicate minerals to form kaolinite.

(3) Sulphide oxidation.

In the unsaturated zone minerals like pyrite or marcasite are oxidized to produce $Fe(OH)_3$(solid). This is one of the most important acid producing reactions in geological systems and is the cause of serious acid drainage problems in mining areas.

(4) *Gypsum precipitation and dissolution.*

Cyclical precipitation and dissolution of gypsum can be important in arid areas. If evaporation generally exceeds precipitation, water that infiltrates in normal years evaporates and deposits a small quantity of gypsum. This gypsum accumulates and exceptional recharge can dissolve it giving rise to very high sulphate concentrations in water moving through the unsaturated zone.

(5) *Cation exchange.*

The most important exchange reactions are where Ca^{2+} and Mg^{2+} in water exchange with sorbed Na^{+} from clay minerals thus softening the water. Divalent ions such as Ca^{2+} and Mg^{2+} replace monovalent ions such as Na^{+} and ions with a smaller hydrated radius such as Ca^{2+} replace ones with a larger hydrated radius such as Mg^{2+}.

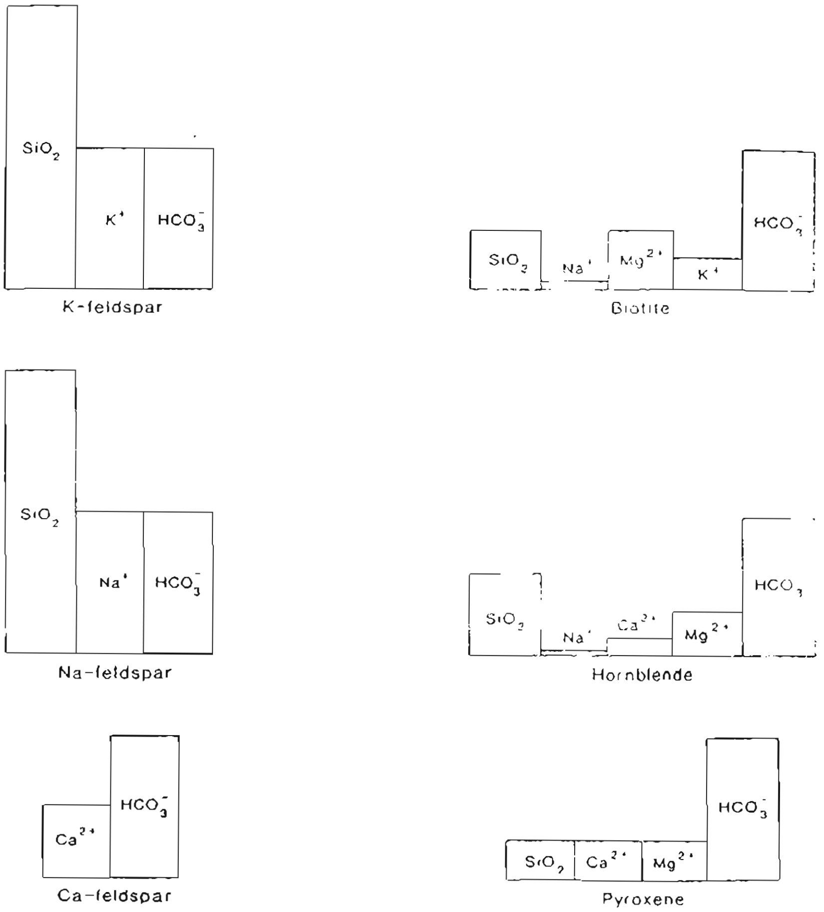


Fig. 6. Water compositions resulting from carbonic acid weathering of some common primary silicate minerals to form Kaolinite (Bricker, 1993).

As well as inorganic reactions important organic reactions also occur in the unsaturated zone. These include:

(1) Dissolution of organic litter.

This is the major source of dissolved organic carbon (DOC) in shallow groundwater. Concentrations of DOC typically fall in a range from 10 to 50 mg/l in the upper soil horizons and less than 5 mg/l deeper in the unsaturated zone. They decline with depth through sorption and oxidation. The most important species are humic and fulvic acids.

(2) Complexation of Fe and Al.

(3) Sorption of organic-metal complexes.

(4) Oxidation of organic compounds.

The oxidation of dissolved organic matter provides a source of CO_2 gas within the unsaturated zone. Through the oxidation of sulphur-containing compounds it can also result in the accumulation of gypsum in shallow soils and in arid areas contribute to recharge with large sulphate concentrations.

REACTIONS IN THE SATURATED ZONE

Processes in the saturated zone are more complex than in the unsaturated zone because of the potential for dispersive mixing and more complicated geological environments. However, most of the same processes are at work and are summarised in Figure 7 (after Domenico and Schartz, 1990).

(1) Weak acid - strong base reactions

If groundwater is not yet in equilibrium with carbonate, silicate and aluminosilicate minerals they will continue to dissolve in the saturated zone. Basically groundwater composition simply proceeds towards equilibrium with those minerals available for dissolution. The extent to which equilibrium is established with respect to particular minerals, such as calcite, dolomite and gypsum can be calculated by a specific computer programme or as part of a larger mineral saturation programme such as WATEQ (Plummer et al, 1976). Measured solute activities (K_{IAP}) can be compared with theoretical equilibrium constants ($K_{mineral}$) to derive a mineral saturation index $SI_{mineral}$.

The dissolution of calcite provides a good illustrative example. Groundwater is saturated with respect to calcite if log K_{IAP}/log $K_{calcite}$ is positive and undersaturated if it is negative. Figure 8 shows data from a suite of groundwaters from the Triassic Sandstones of the English East Midlands where the degree of apparent saturation with respect to the carbonate minerals is related to the various aquifers (Edmunds and Morgan-Jones, 1976). Most groundwaters from the Keuper Sandstone which has a relatively abundant carbonate cement show calcite supersaturation and a few are saturated with respect to dolomite in contrast to groundwaters from the underlying carbonate-poor Bunter Sandstone. The extent of attainment of a particular mineral equilibrium will indicate whether solution or precipitation is occurring in the aquifer which is relevant to problems such as encrustation in boreholes.

(2) Dissolution of soluble salts

When groundwater encounters soluble salts in the saturated zone the impact on groundwater chemistry can be considerable. Common dissolution reactions are given in Figure 7. Where active groundwater flow encounters evaporites, these can dissolve to produce a brine whose composition depends on the particular minerals present. For example in Cumbria UK, where a major hydrogeological investigation of a potential radioactive waste disposal site is currently underway (UK Nirex, 1993), high salinities, up to three times that of seawater, are thought to be derived from the dissolution of halite associated with Permo-Triassic evaporites.

(1) weak acid–strong base reactions, for example,

carbonate minerals + H^+ = cations + HCO_3^-
silicate minerals + H^+ = cations + H_2SiO_3
alumino-silicate minerals + H^+ = cations +
H_2SiO_3 + secondary minerals (for example, clay minerals)

(2) dissolution of soluble salts, for example,

halite: $NaCl(s) = Na^+ + Cl^-$
anhydrite: $CaSO_4(s) = Ca^{2+} + SO_4^{2-}$
gypsum: $CaSO_4 \cdot 2H_2O(s) = Ca^{2+} + SO_4^{2-} + 2H_2O$
carnalite: $KCl \cdot MgCl_2 \cdot 6H_2O(s) = K^+ + Mg^{2+} + 3Cl^- + 6H_2O$
kieserite: $MgSO_4 \cdot H_2O(s) = Mg^{2+} + SO_4^{2-} + H_2O$
sylvite: $KCl(s) = K^+ + Cl^-$

(3) redox reactions, for example,

$$\tfrac{1}{4}O_2(g) + H^+ + e^- = \tfrac{1}{2}H_2O$$
$$\tfrac{1}{2}Fe_2O_3(s) + 3H^+ + e^- = Fe^{2+} + \tfrac{3}{2}H_2O$$
$$\tfrac{1}{2}MnO_2(s) + 2H^+ + e^- = \tfrac{1}{2}Mn^{2+} + H_2O$$
$$\tfrac{1}{8}SO_4^{2-} + \tfrac{9}{8}H^+ + e^- = \tfrac{1}{8}HS^- + \tfrac{1}{2}H_2O$$
$$\tfrac{1}{8}CO_2(g) + H^+ + e^- = \tfrac{1}{8}CH_4(g) + \tfrac{1}{4}H_2O$$
$$\tfrac{1}{4}CO_2(g) + H^+ + e^- = \tfrac{1}{4}CH_2O + \tfrac{1}{4}H_2O$$

(4) cation exchange, for example,

$$\begin{array}{ll} Ca^{2+} & \qquad\qquad\qquad\qquad Ca \\ Mg^{2+} + 2Na\text{-clay} = 2Na^+ + & Mg\text{-clay} \\ Fe^{2+} & \qquad\qquad\qquad\qquad Fe \end{array}$$

Fig. 7. Summary of the important processes affecting ion concentrations in the saturated zone (after Domenico and Schwartz, 1990).

(3) Redox reactions

Redox reactions encountered along a flow system are important in controlling the hydrochemistry of metals such as iron and manganese and carbon, nitrogen and sulphur species. It is often possible to define redox zones in a flow system and an idealized example based on a limestone aquifer is given in Figure 9.

It is possible that many redox reactions are microbiologically mediated. In most groundwater flow systems bacteria are likely to be present, although exact conditions for their active participation may not be present. Thus to demonstrate the presence of sulphate-reducing bacteria is one thing but to show that they are actively reducing sulphate is another (Edmunds, 1986).

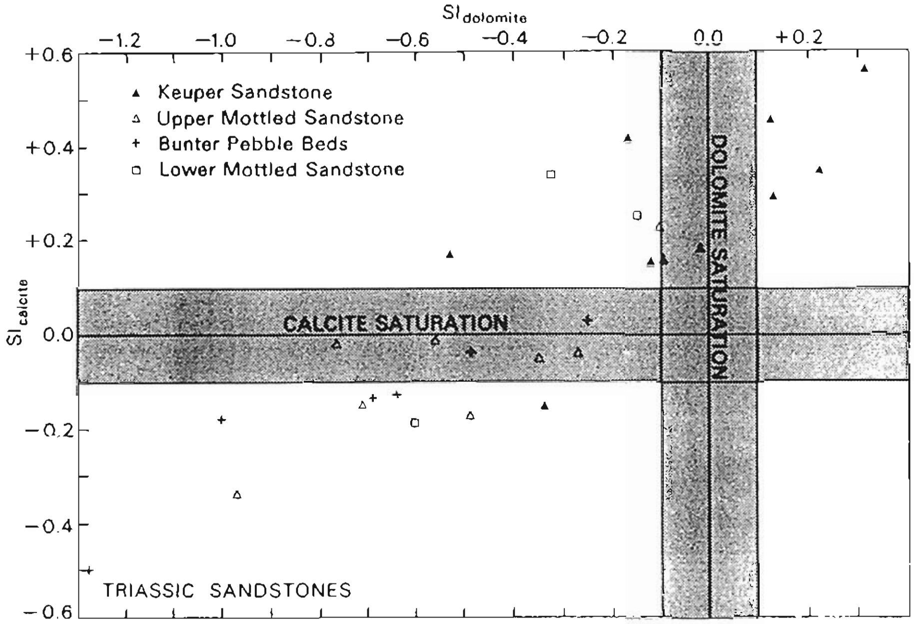


Fig. 8. Plot showing the extent of attainment of equilibrium with respect to calcite and dolomite by groundwaters in the various Triassic aquifers of the English West Midlands (Edmunds and Morgan-Jones, 1976).

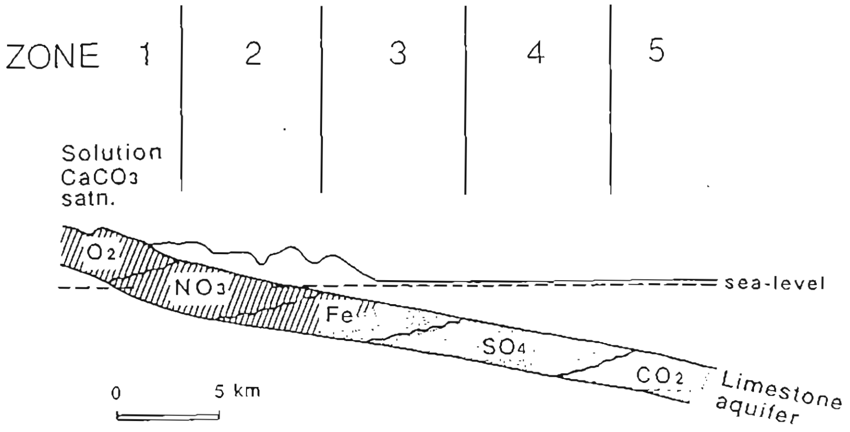
(4) Cation exchange

The most important reactions are those which take Ca^{2+} and Mg^{2+} out of solution and replace them with Na^{+}. The main requirement is a large reservoir of exchangeable Na^{+} which is usually provided by clay minerals deposited in a marine environment. However, nearly all minerals have measurable exchange capacities as a result of crystal imperfections or impurities. For most conditions an affinity series may be defined where an ion with a larger hydrated radius will tend to be displaced from an adsorption site by an ion with a smaller hydrated radius:

$Cs^{+}>Rb^{+}>K^{+}>Na^{+}>NH_4^{+}>Li^{+}$

$Ba^{2+}>Sr^{2+}>Ca^{2+}>Mg^{2+}$

Thus K+ has a tendency to displace Li+ from adsorption sites and so on (Edmunds 1986). In a limestone aquifer one might expect the Mg^{2+}/Ca^{2+} ratio to increase along flow paths as Ca^{2+} displaces Mg^{2+} and for the Sr^{2+}/Ca^{2+} ratio to decrease. However, other factors can complicate the geochemistry and for example in the English Chalk both ratios increase because of a gradual recrystallisation of calcite resulting from the adaption of a marine rock to a freshwater environment (Edmunds et al 1987). In this process, the impurities in the Chalk are gradually released while a fractionally purer low-magnesian calcite is precipitated.



ZONE 1	Oxygen reduction; nitrate present; no iron or H_2S in solution.
ZONE 2	Nitrate reduction; no oxygen, iron or H_2S in solution.
ZONE 3	Iron reduction; no oxygen, nitrate or H_2S in solution; iron present in solution.
ZONE 4	Sulphate reduction; no oxygen or nitrate in solution, iron precipitated as FeS_2; H_2S in solution.
ZONE 5	CO_2 reduction/decarboxylation; methane in solution

Fig. 9. Redox zonation, typical of a limestone aquifer, as groundwater moves from outcrop into confined conditions.

ENVIRONMENTAL ISOTOPES

So far this chapter has concentrated on the elemental composition of groundwater, however isotopic data may also give information about a groundwater system, complementary to that given by elemental analysis.

The stable isotope ratios of oxygen and hydrogen ($\delta^{18}O$, $\delta^{2}H$) and the radioisotope concentration of tritium (^{3}H) give information on the behaviour of the water molecule, in contrast to all other hydrochemical information which refers to solutes in the water. Stable carbon isotopic ratios ($\delta^{13}C$) and radiocarbon (^{14}C) are of value in detailed studies of groundwater carbonate systems and under favourable circumstances can provide information on groundwater residence time. Other radio and stable isotope ratios may be useful for special purposes. For example studies of uranium isotope series disequilibrium ($^{238}U/^{234}U$) can also provide information on groundwater evolution and age.

1. Oxygen and Hydrogen Stable Isotope Ratios

The isotopic ratios $^{18}O/^{16}O$ and $^{2}H/^{1}H$ can be used in hydrochemical studies to provide information on climatic processes, recharge conditions, groundwater origin and the identification of palaeo groundwater. The isotopic composition of water is expressed in comparison to the isotopic composition of sea water. For this purpose an internationally agreed sample of ocean water has been selected, called Standard Mean Ocean Water (SMOW). The isotopic composition of water determined by mass spectrometry, is expressed in per mil ($^{0}/_{00}$) deviations from the SMOW standard as follows:

$$\delta^{18}O‰ = \frac{(^{18}O/^{16}O)\text{ sample} - (^{18}O/^{16}O)}{(^{18}O/^{16}O)\text{ SMOW}}\text{ SMOW} \times 1000$$

and similarly for $\delta D‰$

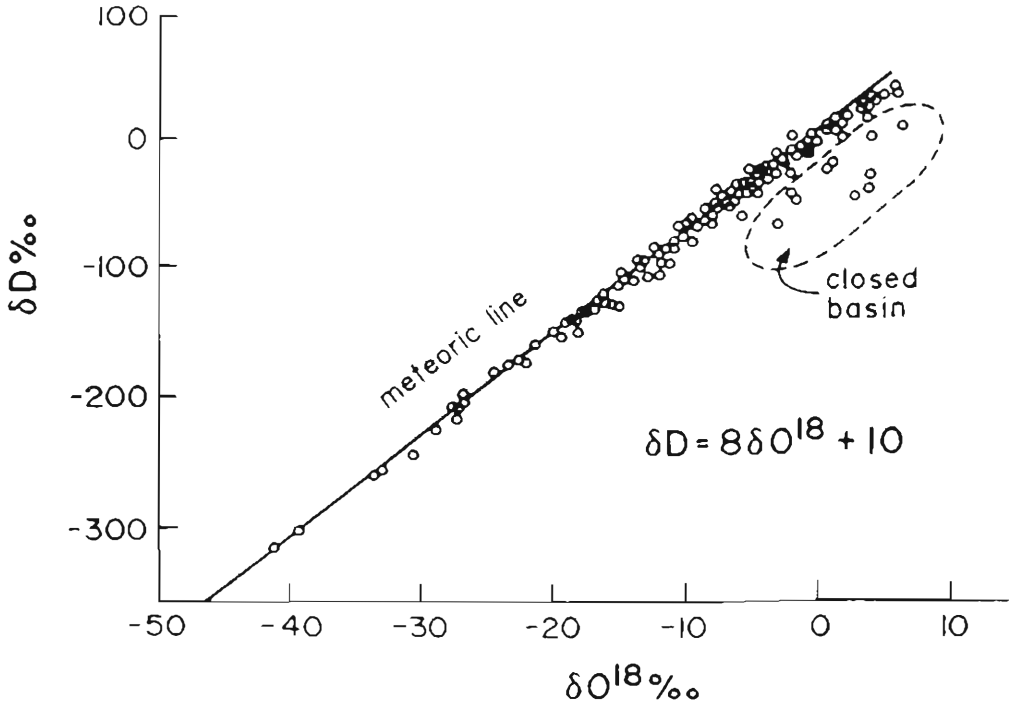


Fig. 10. Isotopic data from about 400 samples of rivers lakes and precipitation from various parts of the world. The best fit line is called the meteoric line. The data in the encircled zone of closed basins is for East African lakes with intensive evaporation (Craig, 1961).

Isotopically light water molecules evaporate more efficiently than the heavy ones and as a result an isotopic fractionation occurs at partial evaporation of water. The vapour is enriched in light water molecules, reflected in relatively negative δD and $\delta^{18}O$ values, whilst in contrast the residual water phase becomes relatively enriched in the heavy isotopes reflected in more positive δD and $\delta^{18}O$ values. Craig (1961) published a diagram based on about 400 water samples of rivers, lakes and precipitation from various countries (Figure 10). The best fit line, called the meteoric line, has been found, with some local variation to be valid over large parts of the world.

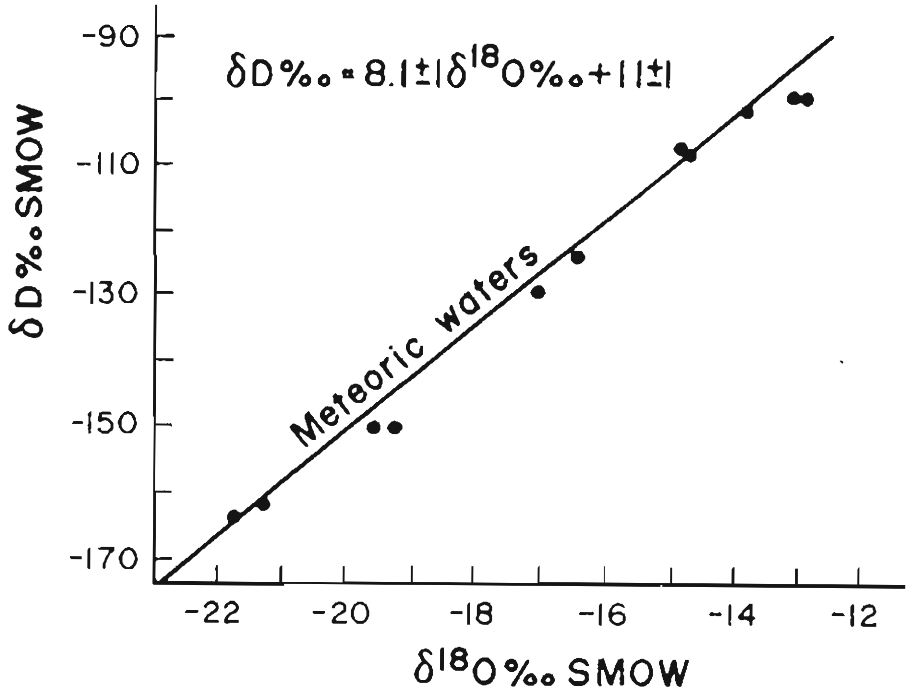


Fig. 11. Isotopic composition of groundwaters from wells in Central Manitoba, Canada. These plot close to the local meteoric line (Fritz, Drimmie and Render, 1974).

The composition of precipitation is reflected, directly or modified, in the composition of groundwater. If groundwater data are plotted along with the local meteoric line deviations from it can be interpreted in terms of secondary processes or as evidence of recharge under a different climate regime. Examples of the use of such techniques are shown in figures 11 and 12. Figure 11 shows the isotopic composition of water sampled from wells in central Manitoba, Canada. These plot clue to the local meteoric line and it was concluded by Fritz, Drimmie and Render, (1974) that evaporation during recharge and isotopic exchange with aquifer rocks were insignificant. This contrasts with the data plotted in figure 12 from Algeria (Gonfiantini, Dincer and Derekly, 1974). Values are scattered along the meteoric line but reveal an internal order, deep groundwaters are isotopically lighter (more negative) than shallow groundwaters. This was taken as an indication that the deep groundwaters are ancient and originated from rains of a different climatic regime, a conclusion supported by the ^{14}C dates (Figure 12).

2. Tritium

Tritium concentrations are reported in terms of tritium units (TU) with 1 TU corresponding to one atom of ^{3}H in 10^{18} atoms of ^{1}H. It occurs naturally in the atmosphere with concentrations of around 5TU in precipitation as a result of cosmic ray bombardment of nitrogen nuclei. It decays with a half-life of 12.35 yr.

However, tritium generated by thermo-nuclear testing in the atmosphere between 1952 and 1963 swamped the natural production of tritium. The concentration of tritium reached a peak in the northern hemisphere in 1963 with values of over 3000 TU reached in the UK.

Ideally, knowing the concentration of ^3H in precipitation and its distribution in groundwater, it should be possible to date the water. Unfortunately it is rarely possible to use tritium data in such a quantitative way. The main application of tritium has been to differentiate pre-1952 water from younger water. If pre-1952 water had an original ^3H concentration of 5TU the concentration in 1993 would be close to the detection limit. Thus, any detectable tritium implies that the water contains a component of more recent or post-1952 water.

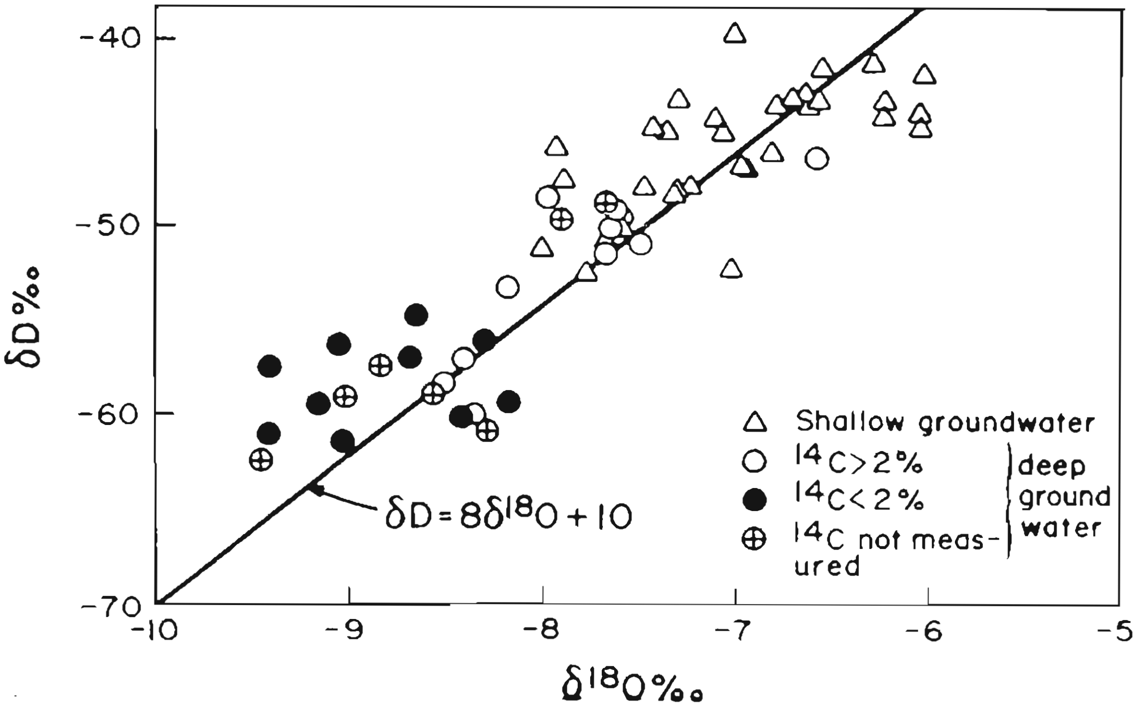


Fig. 12. Isotopic composition of groundwaters from wells in Algeria (Gonfiantini, Dincer and Berekly, 1974).

In the unsaturated zone the bomb tritium peak has been used to provide a marker for monitoring the progress of pollutants such as nitrate towards the water table (Figure 13).

3. Carbon Isotopes

Carbon has three isotopes in Nature, ^{12}C is common and stable; ^{13}C is rare and stable and ^{14}C is very rare and radioactive with a half-life of 5730 years.

Radiocarbon (^{14}C) is continually being produced from nitrogen in the earth's atmosphere and, through a series of exchange, assimilation and other processes of carbon fixation, is incorporated into all living matter. The concentration of ^{14}C is expressed in relation to an international oxalic acid standard. The ^{14}C concentration in the bulk carbon of the standard is defined as 100% modern (pmc) and measurements are reported in the form

$$A = \frac{(^{14}C/^{12}C)\ \text{sample}}{(^{14}C/^{12}C)\ \text{standard}} \times 100\ \text{pmc}$$

In theory it should be possible to use the decay of ^{14}C to establish groundwater age, however there are major complications. If recharging water which dissolved CO_2 in the soil zone reacts with carbonates to form dissolved bicarbonate, the carbon in the bicarbonate will be 50 pmc, according to the equation:

$$\underset{100\ \text{pmc}}{H_2O + CO_2} + \underset{0\ \text{pmc}}{CaCO_2^-} \rightarrow Ca^{2+} + \underset{50\ \text{pmc}}{2HCO_3^-}$$

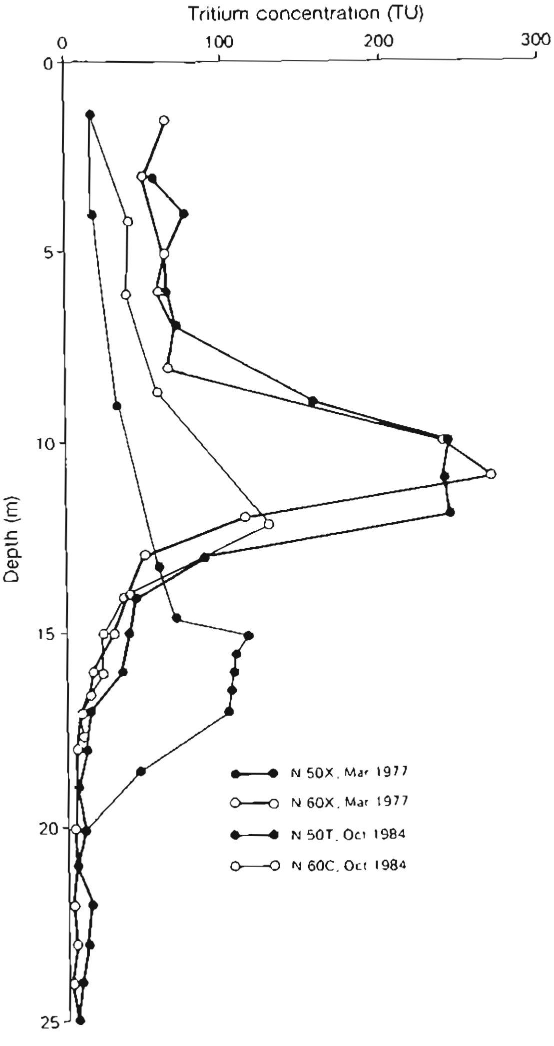


Fig. 13. Sequential tritium profiles from pore waters from the unsaturated zone of the Chalk of west Norfolk, England (Parker, Young and Chilton, 1991).

Similar reactions will occur with silicates, only this time the carbon in the bicarbonate will be 100 pmc because of the reaction:

$$H_2O + \underset{100\ \text{pmc}}{2CO_2} + 2NaAlSi_3O_8 \rightarrow 2Na^+ + Al_2O_3 + 6SiO_2 + \underset{100\ \text{pmc}}{2HCO_3^-}$$

A variety of reactions will occur which cause the initial ^{14}C of the recharged water to vary between 50 pmc (in carbonate rocks) and 100 pmc (in silicate rocks without carbonate). Values of around 60 ±5 pmc seem to be common for recent groundwaters, however, the introduction of ^{14}C into the atmosphere from nuclear bomb tests has increased this concentration up to 200 pmc in very modern waters.

It is possible to use ^{13}C as a means of correcting ^{14}C ages. Materials vary in their $^{13}C/^{12}C$ ratio depending on their origins. Thus marine carbonates have $\delta^{13}C$ of +2 to 0‰ whereas values for organic material and soil are -28 to -20‰. In order to apply the correction it is necessary to know the $\delta^{13}C$ values of the local soil, aquifer carbonates and groundwater. Clearly the composition of ^{13}C in soil materials is likely to vary from place to place and will not be constant over the aquifer. The difficulties mean that ^{14}C cannot really be used to obtain absolute groundwater ages. However, it does provide relative ages, indicate flow velocities and allow mixing systems to be studied.

A typical study from western India is presented in Figure 14 (Borole et al 1979). Figure 14A shows water level contours from the Watrak/Shedi basin from which down-gradient flow and an hydraulic gradient of 0.5 to 0.6 m/km were deduced. ^{14}C dates are presented on Figure 14B as apparent ages assuming an initial value of 100 pmc. Good agreement is obtained between the age contours and the water level contours and an average groundwater velocity of 6 m/y can be calculated.

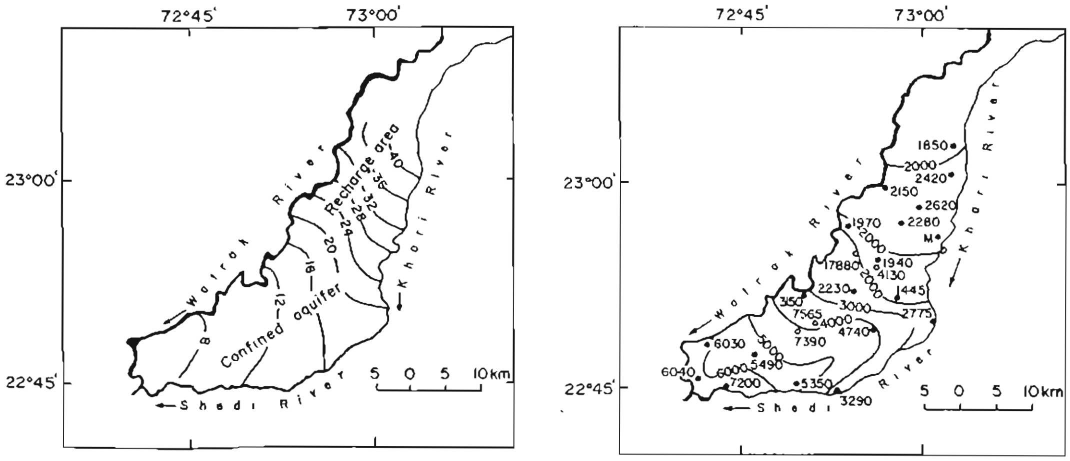


Fig. 14. A. Water-level contours of the Watrak-Shedi Basin, Western India in April 1978. B. Apparent radiocarbon ages in shallow (•) and deeper (o) wells and age contours for the groundwaters of the Basin (Borole et al 1979).

It is also possible to use $\delta^{13}C$ ratios in their own right to identify with the source of dissolved carbon and the extent of chemical reaction with the parent rock.

SALINE GROUNDWATERS

Brackish and saline groundwaters frequently occur in hydraulic continuity with fresh groundwaters and can cause considerable constraints on the exploitation of the fresh water. Because of the various ways in which salinity is imparted to a groundwater, certain chemical signatures can be recognised which may also indicate the origin of the salinity. Most saline waters are dominantly sodium chloride in type so that the major ions are not always of significance in interpretation. The following sources of saline waters are recognised.

(1) Cyclic Salting

A term applied to salt input from precipitation. In areas of low or intermittent recharge cyclic salting and the ensuring salt concentration in the unsaturated zone may lead to significant salination of groundwaters when recharge events occur. Important cyclic salting effects are recorded in a number of arid zone countries.

(2) Modern seawater intrusion

In estuaries and coastal zones modern seawater intrusion occurs either under natural conditions or because of flows induced by abstraction.

(3) Connate groundwaters

Those groundwaters which existed in the aquifer when the materials forming the aquifer were deposited. The presence of truely connate waters is probably rare except in very recent deposits.

(4) Salinity related to hydrogeological entrapment

Such waters may occur as a result of the poor flushing of groundwater systems where saline waters have been entrapped as a result of some geological event such as glacial activity.

(5) Diffusion salinity

Undoubtedly diffusion occurs in transition zones between fresh and saline groundwaters, however, flow-mixing is likely to be dominant and such diffusion effects are not easily recognisable.

GROUNDWATER COMPOSITIONS

Having examined the processes which contribute to the ultimate composition of a groundwater it is now of interest to examine the compositions of groundwaters which result in both carbonate and non-carbonate aquifers. In the UK compositions for a range of aquifers have been discussed by Edmunds et al (1989). One of the areas sampled was the Peak District of Derbyshire in the English Midlands. Here an anticline exposes a limestone inlier (Carboniferous Limestone) beneath shales and sandstones (Millstone Grit). Groundwaters from the two formations have very different geochemical characteristics which are shown on Figure 15.

The absence of carbonate in the Millstone Grit is reflected in the groundwaters with pHs in the range 4.7 to 6.9 and alkalinities of <1 to 88 denoting that there is relatively little acid neutralising capacity. They are well below calcite saturation, with SI calcite between -1.6 and -5.30. In contrast the limestone groundwaters are alkaline and saturated with calcite. The geochemical environment of the Millstone Grit therefore favours the mobilization of metals

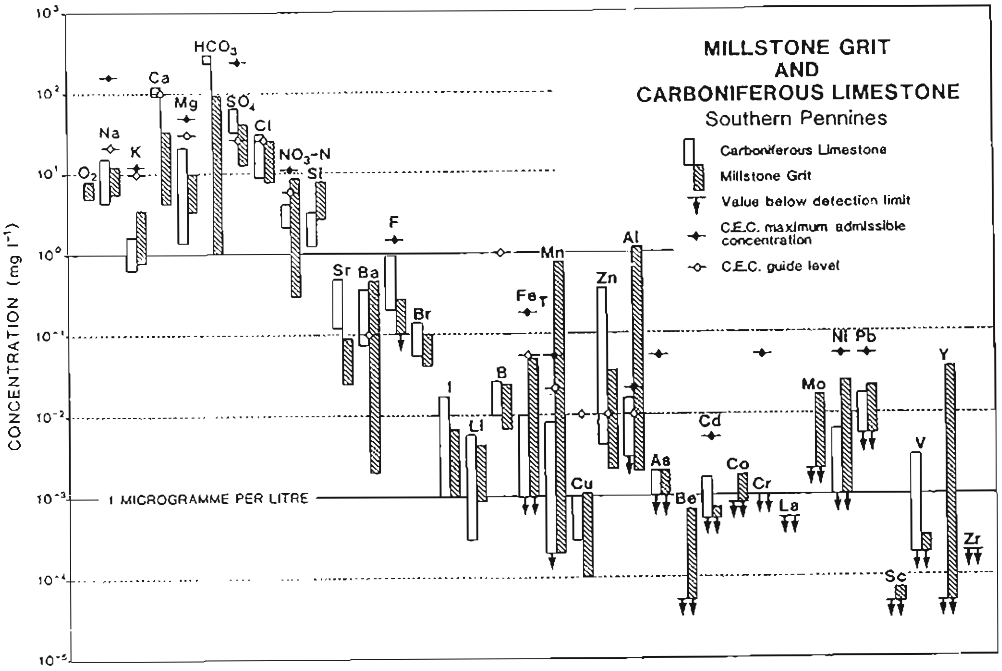


Fig. 15. Element concentrations in the Millstone Grit and the Carboniferous Limestone in order of approximately decreasing abundance (Edmunds et al, 1989).

whereas the more alkaline conditions in the limestone should not give rise to high concentrations. Figure 15 shows the much higher values of Fe, Mn, Al and other metals occurring in Millstone Grit groundwaters.

Perhaps the most studied aquifer in the UK is the Chalk and distinct evolutionary changes can be recognised in the chemistry of its groundwaters as it moves from the outcrop beneath confining deposits. A comparison of the chemical composition of rainfall and groundwater in the outcrop groundwaters is shown in Table 1. Significant amounts of Na, Mg, K, SO_4, Cl and NO_3 in the infiltrating water may be derived from rainfall but some of the NO_3, SO_4 and K are probably derived from fertilizers (Price, Downing & Edmunds, 1993).
The sequence of chemical changes as the groundwater moves downgradient are shown in Figure 16. The dominant change is the conversion of the hard Ca-HCO_3 or Ca-Mg-HCO_3 type water at outcrop to a soft Na-HCO_3 type water in the confined aquifer. At increasing distance downgradient mixing gradually occurs with a Na-Cl water that probably has a marine origin. A redox boundary is present within the confined zone and dissolved iron increases significantly in the aquifer following the onset of reducing conditions (Price, Downing and Edmunds, 1993).

As much of the Norwegian landmass is classified as basement it is important to briefly examine the geochemistry of groundwaters which are likely to occur. In terms of bulk composition, there may be little to distinguish such basement rocks from younger cover and indeed granitic basement is likely to be very similar in composition to many sandstones such as the Millstone Grit discussed above (Edmunds and Savage 1991).

Analyses from the Strath Halladale Granite in Northern Scotland are given in Table 2. Quite high alkalinities are found in the groundwaters which are related to the presence of secondary vein carbonates which although a minor phase dominate the hydrogeochemistry. Such

	Rainfall	Composition of the infiltrate*	Composition of groundwater at outcrop
Calcium	1.4	3.9	100
Magnesium	0.3	0.8	2
Sodium	1.8	5.0	10
Potassium	0.3	0.8	1
Bicarbonate	–	–	280
Sulphate	6.3	17.6	15
Chloride	3.1	8.7	15
Nitrate (as N)	0.4	11.2	20
Ammonia (as N)	0.5	–	–

* Following concentration by evaporation and transpiration

Table 1. Comparison of the chemical composition of rainfall and groundwater in the Chalk as mg/l (Price, Downing and Edmunds, 1993).

dissolved carbonate derived from vein calcite is likely to be a common characteristic of groundwaters from basement terrains. In the shallower environments of the springs, the hydrogeological pathways are in weathered zones, decalcification has occurred and more acid waters are present.

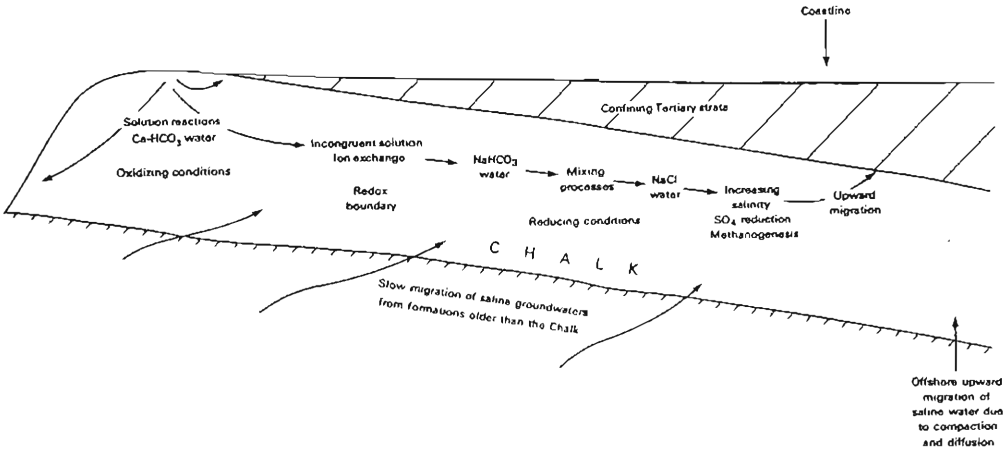


Fig. 16. Schematic diagram of downgradient chemical changes in Chalk groundwaters (Price, Downing and Edmunds, 1993).

In other areas, such as the Carnmenellis Granite in Cornwall, shallow groundwaters and baseflow have much lower pHs (Table 2). Deeper groundwaters in this Granite, represented by water from the South Crofty Mine (Table 2) have high salinity which is considered to have arisen almost exclusively from water-rock interaction within the granite (Edmunds and Savage, 1991). A possible source of chloride is biotite, which in Cornwall may contain up to 1% chloride. Other suggested origins for high salinity in granite groundwaters include migration of sedimentary brines, marine transgressions, residual hydrothermal fluids and breakdown of fluid inclusions.

	(1)	(2)	(3)	(4)	(5)
Number of samples OD	16	1	At 223m below OD	13	At 580m below OD
Temp (^{o}C)	-	-	7.3	-	40
pH	6.19	6.35	7.4	5.44	6.65
Na	14.8	17.1	32	14.7	3520
K	1.38	2.2	1.5	3.18	153
Ca	7.7	21.4	37	12.2	1840
Mg	4.6	12.6	1.4	2.69	55
HCO_3	60	164	201	32.4	60
SO_4	8.8	3.6	10	14.9	129
Cl	20.7	23.6	24	25	9280
NO_3-N	0.09	0.09	<0.04	1.28	<1.0
Sr	0.092	0.280	1.4	0.056	30
Ba	0.043	0.120	0.068	0.006	0.94
Li	<0.2	0.016	0.015	0.0011	107
B	0.03	<0.5	<0.5	0.005	13.9
Si	9.9	12.6	7.9	2.49	16.0
Mn	0.30	1.18	0.17	0.554	3.5
Al	<0.10	<0.1	<0.1	0.122	<0.1
Zn	-	0.086	0.034	0.047	<0.02
Fe_{TOT}	4.74	5.20	0.45	2.00	4.2
F	0.16	<0.2	0.4	-	3.0
Total mineralisation	**133**	**264**	**332**	**112**	**15250**

(1) STRATH HALLADALE SPRINGS
(2) STRATH HALLADALE SHALLOW BOREHOLES
(3) STRATH HALLADALE DEEP BOREHOLE
(4) CARNMENELLIS
(5) SOUTH CROFTY MINE

Table 2. Hydrogeochemical data for groundwaters in granite terrains (modified from Edmunds and Savage, 1991). Unless otherwise indicated concentrations are in mg/l.

ANTHROPOGENIC INFLUENCES

No discussion of groundwater geochemistry would be complete without consideration of the major impacts which urban and industrial developments and modern agricultural practices are having. Incidents can be divided into two broad types on the basis of the source of pollutants. Point-source pollution occurs where pollutants are derived from a discrete, and often readily identifiable source such as a factory, leaking tank or waste disposal site. Diffuse pollution occurs where pollutant input is widely distributed and cannot be pin-pointed precisely. Examples are the widespread application of pesticides and fertilisers to arable land, which may involve the whole outcrop of the aquifer, and acid rain, which may be distributed on a regional or even national scale. In some urban areas, pollutants may leak into groundwater via so many discrete sources that the effect is identical to that of a diffuse pollutant. The major sources of groundwater pollution are outlined in Table 3 and the most important are discussed in this section.

When pollution of a river or storage reservoir occurs the immediate effect on ecosystems can be spectacular but recovery is generally rapid once the source of the pollution has been removed. On the other hand if pollutants leak into aquifers they diffuse into pore spaces and, even when the supply of pollutant is cut off, will still remain until they are flushed out by uncontaminated water. This flushing-out is a long-term process and is well illustrated by the contamination which has resulted from the discharge of mine drainage water into surface lagoons on the Chalk from mines in the East Kent coalfield (Buchan, 1962). After 55 years, drainage from the Snowdown mine, which was pumped on to the suface of the ground for a decade prior to 1935, has still not been flushed out of the aquifer. Discharge from the Tilmanstone mine, which did not cease until 1974, and which has polluted the aquifer over an area of some 27 km^2 is likely to continue to make the aquifer unusable as a groundwater source for a much longer period of time (Headworth, Puri and Rampling, 1980).

Source of Pollution	Type	Principal contaminants causing concern
Agricultural Practices - including the spreading of sewage sludges on land	Diffuse	Nitrates, pesticides and herbicides, biological contaminants
Industry		
a) Landfill	Point source	Range of organic and inorganic contaminants
b) Industrial development particularly associated with the engineering and chemical industries	Multi-point source	Hydrocarbons, chlorinated solvents
c) Mining and resource exploitation	Point source and diffuse	Chloride, sulphate, trace metals
d) Old gasworks	Point source	Phenol
Casual effluent spillages and discharges, including those from septic tanks	Point source	Range of contaminants depending on the source of the spillage
Soakaway drainage from paved areas, roads and airfields	Point source	Hydrocarbons, chloride
Acid Rain	Diffuse	Trace metals

Table 3. Major sources of groundwater pollution (from Mather, 1992).

1. Agricultural pollution

Over the past few decades there has been a continuous increase in agricultural productivity and production. One of the factors contributing to this has been the increasing use of fertilizer nitrogen. During this period the concentration of nitrate in groundwater has increased sharply and a correlation has been made between the two events (e.g. Foster & Crease, 1974). Despite many years of research it is still not clear if groundwater nitrate increases are the result of the leaching of excess fertilizer or because fertilizers have built up the amount of organic matter that microbes can break down to produce nitrate. In the United Kingdom after 1940, substantial areas of old grassland were ploughed, stimulating the microbial population to break down the accumulated organic matter and release nitrate. Other sources of nitrogen include animal manures and slurries and digested sewage which are spread on farmland.

Steadily rising nitrate concentrations are now a feature of many unconfined aquifers and modelling studies predict that they will continue to increase for some time to come. For example in England and France well over 2 million people have been supplied with water from the Chalk containing more than 50 mg/NO_3^- for periods of many years (Foster, 1993). Of importance from the point of view of this chapter are the processes which might attenuate nitrate concentrations. In situ denitrification within the unsaturated zone seems rather unlikely but there is evidence that denitrification occurs in confined conditions (Mariotti et al 1988).

Analysis of groundwater for pesticides was not routine prior to 1989 and the data base available is limited. In the UK the monitoring programme undertaken by Anglian Water Services Ltd (Croll, 1991) has identified a number of herbicides, the triazine herbicides Atrazine and Simazine being by far the most frequently detected. They were found only in younger groundwaters, particularly in aquifer outcrop areas and corresponded well to high nitrate concentrations. It was felt that they originated from use for total weed control on industrial sites, railways and road edges and not from agricultural use.

There is no doubt that pesticides and herbicides are a potential threat to groundwaters. According to Foster, Chilton and Stuart (1991) concentrations in water-supply boreholes are unlikely to approach equilibrium with current pesticide leaching for many years. However, only time and further research will show whether or not impacts similar to those for nitrate will be produced.

2. Industrial pollution

The potential of landfill sites to pollute groundwater is well documented. The main components in the leachates which arise from such sites can be placed in four categories:

a) Major elements or species such a calcium, magnesium, iron, sodium, ammonium, carbonate, sulphate and chloride.

b) Trace metals such as manganese, chromium, nickel, lead and cadmium.

c) A wide variety of organic components, measured as Total Organic Carbon (TOC) or Chemical Oxygen Demand (COD).

d) Microbiological components.

Household waste is reasonably consistent in composition, as is the resulting leachate. Leachate compositions at sites which accept industrial wastes are likely to be much more variable and in general terms leachate composition will reflect the composition of the waste. Thus at the Coatham Stob site in the United Kingdom (Dept of the Environment, 1978), which accepted wastes resulting from the processing of chromite ore, leachates contained up to 11,000 mg/l of chromium and 120 mg/l of vanadium. However, it has been common practice in the UK to dispose of both household and industrial wastes, particularly those described as difficult wastes, at codisposal sites. In many cases the leachate obtained is comparable to that from mature household refuse and it is probable that landfill management and design is as important in determining the composition of leachate as the type of waste deposited (Johanson & Carlson, 1976; Knox 1990). As a result of legislation, the

358

management of landfill sites is now moving towards a total containment option where liners are designed to contain leachates and cappings are designed to reduce infiltration and the formation of leachates. However, many existing landfills still operate using the principle of dilute and disperse where leachates are allowed to migrate in the expectation that natural processes will attenuate leachate components before they reach groundwater abstraction zones.

It is clear from work which has been carried out over many years that landfills produce a highly polluting leachate which can have a significant effect on groundwater. However it is also apparent that many geological environments can be very effective in neutralising and cleansing leachates by geochemical processes. Carbonate aquifers in particular are likely to be much more effective than many sandstone aquifers. This is because of their ability to buffer acidity to a neutral pH, precipitating heavy metals and providing an ideal environment for the development of bacterial populations which attenuate the total organic carbon load (Mather, 1989).

Land contaminated by industrial development gives rise to a range of potential sources of groundwater pollution. These include the disposal of wastes on site, leachate from made ground, leakage from sewers and storage tanks and accidental spillages. Pollution arises either from point sources or, in heavily industrialised areas, from a multitude of point sources such that the whole of an aquifer is subject to contamination by diffuse pollution whose exact source is impossible to identify.

Probably the most common type of incident is spillage of hydrocarbons. As these have a lower density than water they are concentrated at the water table and the capillary fringe and can migrate as a separate liquid phase or dissolved in groundwater. Even more of a problem than these light non-aqueous phase liquids (LNAPLS) are their denser cousins DNAPLS. These dense non-aqueous phase liquids will move through the unsaturated zone but will continue to move vertically through the saturated zone displacing water rather than soil gases. The DNAPLs will move to the base of the permeable unit and will then move downwards following the topography of the basal boundary. This flow direction may be totally different to that of the groundwater. The movement of DNAPLs is difficult to predict in porous environments but the problems will be compounded in hard rock systems. Trying to predict the depth to which the dense fluid will move will be exceedingly difficult, never mind the direction in which fluids will move once they reach that depth.

The light hydrocarbons are more readily degradable and less persistent than the dense solvents. Their presence stimulates the growth of bacteria and in situ biodegradation appears to be limited by the rate of oxygen supply. The solvents are chemically stable and resistant to biodegradation in subsurface environments. Freeze and Cherry (1979) suggested that at sites when DNAPLs were a problem, the local groundwater has terminal cancer and that a cure in the form of returning the aquifer to drinking water standards was not achievable at any cost.

Another form of industrial pollution is that which arises from the mining industry. Groundwater pollution can arise from the extraction process itself, the processing of the commodity or ore, the disposal of spoil and as a result of the abandonment of the mine. Many areas of igneous and metamorphic rocks have metalliferous mineralisation associated with them and are subject to intense mining activity. During operation mines are actively dewatered allowing the access of oxygen. This allows the oxidation of sulphides and other minerals containing metals. Once mining ceases and the workings are allowed to flood, the oxidised phases pass into solution as ferrous sulphate and sulphuric acid. These acid solutions also contain high levels of trace metals and metalloids such as cadmium, mercury, lead and copper.

Once the mine dewatering pumps are turned off, the water table will rise or "rebound" close to its original natural level. The acid mine waters may be discharged through man-made adits or drainage networks or through natural fracture systems resulting in serious water pollution. For example, at Dalquaharran in Ayrshire, Scotland, pumping ceased in 1977 and surface discharge began in 1979. Groundwater discharge rose to 150 l/s with an initial iron concentration of 1200 mg/l, pH 4, aluminium concentration of 100 mg/l and sulphate 6000 mg/l. This water entered the River Girvan from the Dalquaharran No 1 mine entrance and caused irreparable damage to fish stocks and water users downstream. This and a number of other incidents in Scotland are reviewed by Robins (1990).

More recently, in the United Kingdom, serious pollution has arisen following the abandonment of the Wheal Jane Mine in Cornwall. The Mine was the last of a number of tin mines in the area and latterly was dewatering an extensive network of underground workings. Groundwater rebound resulted in acid groundwater, containing high levels of metals, polluting a number of surface water courses. An expensive remediation programme is currently in place to minimise the effects of the acid discharges.

Investigation of such mine water breakouts prior to mine closure is frustrated by the difficulty in predicting the position of discharge points. This can be complicated by changes which may take place in the groundwater flow regime as a result of mining subsidence (Mather et al. 1969).

3. Acid rain

Over the last 20 years there has been widespread concern over the acidification of surface waters in Europe and North America. The pioneering work was carried out in Scandinavia in the 1970's and there has since been much debate over whether acid rain is the major cause or, alternatively, whether the conifer development of moorland is to blame. The scientific debate continues but there is now strong evidence to implicate acidic oxide emissions as a major causal factor (Neal et al. 1991). As well as surface waters, shallow groundwaters are also at risk, particularly those with acid sensitive lithologies. The presence or absence of calcium carbonate is the single most important factor governing acid susceptibility and hard-rock terrains where the geology consists of non-carbonate igneous and metamorphic rocks are likely to be particularly susceptible. For example, shallow groundwaters in granite rocks in Scotland are generally found to be acidic and groundwaters from metasediments frequently give rise to low alkalinity groundwaters (Edmunds & Kinniburgh, 1986).

There are few long term records in existence of water quality in hard rock terrains to demonstrate the effects of acid deposition on shallow systems. Perhaps the best data set is that reported by Hrkal (1992) from the metamorphic rocks and granites of the Bohemian Massif. Bohemia - the western part of the Czech Republic - is one of the most polluted regions of the world. Extensive use of poor quality fossil fuels has led to a reduction in the pH of precipitation and a high rate of atmospheric deposition of nitrogen and sulphur species. Comparison of two groundwater chemistry data sets from the years 1955-69 and 1980-90 respectively, shows a five fold increase in the average concentration of NO_3^- and a halving of the average alkalinity. The most acidified groundwaters were found at the highest altitudes but the most rapid rate of change in the groundwater chemistry was found lower down, where NO_3^- concentrations increased by a factor of 10 over the study period.

Changes in Norway and Sweden appear to be less advanced than those in Bohemia and a study in Varmland, in the crystalline bedrock of southern Sweden, suggests that no regional acidification of importance is in progress (Lång & Swedberg, 1989). However, results suggest that the groundwaters which are most sensitive to acidification are those where discharge from wells in small bedrock aquifers induces the rapid recharge of acidic groundwater from overlying Quaternary deposits.

CONCLUSIONS

Each aquifer will have a characteristic chemistry determined by interaction between recharging groundwater, soil and rock. Due to intensive land use in recharge areas, point-source and diffuse pollution, etc., this chemistry is often masked, and it is in fact difficult at the present day to find any aquifer where the influence of man is still undetectable (Edmunds, 1986).

However polluted an aquifer might become, certain geochemical characteristics of the water are still likely to be determined by reaction with the rock. Thus the Ca^{2+}/Mg^{2+} ratio, and concentrations of Sr^{2+} and Ba^{2+} and carbon isotopic ratios will be geochemically controlled. Other species such as Cl^- and F^-, may be dominated by atmospheric inputs and still others, including SO_4^{2-}, NO_3^-, K^+ and B may be influenced more by anthropogenic activity.

The hydrochemistry will also change across the aquifer due to sequential changes resulting from oxidation/reduction reactions, ion exchange etc. Seasonal fluctuations may occur and longer term changes resulting from drought or overabstraction. Geochemical changes may occur suddenly as a result of a rapid discharge of contaminant or gradually as a result of the incipient leakage of a diffuse pollutant. Thus the composition of any groundwater arises through a complex interaction of factors and is not always directly predictable.

REFERENCES

Borole, D.V., Gupta, S.K., Krishnaswami, S., Datta, P.S. and Desai, B.I., 1979. Uranium isotope investigations and radiocarbon measurements of river-groundwater systems, Sabaramati Basin, Gujarat, India. In: Isotope Hydrology 1978, IAEA Vienna, 1, 181-201.

Bricker, O.P., 1993. The geochemistry of groundwaters in fractured rocks: a geologic perspective. Paper presented at 24th Congress, International Assoc. of Hydrogeologists, Ås, Norway June 1993.

Buchan S., 1962. Disposal of mine drainage water from coal mines into the Chalk of Kent. Proc. Society Water Treatment and Examination, 11, 101-5.

Craig, H., 1961. Isotopic variations in meteoric water Science 133, 1702-3.

Croll, B.T., 1991. Pesticides in surface water and groundwaters. J. Inst. Water Environmental Management, 5, 389-95.

Department of the Environment (DOE) 1978. Cooperative programme of research on the behaviour of hazardous wastes in landfill sites. Final report of the Policy Review Committee. HMSO, London. 169pp.

Domenico, P.A. and Schwartz, F.W., 1990. Physical and Chemical Hydrogeology. J. Wiley and Sons Inc., International Edition.

Edmunds, W.M., 1983. Hydrogeochemical investigations. In: Lloyd, J.W. (editor) Case studies in groundwater resource evaluation. Oxford Univ. Press Chapter 6, 87-112.

Edmunds, W.M., 1986. Groundwater Chemistry. In: T.W. Brandon (editor), Groundwater: occurrence, development and protection, Chapter 3, Inst. Water Eng. Sci., Water Practice Manual 5, 49-107.

Edmunds, W.M., Cook, J.M., Darling, W.G., Kinniburgh, D.G., Miles, D.L., Bath, A.H., Morgan-Jones, M. and Andrews, J.N., 1987. Baseline geochemical conditions in the Chalk aquifer, Berkshire, UK: a basis for groundwater quality management. App. Geochem., 2, 251-274.

Edmunds, W.M., Cook, J.M., Kinniburgh, D.G., Miles, D.L. and Trafford, J.M. Trace-element occurrence in British groundwaters. British Geological Survey Research Report SD/89/3.

Edmunds, W.M. and Kinniburgh, D.G., 1986. The susceptibility of UK groundwaters to acidic deposition. J. Geol. Soc. London, 143, 707-20.

Edmunds, W.M. and Morgan-Jones, 1976. Geochemistry of groundwaters in British Triassic Sandstones. The Wolverhampton - East Shropshire area. Quart. Jl. Engineering Geology, 9. 73-101.

Edmunds, W.M., Savage, D., 1991. Geochemical characteristics of groundwater in granites and related crystalline rocks. In: Applied Groundwater Hydrology, R.A. Downing and W.B. Wilkinson (editors), Clarendon Press, Oxford, 266-282.

Foster, S.S.D., 1993. The Chalk aquifer - its vulnerability to pollution. In: Downing, R.A., Price, M. and Jones, G.P. (editors) The Hydrogeology of the Chalk of North-West Europe, Oxford Science Publications, 91-112.

Foster, S.S.D., Chilton, P.J. and Stuart, M.E., 1991. Mechanisms of groundwater pollution by pesticides. Jl. Inst. Water Environmental Management, 5, 186-93.

Foster, S.S.D. and Crease, R.I., 1974. Nitrate pollution of Chalk groundwater in East Yorkshire - a hydrogeological appraisal. J. Inst. Wat. Engineers, 28, 178-94.

Freeze, R.A. and Cherry, J.A., 1979. Groundwater. Prentice Hall, New York.

Fritz. P. Drimmie, R.J. and Render, F.W., 1974. Stable isotope contents of a major prairie aquifer in central Manitoba, Canada. In: Isotope Techniques in Groundwater Hydrology 1974, IAEA, Vienna, 1, 376-98.

Gonfiantini, R., Dincer, T. and Derekoy, A.M., 1974. Environmental isotope hydrology in the Honda region, Algeria. In: Isotope Techniques in Groundwater Hydrology 1974, IAEA, Vienna. 1, 293-316.

Headworth, H.G., Puri, S. and Rampling, B.H., 1980. Contamination of a Chalk aquifer by mine drainage at Tilmanstone, East Kent, UK. Quart. Journal Engineering Geology, 13, 105-17.

Hrkal, Z., 1992. Acidification of groundwater in the Bohemian Massif. Nor. geol. unders. Bull, 422, 97-102.

Johansen, O.J. and Carlson, D.A., 1976. Characterisation of sanitary landfill leachates. Water Research, 10, 1129-34.

Kinniburgh, D.G. and Edmunds, W.M., 1984. The susceptibility of UK groundwaters to acid deposition. Report to the Department of the Environment by the British Geological Survey, 211pp.

Knox, K., 1990. A review of co-disposal. Proc.1990 Harwell Waste Management Symposium - The 1980's A decade of progress? 54-76.

362

Lång, L.O. and Swedberg, S., 1989. Occurrence of acidic groundwater in Precambrian crystalline bedrock aquifers, Southwestern Sweden. In: Swedberg, S. "Groundwater acidification in Southwestern Sweden; long-term changes in groundwater chemistry", Geologiska Inst., Chalmers Tekniska Högskola/Göteborgs Universitet, Publ. A67, Göteborg, Sweden.

Marriotti, A., Landreau, A. and Simon, B., 1988. Isotope biogeochemistry and natural denitrification process in groundwater; application to the Chalk aquifer of northern France. Geochim. Cosmochim. Acta, 52, 1869-78.

Mather, J.D., 1989. The attenuation of the organic component of landfill leachate in the unsaturated zone: a review. Quart. J. Engineering Geol. 22, 241-6.

Mather, J.D., 1992. The pollution of groundwater by diffuse and point source contanimants. Teaching Earth Sciences, 17(1), 3-12.

Mather, J.D., Gray, D.A. and Jenkins, D.G., 1969. The use of tracers to investigate the relationship between mining subsidence and groundwater occurrence at Aberfan, South Wales. J. Hydrology, 9, 136-54.

Neal, C., Kinniburgh, D.G. & Whitehead, P.G., 1991. Shallow groundwater systems. In: R.A. Downing and W.B. Wilkinson (editors) "Applied Groundwater Hydrology - a British Perspective". Clarendon Press, Oxford, 77-95.

Parker, J.M., Young, C.P. and Chilton, P.J., 1991. Rural and agricultural pollution of groundwater. In: R.A. Downing and W.B. Wilkinson (editors) "Applied Groundwater Hydrology - a British Perspective". Clarendon Press, Oxford, 149-63.

Plummer, L.N., Jones, B.F. and Truesdell, A.H., 1976.WATERQF - A FORTRAN IV version of WATEQ, a computer program for calculating chemical equilibrium of natural waters. U.S. Geol. Survey Water Resour. Invest. 76-13, 61pp.

Price, M., Downing, R.A. and Edmunds, W.M., 1993. The Chalk as an aquifer. In: Downing, R.A., Price, M. and Jones, G.P. (editors). The Hydrogeology of the Chalk of North-West Europe. Oxford Science Publications 33-58.

Raiswell, R.W., Brimblecombe, P, Dent, D.L. and Liss, P.S., 1980. Environmental Chemistry - The Earth-Air-Water Factory. Edward Arnold, London.

Robins, N.S., 1990. Hydrogeology of Scotland. HMSO, London, 90pp.

UK Nirex Ltd. 1993. The geology and hydrogeology of the Sellafield area. Vol. 3. The Hydrogeology. Nirex Report No. 524.

Towards coupling hydrological, soil and weathering processes within a modelling perspective

Colin Neal and Alice J. Robson.
Institute of Hydrology, Maclean Building, Crowmarsh Gifford, Wallingford, OXON, OX10 8BB, UK.

Nils Christophersen
Department of Informatics, University of Oslo, PO Box 1080, Blindern, 0316, Oslo, Norway

INTRODUCTION

General

Computer simulation models have been increasingly relied upon to describe a variety of environmental impacts over the past few decades. These have ranged from limnological models relating nutrient input to eutrophication of freshwaters, to catchment models describing the impact of acid deposition on soil and surface water chemistry. Regional acidification models, such as the RAINS model (Hordijk, 1991), moved to the continental scale and related not just deposition, but energy-use scenarios to acid deposition effects. A further increase in scale is occurring with the extension of general circulation models, which operate at the global level, to assess the effects of greenhouse gases on climate.

In this presentation, the use of models to depict changes at the catchment scale is examined in the context of this course relating to geochemical processes, weathering and recharge in catchments. The theme of how to couple hydrological, soil and weathering processes within a modelling perspective, to provide both a scientific and an applied goal, is examined. Examples are provided of successes and failures over the past decade from stream and catchment acidification research. To do so, I draw upon some of the main findings of our own and our colleagues endeavours in Europe to provide a coherent theme for the presentation and then to link this to associated work. The aim is to show how strong the link must be between observation and modelling and the need for scientific objectivity. For the presentation, the field-based studies considered are centred at Plynlimon in mid-Wales, as the results are characteristic of many acidic and acid-sensitive sites subjected to acidic deposition. Also, the Plynlimon catchments are specifically referred to as they relate to both acidic deposition and conifer afforestation/deforestation - in Scandinavia there has been much debate over what the relative effects have been, while in Britain, both issues are of definite importance. Reference will also be given to the detailed studies at the Birkenes catchment in southern Norway. For the final part of the work, relating to modelling and to what is and what is not achievable, the views expressed are those given by Christophersen *et al.* (1993).

Acidic deposition

From the late 1960's, concerns were raised over acidic deposition of industrially emitted acidic oxides (SO_x in particular) adversely affecting stream and lake ecology in Scandinavia. The environmentally acid sensitive areas were those, ironically, which possessed acidic soils. Indeed the main problem of acidification rests with the lack of weathering reactions of primary rock minerals, to release base cations and promote bicarbonate production from biogenic CO_2 within the soils, to neutralise the acidic inputs. The concerns generated heated political debate across Europe and it lead to considerable scientific research to characterise the nature of the acidification process and to identify potential remedial treatments. Even from the onset, there was much debate over the relative importance of acidic deposition and changing land use. Nowadays, the importance of atmospheric pollution is generally recognised on a continental scale, with, for example, acidification problems being identified across Europe and North America. For Britain, the relative importance of acidic deposition and land use change remains a major issue, and in both aspects will be covered here the UK has a particular problem associated with both aspects. For example, forest development has been the most important land use change in the British uplands this century. A third of the UK is comprised of upland soils that are often acidic and acid sensitive and pollutant deposition levels are as high as those in southern Scandinavia, where acidic deposition has clearly lead to environmental problems.

Acidification in the British Uplands

Considerable evidence has been accumulated which shows that acidification of the UK upland environment has occurred. The first phase of soil acidification probably started several to ten thousand years ago with the development and subsequent loss of birch, hazel, alder and oak forests (Dimbleby, 1952; Taylor, 1974; Pennington, 1984). The deforestation occurred during Neolithic and Bronze age times due to a deteriorating micro-climate about 2700 years ago and local deforestation by man from 5000 years ago up to present day. As a consequence of this, thin acidic moorland soils developed and these are characteristic of much of the British uplands today. In waterlogged areas, where reducing conditions ensured limited breakdown of organic matter, very acidic conditions occurred and peat deposits accumulated. Set against this long term change in the upland environment, further acidification has occurred during the past hundred years and this has an adverse effect on upland ecology.

Major concern has been raised over the deterioration in stream water quality in the British uplands, as associated with this second and ecologically potent phase of acidification. Acidic oxide deposition and conifer planting have been implicated as the major factors in this decline (Stoner and Gee, 1985; UKAWRG, 1988; Whitehead *et al.*, 1988, Jenkins *et al.*, 1990). Both may lead to the generation of more acidic and aluminium bearing stream water and conditions unhealthy for stream biota. However, it is widely believed that, for the UK, as elsewhere, the driving force is the former: conifers enhance the capture of acid pollutants, thereby increasing their impacts; acidification has occurred even for the most extensive tracts of the uplands which had no tree cover. Nonetheless, the relative importance and the degree of interaction of the two factors still needs clarifying (Hornung *et al.*, 1989; Miller, 1985; Jenkins *et al.*, 1990; Rosenqvist, 1978, 1990; Nisbet, 1990; Krug, 1991): for example, under pristine conditions there is evidence both for and against the deleterious effects to stream water of conifer development. In a modelling study of afforestation, Jenkins *et al.* (1990) comments that "afforestation in the absence of acidic deposition, has had a lesser effect on surface water acidification even though the nutrient demands of forest growth have caused significant soil acidification". While acidification of the soils occur by "natural processes", it is the added constraint of changing atmospheric inputs and/or land disturbance that can lead to major changes in the stream. Much of these upland plantations are reaching the harvesting and replanting phases and hence such programmes will accelerate over the next decade (Adamson and Hornung, 1990). These practices could themselves lead to further deterioration in water quality.

THE PLYNLIMON EXPERIMENT, A CASE STUDY

In the spring of 1983, a research programme was initiated to study the hydrogeochemistry of an upland UK environment. The aims were to characterise the effects of forestry management on stream water quality and build on a detailed long term hydrological study at Plynlimon in mid-Wales. From the hydrochemical project's onset, the underlying approach was to study a wide range of chemical constituents. The reasons for this were two fold. First, it was essential to assess hydrological flow pathways at the catchment level. To attain this end, it was decided to use chemical fingerprinting techniques for components with very different chemical properties. The idea, analogous to that used for the development of the periodic table of the elements, was to see if one could find associations of behaviour for the chemical constituents that provide an insight into the underlying hydrogeochemical functioning of the environment. Second, environmental issues concerning the influence of man on the upland environment were being addressed: the monitoring of a wide range of chemical elements was essential - without using this approach, how could one assess those environmental impacts if one was unsure which components were changing under mans influence and which components were detrimental to the environment? Two main sub-catchments of the headwaters of the River Severn in the Hafren forest were selected for study. They were chosen since (1) a deforestation programme was planned for one of the two sub-catchments, (2) detailed hydrological data for these catchments is available since they comprise an integral part of the Institute of Hydrology's water balance studies (Newson, 1976; Kirkby *et al.*, 1991).

The overall project objective was, and remains, to continue and develop further studies on the movement through forested catchments of key elements which directly affect water quality or are relevant to the understanding of processes which affect water quality. Within this context, there are three themes -

1: to understand, at the catchment scale, variations in water quality with hydrology and land use change;

2: to assess changes in stream water chemistry, for a wide spectrum of elements, prior to, during and subsequent to deforestation for one catchment in the Upper River Severn;

3: to develop hydrochemical models of the Plynlimon catchments for process and management applications.

Study area

The work relates to two streams, the Afon Hafren and the Afon Hore, which form the main head water drainage of the River Severn in mid-Wales (Figure 1).

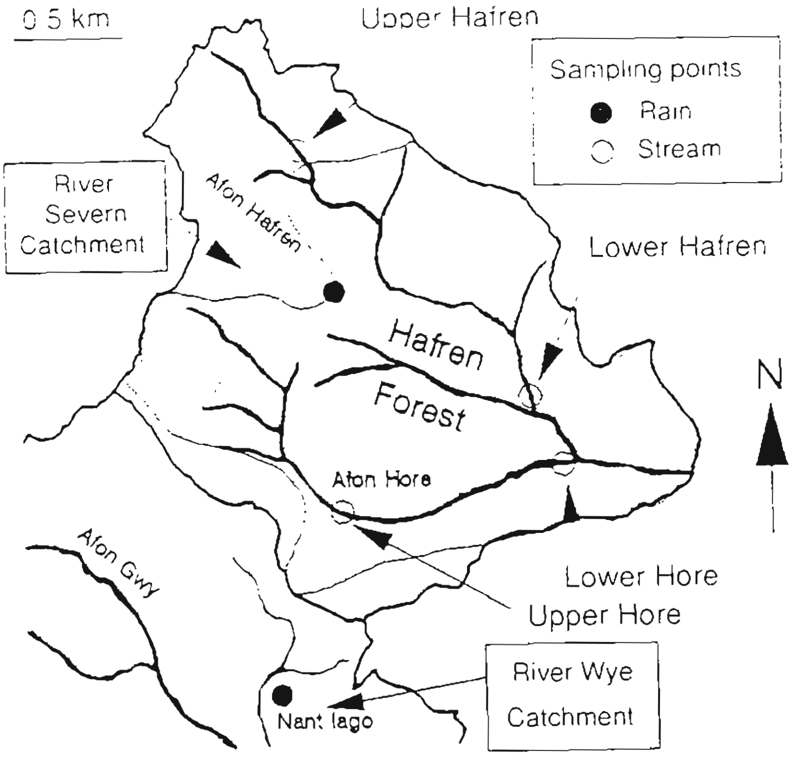


Figure 1. The Plynlimon Catchments

Bedrock geology consists of lower Palaeozoic mudstones, greywackes, sandstones and grits. This is covered by a thin soil, typically 70 cm thick, with organic-rich "L" and "O" horizons (3 to 10 cm). The lower soils consist of a leached "E" horizon (10 to 20 cm) and a fine textured podzolic "B" horizon (about 40 cm thick) merging into a stone "C" horizon. The predominant soil is a stagnopodzol, but peat, brown earth and stagno-gley soils also occur within the two sub-catchments. At the top of each sub-catchment, acid grassland is dominant. On the lower parts of the two sub-catchments, plantation forestry (predominantly Sitka spruce, *Picea sitchensis*) was introduced in various phases between 1937 and 1964, onto acid moorland.

Apart from minor "thinning" of the forest stand, harvesting did not begin until late spring 1985 when clear felling of the lower half of the Afon Hore commenced. This felling took 3 years to complete. In the process, extensive brash and tree stumps were left to decompose in situ. Soon after harvesting, replanting of the slopes with juvenile sitka spruce (< 1 M high) proceeded.

Rainfall averages about 2500 mm/yr with evapotranspirational losses typically amounting to 500 and 650 mm/yr for the Afon Hafren and Afon Hore respectively. Being next to each other, of similar altitude range (360 - 470 m), and similar size (about 340 ha), the streamflow responses to storm events are similar and in phase for the two sub-catchments. The hydrograph response to storms is both rapid and "flashy"; flows vary between 0.01 and 4.5 m^3/s (Newson, 1976; Kirkby *et al.*, 1991).

Hydrochemical behaviour of the Hafren forest

Atmospheric inputs

Plynlimon rainfall is derived from a variety of sources and therefore it has a chemical composition that is highly variable (Table 1). Being close to the Irish sea and the North Atlantic Ocean, westerly winds provide rainfall enriched with sea spray. In contrast, winds from other directions bring air that has passed over agricultural land and industrial areas: this provides rainfall low in sea salt but enriched with lithogenic (eg silica) and pollutant (heavy metals, sulphate, nitrate and ammonia) components. For sodium, magnesium, strontium, bromide and chloride, these components of rainfall come from marine sources; they are transferred from the sea to the atmosphere as sea-spray. Although the rainfall concentrations of these components are very variable, the ratios of their concentrations remain close to those of sea-water. For the non-sea salt components, rainfall concentrations are highly variable, and have low correlations due to their diverse sources: only DOC, ammonia, aluminium and iron are correlated: a composite pattern is observed for components that have both maritime and industrial sources (eg calcium, sulphate and potassium): concentrations are at their lowest at intermediate sea salt levels (Figure 2).

Mist and dry deposition provides an important source of chemicals to the catchment. For example, about 30% of the sea salts and 60% of the sulphate comes from sources other than wet deposition. In mist, concentrations of sea salt and acidic oxide components are high but they are low for the heavy metals.

Table 1. Flow weighted means and ranges for rainfall, cloud water and Afon Hafren and Afon Hore (pre-felling) stream water. All concentrations are given in μEq/l except for pH.

	Rainfall	**Cloud Water**	**Hafren**	**Hore**
Na	2.6 (0−24)	123 (3−419)	3.9 (2.3−5.7)	4.3 (3.0−5.1)
K	0.13 (0−0.9)	4.6 (0.2−16)	0.2 (0.1−1.0)	0.2 (0.1−0.6)
Ca	0.2 (0−2)	5.1 (1−15)	0.8 (0.3−2.0)	1.3 (0.6−4.2)
Mg	0.3 (0−3)	15 (0−51)	0.8 (0.3−1.5)	0.9 (0.6−1.5)
SO_4	1.6 (0−9)	38 (8−106)	4.4 (2−12)	5.3 (3−11)
Si	0.34 (0−20)	0.03 (0−0.3)	1.3 (0.5−4.6)	1.5 (1−5.6)
DOC	0.4 (0−3)	1.6 (0−12)	1.8 (0−4.4)	1.5 (0−5.6)
NO_3	0.8 (0−75)	11 (2−69)	1.5 (0−6)	1.6 (0.4−3.3)
NH_4	0.3 (0−5)	3.3 (1−24)	0.02 (0−1)	0.04 (0−0.8)
PO_4	0.04 (0−2)	0.03 (0−0.02)	0.05 (0−1)	0.06 (0−0.02)
F	0.03 (0−0.23)	0.06 (0−0.41)	0.05 (0−0.2)	0.05 (0−0.2)
Cl	5.1 (0−44)	220 (2−740)	7.3 (4−12)	8.24 (5−19)
pH	4.79 (3.4−7.7)	4.28 (3.5−6.0)	4.46 (4.1−6.8)	4.84 (4.3−7.4)
Alk	−6.7 (−437−201)	−48.1 (−398−31)	−22 (−63−69)	−0.8 (−47−197)

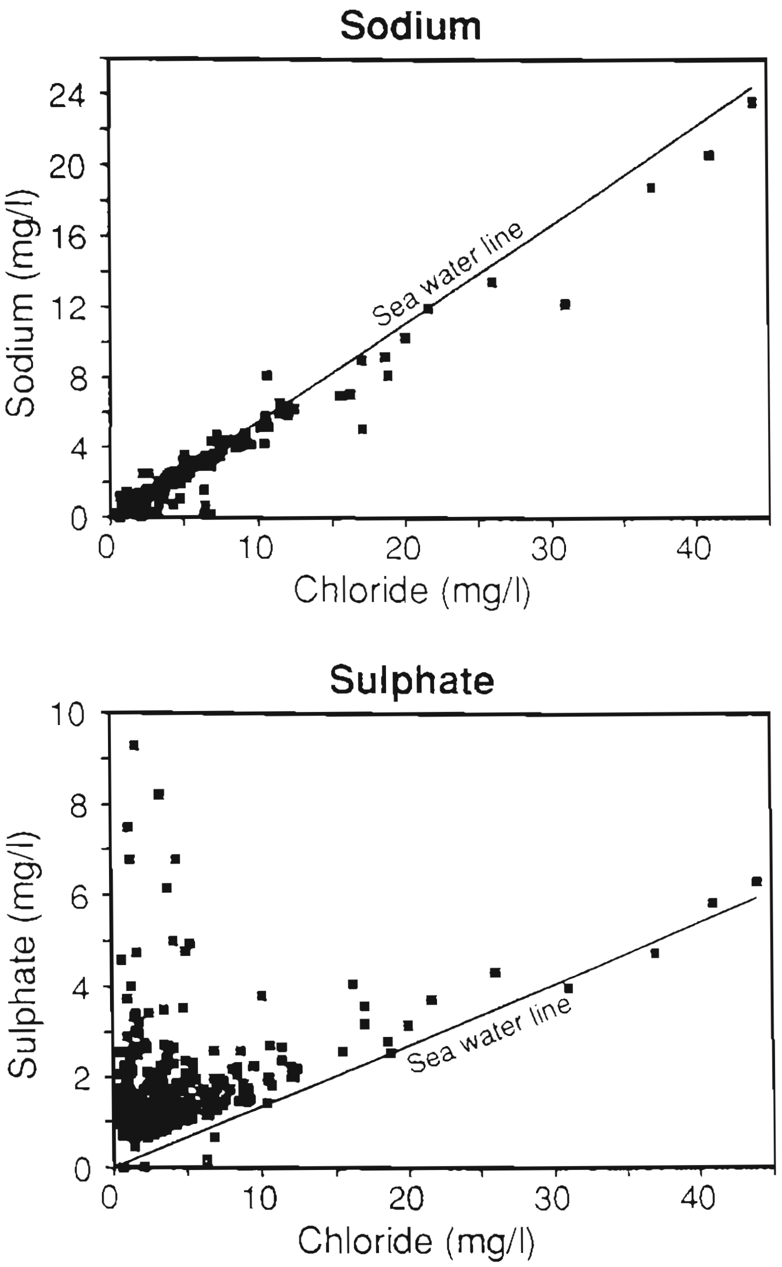


Figure 2. Sodium and sulphate concentration relationships with chloride, in rainfall.

The total pollutant loadings entering the Plynlimon catchments cannot be accurately gauged owing to uncertainties in the amount of dry and mist deposition. In terms of the pollutant inputs of sulphate, nitrate and ammonium as wet deposition, they are relatively low for the UK in terms of concentrations, although the high annual rainfall volume means that the fluxes entering the catchment are relatively high (Table 2). The trace metal concentrations in rainfall at Plynlimon are low compared with many areas of the UK due to the lack of heavy industry in the region. For nitrogen, sulphur and ammonium concentrations are moderate for the UK and are associated with long range transport of industrial and farming emissions. Despite these low concentrations, the high annual rainfall means that the total pollutant load entering the catchment is relatively high.

Table 2. Wet deposited non-marine sulphate, nitrate and ammonium in Plynlimon rainfall.

Concentrations

	Plynlimon	Mid Wales	UK range
Sulphate	18.9	20 - 40	<20 to >100
Nitrate	12.9	10 - 20	<10 to >40
Ammonium	16.7	20 - 30	<10 to >50

Fluxes

	Plynlimon	Mid Wales	UK range
Sulphate	0.76	0.6 - 0.9	<0.3 to >1.2
Nitrate	0.45	0.3 - 0.4	<0.2 to >0.5
Ammonium	0.58	0.3 - 0.6	<0.2 to >0.6

Chemical variations in streams draining undisturbed forest

The Plynlimon catchments experience rainfall with very variable chemistry while the streams exhibit a "flashy" flow response to rainfall inputs. Therefore, one might expect large variations in stream water concentrations that mirror the rainfall signal. This is not usually the case even for components such as chloride and ^{18}O, which are chemically unreactive (Neal and Rosier, 1990). This shows that the catchments have the ability to smooth the rainfall's variable chemical imprint to a very considerable degree. In other words, rainfall does not usually pass directly through the catchment to provide the major volume of water in the stream during the hydrograph response. During the major storm events, the rainfall signal sometimes influences the stream response and affects the response for the next few events

In spite of the weak relationship between rainfall quality and stream chemistry, large fluctuations in many chemical species occur in the stream because of hydrological variations (Table 3). Baseflow waters have higher calcium and silicon concentrations and alkalinities than the corresponding stormflow waters. Stormflow waters, in contrast, have higher aluminium, yttrium, manganese, cobalt, nickel, beryllium and hydrogen ion concentrations than their baseflow counterparts. The major concentration changes occur at relatively low flows: at intermediate to high flows, concentrations remain either constant or decline as flow increases. This difference reflects the large chemical gradients within the catchment. The soil zones, being organic and aluminium oxide/hydroxide rich, produce acidic, aluminium and transition metal rich soil water. The bedrock, consisting of weatherable and acid-soluble inorganic components such as calcite and layer lattice silicates (chlorite and illite), has the capacity to neutralise these acid waters and precipitate the easily hydrolysable transition metals and aluminium. Such reacted water degasses carbon dioxide on its passage to the stream and produces the characteristic low acidity, calcium and silica rich but aluminium and transition metal depleted base flow chemistry. During storm events, when the catchment soils have wetted up and groundwater tables are high, more water entering the stream is derived from the soil zone. With intense and long duration storms, more of the stormflow water is derived from the upper and most acidic soil layers and this change accounts for the declining concentrations of beryllium and aluminium at very high flow values.

Table 3. Major ion chemistries for baseflow and stormflow waters of the Afon Hafren and the upper and lower stretches of the Afon Hore for the period September 1984 to October 1990. All units are mgl^{-1} except for alkalinity (μEql^{-1}) and pH. High flow and low flow correspond with values greater than 0.1 mm/15min and less than 0.01 mm/15 min.

	Afon Hafren		Upper Afon Hore		Lower Afon Hore	
Element	**Base flow**	**Storm flow**	**Base flow**	**Storm flow**	**Base flow**	**Storm flow**
Na	3.9	3.9	3.8	4.2	4.2	4.1
K	0.1	0.2	0.1	0.1	0.2	0.3
Ca	1.0	0.8	3.1	0.7	3.1	1.0
Mg	0.8	0.8	0.9	0.4	1.2	0.8
Total Al	0.1	0.4	0.1	0.5	0.1	0.5
DOC	0.8	2.2	1.0	2.0	1.0	1.9
Cl	6.8	7.2	7.0	8.7	7.5	7.9
SO_4	3.5	4.7	3.1	4.1	4.8	5.1
NO_3	0.8	1.7	0.8	1.2	1.2	2.5
Alkalinity	24.2	-28.3	145.7	-27.7	126.6	-16.0
pH	6.3	4.6	7.1	4.6	7.1	4.8

Under baseflow conditions, water drains mainly from near the stream bank and from the groundwater areas. Variations in the bedrock composition lead to differences in the stream water chemistries of the Afon Hafren and Afon Hore (Reynolds *et al.*, 1986). The presence of calcite and lead/zinc sulphides in the Hore catchment leads to high calcium concentrations and measurable lead and zinc concentrations in the stream (a factor of four or more difference between the two streams). Other differences in bedrock composition lead to contrasting behaviours of strontium and magnesium for the two catchments. Concentrations of lead and zinc in the Afon Hore are lowest under baseflow conditions. Thus either the parent sulphides in the bedrock are not oxidized to soluble forms or there is an unidentified ion exchange or solubility control in the groundwater. Differences in the bedrock are reflected in the correlation structure of the chemical data.

Chemical variations in streams draining disturbed forest

The effects of felling were dominated by the hydrologically induced variability of the system and compounded because the streams are not simultaneously sampled. To resolve this, time series plots were constructed for smoothed data by using a "running mean" covering periods of 4 months. The major changes occurring are as follows (Figure 3),

The sea salts

The stream response to major sea salt rainfall events has changed following felling. The stream values are reduced relative to the control response and over time the difference between the responses widens. Since the decline in flow is relatively small following harvesting, evapotranspirational variations cannot explain this difference. The results fit well with the notion that trees scavenge sea salt components as mist and dry deposition. They suggest that this enhanced deposition corresponds to an accumulation of an extra 20% of catch compared to the rainfall: input-output mass balances gives a value of about 30%, but the loss of trees would not totally remove vegetation/ground scavenging.

The nutrients

Biological activity within the soil affects the concentrations of nutrients in the stream water. Biological activity occurs at three levels (Stevens and Hornung, 1987; Stevens *et al.*, 1989; Emmett, 1989; Hughes *et al.*, 1990). First, decomposition of brash and stumps: this releases potassium and dissolved organic carbon with either a decrease or an increase in nitrate depending upon circumstances. Second, there is a break in the nutrient cycle as there is no longer uptake by the trees: this leads to more potassium and nitrogen being available for leaching. Third, there is increased mineralization of organic matter leading to the soil water being supplied with additional dissolved organic carbon and organic nitrogen. Nitrate supplies also increase due to the nitrification of ammonium. For the control there are small seasonal cycles.

Nitrate and potassium concentrations peak in the winter to spring period before felling and in the autumn to winter after felling (a 3 month phase shift). It seems that the presence of brash on the catchment limits vegetation development and this depresses uptake by the vegetation during the growth period. The breakdown of the brash is insufficiently rapid to remove nitrate by fungal and microbial processes.

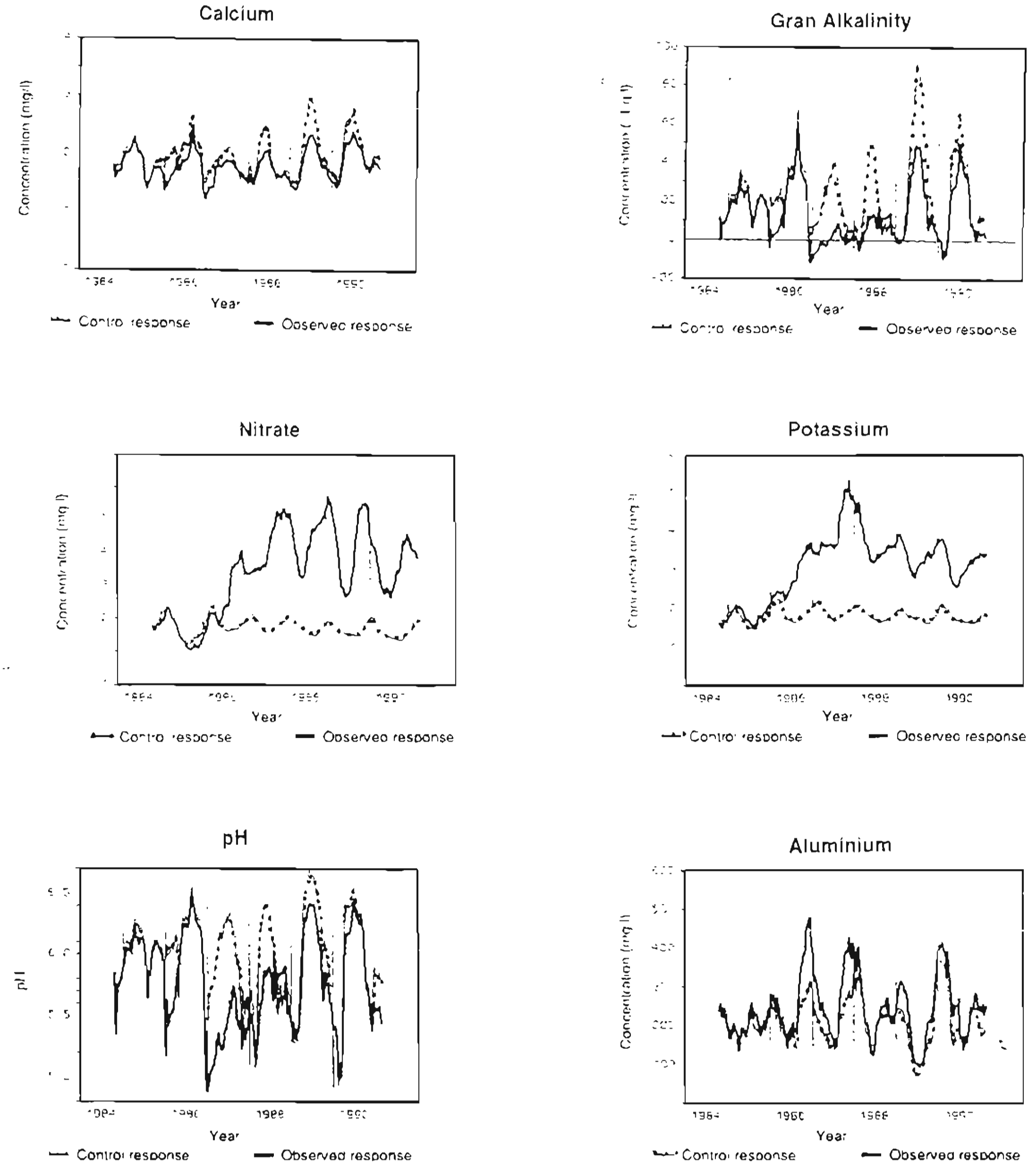


Figure 3. The effects of tree felling on nitrate, potassium, aluminium, calcium, alkalinity and pH: data from Neal et al. (1992b,c).

Variations in dissolved organic matter are linked to several processes: the general disturbance of the land leads to greater solution of organic components; hydrological conditions change so that more water is supplied from the organic-laden upper soil water; the micro climate changes, with the loss of tree shade increasing soil temperature and thus the decomposition rate. For boron, bromide and iodine, the hydrobiogeochemical changes are complex and linked to the variations in the uptake and release rates by micro-organisms and the vegetation as well as to the rate of decomposition.

The inorganic components

The effects of tree harvesting on the inorganic components relate to catchment acidification. In particular there are declines in the concentrations of calcium and alkalinity for the first few years after felling. Surprisingly, these declines occur during the summer, when the baseflow contributions are at their highest. This implies that it is the groundwater which shows the major signs of acidification. There are some signs of soil water acidification as well, but this is primarily observed for aluminium. Thus, seasonal cycling of aluminium is enhanced for the first 2 years following harvesting: the main increases occur during the winter period when aluminium concentrations are at their highest. This indicates that there is a release of aluminium from the soils due to increased nitrate concentrations displacing aluminium from exchangeable cation sites.

Assessing the importance of water source contributions to stream flow

Rationale

The variations in stream water chemistry are linked, at a gross scale, to varying contributions from soil and ground water inputs. To fully interpret the hydrochemical processes operating, there is a need to examine in more detail the chemical hydrograph in order to determine the dynamics of stream water supply. Unfortunately, water mixing relationships are difficult to assess for the streams. Although a neutral bicarbonate-bearing groundwater component is inferred from the stream data, such an endmember has not been sampled until very recently with the introduction of boreholes within the catchment. Most chemical components which exhibit response to flow are chemically reactive and hence conservative mixing formulations cannot be used. Nonetheless, basic inferences can be made which provide some insight into the nature of the hydrological and hydrochemical processes operating. This is considered here by examining the chemical signals in the light of both weekly samples and continuous measurements of pH and conductivity.

Hydrograph separation

Hydrograph separation encompasses a range of techniques which try to elucidate the hydrograph storm response by subdividing the storm waters into two or more components (Neal *et al.*, 1990b,d; Robson and Neal, 1991). Early methods of hydrograph separation were either graphical (Sherman, 1933), or mathematically-based black-box methods

(Hewlett and Hilbert, 1967), both relying entirely on rainfall-runoff data. The results of such exercises are usually a *"slow"* component, assumed to tie in with subsurface flows, and a *"quick"* component, linked to the rain waters. Unfortunately, there is no single unique separation, so that *"quick"* and *"slow"* components are method, not process, dependent. As Hewlett and Hilbert stated, such methods are *"one of the most desperate analysis techniques in use in hydrology"*. Any linkage with physical processes must therefore be considered highly dubious (Beven, 1991).

Chemical separation of stream hydrographs was carried out by Pinder and Jones (1969) using total dissolved solids, and was followed, in the late 1970's, by the use of isotopes (Sklash and Farvolden, 1979). Here, the separation was explicitly made between *"old"* or pre-event water and the *"new"* rainwater. This lead to the unexpected discovery of the dominance of pre-event water in the storm hydrograph response - a feature of the Plynlimon catchments as well. Many other isotope studies have made similar findings (Sklash *et al.*, 1986; Sklash, 1990; Pearce *et al.*, 1986; Rohde, 1981; Hooper and Shoemaker, 1986) indicating the dominance of the pre-event water. However, isotopic separation between rain and pre-event water fails to address many of the variations which are important in terms of stream water quality. Along with many non-isotopic hydrograph separations (Wells *et al.*, 1991; Pilgrim *et al.*, 1979; Hooper and Shoemaker, 1986), the studies neglect the upper soils and any interaction between rain water and soil water; they fail to address within-catchment chemical gradients (Kennedy *et al.*, 1986). Yet, the soil water component has been shown, using 3 component isotope analysis, to be of a different composition to baseflow waters (DeWalle *et al.*, 1988).

Chemical endmember mixing techniques provide a basis which is suited to allowing a subdivision of pre-event waters into distinctive components. Endmember mixing techniques were initially used in the estuarine sciences (see Liss, 1976 for summary). Here, the aim was to explain observed chemical relationships for estuary waters, by comparing selected chemical determinands against a conservative measure, the salinity. The highly saline sea waters and the fresh river waters provided two very distinct water types or endmembers, the relative proportions of which were estimated from the resultant salinity. A straight line relationship between salinity and a determinand was an indication of conservative mixing whereas a non-linear relationship was thought to imply non-conservative mixing (Liss, 1976; Boyle *et al.*, 1974).

The starting point for assessing stream water mixing relationships is to examine concentration changes for the most variable but chemically conservative tracer, acid neutralisation capacity.

The Acid Neutralisation Capacity is a measure of the acidity of a solution and is closely related to pH. Like pH, the ANC shows substantial variations between high and low flows that reflect chemical gradients within the catchment. In contrast to pH, ANC behaves conservatively on mixing; it is unaffected by CO_2-degassing/$Al(OH)_3$-precipitation-solution changes (Reuss and Johnson, 1986), and by precipitation/solution of H-organic and Al-organic ('humic') substances. The primary restriction on the use of ANC is that, during mixing, ion exchange and weathering, reactions involving H^+ and/or inorganic aluminium with other ions are not important. The Acid Neutralization Capacity is defined here (according to Reuss *et al.*, 1986) as

$$\mathrm{ANC} = \Sigma \text{ strong base cations } - \Sigma \text{ strong acid anions.}$$

Thus, for the Plynlimon waters, ANC is approximately given by

$$\mathrm{ANC} \approx [\mathrm{Na^+}] + [\mathrm{K^+}] + 2[\mathrm{Ca^{2+}}] + 2[\mathrm{Mg^{2+}}] + [\mathrm{NH_4^+}] - [\mathrm{Cl^-}] - [\mathrm{NO_3^-}] - 2[\mathrm{SO_4^{2-}}] - [\mathrm{F^-}]$$

or, alternatively, using the charge balance constraint, as

$$\mathrm{ANC} = \Sigma \text{ weak acid anions } - \Sigma \text{ weak base cations}$$

The baseflow is assumed to be representative of the groundwater chemistry. Groundwater has not been directly sampled and average baseflow ANC's have been used to depict the endmember value. Within the soil zones, mean ANC differs for the grassland and forested parts of the catchment and arealy averaged values have been used for the calculations. Under such circumstances the % hillslope (soil) ($\%_{hs}$) in a water sample of measured ANC ($\mathrm{ANC_m}$) is estimated from the simple mixing equation

$$\%_{hs} = 100(\mathrm{ANC_m} - \mathrm{ANC_g})/(\mathrm{ANC_{hs}} - \mathrm{ANC_g}):$$

g and hs denote groundwater and hillslope water values respectively.

Results for this exercise show that as flow increases the proportional contribution of soil water from the hillslope increases and then levels off at high flows (Figure 4). Even under high flow conditions the groundwater component can form a significant contribution to the total flow.

The use of ANC only allows separation of the hydrograph into two end-members. Studying the behaviour of other chemical species provides a means of resolving further end-members even though mixing may no longer be conservative. Different species show increased contributions in particular flow regimes, indicating when chemically distinct soil horizons begin to flow.

Manganese, cobalt and nickel behave fairly conservatively except during drought conditions when they may be adsorbed onto the stream bed. The conservative-mixing nature of these components is most clearly shown when the concentrations are plotted against the acid neutralization capacity since a straight line relationship is observed.

Aluminium concentration changes take place over a wide range of flows. When concentrations are plotted against ANC, a curvilinear graph results. This behaviour is explained by non-conservative mixing involving aluminium solution/precipitation either in the stream or in the near stream bank areas and this is discussed later in this paper.

Hydrograph separation: continuous data

Over the past three years, a considerable amount of data has been collected on changes in rain and stream water pH and conductivity (Figure 5). At present this data has not been fully evaluated given the complexity of the relationships being seen. Assessment has to be

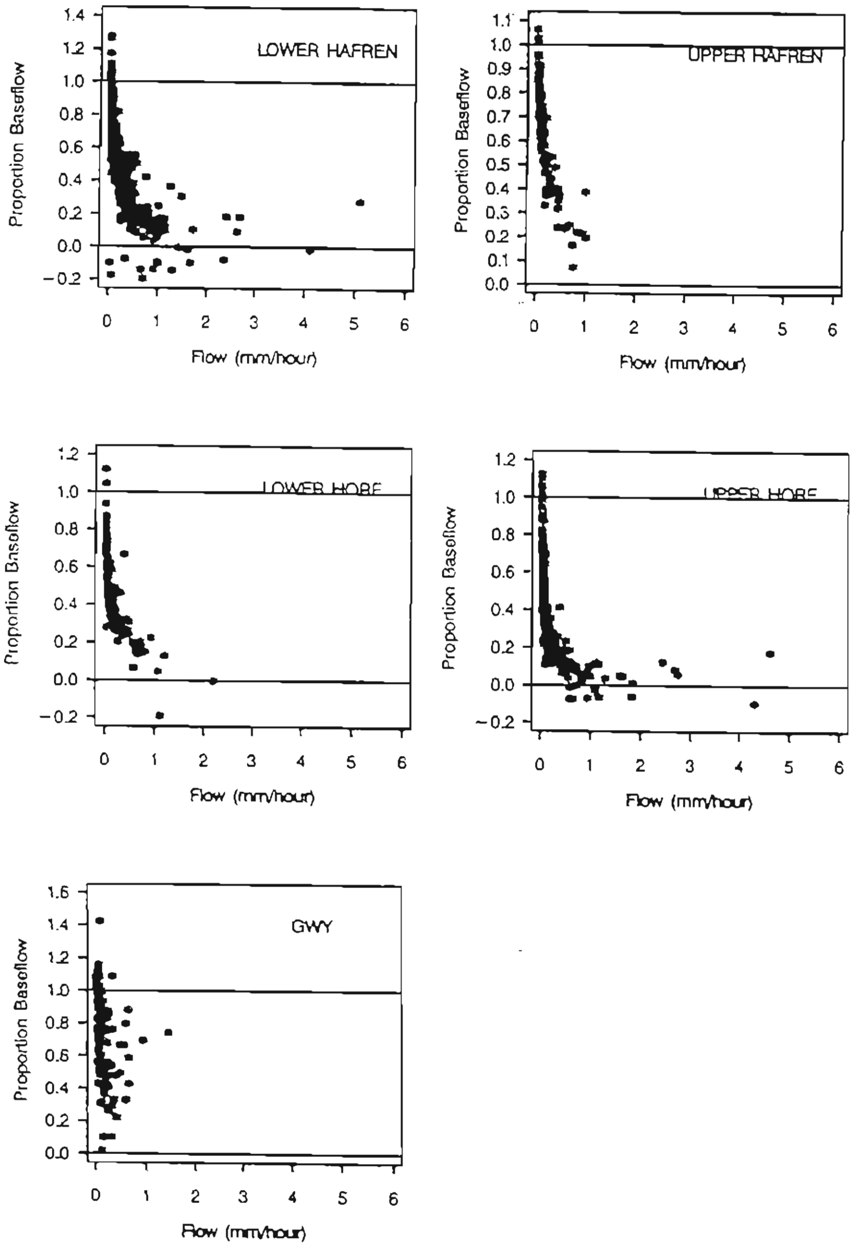


Figure 4. The ANC hydrograph split for the Plynlimon catchments: data from Robson 1993.

made on a storm to storm basis and the statistical and modelling work has yet to be fully established. Nonetheless, the early results show four main features. First, the short term changes in stream pH are very similar to those described for the weekly data. There is some hysteresis between the hydrograph rise and decline, but pH levels are generally quite closely linked with flow and this relationship does not change significantly with time. Second, there is a broad relationship between conductivity and flow on an event basis: as flow increases so too does conductivity. The rainfall signal is not usually seen within the stream hydrograph response. Such behaviour fits in well with the two component mixing model in which rainfall contributions are assumed to be small. However, for large storms or when the catchment is very wet, the rainfall signal is, in part, rapidly transferred to the stream. Third, despite the broad relationship between conductivity and flow, examination of a series of storm events reveals that the response to flow changes with rainfall inputs to the catchment. These changes indicate that the soil water endmember does not have a constant chemical composition. When the catchment is supplied by quantitatively significant amounts of rainfall possessing either low or high salt contents, the soil water end member has a correspondingly low or high conductivity. Fourth, with the onset of a storm, the conductivity sometimes decreases before it increases. This pattern is most pronounced for the Afon Hore, although it is also observed on the Afon Hafren. This indicates the presence of a second groundwater endmember for the Afon Hore and a differential behaviour between the Afon Hafren and the Afon Hore. Given the rapidity and the short lived nature of the response, the results indicate the input of a water from near the stream bank. Such differences could not be spotted with the weekly data: they are minor and occur over a very short time period.

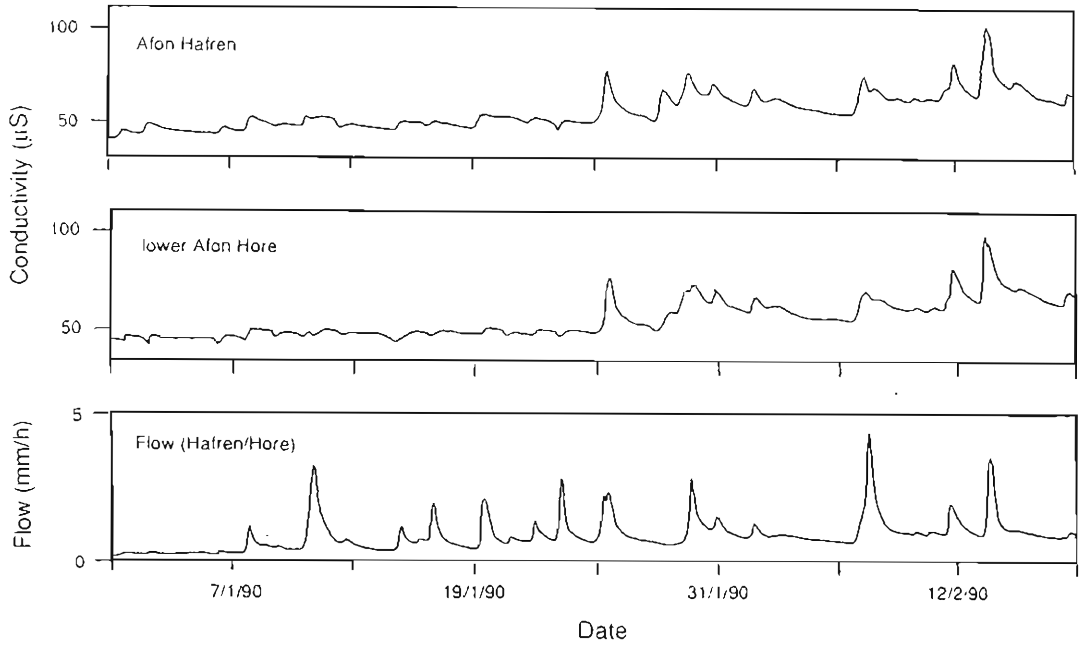


Figure 5. Short term conductivity changes in Afon Hafren and Afon Hore stream water: data Robson 1993.

Conclusions concerning water mixing relationships

A broad explanation of the chemical variations within the Plynlimon stream waters can be formulated using weekly spot sample data, but the system is of sufficient complexity that more studies are required to describe mechanistically the detailed changes occurring. The results indicate that the water supplied to the stream comprise a mixture of at least two types of ground water and two types of soil water as well as rain water.

Chemical mixing, speciation and solubility controls for aluminium

Rationale

An important criteria for a healthy stream ecology is that inorganic forms of aluminium remain low (UKAWRG, 1988; Egglishaw *et al.*, 1986; Stoner and Gee, 1985). The environmentally harmful forms of aluminium within the streams are associated with trivalent aluminium, hydroxy-aluminium complexes and freshly formed polymeric hydrolysis products (Muniz and Leiverstad, 1980). During the acidification process, concentrations of these environmentally harmful forms increase in the streams (UKAWRG, 1988). Reliable estimates of the levels of the different aluminium complexes within solution are needed as some forms present in upland waters, such as the organic, fluoride and silicate complexes, reduce the harmful effect (eg. Chappell and Birchall, 1988; Birchall *et al.*, 1989).

In this section the subject of aluminium solubility and speciation controls is addressed. This is provided to indicate the extent of environmentally harmful and inert levels of aluminium complexes in Plynlimon stream waters. To achieve this objective, it has been necessary to simplify the analysis as the speciation is influenced by a complex series of interactive factors such as hydrology, variable source area and temperature. This has been achieved by considering an averaged situation where the stream water is considered as a mixture of soil and ground water end-members of fixed composition. Subsequently, the speciation of Plynlimon stream waters are examined to characterize the overall behaviour.

Modelling chemical speciation

The two component mixing model used here is an extension to that described above. During the mixing of soil and ground water, microcrystalline gibbsite ($Al(OH)_3$) precipitates if the mixed waters become oversaturated with respect to this phase. No allowance is made for solution of microcrystalline gibbsite under the conditions where the mixed waters are undersaturated with respect to this phase. Within the calculation, account is taken of aluminium complexation with fluoride, sulphate, hydroxide, organics and silicate. Following mixing and on passage to the stream, water chemistry is modified by degassing of carbon dioxide and the extent of aluminium complexation varies in response to pH change. Partial pressures of carbon dioxide were set to constant values: 3 and 30 times the atmospheric value for the stream and soil zones, respectively.

The detailed speciation of aluminium is evaluated using the latest version of the programme ALCHEMI (Schecher and Driscoll, 1988). This involves using (a) output data, on inorganic aluminium, pH and bicarbonate, from the mixing model described above, (b) mean concentrations for the other components needed in the ALCHEMI computations (as determined from the hydrochemical data set) and (c) an average value for organically bound aluminium in the stream.

Mixing model results

There is a strong relationship between hydrogen ion and total dissolved aluminium concentrations for Afon Hafren stream water. The relationship is non-linear (Figure 6). At low to moderate hydrogen ion concentrations total aluminium concentrations increase sharply with increasing hydrogen ion concentration: at higher hydrogen ion concentrations total aluminium concentrations increase only moderately with increasing hydrogen ion concentration. This pattern is well characterised by the mixing model. Four effects produce the curve shown. First, microcrystalline gibbsite is precipitating at low hydrogen ion concentrations - this produces a cubic relationship between trivalent aluminium and hydrogen ion concentrations. Second, trivalent aluminium is supplemented by the presence of aluminium hydroxide and fluoride complexes at low hydrogen ion concentrations. Third, on entering the stream, the mixed waters degas carbon dioxide and this causes an increase in pH but no change in the total aluminium concentration. The pH change is most noticeable at the lowest hydrogen ion concentrations when bicarbonate buffering of acidity is most important. The first term gives the main structure to the upper portion of the curve while the other two terms control the lower portion. Fourth, at moderate to high hydrogen ion concentrations, both hydrogen ions and total aluminium are unreactive and a simple linear mixing pattern results - carbon dioxide degassing does not significantly change the pH while aluminium hydroxide precipitation does not come into play.

The analysis shows the following:

1: the relative concentrations of trivalent aluminium and aluminium complexes with hydroxide, fluoride, dissolved organic matter and silicate are all significant for the water end-members and their mixtures: they comprise over 15% of the total aluminium in solution in part of the pH range (Figure 7): aluminium sulphate and aluminium hydroxy fluoride complexes make up less than 1 % of the total aluminium concentration;

2: the highest concentrations of trivalent aluminium occur at the lowest pH considered; concentrations decline to insignificant levels at pH's greater than 5;

3: at intermediate pH's a mixture of complexes occur in similar proportions;

4: in the higher pH range hydrolysis of trivalent aluminium becomes more prevalent and aluminate starts to dominate;

5: the proportions of the various aluminium complexes vary markedly as a function of temperature, although the same basic patterns are maintained. This variation is due to the pronounced temperature dependence of the equilibrium constants for aluminium hydrolysis and is most important for the aluminate ion ($Al(OH)_4$).

Consequently, the highest deviations occur at the highest pH's.

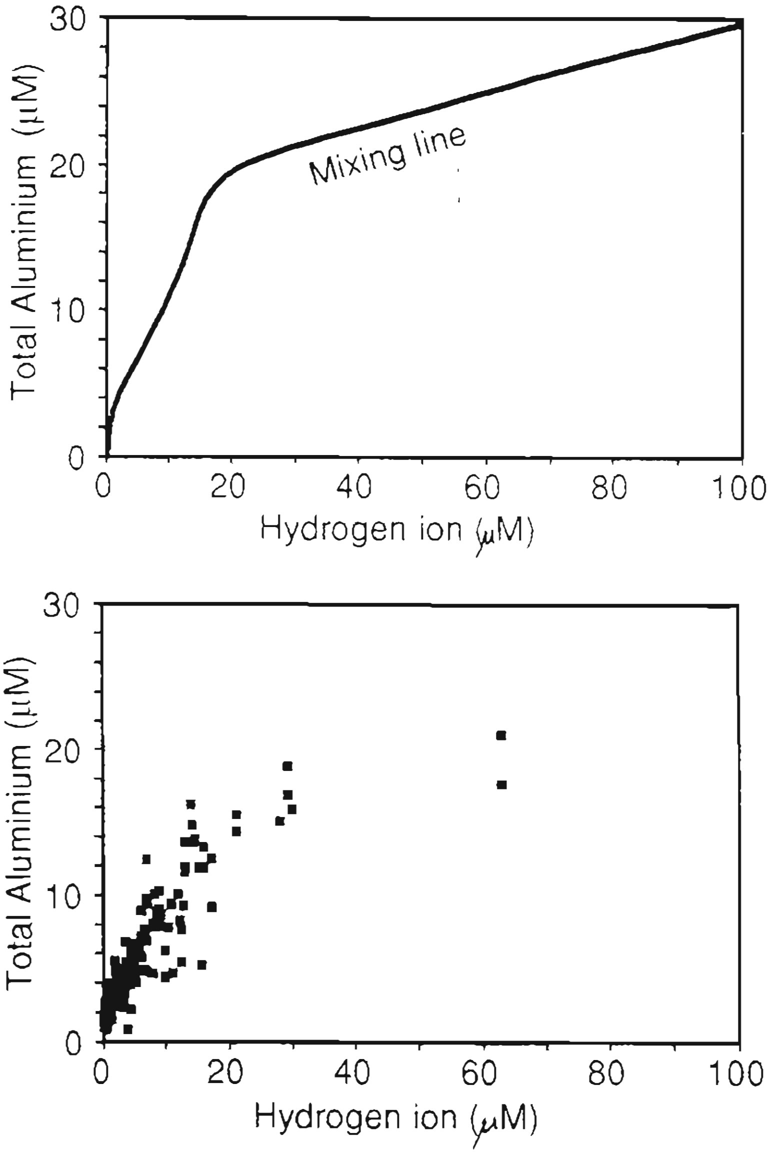


Figure 6. Concentration relationships between total aluminium and hydrogen ions for the Afon Hafren: observed and modelled values: data from Neal (in press).

Field results

Analysis of the stream water data show similar features of aluminium complexation described in the previous section. The results show:-

1: trivalent aluminium predominates at low pH;

2: although most fluoride is bound to aluminium at low pH, there is insufficient total fluorine present to complex large amounts of aluminium present at low pH;

3: aluminium hydroxide complexation increases in importance at intermediate to high pH;

4: trivalent aluminium and aluminate while important at intermediate pH's, do not dominate the system in this range;

5: while organic aluminium concentrations vary with pH, the variation is small - at low pH the organic acids are essentially undissociated while at high pH trivalent aluminium is extremely low.

The field results exhibit considerable scatter for the various aluminium species. This scatter represents the variability in the stream water chemistry and temperature. Temperature related scatter occurs for the higher pH waters: the greatest temperature dependent equilibrium constants are for those species most prevalent at higher pH; baseflow occurs during both the summer and winter periods and therefore temperature is at it's most variable (2 to 20°C).

Aluminium regulation within Plynlimon soils

One of the features of the acidification research has been the assumption that the soils undergo simple cation exchange reactions and that the transfer of acidity and aluminium to the stream are linked to the movement of "mobile anions" such as sulphate, nitrate and chloride. However, hydrogen ions and trivalent aluminium concentrations conform neither to the theoretical cubic relationship nor the associated temperature dependent co-relationship as discussed in more detail later in this paper (see also Neal, 1988a,b, 1990c; Mulder *et al.*, 1989; McMahon and Neal, 1990). These findings are reinforced by hydrochemical features associated with the catchment scale perturbations. The simple ion exchange equations commonly employed, indicate that all the base cations increase as the anion concentration increases (Reynolds *et al.*, 1988; Neal *et al.*, 1989); this is not the case with calcium and sodium as their concentrations decrease in the stream under high flow conditions following felling. Furthermore, the sea salt events which occasionally characterize the stream chemistry are not possessed of abnormally high acidity and aluminium levels in contradiction to the simple cation exchange formulation. Consequently, one of the most important features of the Plynlimon study has been the recognition that, despite over a decade of intensive research world-wide, the mechanisms determining hydrogen ion and aluminium mobility, remain obscure.

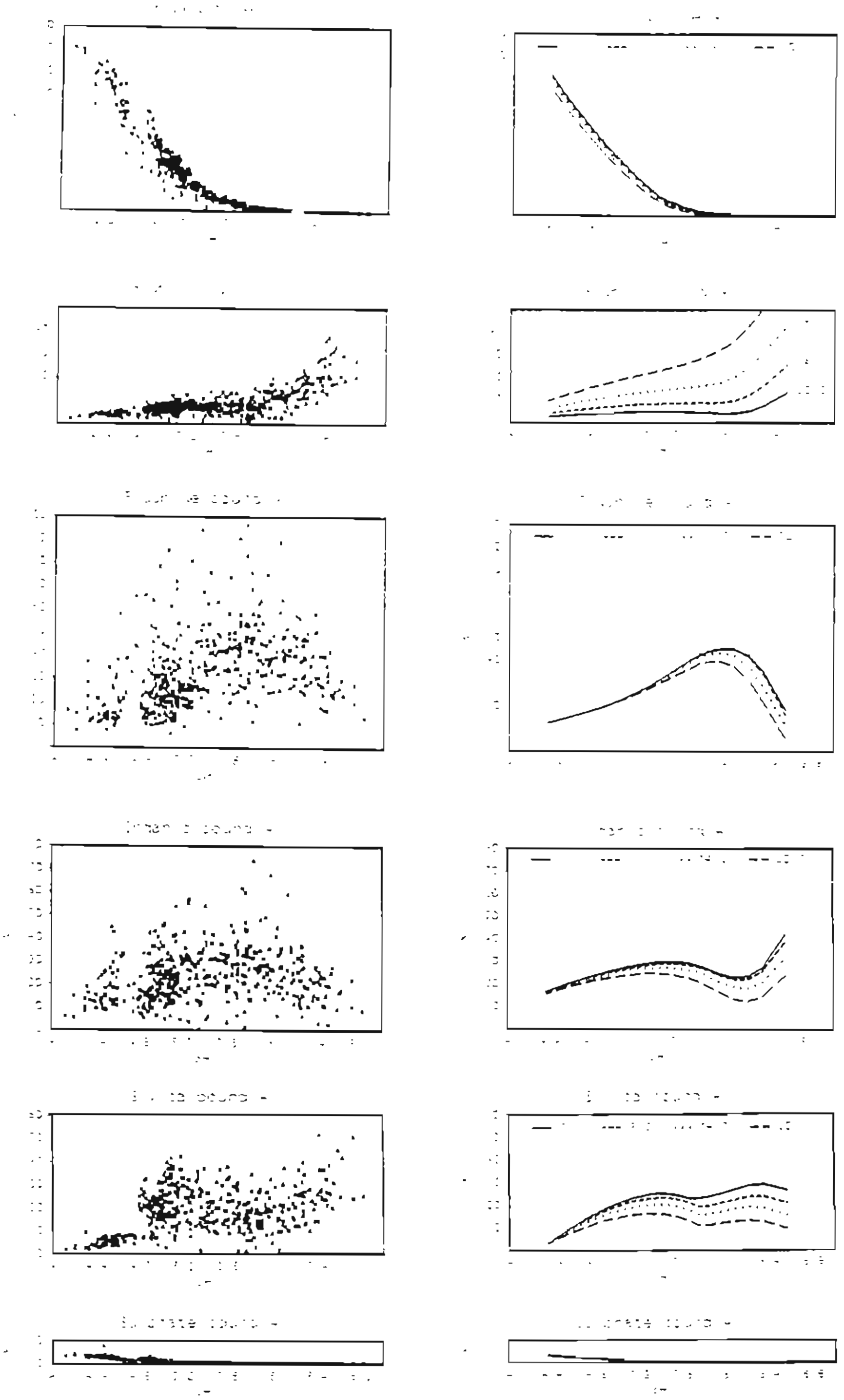


Figure 7. Aluminium speciation in Afon Hafren stream water: chemical mixing model results and field data: data from Neal (in press).

Modelling studies

Modelling background

Modelling work has been undertaken for the Plynlimon catchments to describe both the short and long term variations in stream water quality. These modelling studies constituted part of the "Surface Water Acidification Programme" and the European Economic Communities "Encore" project. A brief summary of the findings so far are presented and the future direction of the work is considered.

Short term modelling studies

The Birkenes model

As one outcome of the extensive acid deposition modelling effort, the Birkenes model was formulated (Christophersen *et al.*, 1982). The model was conceptually simple (Figure 8). It comprised a two reservoir hydrological model (plus a top snow melt layer) operating on a daily time-step upon which has been imposed what was at the time deemed the important chemical processes that control the acidification of catchments. Inputs to the model are precipitation, mean daily soil and air temperature and sulphate deposition rates. The model represented stream flow as a mixture of soil water from the top reservoir and groundwater areas by the lower reservoir. The importance of the model was that it forever dispelled the myth that acid rain passed straight through catchments to give acid surges in streams, thus destroying fish and other aquatic life. It provided the main sounding board for recognising that within catchment processes (cation exchange and mineral weathering) coupled with hydrological flow pathways are critical in determining runoff chemistry.

Each of these reservoirs obeyed simple equations that were representative of classic hydrological rainfall-runoff modelling and chemical reactions were depicted using equations that conformed with the classic soil science literature. Within the model, the processes depicted for the soil zone (the upper reservoir) were -

1) Water - precipitation, evapotranspiration, infiltration to the lower reservoir and discharge to the stream.
2) Sulphate - wet plus dry deposition, adsorption/desorption and mineralization.
3) calcium and magnesium - cation exchange.
4) hydrogen ion - cation exchange, equilibrium with $Al(OH)_3$ (gibbsite) and bicarbonate.
5) Trivalent aluminium - equilibrium with $Al(OH)_3$ (gibbsite).
6) Inorganic carbon - equilibrium between H_2CO_3 and HCO_3^- in the soil solution and seasonally varying CO_2 in the soil atmosphere.

Within the groundwater areas (the lower reservoir), the reactions were deemed to be,

1) Water - infiltration, evapotranspiration, infiltration to the lower reservoir, piston flow and discharge to the stream.

2) Sulphate - adsorption/desorption and reduction.
3) calcium and magnesium - release by weathering.
4) hydrogen ion - consumption by weathering and equilibrium with $Al(OH)_3$ (gibbsite) and bicarbonate.
5) Trivalent aluminium - equilibrium with $Al(OH)_3$ (gibbsite).
6) Inorganic carbon - equilibrium between H_2CO_3 and HCO_3^- in the soil solution and the soil atmosphere.

The Birkenes Model

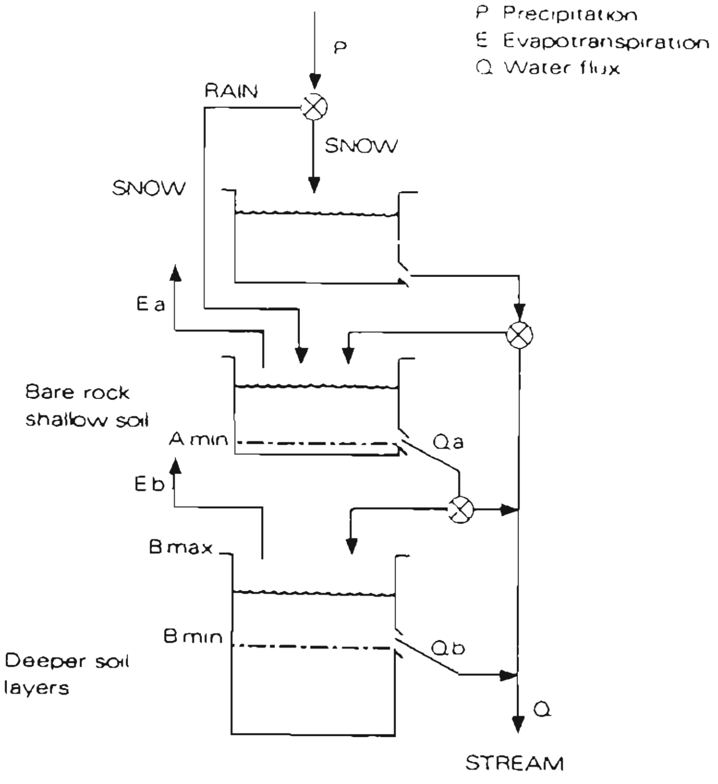


Figure 8. The Birkenes Model.

Birkenes model application

Initially it was intended that the Birkenes model would be applied to the full suite of major element chemistries. At the time of application, Birkenes type modelling was mainly confined to flow and ions typically associated with the acidification process (hydrogen, base cations, inorganic aluminium and sulphate). However, here, the first step taken was to examine chloride. This component can be considered to be a conservative tracer and the large variations in chloride in rainfall provides a good chemical tracer to test the model.

The model was first calibrated for hydrology using data for the period April 1983 to May 1985. The results of this exercise showed an extremely good fit (Figure 9). Indeed early simulations showed a poor fit for two months during the calibration period and further examination showed the original hydrological data to be incorrect at the anomalous time. However, when the model was run for chloride using weekly bulked rainfall data with weekly grab sampling of stream water, serious discrepancies occurred (Figure 9). Thus the model predicted increases in chloride concentrations during baseflow periods and large fluctuations in concentrations due to variations in atmospheric inputs and flow, which were not observed. In order to explain these differences within the current model structure, modifications are needed. On a qualitative level, these could include -

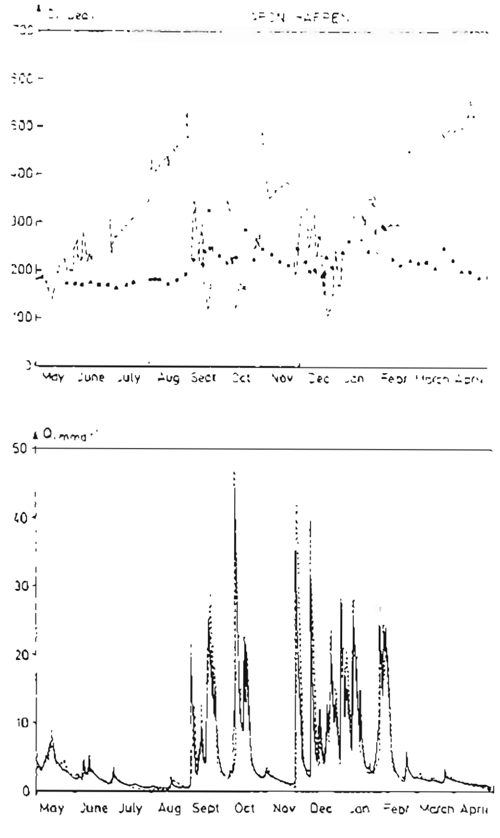


Figure 9. Flow and chloride simulations of the Afon Hafren using the Birkenes model: data from Neal et al. 1988.

1) increasing the amount of water in each reservoir or adding a partially mobile chloride store in the upper, soil, reservoir: this would increase the damping response to the rainfall signal.

2) decreasing evapotranspiration from the lower reservoir - this would reduce the modelled increase of stream chloride concentration during the summer months.

3) Increasing the dry deposition of sea salts during the summer months or introducing bypass mechanisms whereby rainwater passes directly through the upper soils, without chemical mixing, before entering the groundwater zone. Again, this will negate the modelled stream chloride concentration increase during the summer.

Despite the elegance and simplicity of the Birkenes model, in order to improve the model fit, extra processes needed to be included to match prediction and field observation. These extra processes have not been found conclusively from field evidence. During the analysis it was realized that the depiction of catchments in terms of upper and lower reservoirs was incorrect in that the Plynlimon catchment being modelled is both long (one kilometre and thin (one metre) and that water on travelling through the catchment to the stream has many opportunities to mix through the various soil layers by stochastic flow routing processes and this may well produce the required damped chloride response. In contrast, there are considerable vertical chemical gradients that can regulate other components such as hydrogen, base cations and aluminium (see earlier sections on the hydrogeochemical functioning of the Plynlimon catchments). Thus there is no reason why the chemically responsive components should be damped during a hydrograph response to rainfall.

Topmodel

Modelling studies of chemical changes occurring in the Afon Hafren and Afon Hore during storm events has centred on the application of a semi-distributed physically based hydrological model - TOPMODEL (Beven and Kirkby, 1979; Beven *et al.*, 1984) by Dr Alice Robson (1993). TOPMODEL represents evapotranspiration, interception,and transpiration losses coupled with estimates of saturated sub-surface flows, unsaturated flow and quick flow. The essence of the hydrological structure in TOPMODEL lies in the exponential transmissivity curve which determines the saturated flow characteristics. Catchment topography is examined by means of a topographic index, $\log_e(a/\tan\beta)$, Figure 10; a is the area draining through a grid square per unit length of contour and $\tan\beta$ is the average outflow gradient from the square. The index is determined from a digital terrain map across a grid covering the catchment: a high index indicates a wet part of the catchment (a large contributing area or very flat slopes); areas of low index are usually drier resulting from either steep slopes or a small contributing drainage area. This model was chosen for two reasons. First, it takes into account the catchment feature most important for determining the hydrological response - the topography and allows for different parts of the catchment to wet up and dry out over time with changes in rainfall amount. Second, it is an intermediate complexity model lying between that of the Birkenes case described above and much more elaborate fully distributed models.

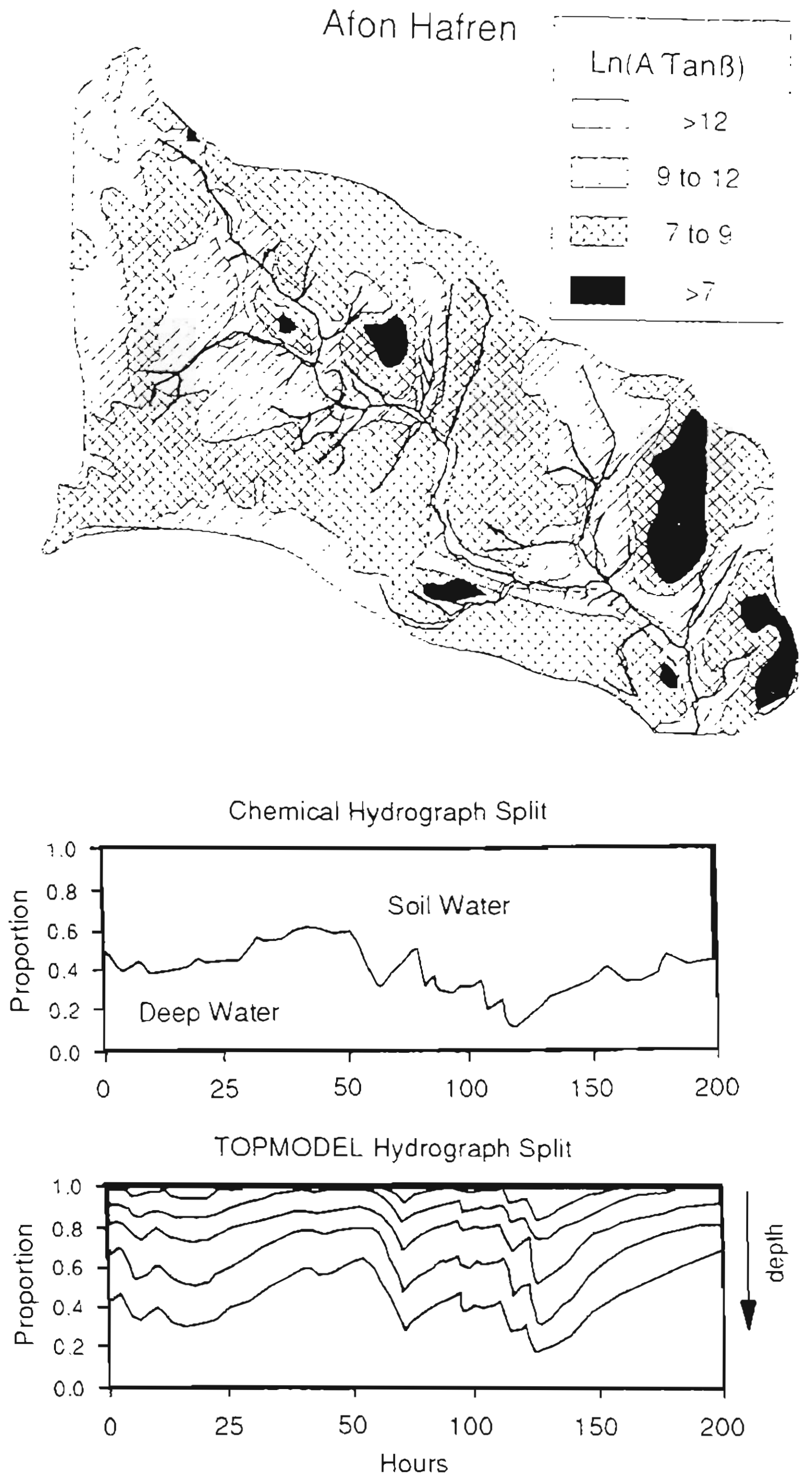


Figure 10. TOPMODEL application: the topographic index and hydrograph split comparisons between model and ANC evaluations: data from Robson 1993.

The work has focused on trying to link the hydrological and the chemical response of the stream on an episodic basis. TOPMODEL was adapted so that the contributions to the stream from different soil horizons could be estimated. To do so, TOPMODEL was used to obtain contributions from different soil layers within the catchment by estimating the depth to the water table in the banks and the flow contributions were integrated along the length of the stream. TOPMODEL was then calibrated using hourly rainfall and flow records. TOPMODEL estimates two components of streamflow - (1) quickflow, thought to move by macropore-flow, overland-flow or piston displacement and (2) a subsurface flow.

pH/ANC and conductivity signals have been modelled for the Hafren stream. For ANC, TOPMODEL was used to predict components of flow, and these were related to the chemistry of the soils. The model structure was found to be capable of explaining stream chemistry assuming mixing relationships to be appropriate (Figure 10). The results suggested that at least four endmembers are required to explain the lower and upper Hafren stream chemistry. In general, the parameters were not particularly well defined, and cross validation did not produce very good results. It is not possible to judge exactly why such problems arise, because of the problems with pH measurements. Although ANC is modelled successfully here, the results are not considered to provide evidence for mixing hypotheses - the ANC is very highly correlated with flow and this may be the real reason why ANC predictions have been possible (as in the Birkenes case presented above).

The conductivity data have been modelled using a layered structure within a simplified version of TOPMODEL, TOP1D. TOP1D is a one dimensional simplification of TOPMODEL. It retains the hydrological features of TOPMODEL but it is semi-distributed: it uses the exponential transmissivity curve as its base. This structure allows the rainfall conductivity signal to be traced as it moves through the catchment. The model is able to reproduce the broad features of the stream conductivity response - although it fails to match all of the short-term dynamics. The model implies that there are significant chemical stores with which rainwater mixes in the soil zones. Again, the results mimic those of the Birkenes case provided above.These stores account for the high damping of the rainfall inputs. The conductivity results suggest that the ranges of residence times for rainwater are very wide. The oldest stores of water, seen mainly at baseflows, change in composition only very slowly. The upper soil waters change more rapidly, but are still highly damped relative to rainfall. Estimates of the cumulative percentage of the tracer flux leaving the catchment give values of only 2/3 of the rainfall input as leaving the catchment within three months of the rainfall event. The results indicate that water within the catchment cannot be treated as having a uniform chemistry. This means that isotope separations using rainwater and baseflow components could be in error.

Long term modelling studies

Long term atmospheric deposition of acidic pollutants can cause gradual changes in the chemical stores of the soils and ground waters of sensitive upland catchments (Whitehead *et al.*, 1988). Superimposed on these long term changes are the episodic acidic pulses occurring during storm events. These episodic chemical pulses, during which hydrogen and aluminium concentrations may increase by orders of magnitude, are critical in determining whether streams can sustain healthy fish and stream invertebrates (UKAWRG,

390

1988). Fish kills may occur not only from long term chemical changes (chronic effects), but also as a result of short term variations in the stream chemistry during storm events (acute effects: Haines, 1986; Turnpenny *et al.*, 1987; UKAWRG, 1988). To understand the effects of acidification, the short term episodic response and the slowly varying long term trends must be assessed alongside one another.

A long term modelling approach is provided to consider both the average changes and the hydrologically induced short term changes. To do so the MAGIC model (Model of Acidification of Groundwaters In Catchments) is employed here.

MAGIC

MAGIC is a model designed to examine the long term changes in stream and soil waters occurring in response to acid inputs to the system. MAGIC assumes that atmospheric deposition, mineral weathering and cation exchange processes in the soil zone are responsible for the observed stream water chemistry in a catchment. Full descriptions of the conceptual basis of the model and the equations on which it is based can be found in Cosby *et al.* (1985a,b). The model comprises one or two reservoirs. In the simplest form, a one box representation, the catchment is considered as a single soil/groundwater unit in which cation exchange and weathering zones are combined. In the two reservoir case, the structure can either be like that for the Birkenes model or as two adjacent sub-catchments. Unlike the Birkenes model, the hydrological sub-model operates over monthly or yearly time steps and does not incorporate individual storm events. For the MAGIC model there is also an above-ground compartment to allow for element uptake by and release from the biomass. A sequence of atmospheric deposition and mineral weathering is assumed in the model. Current deposition levels of base cations, sulphate, nitrate and chloride are needed, along with some estimate of how these levels have varied historically. Historical deposition variations may be scaled to emissions records or may be taken from other modelling studies of atmospheric transport into a region. Weathering estimates for base cations are extremely difficult to obtain. None the less, it is the weathering process that controls the long term response and recovery of catchments to acidic deposition and some estimate is required. Several chemical, biological and hydrological processes control stream water chemistry. These are often interactive and not easily identifiable from field observation and as a consequence of this not all the factors are determinable from field measurements. Nb, recent developments with kinetic formulations of weathering may well allow direct estimates of weathering, given information on the mineral composition of the soil/groundwater-matrix, water residence time and moisture content (Sverdrup, 1990). A summary of the main processes represented in the model are as follows :-

(1) Strong acid anion concentrations are calculated for the soils and stream waters. Sulphate is assumed to follow the Langmuir isotherm

$$E_s = E_{mx} (SO_4^{2-})/[C + (SO_4^{2-})]$$

where E_{mx} is the maximum adsorption capacity of the soils (mEq/kg) and C is half the saturation concentration (mEq/m^3). This allows for the lags which can occur between atmospheric deposition and the resultant changes in stream water sulphate concentrations. Chloride, nitrate and fluoride are assumed not to have an adsorbed phase. Loss of ammonia and nitrate to the vegetation is included within the model by estimating a percentage uptake.

(2) Cation exchange processes involving aluminium, sodium, calcium, magnesium and potassium are assumed to be operative in the soils. The general form for the exchange reactions between the ions M^{m+} and N^{n+} and their adsorbed states MX_m and NX_n, is

$$n\,M^{m+} + m\,NX_n = m\,N^{n+} + n\,MX_m$$

The total cation exchange capacity is then defined as

$$CEC = \Sigma \text{ Exchangeable cations } (mEq/m^3)$$

Equilibrium expressions for the cation reactions are approximated by

$$S_{MN} = \{N^{n+}\}^m E_M^n / \{M^{m+}\}^n E_N^m$$

where S_{MN} is the selectivity coefficient between M and N, { } refers to the chemical activity of an ion in solution, and E is the exchangeable fraction for the appropriate ion on the soil complex, e.g.

$$E_N = X / CEC$$

with X equalling the amount of adsorbed cation N (mEq/m^3): the cation exchange aspect is covered in more detail later in the paper.

(3) Within the soil, aluminium concentrations are assumed to be in equilibrium with the solid phase $Al(OH)_3$.

$$3\,H^+ + Al(OH)_3(s) = Al^{3+} + 3\,H_2O$$

with equilibria constant

$$K_{Al(OH)3} = \{Al^{3+}\}/\{H^+\}^3$$

Sulphate and fluoride complexes with aluminium in solution are also included.

(4) Allowance is made for dissolution of carbonic acids caused by the elevated pCO_2 levels in the soils, and for the effects of degassing as water moves from the soils to the stream.

(5) It is assumed that there are constant long term net inputs of base cations from mineral weathering. These are difficult to measure and must therefore be estimated.

(6) Organics are modelled within MAGIC using a triprotic representation.

The MAGIC model application to describe average changes in stream water

chemistry in relation to atmospheric deposition and conifer afforestation

Early attempts to model the forestry system were confined to a one box application of the MAGIC model: the upper soil and groundwater areas were lumped together. This application, for the Dargall Lane catchment in south west Scotland modelled the effects of changing atmospheric pollution levels of catchments of moorland, afforested moorland and deforested plantation forest. For the simulations only the effects of changing atmospheric scavenging of pollution by the change in vegetation and the associated changes in evapotranspiration were examined: planting trees enhances the scavenging of pollutants from the atmosphere and increases the degree of evapotranspiration. The study showed (Figure 11 and 12):-

1) Acidic oxide deposition results in stream water acidification for both moorland and forested areas: the degree of the acidification increased with increasing acidic oxide deposition.

2) The introduction of trees enhanced the acidification effect because of the scavenging of atmospheric pollutants.

3) Harvesting of trees in an environment polluted by atmospheric deposition results in a reduction in the stream acidity and an increase in stream pH (provided that biological effects are small - i.e. nitric acid is not generated).

4) Afforestation does not cause acidification unless there are pollutant inputs to the catchments (unless base cation uptake from the growing second generation forest depletes the exchangeable cation store within the soil).

The MAGIC model application on a short time base

The modelling approach is applied to the Gwy, a typical example of a British acid-moorland upland system. Here a more detailed account of the modelling work is presented as it contains elements of the important water tracing, hydrograph splitting and hydrogeochemical framework described above. The site is selected because of its moorland vegetation; this allows investigation of the long term effects of acidic precipitation without the added complexity of examining the effects of afforestation, felling and reafforestation. The results presented here update earlier work presented in Robson *et al.* (1991) and Neal *et al.* (1992) in the light of further data (Robson, 1993). The analysis is enhanced by the inclusion of more detail on the organics. Plynlimon streams and soils. Here, the MAGIC model is used to estimate these gradual changes in the endmembers caused by acidic rainfall inputs.

The aluminium hydrogen relationship used in MAGIC does not allow directly for exchangeable hydrogen [HX], even though this may be a significant component of the exchange fraction. However, in practice [AlX_3] is used to represent both the exchangeable hydrogen and the exchangeable aluminium - since [AlX_3] is taken to be the difference between the CEC and the exchangeable base cations. This simplification is reasonable because the relative proportions of exchangeable hydrogen and aluminium will remain fairly constant; the $\{Al^{3+}\}/\{H^+\}^3$ ratio in solution is defined to be constant.

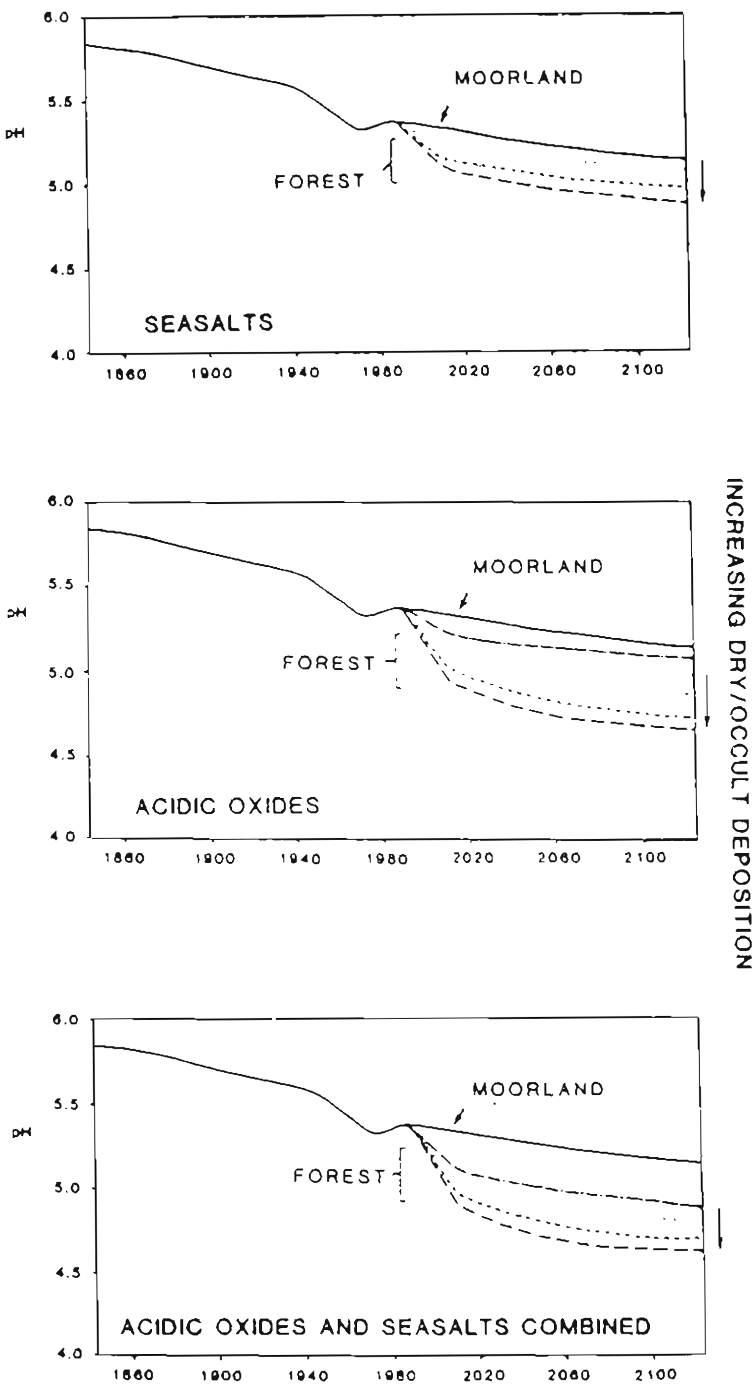


Figure 11. MAGIC simulation of the pH response from the Dargall Lane catchment, comparing moorland and afforested moorland responses to changing long term atmospheric deposition patterns appropriate for Scotland: data from Neal et al. 1986.

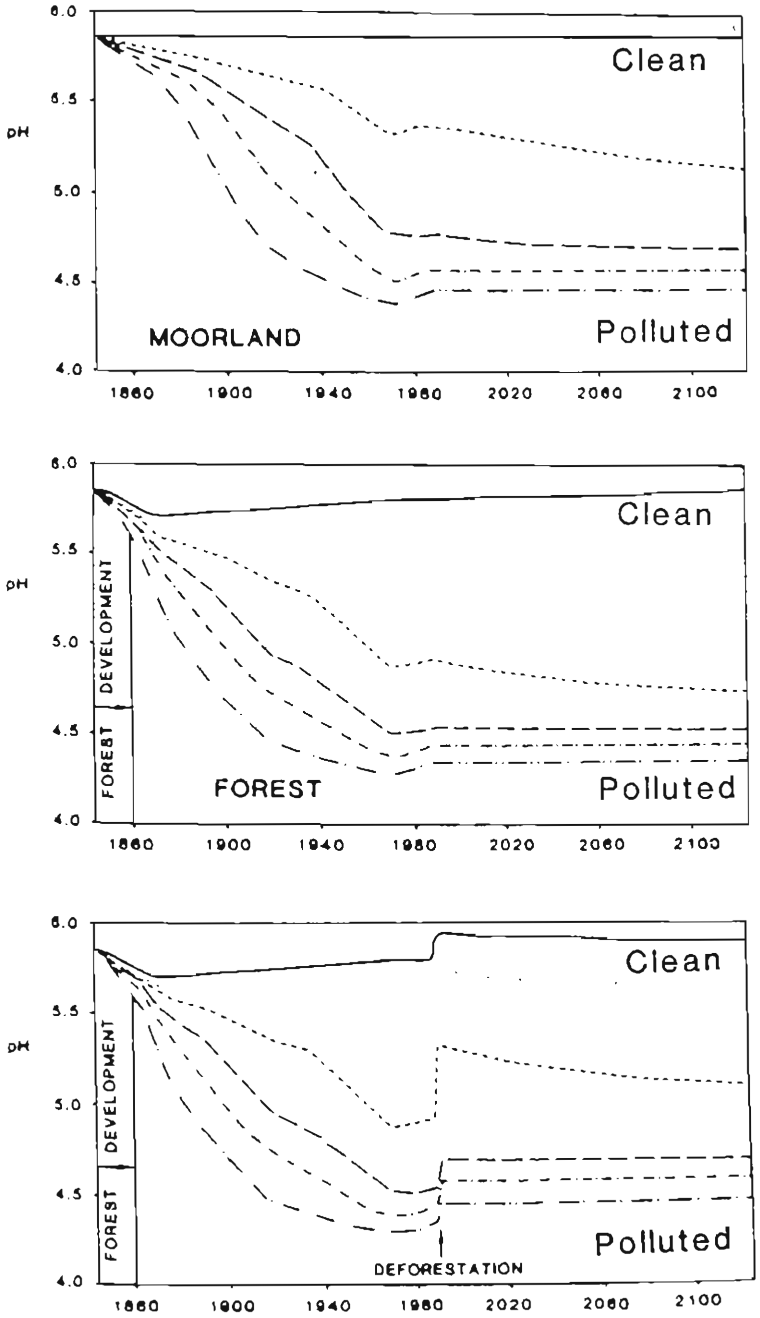


Figure 12. MAGIC simulation of the pH response from the Dargall Lane catchment, comparing moorland, afforested moorland and deforestation responses to changing long term atmospheric deposition patterns covering no to high pollution loadings: data from Neal et al. 1986.

A two layer version of MAGIC (Jenkins and Cosby, 1989) is used in the application, enabling the simulation of two chemically distinct waters. The layers are chosen to correspond with the flow components identified from the mixing considerations outlined earlier; the top layer represents the lumped O, E and B and C soil horizons, whilst the bottom layer represents the deeper till layers and other water stores which produce baseflow. The flow proportions included in the model are those suggested by the mixing approach. Concentrations of chloride and sulphate in the rainfall were adjusted for occult and dry deposition of sea-salts and anthropogenic sulphur compounds. Chloride is assumed to be conserved within the catchment, whereas sulphate is currently assumed to be in near-equilibrium, so that the total sulphate input is equal to the stream output. Differences between the rainfall concentrations and the stream outflow concentrations (allowing for evaporation effects) are assumed to result from dry and occult deposition, and these components are incorporated within the adjusted rainfall. The two layers in the model are conceptualised such that different reaction mechanisms dominate in each layer. The water from the upper soils is known to be acidic and organic and aluminium rich, indicating that ionic exchange mechanisms and organic acid deprotonation reactions are the most important influence on soil water chemistry. Water from the deeper sources is rich in base cations, but low in exchangeable cations, as it comes from a high weathering zone. In this application, it is therefore assumed that ion-exchange occurs at a significant level only in the top layer. On the other hand, weathering is assumed to take place predominantly in the deeper layers. The depth of the upper soil (0.9 m) corresponds to the average combined depth of the O, E and B and C horizons and of the peats. An average depth of 1 m is assumed for the lower layer. The bulk density and cation exchange of the upper soils is calculated from field data for the Gwy soils. The bottom layer is represented as being denser and of lower porosity. The partial pressure of carbon dioxide is assumed to be 30 times atmospheric pressure for both layers. For the stream, the pCO_2 is set to 2.5 times atmospheric pressure, in line with average observed values for the Plynlimon streams (Neal and Hill, 1993).

Very little is known about the sulphate characteristics of the Gwy soils. The selected constants which describe the isotherms were chosen to be typical of values used in earlier MAGIC applications to the uplands (Jenkins and Cosby, 1989; Whitehead *et al.*, 1988). The same sulphate absorption isotherm was used in the two layers. The model is calibrated to the present day chemistry of the stream water, the soil water and the deeper groundwater by adjusting weathering rates, uptake rates and the initial soil base saturation. Given the limitations of the assumptions required within the mixing approach, and the analytical error in the chemical measurements, a perfect match between the modelled and the observed chemical species in the stream is unlikely. The ANC gives a good indication of the overall composition of a water sample and, whereas some deviations in individual determinands are acceptable, it is important that the modelled ANC is accurate.

The model was run for the period 1844 to 1984. The selectivity coefficients and initial base saturation were calibrated so as to match present day base saturation characteristics of the upper soils and peats. Weathering/uptake rates were then adjusted for the two layers in order to match stream and endmember chemistries as closely as possible. Nitrate and ammonia were modelled by calibrating catchment uptake to match the difference between the inputs and the outputs. Weathering-inputs/biological-uptake-rates were used to match observed base cation concentrations. Optimisation was performed manually, by adjusting parameters so as to match the present day stream chemistry as closely as possible.

The model has also been used to estimate future changes in stream, soil and deep water components. This was done by continuing the already calibrated run for 1844 to 1984, using two future sulphate deposition reduction scenarios; a linear reduction in non-marine sulphate to either 30% or 60% of present day levels, starting in 1984 and being completed by the year 2000, and held at a constant level thereafter. The time trends of reconstructed soil, ground and stream water chemistry from 140 years ago, that is pre-acidification, up to 2124, are shown in Figure 13. The ANC shows a decline from 1844 to the 1980's and recovery thereafter. Recovery under 60% reduction is more substantial than with 30% reduction although the recovery takes about the same time to plateau. Initial recovery is fast and is followed by a very gradual long-term improvement in ANC. The two soil layers and the stream ANC approximately parallel one another throughout the modelled record. Stream pH falls rapidly as deposition increases in the 1950's, levels out in the 1980's, and then rises in the future in response to decreases in atmospheric deposition. The stream chemistry follows a similar pattern to the soils and groundwater, but the variations are greater owing to the lower partial pressure of carbon dioxide in the stream. As a consequence, CO_2 degassing occurs as water moves from catchment to stream, and the pH of the stream is not always bounded by the two endmember pH values, especially at higher pH levels. Sulphate concentrations in all three model compartments change in line with the assumed changes in sulphate deposition. The baseflow endmember takes longer to adjust to changes in sulphate concentration because it is replenished from the upper box and is therefore only indirectly linked to the rainwater.

The response of the base cation concentrations to sulphate reduction is somewhat complex. Calcium and magnesium concentrations generally parallel changes in sulphate input. A 30 or 60 % reduction in sulphate input causes fewer cations to be exchanged from the soils, and as a result cation concentrations decrease rapidly in both endmembers with the greatest decrease for 60% reduction. For magnesium, some long term recovery is predicted, more so for the 60% than for the 30% reduction. By 2124, the predicted stream concentrations of magnesium are very similar for both 60% and 30% reductions. For calcium there is no long term recovery. These changes are linked to the changes in soil exchangeable bases (Figure 14). Exchangeable magnesium both decreases and recovers most quickly, relative to the other exchangeable cations. A 60% reduction in sulphate input produces a significant improvement in base saturation for magnesium, but there is little recovery for calcium under either scenario.

To simulate the short term chemical changes in the stream waters, the predicted endmembers are recombined according to the proportions given by hydrograph separation. To simplify matters, it is taken that the catchment remains hydrologically unaltered (i.e. flow pathways do not change significantly), and that future rainfall patterns remain comparable to those of the 1980's.

Variations in flow are matched by altered proportions of the endmembers in the streams, and thus by altered stream chemistry. The chemical mixing reactions for the endmembers are complex because non-conservative processes come into play for determinands such as hydrogen and aluminium as shown earlier in this paper (Neal and Christophersen, 1989). The ALCHEMI chemical speciation programme has been used, again, to assess the chemical changes occurring when the endmember waters are mixed. The endmember compositions were then mixed in varying proportions (0 to 100% in 1% steps) and the resulting equilibrium chemical composition was found using ALCHEMI: conditions were set so that dissolved carbon dioxide levels corresponded with those observed in the stream

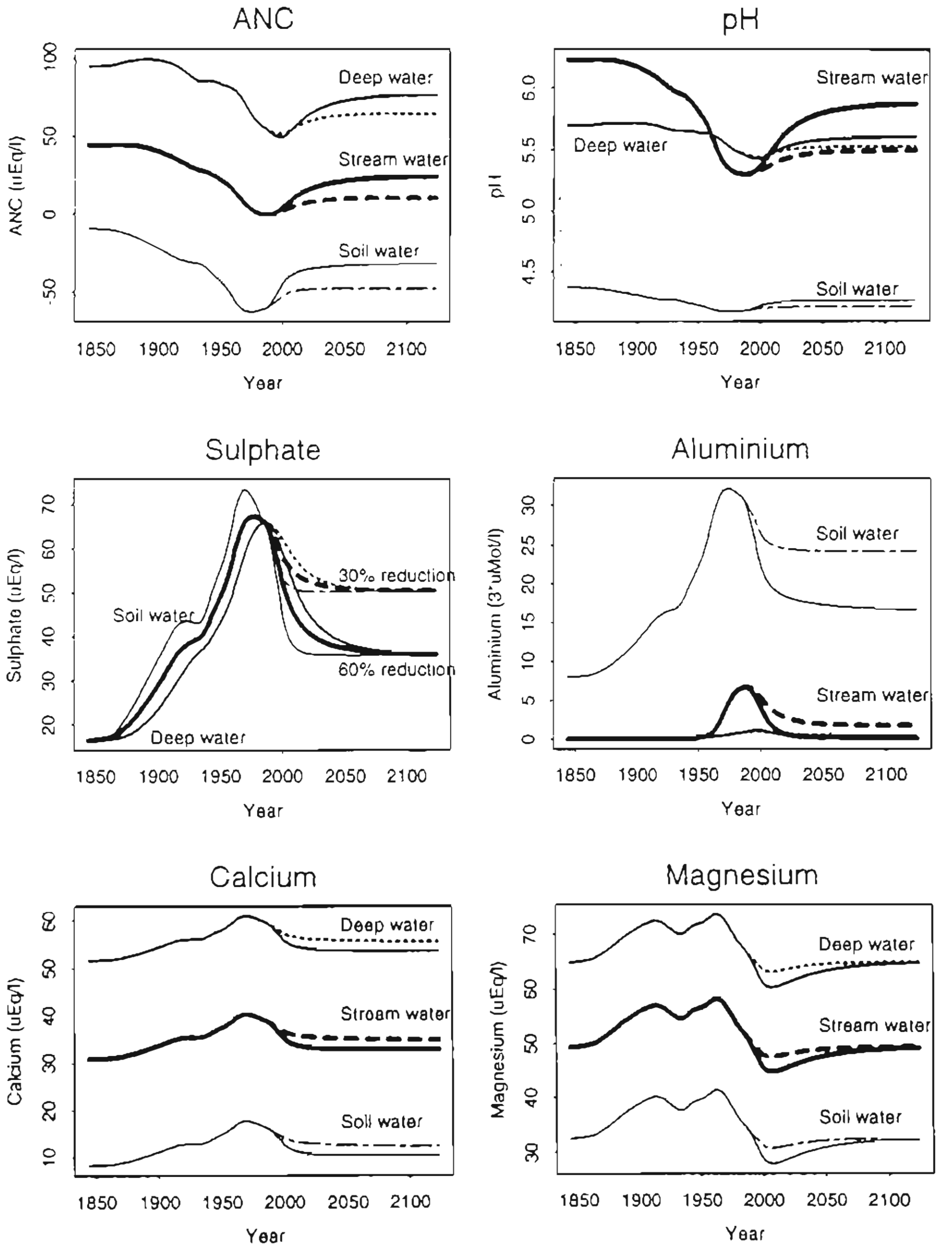


Figure 13. Time series trends in stream and soil chemistry for the two box MAGIC application for the Gwy catchment: data from Robson 1993.

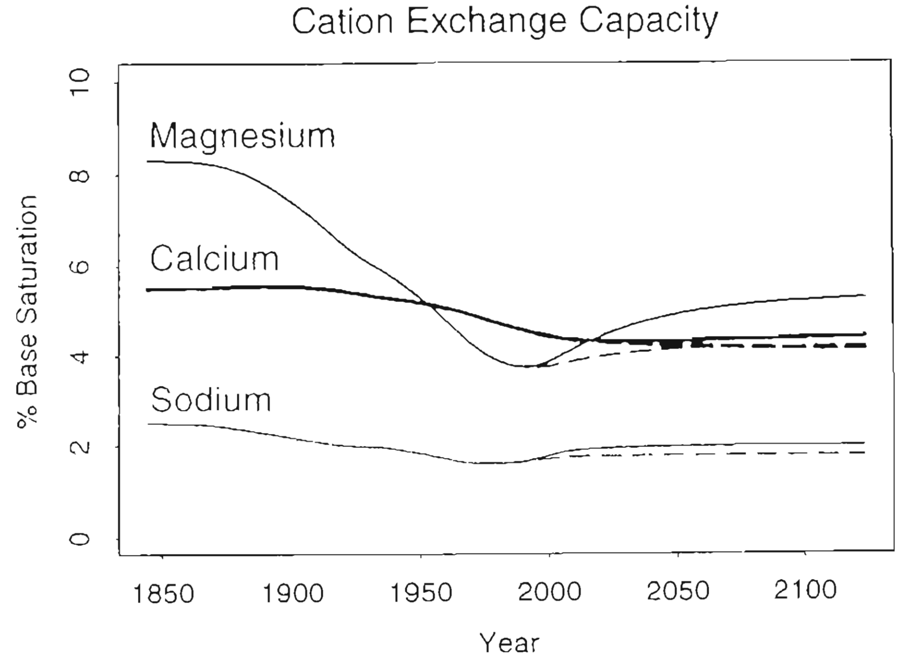


Figure 14. Time series trends in soil exchangeable base cation chemistry for the two box MAGIC application for the Gwy catchment: data from Robson 1993.

(approximately 2.5 times the atmospheric value). Within the calculation, allowance was made for aluminium hydroxide to precipitate from stream waters when they become oversaturated with respect to natural gibbsite following CO_2 degassing.

The results obtained from the mixing analysis allow the reconstruction/prediction of time series of estimated concentrations. These are produced using a time series of proportions of mix, estimated from current day continuous stream chemistry. The endmembers are combined using these proportions according to the ALCHEMI results. In general, the basic patterns of chemical storm response remain essentially the same, but the peak and trough values vary. For ANC, the time series for the different years are parallel to one another. The most non-linear response of concentration variation with flow is the total aluminium charge which is low at baseflows in all years, but at high flows, is very different in each of the years considered. Figure 15 shows the model assessment of acid events with predictions for future scenarios involving different atmospheric pollutant loadings.

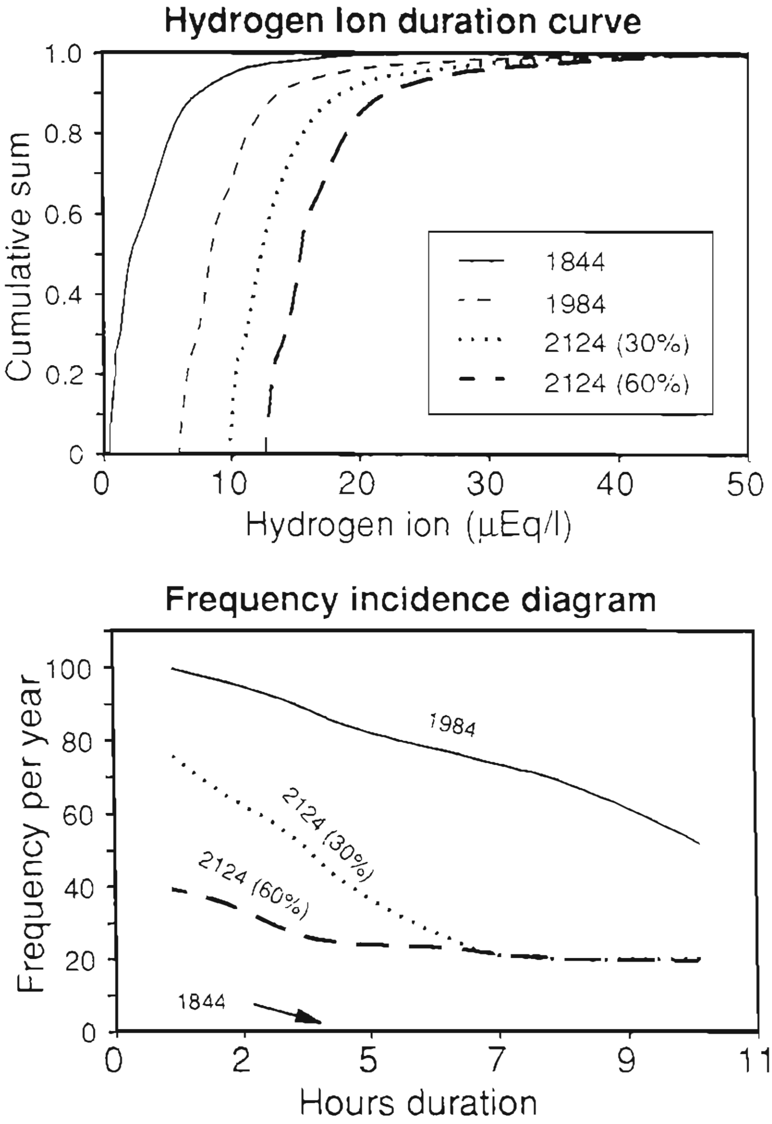


Figure 15. MAGIC model assessment of the frequency of acid events for the Gwy catchment: data from Neal et al. 1992.

The changes predicted by MAGIC suggest that the catchment endmember chemistry is likely to improve relative to present conditions, provided sulphate reduction is enforced. A 60% reduction is likely to result in a much greater improvement than 30% reduction, in the sense that the stream water chemistry will be more ecologically favourable especially at high flows. For example, the model predicts that whereas the difference in cation concentrations at high and low flows will remain the same, relative to each other, the predicted recovery in the upper soil water chemistry, will lead to a decrease in the peak concentrations of hydrogen and aluminium seen during high flows. Baseflow concentrations of both hydrogen and aluminium are currently low, and this situation will continue.

The approach of combining MAGIC and mixing techniques has been repeated in a similar from for Panola, U.S.A. (Hooper and Christophersen, 1992). Here, the model was calibrated to soil water chemistry alone and was able to explain stream water chemistry on this basis. Three endmember were used, one of which was assumed constant over time. A similar feature of both applications was the need to incorporate a base cation source into the upper soil layer.

TOWARDS LUMPED INTEGRATED MODELS OF COMPLEX HETEROGENEOUS ENVIRONMENTAL SYSTEMS

So far, the model development has been made in a step wise fashion based on the regular patterns of water quality response seen in the stream in response to rainfall events. However, in the process the wide variability in water quality seen on or within the catchment has not been considered. This additional consideration is now made here as it has many ramifications to how environmental systems can be modelled. In reality, the natural environment is complex and heterogeneous. For example, rain falling onto a region varies in amount, intensity and chemical composition in time and space in a way that cannot be predicted in a detailed mechanistic way. Indeed, this complexity can be observed within the stream as exemplified by the aluminium speciation studies described earlier in this paper. To understand the complex and variable nature of environmental systems, it is necessary to average data in a way that simplifies and mimics the broader features of environmental response. Unfortunately, in doing so, there is a major inconsistency (Neal, 1992): most of the chemical reactions are non-linear and Average(F(x)) seldom equals F(average(x)). This problem has been almost universally ignored by modellers who use lumped models and it has not been treated above in any serious way above, except for the case of the Birkenes model application to Plynlimon, and even there, the subject was only just touched upon.

In this section, chemical variability within acidic soil waters, field information and a modelling exercise is used to assess whether simple chemical theory pertaining to cation exchange may apply at the averaged level. This theory underlies both the Birkenes and MAGIC models. These data bring together and summarise a series of modelling and field based studies (Neal, 1992; Neal and Robson, 1994a,b; Neal *et al.*, 1994).

Integrating variations of soil adsorbed cations into a cation exchange model

Soils exhibit large spatial variation, in physical make up, and in their chemical and hydrological properties (Mulder *et al.*, 1991; Neal, 1992; Taugbøl and Neal, 1994; Avila *et al.*, 1994). Hydrochemical models of catchments standardly "lump" such variability so that the model structure remains manageable (Christophersen *et al.*, 1982; Cosby *et al.*, 1985a,b). This means that the soils/soil-waters are assumed to be of constant and homogeneous composition at a given time or over a given time span. For catchment-based hydrochemical models, it is essential to establish how the soil variability influences the stream chemistry. This is required because water draining from the soil provides a chemically and hydrologically important component of storm runoff. To tackle this issue, the effects of heterogeneity within the catchment need to be examined. Synthetically generated data are used to examine how heterogeneity affects the properties of averaged quantities. This is exemplified by the detailed and extensive hydrochemical work at the Birkenes catchment in southern Norway (Christophersen *et al.*, 1982, 1984, 1990a,b). This Norway spruce catchment on acidic and acid-sensitive soils has experienced acidic deposition typical of many parts of southern Scandinavia. The variability of the cation exchange fraction in localized areas of the soil (Mulder *et al.*, 1991) is incorporated into the analysis to provide a field-based measure of the actual heterogeneity. The Birkenes soil data consists of measurements of the organic horizons of a matrix of podzols, peat and shallow organic soils collected from a 200 by 200 m grid. These organic horizons are common to many acidified areas: the results are therefore likely to be representative of many acidic and acid sensitive regions of environmental concern. Only the principal cations which make up the soil water and soil exchangeable fractions are considered here; i.e. Na^+, H^+, Mg^{2+}, Ca^{2+} and Al^{3+}.

For a soil region in which simple cation exchange equilibrium processes operate, then for any cation pair -

$$mAds(N) \;+\; nM^{m+} \;=\; nAds(M) \;+\; mN^{n+} \qquad \text{eqn1.}$$

M and N refer to two cations M and N with respective charge m and n. "Ads()" refers to an adsorbed phase.

The cation concentration relationships are described by the equation -

$$\underline{K}_{M,N} \;=\; (Ads^n(M)/Ads^m(N))*(\{N^{n+}\}^m/\{M^{m+}\}^n) \qquad \text{eqn2.}$$

where $K_{M,N}$ is an equilibrium constant, {} refers to ion concentrations (in equivalents units), and "Ads()" is dimensionless.

To distinguish between the variations in solution and in the soil cation exchange stores, eqn 2 is recast by introducing two terms, the <u>I</u>on <u>A</u>ctivity <u>R</u>atio (IAR) and the <u>A</u>dsorbed <u>C</u>ation <u>R</u>atio (ACR). These terms are defined by

$$IAR_{N,M} \;=\; \{M^{m+}\}^n/\{N^{n+}\}^m \qquad \text{eqn3.}$$

$$ACR_{N,M} \;=\; Ads^n\{M\}/Ads^m\{N\} \qquad \text{eqn4.}$$

and

$$\mathrm{IAR}_{M,N} = \underline{K}_{M,N} * \mathrm{ACR}_{M,N} \qquad \text{eqn5.}$$

To examine how the variations in micropore water chemistry (areas in equilibrium with the soil cation exchange sites) relate to the average chemistry of the soil water being transported to the stream (here termed average or macropore water), there is a need to (1) assess how within the catchment variability in ACR maps to the corresponding variation in IAR, and (2) examine how cation concentrations behave within the averaged soil water. Here, only one case is considered: K is assumed to be a constant, average, value throughout the soil, and anionic concentrations are constant within the soil micropores. Thus, at each localized micropore area of the soil, IAR varies only in response to the localized variability in ACR (Eqn 5). Average K is determined from the average soil water chemistry determined in the field and from average exchangeable cation data. Inevitably there will be a bias owing to the heterogeneous nature of the situation being represented. None the less, the bias is unimportant for the theme being developed in this presentation. The analysis is described in detail by Neal and Robson (1994a,b), but, in brief,

1) Average K is estimated for Na-H, Na-(Ca+Mg) and Na-Al exchange.

2) Localized ACR values are approximated using the exchangeable cation data from the individual sites and the corresponding local IAR values are found using eqn 5.

3) The average catchment anion concentration is varied (as for example during the course of a storm event): all points in the catchment are assumed to have the same anion concentration at any one time and each contributes equally to the stream: ACR and IAR values at each point in the catchment are taken as constant. Cation concentrations are estimated for selected levels of total anion concentrations ($\{A^{\Sigma-}\}$) in the soil micropores using the IAR values. These concentrations are determined using equations 2 to 5 and the charge balance equation.

Results of the modelling exercise

Variation in field ACR values introduces a large spatial variation in local soil-micropore water chemistry even when the anion concentration is fixed and uniform across the whole catchment. The results show large scatter at each anion concentration and many of the patterns, if not random, are near random (Figures 16 and 17). For the monovalent cations, different features are seen at low and high $\{A^{\Sigma-}\}$. When $\{A^{\Sigma-}\}$ is low, there is a negative and linear relationship between sodium and hydrogen-ion: precisely the opposite of the theoretical relationship for the stream and a consequence entirely of these cations making up the main proportion of the positive charge at low $\{A^{\Sigma-}\}$. In contrast, when $\{A^{\Sigma-}\}$ is very high, there will be no statistically significant relationship between sodium and hydrogen ion. This is because at high ionic strength, divalent and trivalent cations dominate the positive charge. As $\{A^{\Sigma-}\}$ increases, the minimum, average and maximum cation concentrations also increase. The relationships between averaged cation and anion concentrations modelled compare reasonably well with what would be expected from a homogeneous cation exchange system which is representative of average conditions. In other words, the curvature of the average lines is near to that for the homogeneous system. For the averaged (or macropore) data to match the homogeneous cation exchange case exactly, then plots of $\{Na^+\}$ against $\{H^+\}$, $\{Na^+\}^2$ against $\{M^{2+}\}$ and $\{Na^+\}^3$ against $\{Al^{3+}\}$, should give straight lines which pass through the origin (as ACR is held constant, then IAR is also constant). To a first approximation, this appears to be the case (Table 4: the intercepts are statistically insignificant from zero at the 95% level). These results indicate that the properties of the averaged data are such that they can be represented by homogeneous cation exchange - provided the equilibrium constants for cation exchange do not vary considerably within the soil (Neal and Robson, 1994a,b). Thus, for the homogeneous reaction given by eqn 2, the analogous equation for the heterogeneous case is

$$\underline{\underline{K}}_{M,N} = (Ads^n(M_{Average})/Ads^m(N_{Average}))*([N_{Average}{}^{n+}]^m/[M_{Average}{}^{m+}]^n)$$

in other words, for the gross lumping case presented here average(F(x)) approximately equals F(average(x)).

Implications of the modelling exercise

The analysis of the Birkenes soil data indicates that, with the high degree of variation in ACR, large variations in soil water chemistry occur even in localized regions of the soil. For the averaged data, the relationships between ions may well relate in form to equations for homogeneous cation exchange. The power relationships between cation concentrations in averaged waters generally lie between the ratio of the cation charges and unity. Whatever the type of heterogeneity, the effect of gross averaging is to shift the relationships nearer towards linearity (Neal, 1992; Neal and Robson, 1994a,b). The size of the deviation and the degree of data scatter for averaged data is likely to be affected by the precise distribution of the heterogeneous features and by the effective degree of averaging. In practice, further averaging effects are likely at the catchment scale, so deviations from theory may be still lower.

With regard to understanding field data, the samples collected using lysimeters will have chemical compositions dependent upon the extent of the chemical variability within the soil micropores and the volumetric contribution of each water type. Lysimeters may thus collect water from micropores which all have very similar water types, or may collect a mixture of radically different micropore waters. Thus, the lysimeter sample could have properties similar to micropore waters, or properties similar to averaged water types, or something in between. This type of unpredictable behaviour is what is actually observed in the field (Neal, 1992; Taugbϕl and Neal, 1994; Avila *et al.*, 1994). Unfortunately, such patterns are seldom published owing to the reluctance of many authors and reviewers alike to publish "messy" data (Taugbϕl and Neal, 1994). With such field data, the relationships between cations in the lysimeters can vary from straight lines, to curves, to random patterns. By acknowledging the heterogeneity of the system, it becomes obvious that such data may still be consistent with a micropore cation exchange system - soil lysimeter data cannot be used as a test for cation exchange theory.

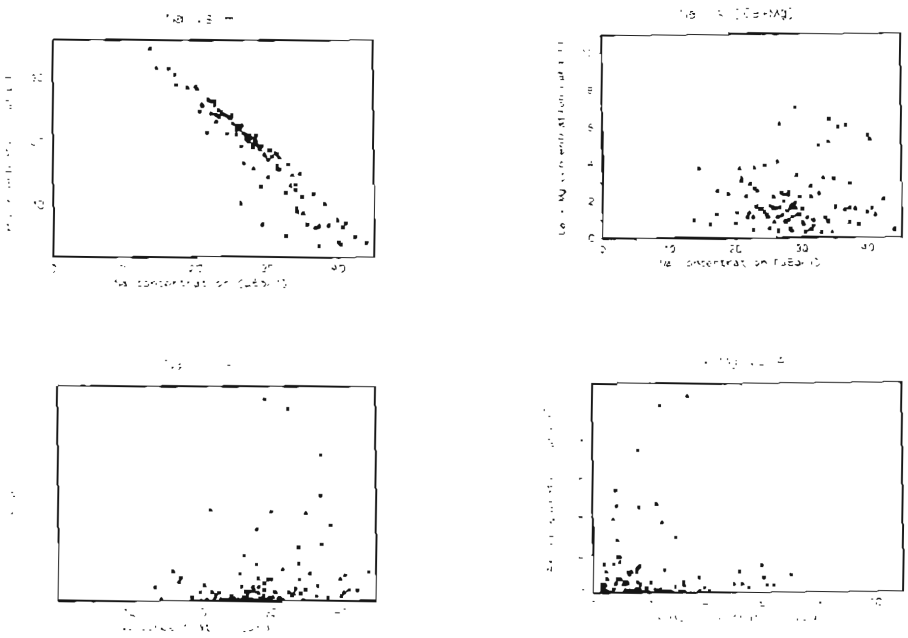


Figure 16. Cation concentration relationships for micropore waters with anion concentrations set to 50 µEq/l: data from Neal and Robson 1994b.

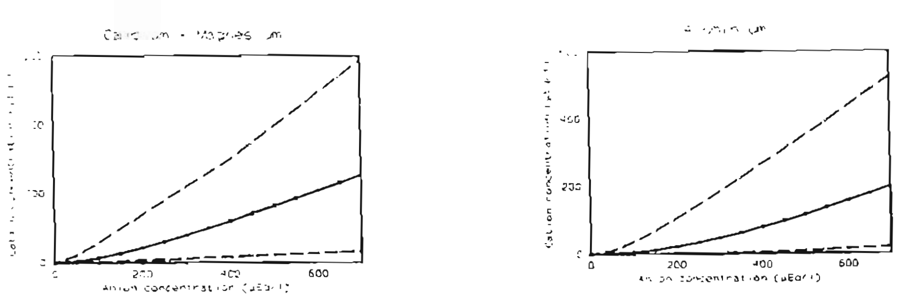


Figure 17. Micropore cation concentration variations with changing anion concentration: minimum, average and maximum changes: data from Neal and Robson 1994b.

Table 4. Comparison of theoretical and modelled power relationships for the Birkenes soil water study. Average cation concentrations at each anion concentration level are logged and regressed against one another. The regressions assess the appropriateness of the underlying theoretical cation exchange reactions to the averaged data. For the lumping theory to hold, p_{calc}*, the gradient of the linear regression, should correspond to* p_{theory}*, the theoretical power for homogeneous cation exchange. For each regression, the left cation in column one is the independent variable and the right cation is the dependent variable.*

	p_{calc}	p_{theory}	r^2
Na-H	1.03 ± 0.00	1	0.999
Na^2-Ca	2.01 ± 0.01	2	0.999
Na^3-Al	2.58 ± 0.01	3	0.999
Ca^3/Al^2	1.28 ± 0.01	1.5	0.999

The hydrochemical models presently in use, address either the short term episodic variability of stream water chemistry or the long term year-to-year trends, concerned primarily with the effects of changing cation exchange stores in the soil and in the prediction of average stream water chemistry. The results of this study support the use of averaging techniques for the short term models, providing the underlying cation exchange formulation is correct and the heterogeneity considered here is not excessively large. For the long term there is more difficultly. Localized differences in hydrological flux, chemical input and weathering rates will influence the changes through time of the adsorbed cation stores. Quantifying such changes is nearly impossible even though there may be systematic elements to these changes. If such changes result in ACR distributions which are strongly skewed, there could be marked deviations away from idealized behaviour. For example, Hooper and Christophersen (1992) showed clearly that a small surface horizon was critical in the evolution of stream water quality in response to acidic deposition.

Field testing of the lumped equations

The theory described above is based on the assumption, used almost universally in modelling of acidic catchments, that simple cation exchange reactions provide an appropriate description of the reactions occurring in soil micropores. However, is this assumption valid? And what, really, are the underlying mechanisms? There is little evidence for suggesting that the theoretical power relationships between cations in the soil solution hold and that soil and stream waters are in equilibrium with solid aluminium hydroxide, not just at Plynlimon as mentioned above, but for many catchments (Neal *et al.*, 1990a). Also, there is an unproven assumption that the soil system can be described simply by inorganic reactions: the soil cation exchanger is dominated by organic exchange sites (which are not of fixed charge) and disassociation of organic acids in the soil solution affects the soil water acidity (Krug and Frink, 1983; Hendershot *et al.*, 1991).

To test the validity of the cation exchange equilibria as applied on a catchment scale, standard techniques (eg collecting waters from soil lysimeters) are inappropriate. Rather, measurements need to be made at a more appropriate and better focused level. This requires catchment based experiments involving field sites where drainage is primarily from the soil zone. Consider such a case for a small drainage are in mid-Wales.

Since 1988, weekly sampling of a first order ephemeral stream draining a spruce forested hill-slope was undertaken at Plynlimon (Neal *et al.*, 1994). This stream, south-2-Hore, drains thin stagnopodzol acid moorland soils (< 1 metre) overlying slates and shales. The area, afforested in the 1930's with sitka spruce (*Picea Sitchensis*) was clearfelled during 3 months in the autumn of 1989.

The south-2-Hore streamlet run-off is characterised by acidic and aluminium-bearing waters similar to those described above for Birkenes and,except for baseflow, the main Plynlimon forest streams (Neal *et al.*, 1994). For Na^+, Ca^{2+}, Mg^{2+}, H^+, total-aluminium and trivalent aluminium there is a variable inter-correlation (Table 5). The major cations (Na^+, Ca^{2+} and Mg^{2+}) are highly correlated and these ions are also highly correlated with the total anion concentrations ($r^2 > 0.7$, $N = 205$). These patterns are essentially the same for the periods prior to felling, post felling and during felling. There are long term patterns of change. For example, during the winter of 1989 concentrations of chloride and cations were at their maximum. The changes observed do not relate directly to varying yearly inputs of the sea salts as, for example, there is no relationship between chloride concentrations in rainfall and chloride concentrations in stream water ($r^2 = 0.01$, $N = 195$). Rather, they reflect hydrological changes and possibly a reduction in the scavenging of sea salts with tree harvesting. In contrast, there is a poor correlation between the major cations with H^+ and total-aluminium (Table 5). It seems that H^+ variations are related to the deforestation phase which seems to lead to increased acidification. The relationship between total aluminium and the other cations is unclear: it may be related to deforestation and nitrate production but data scatter is high and the relationship is weak.

The results, even for the major cations, do not conform with cation exchange theory when examined in detail: the power relationships observed between cations of different charge are well away from those given by theory for homogeneous cation exchange (Table 5). In the case of H^+ and Al^{3+} no simple relationship occurs between each other or the major cations and the data is highly scattered and, as with the major cations, the power relationships differ substantially from those given by theory (Table 5). While there is a broad relationship between the cation concentrations and the sum of the anions, there is no proper compatibility with the cation exchange formulations used in most acidification models. The Plynlimon results are hardly exceptional because nearly identical results are found for an acidified moorland site in southern Norway (Storgama, Table 5) and for a Holm oak catchment in north-eastern Spain (Avila *et al.*, 1994). For the major cations, the power relationships describing their co variations lie between the theoretical values and unity at both sites. This is compatible with the theory described above for spatially and temporally heterogeneous cation exchange. However, the decoupling of H^+ and Al^{3+} from the major cations, implies that these components are not determined primarily by simple cation exchange reactions in the soil. Indeed, previous work at Plynlimon indicates that aluminium release to drainage water can be associated with kinetic "weathering" type reactions involving proton consumption (Reynolds and Hughes, 1986; Reynolds *et al.*, 1989). These findings together with the evidence from long term trends of a decoupling of

Table 5. Linear regression data for logged concentration data for the South-2-hore and Storgama streams. The gradients correspond to the power relationships between ions and the bracketed terms are twice the standard error: n = 205.

South-2-Hore

	r^2	**Gradient (regression)**	**Gradient (theory)**
$Na^+ - Ca^{2+}$	0.422	0.73 ± 0.12	1/2
$Na^+ - Mg^{2+}$	0.828	0.78 ± 0.05	1/2
$Na^+ - Al^{3+}$	0.123	0.54 ± 0.20	1/3
$Na^+ - H^+$	0.080	0.24 ± 0.11	1/1
$Ca^{2+} - Mg^{2+}$	0.641	0.61 ± 0.06	1/1
$Ca^{2+} - Al^{3+}$	0.007	0.12 ± 0.20	2/3
$Ca^{2+} - H^+$	0.018	0.10 ± 0.11	2/1
$H^+ - Al^{3+}$	0.271	0.94 ± 0.21	1/3

Storgama

	r^2	**Gradient (regression)**	**Gradient (theory)**
$Na^+ - Ca^{2+}$	0.671	0.78 ± 0.04	1/2
$Na^+ - Mg^{2+}$	0.739	0.86 ± 0.04	1/2
$Na^+ - Al^{3+}$	0.200	0.89 ± 0.14	1/3
$Na^+ - H^+$	0.039	0.20 ± 0.04	1/1
$Ca^{2+} - Mg^{2+}$	0.753	0.91 ± 0.02	1/1
$Ca^{2+} - Al^{3+}$	0.280	1.10 ± 0.14	2/3
$Ca^{2+} - H^+$	0.013	0.12 ± 0.08	2/1
$H^+ - Al^{3+}$	0.466	1.33 ± 0.11	1/3

the base cations from hydrogen ion and aluminium concentrations (Dillon *et al.*, 1987; Driscoll *et al.*, 1989; Christophersen *et al.*, 1990a), cast doubt upon the value of the commonly used cation exchange equations. Thus the mobile anion theory which is almost universally evoked in acidification models, seems irrelevant.

Summary findings

The above studies illustrate the difficulties of dealing with heterogeneous chemical systems. For the soils case, heterogeneity introduces problems both in the collection of appropriate data to determine the underlying chemical processes and in the production of lumped models. Despite this, the use of synthetic data has explained why field data may not match the underlying theory and helped to formulate the experimentation to test both the theory and the applicability of the lumped models. The second case provides a very disturbing example of incorrect conclusions resulting from averaging procedures and of how the acquisition of detailed experimental information (in this case thermodynamic details of the inorganic carbon system) is of no use for the modelling purpose.

DISCUSSIONS

In terms of the objectives of this presentation on modelling the main message of this paper is that, while systems may be modelled in relatively simple ways by approaches and concepts which seem eminently sensible at the time, wrong conclusions may well result.

The problem for both hydrologic and chemical modelling is our inability to measure the key properties at the same scale that is needed to model the catchment. For example, the chemical models include non-linear equations, such as chemical equilibria, that have been derived for homogeneous environments, not for lumped averages over heterogeneous conditions (as discussed above and elsewhere (Neal, 1992; Neal and Robson, 1994a,b; Rastetter *et al.*, 1992) but needs to be more fully explored in quantitative terms. For hydrology, a description of streamflow generation, starting from first principles, is not available because the simplifications generally used for describing flow through porous media cannot be applied to the shallow soils in upland catchments of most interest in acidification studies. For example, these soils are anisotropic and contain macropores as important conduits. Such complications are solvable in principle if the properties of the porous media could be measured. But, the lack of definitive in situ measurements of soil properties, such as unsaturated conductivity, at the appropriate scale has so far made the problems insurmountable. There are long-standing and ongoing fundamental questions in hydrology as to how water is moving in a catchment (Dooge, 1986; Beven, 1989). For the chemical aspects there is no adequate thermodynamic theory that partitions the chemical transfers between the mobile and immobile soil water environments.

The types of averaged properties identified here have wide and profound ramifications. This work demonstrates that averaging procedures may distort the picture of possible environmental impacts; simple synthetic data exercises can highlight the nature and direction of potential distortion. For example, terms such as the classical mass action equations which apply in a homogeneous situation, no longer apply in a heterogeneous world: only the mass and charge balance terms remain strictly applicable. The work on testing the cation exchange models at a representative level highlights a missing or incorrectly posed process fundamental to acidification research. In some ways, the methods used here resemble those associated with a very active and developing area of scientific research called complexity theory. Here one is concerned with how properties at the local scale determine global characteristics. Recent research has highlighted many areas of the environmental sciences where simple patterns occur at the macroscopic scale, but where immense complexity is observed at the microscopic level even though the determining mechanisms may well be described by simple rules (Cohen and Stewart, 1994). However, the simple emergent rules identified, such as the straight line relationships shown in this paper, must be determined from field observations: theory simply explains the broad reasons why a particular phenomenon occurs. Thus, given a mismatch between field observation and modelling work, the model must be rejected or changed or challenged. This is the opposite path to the way the acidification research has developed: a key issue for environmental impact assessment must be to decide when a model can or cannot be used with confidence.

There is no easy route to a next generation of catchment models based more on field evidence. More data, *per se*, does not solve the problem. Although future advances in instrumentation may provide measurements at a larger scale (such as the effective conductivity of a hillslope), there is much that can be done today on the modelling side to bridge the difference in scale between models and observation. Here two approaches are suggested that should be useful in this respect and this is followed by considering modelling of environmental systems in a larger setting.

In the first approach, an *ad hoc* technique is presented to make catchment modelling more measurement based. This approach, however, is specific to catchment research. The second approach comprises synthetic data analyses that can be used to test how the variables and concepts in a model can be related to measurable quantities and to explore how measurements, taken at one scale, may be related to the physics and chemistry on other scales. Although discussed in the context of catchment acidification, this approach applies generally to models of environmental systems.

The *ad hoc* approach correlates observed temporal variations in streamwater chemistry with the chemistry of observed soil water solutions. Soil waters often have a distinct chemical fingerprint because they arise from horizons that have chemically different solid phases due to natural soil formation. At some catchments, streamwater can be described as a dynamic mixture of these soil solutions, which are called end-members (Christophersen *et al.*, 1990b; Hooper *et al.*, 1990; Neal *et al.*, 1990b,d).

By viewing soil solutions relative to streamwater, a measure for their variation can be established. Suppose the end-member chemical variations in time and space are small relative to the temporal variations in streamwater, and that the end-members are well separated in a chemical sense. Then, an end-member and the soil environment from which it arises, can be considered homogeneous enough from the point of view of explaining the origins of streamwater. The end-member units may then form the basis for a hydrological model as the relative contributions from the end-members to the stream can be computed and used to infer the temporal changes in hydrological routing through the catchment (Neal *et al.*, 1992a; Hooper and Christophersen, 1992). The hydrological structure is constrained by measured soil solution chemical properties, instead of imposing a priori a "conceptual" lumped hydrological model on the catchment. Information on the hydrology is therefore obtained implicitly by using the chemical signals. An analogy from hydrology is the use of stochastic analysis for river basin planning (Loucks *et al.*, 1981). In both cases, a process-based hydrological model was unnecessary to solve the problem at hand.

In the end-member approach, nothing is stated about how the end-member chemistry arises or how it will evolve through time. However, because these soil environments are defined through measurements, their actual physical locations are known and they can be examined to determine the within-end-member variability, and hence the appropriateness of applying an equilibrium chemical model. The end-member approach has only been applied to catchments with areas on the order of 1 square kilometre. Its applicability to larger catchments is not known, although it is an active area of research. Nonetheless, this approach provides an example of a measurement-based method that gives information on the internal catchment hydrochemistry at the appropriate scale.

The second approach uses synthetic data to establish more scientifically based models. In hydrology, for example, the approach could involve generation of an "artificial" catchment as defined by the model under investigation for a given set of parameters. From the model, one could then generate "observations" comprising, for example, runoff and water table elevations. Binley and Beven (1989) provide an example of this approach for a headwater hillslope. The necessary, but not sufficient, condition from the scientific viewpoint, is that the model is well enough posed so that all the parameters to be calibrated can be determined unambiguously from the "observations" (properly error corrupted) given another set of initial calibration parameter values. The synthetic data exercise determines the type and density of measurements required. The practicality of implementing the required field program can be assessed, or, conversely, the utility of the information gained from a certain level of effort in the field can be assessed. After passing this test, the hydrological model is applied to the field. One is then in a much better position to perform a more stringent test of the relevance of the model structure and its parameters to the real environment.

A synthetic data exercise for the hydrological part of the Birkenes model along these lines was carried out by Hooper *et al.* (1988). Here, the model could not pass the first test since some of the parameters, even in this simple model, turned out to be ambiguous given only the runoff and chemically conservative tracer signals. Synthetic data have been used in hydrology elsewhere (Ibbitt and O'Donnell, 1971; Gupta and Sorooshian, 1985; Wheater *et al.*, 1986). Our proposal is simply that environmental models routinely be subjected to such tests before they are applied and that efforts be directed towards development of the necessary software tools including parameter identification methods.

On the chemical side, synthetic data have proved useful in assessing the methods for determining the presence or absence of chemical equilibria in natural waters as shown above (Neal, 1988 a,b; Christophersen and Neal, 1989). In addition, such procedures are beginning to be used to assess the reliability of a lumped representation by investigating the properties of measurements representing averages over a non-homogenous environment. One example is the chemical properties of soil water collected in lysimeters (Neal, 1992; Neal and Robson, 1994a,b; Taugbøl and Neal, 1993). These devices (typically on the decimeter scale) are made of a porous medium that allows sampling of the soil solution by applying suction. If the soil is not homogenous on the decimeter scale, the solution sampled will be drawn from a heterogeneous environment. By assuming it is a mixture of water from homogenous micro-environments, each obeying equilibrium chemistry, one can investigate the properties of the average using the computer. In such situations, the results so far do provide some basis for the lumping procedure used to date whereby average soil properties are related to average soil water chemistry through the same type of equilibrium equations used in the homogenous case (Neal, 1992; Neal and Robson, 1994a,b; Taugbøl and Neal, 1993). The results of Kirchner (1992) and Kirchner *et al.*, (1992), using real data, point in a similar direction. These authors have tested the runoff chemistry from several catchments, and state that chemical equilibrium relationships for homogenous systems could still be applicable at the catchment scale. The use of large scale field studies perturbing catchments, will also be useful in assessing these questions. At present, however, more work is needed before one can say from a process based viewpoint how lumped chemical properties should be related at the catchment scale.

The difficulties discussed here are not limited to hydrochemistry. Peters (1990) gives an account of related problems in limnology. In that area, long term field studies and modelling efforts aimed at quantifying the processes behind lake eutrophication have been carried out. The development of process based models is problematic for many of the same reasons as discussed here. However, from the management viewpoint, one has in limnology reasonably reliable methods for predicting lake response to nutrient (i.e. phosphate) inputs. These techniques, developed by Vollenweider and others, are empirically based (cf. OECD, 1982). Given enough observations, empirical models will often be sufficient from a management viewpoint - after all, the methods used in managing the complex systems encountered in agriculture and forestry are empirical in nature.

Within acidification research, a related empirical approach is due to Henriksen (1980). However, in the acidification case, the important processes take place on larger spatial and temporal scales than in limnology. If one has to wait some years to see the effects of a change in phosphate inputs to a lake, the response time of a stream, regarding a change in sulphate deposition, is rather measured in decades. This implies that considerably more time and effort are needed to arrive at useful empirical relationships. Considering global climatic change, these concerns are again magnified - see Hauhs (1990) for a discussion of empirical versus process based approaches to the problem of ecosystem management.

Given this backdrop, one might well ask to what degree process based models will be able to quantify the behaviour of lakes, catchments and other environmental systems at the scale of interest. There could be fundamental limitations that cannot be overcome, given the measurements one can perform - i.e. How modellable is the environment? Probing such questions is vital to the future of environmental modelling.

RECOMMENDATIONS

From this presentation and associated work (Christophersen *et al.*, 1993), the following recommendations are advanced:

a) When fitting a model to observations, the limited information content of the calibration signals must be recognized. For hydrology, this implies that modelling advances can only be made through the incorporation of chemical signals, groundwater elevations or other "internal" state variables in addition to runoff. Thus fitting a model to the output signal of a complex environmental system (e.g., runoff measurements for catchments) is a necessary, but weak, condition and does not give scientific credibility to the mechanisms contained within the model. Such a fit only represents a starting point in model development.

b) The internal structure of current hydrochemical models has not been tested at the catchment scale. Recognition of this may serve to focus research efforts to advance the understanding of how ecosystems function. Extensions of current models without addressing their internal structure (for example, adding nitrate cycling to existing acidification models) should not be construed as scientific advances.

c) To progress within catchment science, studies of intensively and extensively monitored sites are needed where the internal structure of the catchment is examined.

d) To progress within environmental modelling, fuller use of synthetic data must be made both for tying models to observations and for exploring the observations themselves. The computer simulations will offer important learning tools in their own right.

e) Improved modelling of environmental systems is not a technological problem in that more data will necessarily give us better models. Rather there could be fundamental limitations to the modelling of these systems that are critical to unravel.

f) The fundamental equations depicting the acidification of catchments needs substantial re-examination.

g) And finally, environmental impacts must be related to biological response. And yet none of the present models address such an issue directly. Indeed, as with many biological processes, environmental impacts can only be addressed in a descriptive sense. Biological systems change in response to environmental pressures often in an obscure way. For example, at Plynlimon, deforestation has led to a deterioration in water quality characteristic of acidification, and yet acid sensitive may-flies have returned to the stream. Perhaps the greatest challenge within the environmental modelling studies is to produce an environmental impact model of direct relevance to the biology.

REFERENCES

Adamson, J.K. and Hornung, M. 1990. The effect of clearfelling a sitka spruce (*Picea sitchensis*) plantation on solute concentrations in drainage waters. Journ. Hydrol., 116, 287-298.

Avila, A., Bonilla, D., Rodà, F., Pinõl, J. and Neal, C. in press. Variations in soil solution chemistry for a holm oak (*Quercus ilex*) forest: inferences on chemical processes for a montane Mediterranean area. J. Hydrol.

Beven, K.J. 1989. Changing ideas in hydrology: the case of physically based models. J. Hydrol., 105, 157-172.

Beven, K.J. 1991. Hydrograph separation? BHS Third National Symposium, Southampton University, September 1991. Institute of Hydrology., 3.1-3.8.

Beven, K.J. and Kirkby, M.J. 1979. A physically based variable contributing area model of basin hydrology. Hydrol. Sci. Bull., 24 (1), 43-69

Beven, K.J., Kirkby, M.J., Schoffield, N. and Tagg, A. 1984. Testing a physically based flood forecasting model TOPMODEL for three UK catchments. J. Hydrol., 69, 119-143.

Binley, A. and Beven, K.J. 1989. Modelling heterogenous Darcian headwaters, in British Hydrological Society, 2nd National Symposium, Univ. of Sheffield, 4-6 Sept. Published by Inst. of Hydrol., Wallingford, Oxon, U.K., pp. 1.17 - 1.22.

Birchall, J.D., Exley, C., Chappell, J.S. and Phillips, M.J. 1989. Acute toxicity of aluminium to fish eliminated in silicon-rich waters. Nature, 338, 146-148.

Boyle, E., Collier, R., Dengler, A.T., Edmond, J.M., Ng, A.C. and Stallard, R.F. 1974. On the chemical mass-balance in estuaries. Geochimica et Cosmochimica Acta, 28, 1719-1728.

Chappell, J.S. and Birchall, J.D. 1988. Aspects of the interaction of silicic acid with aluminium in dilute solution and its biological significance. Inorg. Chim. Acta, 153, 1-4.

Christophersen, N. and Neal, C. 1989. A rational approach to the assessment of aluminium solubility controls in freshwaters. Sci. Tot. Environ., 84, 91-100.

Christophersen, N. and Neal, C. 1990. Linking hydrological, geochemical, and soil chemical processes on the catchment scale: an interplay between modelling and field work. Water Resources Research, 26, 3077-3086.

Christophersen, N., Seip, H.M. and Wright, R.F. 1982. A model for streamwater chemistry at Birkenes, Norway. Water Resour. Res., 18, 977-996.

Christophersen, N., Rustad, S. and Seip, H.M. 1984. Modelling streamwater chemistry with snowmelt. Philos. Trans. Royal, Soc., London, ser. B, 305, 427-439.

Christophersen, N., Robson, A.J., Neal, C., Whitehead, P.G., Vigerust, B. and Henriksen, A. 1990a. Evidence for long-term deterioration of streamwater chemistry and soil acidification at the Birkenes catchment, southern Norway. Journ. Hydrol., 116, 63-76.

Christophersen, N., Neal, C., Hooper, R.P., Vogt, R.D. and Andersen, S. 1990b. Modelling streamwater chemistry as a mixture of soilwater end-members - A step towards second generation acidification models. J. Hydrol., 116, 307-320.

Christophersen, N., Neal, C. and Hooper, R.P. 1993. Modelling the hydrochemistry of catchments: a challenge for the scientific method. Journ. Hydrol., 152, 1-12.

Cohen, J. and Stewart, I. 1994. The Collapse of Chaos: Discovering simplicity in a complex world. Viking (London, UK), 495pp.

Cosby, B.J., Wright, R.F., Hornberger, G.M. and Galloway, J.N. 1985a. Modelling the effects of acidic deposition: assessment of a lumped-parameter model of soil water and stream water chemistry. Water Resour. Res., 21, 51-63.

Cosby, B.J., Wright, R.F., Hornberger, G.M. and Galloway, J.N. 1985b. Modelling the effects of acid deposition: estimation of long-term water quality responses in a small forested catchment. Water Resources Res., 21, 1591-1601.

Dillon, P.J., Read, R.A. and deGrosbois, E. 1987. The rate of acidification of aquatic ecosystems in Ontario, Canada. Nature, 329, 45-48.

Dimbleby, G.W. 1952. Soil regeneration on the north-east Yorkshire moors. J. Ecol., 40, 331-341.

Dooge, J.C.I. 1986. Looking for hydrologic laws. Water Resour. Res., 22, 46S-58S.

Driscoll, C.T., Likens, G.E., Hedin, L.O., Eaton, J.S. and Bormann, F.H. 1989. Changes in the chemistry of surface waters. Environ., Sci. and Technol., 23, 137-143.

Egglishaw, H.J., Gardener, R. and Foster, J. 1986. Salmon catch decline and forestry in Scotland. Scot. Geol. Mag., 102, 57-61.

Emmett, B. 1989. The effects of harvesting intensity on soil nitrogen transformations in Sitka spruce (*Picea sitchensis* (Bong.) Carr.) plantation at Beddgelert Forest (N. Wales). PhD thesis, University of Exeter, UK.

Gupta, V.J. and Sorooshian, S. 1985. The automatic calibration of conceptual catchment models using derivative-based optimization algorithms. Water Resour. Res., 21, 473-485.

Haines, T.A. 1986. Fish population trends in response to surface water acidification. In: Acid Deposition, Long-term trends. Nat. Acad. Press, Washington, DC, 300-334.

Hauhs, M. 1990. Ecosystem modelling: Science or technology?. J. Hydrol., 116, 25-34.

Hendershot, W.H., Warfvinge, P., Courchesne, F. and Sverdrup, H.U. 1991. The mobile anion concept - time for a reappraisal? Journ. Environ. Qual., 20(3), 505-509.

Henriksen, A. 1980. Acidification of freshwaters - a large scale titration, In Drabløs, D. and Tollan, A. (Eds.), Ecological Impact of Acid Deposition. Norwegian Institute for Water Research, 68-74.

Hewlett, J.D. and Hilbert, A.R. 1967. Factors affecting the response of small watersheds to precipitation in humid areas. In: Forest Hydrology, Sopper, W.E. and Lull, H.W. (eds). Pergamon, Oxford, UK, 275-290.

Hooper, R.P. and Shoemaker, C.A. 1986. A comparison of chemical and isotopic hydrograph separation. Water Resources Res., 22 (10), 1444-1454.

Hooper, R.P., Stone, A., Christophersen, N., deGrosbois, E. and Seip, H.M. 1988. Assessing the Birkenes model of stream acidification using a multisignal calibration methodology. Water Resour. Res., 24, 1308-1316.

Hooper, R.P., Christophersen, N. and Peters, N.E. 1990. Modelling streamwater chemistry as a mixture of soilwater end-members - An application to the Panola Mountain catchment, Georgia, USA. J. Hydrol., 116, 321-343.

Hooper, R.P. and Christophersen, N. 1992. Predicting episodic stream acidification in the Southeastern United States: Combining a long-term acidification model and the end-member mixing concept. Water Resour. Res, 28, 1983-1990.

Hordijk, L. 1991. Use of the RAINS model in acid rain negotiations in Europe. Environ. Sci. and Tech., 25, 596-603.

Hornung, M., Adamson, J.K., Reynolds, B. and Stevens, P.A. 1989. Impacts of forest management practices in plantation forests. Air Pollution Research Report 13: Effects of land use in catchments on the acidity and ecology of natural surface waters, 91-106. Commission of the European Communities, CEC-DG XII/E-1, 200 Rue de la Loi, B-1049, Brussels, Belgium.

Hughes, S., Reynolds, B. and Roberts, J.D. 1990. The influence of land management on concentrations of dissolved organic carbon and its effects on the mobilisation of aluminium and iron in podzol soils in mid-Wales. Soil Use and Management, 6, 137-144.

Ibbit, R.P. and O'Donnell, T. 1971, Fitting methods for conceptual catchment models. J. Hydraul. Div. Am. Soc. Civ. Eng., 97, 1331-1342.

Jenkins, A. and Cosby, B.J. 1989. Modelling surface water acidification using one and two soil layers and simple flow routing. In: Regional acidification models, Kamari, J., Brakke, D.F., Jenkins, A., Norton, S.A. and Wright, R.F. (eds). Springer-Verlag, Berlin, Heidelberg, Chapter 19.

Jenkins, A., Cosby, B.J., Ferrier, R.C., Walker, T.A.B. and Miller, J.D. 1990. Modelling stream acidification in afforested catchments: an assessment of the relative effects of acid deposition and afforestation. Journ. Hydrol., 120, 163-181.

Kennedy, V.C., Kendall, C., Zellweger, G.W., Wyerman, T.A. and Avanzino, R.J. 1986. Determination of the components of stormflow using water chemistry and environmental isotopes, Matolle river basin, California. J. Hydrol., 84, 107-140.

Kirby, C., Newson, M.D. and Gilman, K. 1991. Plynlimon research: the first two decades. Institute of Hydrology Report 109, 1-187.

Kirchner, J.W. 1992. Heterogenous geochemistry of catchment acidification. Geochim. et Cosmochim. Acta, 56, 2311-2328.

Kirchner, J.W., Dillon, P. and LaZerte, B.D. 1992. Catchment geochemical buffering predictions tested with acid episodes. Nature, 358, 478-482.

Klemes, V. 1986. Dilettantism in hydrology: Transition or destiny. Water Resour. Res., 22, 177S-188S.

Krug, E.C. 1991. Review of acid deposition-catchment interaction and comments on future research needs. J. Hydrol., 128, 1-27.

Krug, E.C. and Frink, C.R. 1983. Acid rain on acid soil: a new perspective. Science, 122, 520-525.

Liss, P.S. 1976. Conservative and non-conservative behaviour of dissolved constituents during estuarine mixing. In: Estuarine chemistry, Burton, J.D. and Liss, P.S. (eds). Academic Press, London, 93-130.

Loucks, D.P., Stedinger, J.R. and Haith, D.A. 1981. Water Resource Systems Planning and Analysis. Prentice-Hall, Inc., Englewood Cliffs, NJ, 559 pp.

McMahon, R. and Neal, C. 1990. Aluminium dis-equilibrium solubility controls in Scottish acidic catchments. Hydrol. Sci. Bull., 35 (1), 21-28.

Miller, H.G. 1985. The possible role of forests in stream acidification. Soil use and management, 1, 28-35.

Mulder, J., van Breemen, N. and Eijck, H.C. 1989. Depletion of soil aluminium by acid deposition and implications for acid deposition. Nature, 337, 247-249.

Mulder, J., Pijpers M. and Christophersen, N. 1991. Water flow paths and the spatial distribution of soils and exchangeable cations in an acid rain-impacted and a pristine catchment in Norway. Water Resour. Res., 27, 2919-2928.

Muniz, I.P. and Leivestad, H. 1980. Acidification effects on freshwater fish. In: Drablos, D. and Tollan, A., (eds). Ecological impacts of acid precipitation. Oslo-As:SNSF, 84-92.

Neal, C. 1988a. Aluminium solubility in acid waters. Earth Planetary Sci. Lett., 86, 105-112.

Neal, C. 1988b. Aluminium solubility relationships in acid waters; a practical example of the need for a radical reappraisal. J. Hydrol., 104, 141-159.

Neal, C. 1992. Describing anthropogenic impacts on streamwater quality: the problem of integrating soil water chemistry variability. Sci. Tot. Environ., 115, 207-218.

Neal, C. in press. Aluminium speciation variations in an acidic upland stream draining the Hafren spruce forest, Plynlimon, mid-Wales; J. Hydrol.

Neal, C. and Robson, A.J. 1994a. Integrating soil water chemistry variations at the catchment level within a cation exchange model. Sci. Tot. Environ. 144, 93-102.

Neal, C. and Robson, A.J. 1994b. Integrating soil water chemistry variations within a cation exchange model. Sci. Tot. Environ., in press.

Neal, C., Whitehead, P.G., Neale, R. and Cosby, B.J. 1986. Modelling the effects of acidic deposition and conifer afforestation on stream acidity in the British Uplands. J. Hydrol., 86, 15-26.

Neal, C., Christophersen, N., Neale, R., Smith, C.J., Whitehead, P.G. and Reynolds, B. 1988. Chloride in precipitation and streamwater for the upland catchment of River Severn, Mid-Wales; Some consequences for hydrochemical models. Hydrol. Proc., 2, 155-165.

Neal, C., Reynolds, B., Stevens, P. and Hornung, M. 1989. Hydrogeochemical controls for inorganic aluminium in acidic stream and soil waters at two upland catchments in Wales. J. Hydrol., 106, 155-175.

Neal, C., Mulder, J., Christophersen, N., Neal, M., Waters, D., Ferrier, R.C., Harriman, R. and McMahon, R. 1990a. Limitations to the understanding of ion exchange and solubility controls for acidic Welsh, Scottish and Norwegian sites. J. Hydrol., 116, 11-23.

Neal, C., Robson, A.J., and Smith, C.J. 1990b. Acid Neutralization Capacity variations for Hafren Forest streams: inferrences for hydrological processes. J. Hydrol., 121, 85-101.

Neal, C., Neal, M., Waters, D., Mulder, J., Christophersen, N., Ferrier, R.C., McMahon, R and Harriman, R. 1990c. Limitations to the understanding of ion-exchange and solubility controls for acidic Welsh, Scottish and Norwegian sites. Journ. Hydrol., 116, 11-24.

Neal, C., Smith, C.J., Walls, J., Billingham, P., Hill, S. and Neal, M. 1990d. Hydrogeochemical variations in Hafren forest stream waters, mid-Wales. Journ. Hydrol., 116, 185-200.

Neal, C., Robson, A.J., Reynolds, B. and Jenkins, A. 1992a. Prediction of future short-term stream chemistry - a modelling approach. J. Hydrol., 130, 87-103.

Neal, C., Fisher, R., Smith, C.J., Hill, S., Neal, M., Conway, T., Ryland, G.P. and Jeffrey, H.A. 1992b. The effects of tree harvesting on stream water quality at an acidic and acid sensitive spruce forested area: Plynlimon, mid-Wales. J. Hydrol., 135, 305-319.

Neal, C., Reynolds, B., Smith, C.J., Hill, S., Neal, M., Conway, T., Ryland, G.P., Jeffrey, H.A., Robson, A.J. and Fisher, R. 1992c. The impact of conifer harvesting on stream water pH, alkalinity and aluminium concentrations for the British Uplands: an example for an acidic and acid sensitive catchment in mid-Wales. Sci. Total Environ. 126, 75-87.

Neal, C., Neal, M., Ryland, G.P., Jeffery, H.A., Harrow, M., Hill, S. and Smith, C.J. 1994. Chemical variations in near surface drainage water for an acidic spruce forested UK upland area subjected to timber harvesting: inferences on cation exchange processes in the soil. Science of the Total Environment, in press.

Newson, M.D. 1976. The physiography, deposits and vegetation of the Plynlimon catchments. Inst. of Hydrology Report., 30, 1-59.

Nikolaidis, N.P., Rajaram, H., Schnoor, J.L. and Georgakakos, K.L. 1989. A generalized soft water acidification model. Water Resour. Res., 24, 1983-1996.

Nisbet, T.R. 1990. Forests and surface water acidification. Forestry Commission Bull., 86, 1-7.

OECD, 1982. Eutrophication of waters. Monitoring, assessment and control. OECD, Paris, 154pp.

Pearce, A.J., Stewart, M.K. and Sklash, M.G. 1986. Storm runoff in humid headwater catchments 1. Where does the water come from? Water Resources Res., 22 (8), 1263-1272.

Peters, R.H. 1990. Pathologies in limnology, in de Bernardi, R. and Giussani, G., and Barbanti, L. (eds.), Scientific Perspectives in theoretical and applied limnology. Mem. Ist. Ital. Idrobiol., 47, 181-217.

Pennington, W. 1984. Long term natural acidification of upland sites in Cumbria: evidence from post-glacial lake sediments. Rep. Freshwater. Biol. Ass. No. 52, 28-46.

Pilgrim, D.H., Huff, D.D. and Steele, T.D. 1979. The use of specific conductance and contact times relations for separating flow components in storm runoff. Water Resources Res., 15 (2), 329-339.

Pinder, G.F. and Jones, J.F. 1969. Determination of the groundwater component of peak discharge from the chemistry of total runoff. Water Resources Res., 5, 438-445.

Rastetter, E.B., King, A.W., Cosby, B.J., Hornberger, G.M., O'Neill R.V. and Hobbie, J.E. 1992. Aggregating fine-scale ecological knowledge to model coarser-scale attributes of ecosystems. Ecol. Appl., 2, 55-70.

Reuss, J.O. and Johnson, D.W. 1986. Acid Deposition and the acidification of soils and waters. Ecological studies 59. Springer Verlag. 1-119.

Reuss, J.O., Christophersen, N. and Seip, H.M. 1986. A critique of models for freshwater and soil acidification. Water Air Soil Poll., 30, 909-931.

Reynolds, B. and Hughes, S. 1989. An ephemeral forest drainage ditch as a source of aluminium to surface waters. Sci. Tot. Environ., 80, 185-193.

Reynolds, B., Hornung, M. and Hughes, S. 1989. Chemistry of streams draining grassland and forest catchments at Plynlimon, mid-Wales. Hydrol. Sci. J., 34 (6), 667-686.

Reynolds, B., Neal, C., Hornung, M. and Stevens, P.A. 1986. Baseflow buffering of stream water acidity in five mid-Wales catchments. Journ. Hydrol., 87, 167-185.

Reynolds, B., Neal, C., Hornung, M. and Stevens, P.A. 1988. Impact of afforestation on the soil solution chemistry of stagnopodzols in mid-Wales. Water, Air and Soil Pollution, 38, 55-70.

Robson, A.J. 1993. The use of continuous measurement in understanding and modelling the hydrochemistry of the uplands. PhD dissertation, University of Lancaster, Lancaster, UK, 278pp.

Robson, A.J. and Neal, C. 1991. Chemical signals in an upland catchment in mid-Wales - some implications for water movement. In: BHS conference proceedings, Southhampton University. Institute of Hydrology. 3.17-3.24.

Rohde, A. 1981. Spring flood. Meltwater or groundwater? Nordic Hydrol., 43, 45-65.

Rosenqvist, I.T. 1978. Alternative sources for acidification of river water in Norway. Sci. Tot. Env., 10, 39-49.

Rosenqvist, I.T. 1990. From rain to lake; pathways and chemical change. J. Hydrol., 114, 3-10.

Schecher, W.D. and Driscoll, C.T. 1987. An evaluation of uncertainty associated with aluminium equilibrium calculations. Water Resour. Res., 23, 525-534.

Sherman, L.K. 1932. Stream flow from rainfall by the unit-hydrograph method. Engineering news record, 108, 501-505.

Sklash, M.G. 1990. Environmental isotope studies of storm and snowmelt runoff generation. In: Process studies in Hillslope Hydrology, M.G. Anderson and T.P. Burt (eds). Wiley, Chichester, 401-435.

Sklash, M.G. and Farvolden, R.N. 1979. The role of groundwater in storm runoff. J. Hydrol., 43, 45-65

Sklash, M.G., Stewart, M.K. and Pearce, A.J. 1986. Storm runoff generation in humid headwater catchments. 2 A case study of hillslope and low order stream response. Water Resources Res., 22, 1273-1282.

Stevens, P.A. and Hornung, M. 1987. Nitrate leaching from a felled sitka spruce plantation in Beddgelert Forest, north Wales. Soil Use and Management, 4, 3-8.

Stevens, P.A., Hornung, M. and Hughes, S. 1989. Solute concentrations, fluxes and major nutrient cycles in a mature Sitka spruce plantation in Beddgelert Forest, North Wales. Forest Ecology and Management, 27, 1-20.

Stoner, J.H. and Gee, A.S. 1985. Effects of forestry on water quality and fish in Welsh rivers and lakes. Journ. Inst. Water Eng. Sci., 39, 27-45.

Sklash, M.G. and Farvolden, R.N. 1979. The role of groundwater in storm runoff. J. Hydrol., 43, 45-65.

Sverdrup, H.U. 1990. The kinetics of base cation release due to chemical weathering. Chartwell Bratt Ltd., Bromley, Kent, UK, 246pp.

Taugbøl, G. and Neal, C. 1993. Soil and streamwater chemistry variations on acidic soils. Application of a cation exchange and mixing model at the catchment level. Sci. Tot. Environ. (in press).

Taylor, J.A. 1974. Organic soils in Wales. In: Adamson, W.A., (ed) Soils in Wales. Welsh soils discussion group. Report 15, 30-43.

Turnpenny, A.W.H., Sadler, K., Aston, R.J., Milner, A.G.P. and Lyman, S. 1987. The fish population of some streams in Wales and Northern England in relation to acidity and associated factors. J. Fish. Biol., 31, 415-434.

UKAWRG, 1988. United Kingdom Acid Waters Review Group, second report. Acidity in United Kingdom fresh waters. Her Majesty's Stationary Office, London, 1-61.

Wells, C., Cornett, R.J. and Lazerte, B.D. 1991. Hydrograph separation: A comparision of geochemical and isotopic tracers. J. Hydrol., 122, 253-274.

Wheater, H., Bishop, K.H. and Beck, M.B. 1986. The identification of conceptual hydrologic models for surface water acidification. Hydrol. Proc., 1, 89-108.

Whitehead, P.G., Bird, S., Hornung, M., Cosby, B.J., Neal, C. and Paricos, P. 1988. Stream acidification trends in the Welsh Uplands: a modelling study of the Llyn Brianne catchments. Journ. Hydrol., 101, 191-212.

Chemical Analysis of Rocks and Soils

Magne Ødegård
Laboratories Division,
Geological Survey of Norway

Introduction

Information on chemical composition of geological materials is often of fundamental importance within many branches of modern geoscience, and applies to geomaterials in their widest sense, as solid rocks and ores, sediments, water and air.

Analytical chemistry is therefore an important element within all fields of geology and related activities. The ever increasing demand for geoanalyses is mainly concentrated within three areas, namely:

1. Basic research within geology and geochemistry
2. Exploitation and use of georesources
3. Environmental surveillance

Basic research within many fields of geology and geochemistry has created an extended interest for chemical data, not only as regards element coverage and concentration level, but also as regards type of geological material. The same applies to the vast group designated exploitation and use of georesources, a group which also includes the rapidly developing area of material science. Even if the increasing demand for chemical analysis of geomaterials primarily is concentrated within these two groups, considerable geoanalytical service is also attached to environmental surveillance and control, an activity of vital importance and with high political priority.

Todays extensive use of analytical data on geological materials at all concentration levels had not been possible without an intense research in analytical inorganic chemistry, research which has led to the development of rapid and accurate instrumental analytical methods with good element coverage and detection capabilities.

Historical development

In the early days of the development of mineralogy and petrology as sciences, geological materials like rocks, ores and minerals were exclusively analysed with gravimetric and titrimetric wet chemical methods. These classical methods were largely suitable only for major elements, i.e. concentration levels above 0.1%, and used mainly for the characterization and description of minerals and rocks. If minor and trace elements were to be determined, often complicated and time-consuming separation and/or enrichment procedures had to be used. Many wet chemical methods are very accurate, and are even today important supplements to modern instrumental methods.

The great breakthrough in the field of inorganic chemical analysis came with the development of the spectrochemical analytical methods X-ray fluorescence and optical emission spectrometry. From the time around 1930 and up to now, these techniques have

developed to be indispensable analytical methods with high precision and great capacity.
An essential part of the inorganic chemical analyses being done today on geological materials are done by either of these two methods, or by atomic absorption spectrometry.
Mass spectrometry, which is of particular significance within geology, is also a relatively old analytical technique originally applied on solid samples. The latest development in this field is the so-called plasma mass spectrometer, where an inductively coupled plasma is used as ionization source. This technique is therefore mostly used on solutions, and is a very sensitive and valuable method for many purposes, e.g. environmental analyses at low concentration levels.

Another recognized analytical technique with a fairly long history is neutron activation analysis, but since this technique normally is dependant on an atomic reactor for irradiation, its use is relatively limited.

Determination of inorganic anions were originally done by gravimetric and titrimetric methods, but these methods are largely abandoned. The most used methods for anions today are ion chromatography and ion selective electrodes.

In the field of inorganic chemical analysis there exist today several refined combinations of analytical techniques with excellent detection capabilities, e.g. combination of chromatography and plasma emission spectrometry or plasma mass spectrometry. By this technique elements or groups of elements are chromatographically isolated and enriched before they are determined by either of the two plasma techniques.

Total analysis versus partial analysis

There are in principle two different ways of analysis being used today for geological materials, namely:

1. Analysis of total content.
2. Partial analysis based on some form of extraction.

Analysis of the total element content of a sample material is a well defined task which does not need further comments.

Partial analysis, on the other hand, is a way of analysis which needs several comments. This way of analysis is mainly used in two cases, namely:

1. Environmental studies.
2. Geochemical prospecting for ores and minerals.

Both environmental pollution studies and geochemical prospecting for ores and minerals are activities which have three basic elements in common, namely:

1. A primary source for supply of chemical elements.
2. A mechanism for dispersing the elements.
3. A medium for picking up or collecting the elements.

By both activities the main aim will be to determine quantitatively and selectively the elements or element compounds which secondarily have been supplied to the geological medium from an external source. This selective task can never be totally fulfilled.

There are two factors related to the extraction technique which are of particular importance, namely:

1. The extraction attack must be strong enough to bring the secondarily supplied elements or element compounds to be studied in solution.
2. The acid attack should influence the primary sample minerals as little as possible to avoid introduction of unwanted elements.

Unwanted elements in this context are first of all potential analyte elements primary present in the geological medium, and which at the same time are completely or partly released during the extraction. This can in extreme cases disturb or even over-shadow effects which are to be studied, the danger being especially great when working at low concentration levels and with marginal effects. This is often the situation when studying the effects of moderate or marginal environmental loads.

Unwanted elements are also all other matrix elements which go in solution during extraction. This increases the content of dissolved solids in the analysis solutions, which in most cases complicates the subsequent analysis.

The degree of which unwanted elements of the two categories are introduced from geological materials during acid extraction is dependant on the mineralogical composition of the material. Information on general solubility of rock-forming silicate minerals is very scarce in the literature, the most systematic and comprehensive work being that of Graff et al. (1985). These studies show that a lot of silicate minerals, which traditionally have been considered as insoluble in mineral acids, except HF, have substantial solubility in mineral acids as HCl and HNO_3, in some cases up to 100%. Studies by Faye (1982) show that many common accessory minerals of rocks, and which also have been regarded as very resistant against acid attacks, have considerable solubility.

The most important aspects of these studies, especially in relation to environmental studies and geochemical prospecting, will be the possibility of releasing analyte elements bound in the lattice of the primary silicate minerals. Evaluation of such contributions requires:

1. Knowledge of the composition of the rock material.
2. Knowledge of the basic geochemical distribution laws.
3. Knowledge of the general solubility of rock-forming silicate minerals.

It is obvious that a total analysis and a partial analysis in most cases give entirely different information, and the choice of method must therefore be carefully based on the purpose of the analysis. This choice influences only on the sample preparation stage, and does not in principle influence the subsequent instrumental analysis.

Analytical methods

The analytical methods which will be treated in this paper are confined to the methods which are in use at the Geological Survey of Norway (NGU).

Some basic features are common to all spectrochemical analytical techniques. All techniques are based on measurement of electromagnetic radiation, either emitted or absorbed, and all techniques are based on a calibration system where a relationship between radiation intensity and concentration is established. The techniques are further dependent on an excitation source, a device for isolating the used wavelength, and a device to detect the radiant energy. These main elements of a spectrochemical analysis method are illustrated in Fig. 1. Emission spectroscopy is chosen as example, and the figure also shows the historic development which has taken place within this technique.

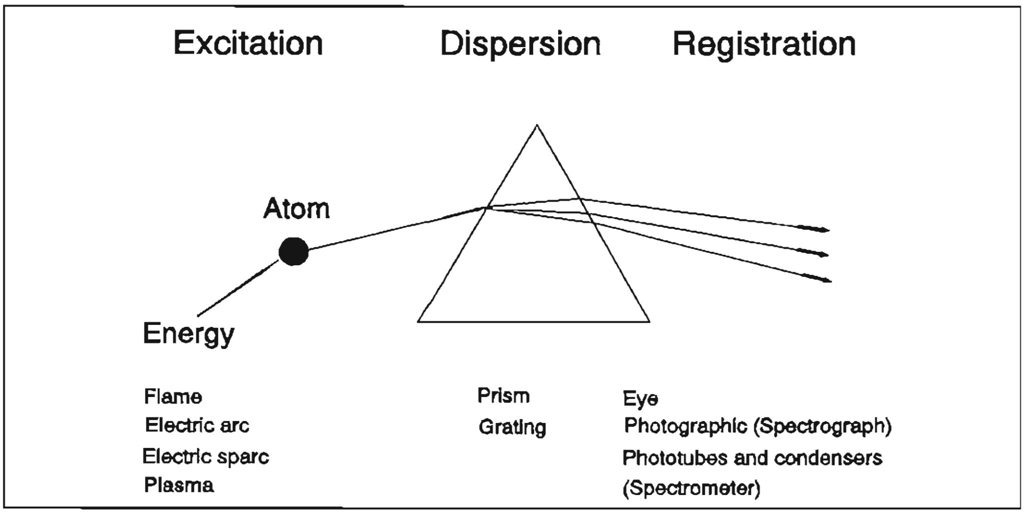


Fig.1. Simplified outline of emission spectroscopy

X-ray fluorescence

X-ray fluorescence (XRF) represents one of the few instrumental analytical methods which normally ~~applies~~ is used on solid samples. The principle of XRF is in short that a representative surface of the sample material is irradiated by X-rays from an X-ray tube, whereby the atoms of the sample are excited and emit fluorescence radiation. This radiation is normally isolated according to wavelength by means of an analysing crystal and finally recorded by a counter (wavelength dispersive XRF). In some cases the radiation is isolated according to energy (energy dispersive XRF). The basis for both techniques is that the intensity of the fluorescence radiation is a function of concentration.

XRF is one of the most important and accurate analytical methods for geological materials, both as regards major elements and many elements at minor and trace level. This is first of all due to very stable instrumentations with high grade of precision.

There are two major problems related to XRF, which have to be overcome. These are matrix effects and grain size effects. Matrix effects are due mainly to absorption of primary and secondary X-rays by the elements of the sample. Less important are enhancement effects. Grain size effects are physical effects which are related to the size of the sample particles in relation to the penetration depth of the X-rays.

Major elements

A major rock element analysis includes the elements Si, Al, Fe, Ti, Mg, Ca, Na, K, Mn and P, and is today most often done by XRF. The characteristic fluorescence radiation from the light rock elements is relatively long, and therefore sensitive to grain size effects. Most geological materials are crystalline, and even after prolonged grinding grain size effects will normally be considerable. Analysis of major rock elements are therefore mostly based on fusion of the samples by a flux material. This is a sample preparation technique which completely eliminates grain size effects, and a technique which also in other respects are nearly ideal. By fusion of the crystalline rock sample with a flux the

homogeneous material. This isoformation technique also has the great advantage that synthetical standards of high accuracy can be prepared from oxides with known stochiometry, and used in the same way as natural standard reference materials for instrumental calibration. With this nearly ideal sample preparation technique combined with mathematical correction for matrix effects, the XRF method is today without competition in major rock element analysis, both as regards quality and speed.
The sample preparation method in use at NGU applies a sample weight of 0.8g and a flux weight of 5.6g $Li_2B_4O_7$. The material is fused in Au/Pt-crucibles in an automatic fusion machine (Claisse fluxer), and finally cast to beads with flat surface. The X-ray analyses are carried out on a Philips X-ray spectrometer model PW 1480. Table 1 gives an overview of analytical range of the different components together with analytical accuracy.

Table 1. X-ray fluorescence analysis of major rock elements based on fusion with $Li_2B_4O_7$.

Component	Concentration range (% wt.)	Accuracy
SiO_2	0.5-5	± 15% rel.
	5-25	± 10% rel.
	25-100	≤ 5% rel.
Al_2O_3	0.05-0.5	± 0.05% wt.
	0.5-1.5	± 0.1% wt.
	1.5-5	± 0.2% wt.
	5-10	± 0.4% wt.
	>10	≤ ±0.8% wt.
Fe_2O_3	0.03-0.1	± 0.03% wt.
	0.1-7.5	± 15% rel.
	7.5-25	± 10% rel.
	>25	≤ ± 5% rel.
TiO_2	0.005-0.05	±0.005% rel.
	0.05-0.3	± 15% rel.
	0.3-0.75	± 10% rel.
	0.75-2.5	± 5% rel.
	>2.5	≤ ± 2.5% rel.
MgO	0.03-0.5	± 0.03% wt.
	0.5-5	± 10% rel.
	5-25	± 5% rel.
	>25	≤ ± 2.5% rel.
CaO	0.007-0.05	± 0.007% wt.
	0.05-1	± 15% rel.
	1-5	± 10% rel.
	5-20	± 5% rel.
	>2.5	≤ ± 2% rel.
Na_2O	0.03-0.1	± 0.03% wt.
	0.1-0.5	± 0.07% wt.
	0.5-2	± 15% rel.
	2-5	± 10% rel.
	>5	≤ ± 5% rel.
K_2O	0.005-0.02	± 0.005% wt.
	0.02-0.5	± 15% rel.
	0.5-2.5	± 10% rel.
	2.5-7.5	± 5% rel.
	>7.5	≤ ± 2.5% rel.
MnO	0.004-0.02	± 0.004% wt.
	0.02-0.2	± 15% rel.
	0.2-0.5	± 10% rel.
	>0.5	≤ ± 5% rel.
P_2O_5	0.03-0.1	± 0.03% wt.
	0.1-0.5	± 15% rel.
	0.5-1	± 10% rel.
	>1	≤ ± 5% rel.

Minor and trace elements

Analysis of minor and trace elements in geological materials is normally based on finely ground samples which are mixed with a binder and finally pressed to pellets. The sample preparation technique used for major elements gives a dilution of the sample by a factor of 8, a dilution which one normally cannot afford for trace elements at low concentration levels. Grain size effects are less dominant for the heavier elements due to their relatively short wavelengths.

The preparation technique used at NGU applies a sample weight of 5.4 g which is mixed with 1.2 g of wax. The samples are pressed in an hydraulic press to tablets with a pressure of 20 kN.

Instrumental calibration is based on international reference materials which are prepared in the same way as the samples. Correction for matrix effects are done either by referring the net line intensities to a Compton peak or to a chosen background position. Below are listed the minor and trace elements which are analysed, together with their detection limits in ppm (3σ-level).

Element:	LLD/ppm:	Element:	LLD/ppm
Pb	9	U	7
Th	10	RB	4
Y	3	Nb	4
Sr	4	Zr	3
Mo	4	As	6
V	3	Cr	4
Sc	8	S	6
Cl	13	F	300
Ba	9	Sb	8
Sn	8	Cd	6
Ag	8	Ga	5
Zn	3	Cu	6
Ni	5	Yb	16
Co	10	Ce	12
Nd	8	La	11
W	10		

Precision tests are done by repeated measurements over a time period of 6 months of a natural rock sample.

Element	Conc/ppm	σ/ppm
Pb	32.8	4.3
U	19.6	4.0
Th	<LLD	1.6
Rb	176	2
Y	46.9	1.5
Nb	20.9	1.0
Sr	86.7	1.0
Zr	291	3
Mo	<LLD	1.5

Element	Conc/ppm	σ/ppm
As	<LLD	1.7
V	173	2.8
Cr	88.8	3.1
Sc	15.9	2.4
S	<LLD	180
Cl	<LLD	10
F	800	80
Ba	611	11
Sb	<LLD	2.2
Sn	6.8	2.8
Cd	<LLD	2.4
Ag	<LLD	2.4
Ga	20.4	2.5
Zn	117	2
Cu	52.0	2.5
Ni	47.9	1.1
Yb	<LLD	4,7
Co	20.9	1.9
Ce	131	10
Nd	48.6	3.0
La	67.1	4.1
W	<LLD	4.6

Analytical accuracy which can be defined as the difference between correct concentration and the found value, is a function of concentration. The absolute accuracy in ppm at concentrations near the detection limit is of the order of 5 to 20 ppm for the element group Nb, Y, Rb, Pb, Th, U, Sr, Mo, Sc, As, Sr, V, Zr, Sn, La, Ni, Ga, Sb, Cd, Nd, Ce, Ag, Co, Yb, Zn and W. The accuracy for Cr and Ba is correspondingly in the range 30 to 50 ppm.

Even if X-ray fluorescence applies solid samples without any dilution its detection capabilities are often insufficient for many purposes, e.g. environmental analyses and geochemical studies. For concentrations above around 10 ppm X-ray fluorescence is a valuable alternative with high grade of accuracy.

Inductively coupled plasma atomic emission spectrometry

Emission techniques with flame, electric ark or spark excitation have long traditions in the analysis of geological materials, especially for minor and trace elements. This is mainly due to low detection limits and excellent multielement capability. Electric arc and spark excitation were mostly used on solid samples, but for some purposes high voltage spark excitation was used on solutions.

A relatively new excitation source for solutions with superior excitation properties is the so-called inductively coupled plasma (ICP). The principle for the ICP is in short that argon gas is heated to very high temperature, up to 10 000°C, by an induction coil carrying an alternating current of high frequency, mostly of 27.12 MHz. The sample to be analysed passes a nebulizer where the solution is transformed to a mist of small droplets before entering the plasma where the atoms are excited.

The development of inductively coupled plasma atomic emission spectrometry (ICP-AES) represented in many ways an important breakthrough in the emission analysis of solutions, and this technique is today well established as a multielement technique with low detection limits and high precision. At NGU the method has been in use since 1978 for analyses related to geochemical prospecting and mapping, and in later years for environmental investigations. The instrumentation is a Thermo Jarrell Ash ICAP 61.

Since the method works with solutions, solid samples must be brought in solution prior to analysis. Depending on the purpose of the analysis this can be done in different ways, either as total dissolution or as a partial dissolution with a weak extrant.

A standard procedure used at NGU is autoclave extraction in 7N HNO_3 in accordance with Norsk Standard NS 4770. By this technique a sample weight of 1 g is used and an analysis volume of 100 ml. This means that the samples are diluted by a factor of 100, which must be born in mind when considering detection limits. Table 2 shows the obtained lower limits of determination.

Table 2. Lower limits of determination of solid samples by ICP-AES and autoclave extraction according to NS 4770.

Element	Blanks (Values based on standard dilution 1:100)		Lower limits of determination (ppm)
	Average (ppm)	Standard deviation (ppm)	
Si	27.20	12.46	100.-
Al	5.16	2.19	20.-
Fe	2.60	0.67	5.-
Ti	0.11	0.15	1.-
Mg	9.54	10.24	100.-
Ca	54.82	43.51	200.-
Na	58.06	54.96	200.-
K	12.44	21.10	100.-
Mn	0.09	0.05	0.2
P	2.12	3.03	10.-
Cu	0.10	0.26	1.-
Zn	0.45	0.49	2.-
Pb	-3.42	1.34	5.-
Ni	0.002	1.01	2.-
Co	-0.02	0.20	1.-
V	0.35	0.30	1.-
Mo	-1.48	0.24	1.-
Cd	-0.16	0.11	1.-
Cr	0.17	0.25	1.-
Ba	0.25	0.18	1.-
Sr	0.44	0.40	2.-
Zr	-1.8598	0.2067	1.-
Ag	-0.1202	0.1192	1.-
B	1.2525	0.6389	5.-
Be	0.0148	0.0208	0.2
Li	0.2593	0.1868	1.-
Sc	0.0344	0.0507	0.2
Ce	3.8075	2.3758	10.-
La	0.3201	0.2554	1.-
Y	0.0334	0.0259	0.2
Sb	0.7275	1.2231	10.-
As	-0.8945	2.1192	10.-

The ICP-AES method can of course be used on solutions which have not been prepared by extraction, and the lower limits of determination will in this case be lower. Obtained limits of determination are given in table 3.

Table 3. Lower limits of determination of solutions by ICP-AES.

Element	Standard deviation on blank in ppm	Lower limits of determination in ppm
Si	0.0065	0.02
Al	0.0076	0.02
Fe	0.0014	0.01
Ti	0.00033	0.005
Mg	0.015	0.05
Ca	0.0033	0.02
Na	0.0080	0.05
K	0.21	1.00
Mn	0.00034	0.002
P	0.036	0.20
Cu	0.00074	0.005
Zn	0.00043	0.002
Pb	0.013	0.05
Ni	0.0090	0.05
Co	0.0020	0.01
V	0.00097	0.005
Mo	0.0012	0.01
Cd	0.0020	0.01
Cr	0.0015	0.01
Ba	0.00029	0.005
Sr	0.00012	0.005
Zr	0.00075	0.005
Ag	0.00084	0.01
B	0.0053	0.05
Be	0.00013	0.002
Li	0.00083	0.005
Sc	0.00012	0.001
Ce	0.0079	0.05
La	0.0013	0.01
Y	0.00021	0.001
Sb	0.043	0.20
As	0.021	0.10

Determination of precision of an analytical method must include all steps related to the analysis, in this case both extraction and the subsequent analysis. Table 4 shows precision of the instrument and precision on complete analysis, based on reference sample PACS-1.

Table 4. Instrumental precision and precision of complete method.

PACS-1				
Element	10 analyses of one solution (extract)		Single analyses of 10 solutions (extracts)	
	Average	Standard deviation	Average	Standard deviation
Si	190.7 ppm	1.7 ppm	98.0 ppm	17.0 ppm
Al	2.33 %	138.0 ppm	2.42 %	0.10 %
Fe	4.53 %	324.- ppm	4.49 %	0.11 %
Ti	0.13 %	10.- ppm	0.13 %	23.9 ppm
Mg	1.06 %	73.- ppm	1.04 %	186.6 ppm
Ca	1.04 %	64.- ppm	1.01 %	186.8 ppm
Na	1.83 %	138.- ppm	1.81 %	353.- ppm
K	0.39 %	24.- ppm	0.39 %	56.8 ppm
Mn	302.1 ppm	1.9 ppm	298.5 ppm	5.8 ppm
P	887.9 ppm	11.7 ppm	887.4 ppm	17.4 ppm
Cu	442.6 ppm	3.2 ppm	454.1 ppm	7.7 ppm
Zn	772.8 ppm	4.4 ppm	765.9 ppm	3.2 ppm
Pb	377.1 ppm	3.9 ppm	375.2 ppm	6.0 ppm
Ni	36.5 ppm	0.8 ppm	38.3 ppm	2.5 ppm
Co	16.0 ppm	0.2 ppm	16.2 ppm	0.4 ppm
V	89.4 ppm	0.6 ppm	88.2 ppm	1.6 ppm
Mo	9.1 ppm	0.3 ppm	9.0 ppm	0.4 ppm
Cd	2.8 ppm	0.2 ppm	3.0 ppm	0.3 ppm
Cr	55.9 ppm	0.4 ppm	55.4 ppm	0.9 ppm
Ba	376.3 ppm	3.- ppm	371.2 ppm	3.2 ppm
Sr	85.4 ppm	0.8 ppm	79.6 ppm	1.5 ppm
Zr	18.8 ppm	0.1 ppm	18.6 ppm	0.3 ppm
Ag	1.3 ppm	0.1 ppm	0.4 ppm	0.1 ppm
B	47.1 ppm	0.8 ppm	44.2 ppm	1.6 ppm
Be	6.3 ppm	0.1 ppm	5.9 ppm	0.1 ppm
Li	26.6 ppm	0.3 ppm	26.1 ppm	0.5 ppm
Sc	7.1 ppm	0.0 ppm	7.- ppm	0.1 ppm
Ce	21.1 ppm	1.0 ppm	20.4 ppm	0.9 ppm
La	13.1 ppm	0.3 ppm	14.5 ppm	0.3 ppm
Y	8.9 ppm	0.1 ppm	8.7 ppm	0.2 ppm
Sb	26.3 ppm	0.9 ppm	27.6 ppm	1.6 ppm
AS	167.1 ppm	3.3 ppm	155.5 ppm	4.7 ppm

Analytical accuracy is like precision dependant both on the instrument and on the extraction. It is therefore important that both steps are included in the tests on analytical accuracy. However, many samples are in the form of solutions, and do not therefore undergo extraction. Table 5 shows analytical accuracy based on direct analysis of sample solutions. For this test so-called quality control standards have been used.

Table 5. Accuracy obtained by ICP-AES analysis of solutions based on quality control standards QC-24 og QC-7.

Element	Obtained result	Correct values	Relative deviation (%)
Si	-	-	-
Al	25.60 μg/ml	25.00 μg/ml	2.4
Fe	25.25 "	25.00 "	1.0
Ti	5.26 "	5.00 "	5.2
Mg	24.79 "	25.00 "	0.8
Ca	24.82 "	25.00 "	0.7
Na	25.42 "	25.00 "	1.7
K	50.08 "	50.00 "	0.2
Mn	5.00 "	5.00 "	0.0
P	5.06 "	5.00 "	1.2
Cu	5.08 "	5.00 "	1.6
Zn	4.82 "	5.00 "	3.6
Pb	4.87 "	5.00 "	2.6
Ni	5.06 "	5.00 "	1.2
Co	4.93 "	5.00 "	1.4
V	4.86 "	5.00 "	2.8
Mo	5.01 "	5.00 "	0.2
Cd	0.495 "	0.500 "	1.0
Cr	4.98 "	5.00 "	0.4
Ba	4.90 "	5.00 "	2.0
Sr	4.45 "	5.00 "	11.0
Zr	5.20 "	5.00 "	4.0
Ag	-	-	
B	5.03 "	5.00 "	0.6
Be	5.11 "	5.00 "	2.2
Li	5.03 "	5.00 "	0.6
Sc	4.97 "	5.00 "	0.6
Ce	5.05 "	5.00 "	1.0
La	5.00 "	5.00 "	0.0
Y	4.98 "	5.00 "	0.4
Sb	5.18 "	5.00 "	3.6
As	5.06 "	5.00 "	1.2

Accuracy of complete procedure involving extraction and analysis are based on international reference materials. Table 6 and 7 show results on respectively PACS-1 and CRM-277. The reference materials are certified for total element content, while the ICP-AES analysis is based on partial extraction. The agreement is therefore expected to be good only for easily extractable elements. Analytical accuracy is in such cases best controlled by crosstesting based on identical extraction procedures.

Table 6. ICP-AES analysis of PACS-1 based on extraction according to NS4770.

Element	Obtained result	Certified values (total)
Si	98.0 ppm	55.7 ± 0.5 % SiO_2
Al	2.42 %	12.23 ± 0.22 % Al_2O_3
Fe	4.49 %	6.96 ± 0.12 % Fe_2O_3
Ti	0.13 %	0.703 ± 0.011 % TiO_2
Mg	1.04 %	2.41 ± 0.09 % MgO
Ca	1.01 %	2.92 ± 0.13 % CaO
Na	1.81 %	4.40 ± 0.11 % Na_2O
K	0.39 %	1.50 ± 0.09 % K_2O
Mn	298.5 ppm	470 ± 12 ppm
P	887.4 ppm	0.253 ± 0.018 % P_2O_5
Cu	454.1 ppm	452 ± 16 ppm
Zn	765.9 ppm	824 ± 22 ppm
Pb	375.2 ppm	404 ± 20 ppm
Ni	38.3 ppm	44.1 ± 2.0 ppm
Co	16.2 ppm	17.5 ± 1.1 ppm
V	88.2 ppm	127 ± 5 ppm
Mo	9.0 ppm	12.3 ± - ppm
Cd	3.0 ppm	2.38 ± 0.20 ppm
Cr	55.4 ppm	113 ± 8 ppm
Ba	371.2 ppm	-
Sr	79.6 ppm	277 ± - ppm
Zr	18.6 ppm	-
Ag	0.4 ppm	-
B	44.2 ppm	-
Be	5.9 ppm	-
Li	26.1 ppm	-
Sc	7.- ppm	-
Ce	20.4 ppm	-
La	14.5 ppm	-
Y	8.7 ppm	-
Sb	27.6 ppm	171 ± 14 ppm
As	155.5 ppm	211 ± 11 ppm

Table 7. ICP-AES analysis of CRM-277 based on extraction according to NS 4770.

Element	Obtained results	Certified values (total) *	
Si	102.2 ppm	(23.0 %)	**
Al	2.60 %	(4.8 %)	
Fe	4.67 %	(4.55 %)	[4.17 %]
Ti	262.9 ppm	(0.30 %)	
Mg	0.88 %	(1.0 %)	
Ca	3.63 %	(6.0 %)	
Na	0.81 %	(1.2 %)	
K	0.64 %	(1.6 %)	
Mn	0.14 %	(0.16 %)	[0.162 %]
P	0.40 %	(0.41 %)	
Cu	101.9 ppm	101.7 ± 1.6 ppm	[97.2 ppm]
Zn	508.0 ppm	547 ± 12 ppm	[557 ppm]
Pb	129.8 ppm	146 ± 3 ppm	[137.5 ppm]
Ni	40.3 ppm	43.4 ± 1.6 ppm	[34.9 ppm]
Co	14.4 ppm	(17 ppm)	
V	72.1 ppm	(102 ppm)	
Mo	1.2 ppm	(1.5 ppm)	
Cd	11.0 ppm	11.9 ± 0.4 ppm	[10.8 ppm]
Cr	138.2 ppm	192 ± 7 ppm	[145.6 ppm]
Ba	117.9 ppm	(329 ppm)	
Sr	241.8 ppm	-	
Zr	18.7 ppm	-	
Ag	0.1 ppm	(3.3 ppm)	
B	47.5 ppm	-	
Be	6.9 ppm	(1.6 ppm)	
Li	30.4 ppm	-	
Sc	5.9 ppm	9.0 ± 0.12 ppm	
Ce	78.0 ppm	-	
La	36.6 ppm	(45 ppm)	
Y	29.2 ppm	-	
Sb	19.9 ppm	(4.0 ppm)	
As	11.9 ppm	47.3 ± 1.6 ppm	

* Values in brackets are not certified.
** Figures are uncertified values based on aqua regia dissolution.

Atomic absorption spectrometry

Atomic absorption spectrometry (AAS) is a well established analytical technique which in many cases can be an alternative to ICP-AES, since detection capabilities with some exceptions are comparable, and both techniques most often work with solutions.

The main principle for AAS is that free atoms in the ground state absorb resonance radiation which is emitted from a lamp and isolated in a monochromator. Absorption of light is a function of concentration. Because the temperature in a flame is much lower than in a plasma, the detection limits for easily ionizable elements are often better by AAS than by ICP-AES. Elements forming refractory compounds in the flame have as a rule far lower detection limits by ICP-AES.

At NGU AAS is today mainly used for analysis of Cd and Pb by graphite furnace

technique, and for analysis of Hg by cold vapour technique. The standard extraction procedure used for ICP-AES is also used for the AAS analyses. Since Hg routinely is determined in most samples, the material is dried at a temperature of 50°C for 48 hours to avoid loss by evaporation.
The instrumentation used for Pb and Cd is a Perkin Elmer spectrophotometer model 5000, and for Hg a Perkin Elmer model 403.

The lower limit of determination of Cd and Pb in solid geological materials are:

Cd: 5.0 ppb
Pb: 50 ppb

Tests show that precision normally is below 10% rel.

Tests on accuracy are based on reference materials certified for total content. It is expected that most Cd and Pb dissolve during extraction, and analytical accuracy seems to be within ±10 % rel. for both elements in most cases.

The lower limit of determination of Hg in geological materials is:

Hg: 10 ppb

Precision has been tested to be around 10% rel..

Accuracy has been found to be in the order of ±10% rel., with the same reservation that the reference materials are certified for their total element content.

Quality control

The quality system being used in the accredited laboratories of NGU is in accordance with the requirements in NS-EN 45001.
One of the most important elements in this control is the use of so-called X-diagrams. This is a system where a homogeneous natural sample is analysed with regular intervals and used for controlling analytical quality. Based on a certain number of determinations, for instance 10, over a certain time period, average values and standard deviations are calculated for the actual elements. These informations are used as criteria for accepting future analytical values of this control sample. The system operates with two limits, one being an alarm limit which is average value $\pm 2\sigma$, and the other being an action limit which is average value $\pm 3\sigma$.

References

Augustitis, S. S.(editor) 1983: Leaching and diffusion in rocks and their weathering products. Theophrastus Publications S.A. Athens 1983.

Faye, G. Chr. & Ødegård, M. 1975: Determination of major and trace elements in rocks employing optical emission spectroscopy and X-ray fluorescence. Norges geol. unders. 322, 35-53.

Faye, G. Chr. 1982: Metodestudier i geokjemi - HNO_3-ekstraksjon av geokjemiske prøver. Nor. geol. unders. report 1687 C.

Graff, P. R. & Røste, J. R. 1985: Utluting av silikatmineraler med mineralsyrer. Nor. geol. unders. report 85.105.

Løbersli, E. M., Steinnes, E. and Ødegård, M. 1990: A Historical Study of Mineral Elements in Forest Plants from South Norway. Environm. Monitoring and Assessment. Vol.15, pp. 111-129.

Olsen, S. D., Bøhmer, R. G. og Ødegård, M. 1985: Induktivt koblet plasma som spektroskopisk eksitasjonskilde-De mange muligheters metode. Kjemi nr. 4, 1984.

Steinnes, E., Johansen, O., Røyset, O. and Ødegård, M. 1993: Comparison of different multielement techniques for analysis of mosses used as biomonitors. Environm. Monitoring and Assessment. Vol.25, pp 87-97.

Sæther, O.M., Misund, A., Ødegård, M., Andreassen, B.Th. and Voss, A. 1992: Groundwater contamination at Trandum landfill, Southeastern Norway. Nor. geol. unders. Bull. 422, 83-95.

Ødegård, M. 1979: Determination of Major Elements in Geological Materials by ICAP Spectroscopy. Jarrell-Ash Plasma Newsletter, Vol.2. No.1.

Ødegård, M. 1980: Plasmaspektrometri- En viktig analysemetode for geologiske materialer. NGU's årsmelding 1980.

Ødegård, M. 1981: The use of inductively coupled argon plasma (ICAP) atomic emission spectroscopy in the analysis of stream sediments. J. Geochem. Explor., Vol.14, 119-130.

Ødegård, M. 1983: Utvidet program for analyse av geologiske materialer basert på syreekstraksjon og plasmaspektrometri. Nor. geol. unders. report 2113.

Ødegård, M. and Andreassen, B. Th. 1986: Methods for water analysis at the Geological Survey of Norway. The Norwegian Academy of Science and Letters.

Ødegård, M. 1990: Den analytiske kjemis plass i geoinstitusjonen. Geonytt. Årgang 17. Nr. 2.

Ødegård, M. 1990: Reiserapport fra studietur USA, okt. 1990. NGU-rapport 90.016.

Ødegård, M. 1991: Rapport fra møte i "ATV-Komiteen Vedrørende Grundvandsforurening" - København 25 september 1991. NGU- rapport 91.016.

Chemical analysis of water

John Mather
Lyell Professor of Geology, Royal Holloway, University of London, Egham, Surrey, UK.

INTRODUCTION

Procedures are available for the analysis of waters of a wide range of quality, including water suitable for domestic supply, surface water, groundwater, cooling or circulating water, boiler water, treated or untreated effluents and sea water. This presentation will concentrate on the analysis of groundwaters, but even here the total range of constituents that could be determined is very large. Each constituent may give some information about the nature of the aquifer, residence time, amount of contamination and so on. It will generally be necessary for reasons of economy to restrict the number of parameters measured.

Historically, the type of chemical analysis carried out by the water industry has not been adequate for geochemical or hydrogeological analysis, giving only potability criteria. The minimum set of data useful for geochemical study is the mineral analysis for the major ions - Ca^{2+}, Mg^{2+}, Na^{+}, K^{+}, $SO_4^{2-}Cl^{-}$, HCO_3^{-} and NO_3^{-} - and pH value. Using only these parameters together with temperature, it is possible to infer a lot about the history of the water and the rocks through which it has passed. This should form the basic unit of analysis and interpretation: minor elements, trace elements, organic species and isotopic studies can provide useful additional information.

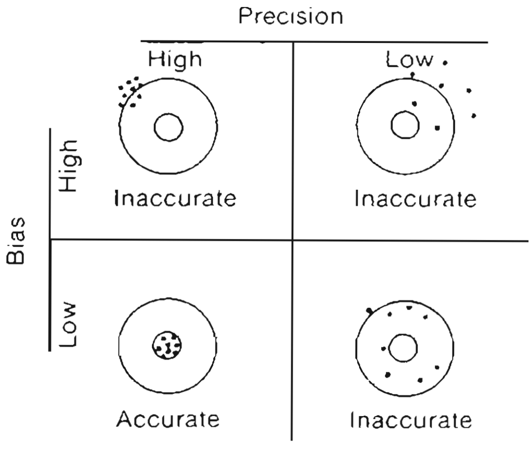


Fig. 1. Definition of accuracy (from Clesceri, Greenberg and Trussell, 1989).

DATA QUALITY

The principal indicators of data quality are its bias and precision which, when combined express its accuracy. Bias can be defined as "consistent deviation of measured values from the true value, caused by systematic errors in procedure" and precision as "measure of the degree of agreement among replicate analyses of a sample, usually expressed as the standard deviation". Accuracy combines bias and precision to reflect the closeness of a measured value to a true value (Clesceri, Greenberg and Trussell, 1989). The relationship between these terms is shown in Figure 1. Of the four possible outcomes only the condition of low bias and high precision is accurate.

Other data quality indicators are method detection limit and representativeness. The detection limit is the smallest amount that can be detected above the noise in a procedure and within a stated confidence limit. Representativeness can relate both to the sample itself and to the sampled population. A method may be very accurate, but if the results do not represent the population the data are useless. These problems are likely to be less significant in the analysis of water samples than in analyses of rocks and soils.

The correctness of a groundwater analysis can be checked by carrying out a cation-anion balance. The cation and anion sums, when expressed as millequivalents per litre, must balance because all groundwaters are electrically neutral. The criterion for acceptance is that

$$\frac{\Sigma \text{ cations} - \Sigma \text{ anions}}{\Sigma \text{ cations} + \Sigma \text{ anions}} \times 100 \text{ is within } \pm 5\%$$

Most laboratories have also set up a set of operating principles that, if strictly followed during sample collection and analysis will produce data of known quality. Such quality assurance (QA) includes both quality control and quality assessment. Quality control will include such things as certifying operator competence, analysis of externally supplied standards and reagent blanks, calibration with standards, analysis of duplicates etc (Clesceri, Greenberg and Trussell, 1989). Quality assessment is the process of using both internal and external quality control measures to determine the quality of data produced by a laboratory. It includes such things as performance evaluation samples, laboratory intercomparison samples and performance audits as well as internal quality control. These are applied to test the recovery, bias, precision, detection limit and adherence to standard operating procedures. Adequate quality assurance procedures are now a prerequisite before a laboratory can tender for much contract work in the UK, particularly contract work related to the radioactive waste programme.

SAMPLE COLLECTION AND ANALYSIS

Analysing a sample of groundwater should not represent a serious challenge given that the species to be determined are already in solution. Unfortunately the position is not quite as simple as it appears and there are many problems to frustrate the analyst.

The first difficulty is that the analysis should represent as closely as possible the in situ aquifer conditions. Some species, especially those involved in carbonate and redox reactions will undergo changes on reaching the atmosphere and need to be analysed in the field at the well head. Redox changes lead to other difficulties such as the precipitation of metals and thus samples need to be stabilised to keep metal ions in solution. A protocol for the collection of groundwater samples is given in Figure 2.

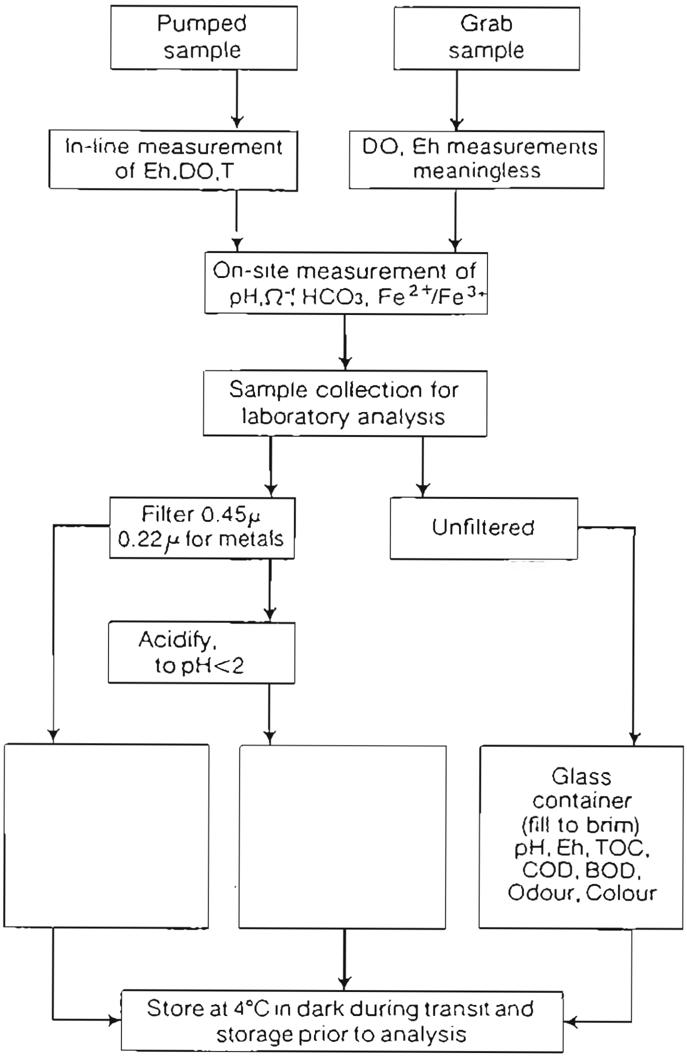


Fig. 2. Flowchart for the collection and analysis of groundwater samples.

(1) Field parameters

The field parameters normally measured are pH, conductivity, alkalinity, temperature, redox potential and dissolved oxygen. Some of these require in-line measurements, whereas others may be analysed on site but not necessarily in-line.

pH

When groundwater is discharged from an aquifer the physical controls governing H^+ ion activity are changed and it is very important that pH is measured at the well head. The pH is determined using a glass electrode compared with a reference electrode of a know potential by means of a pH meter or other potential-measuring device. Using modern equipment pH measurements in the field can realistically be made to ±0.01 of a unit.

Conductivity

This provides a valuable method for the rapid estimation of total mineralisation and reliable equipment is readily available. Conductivities in groundwater are normally reported as micromhoes per centimetre or in SI units, microsiemens per centimetre (μScm^{-1}). As ionic activity is affected by temperature, electrical conductivity increases with temperature and most measurements are reported at 25°C.

Alkalinity

Measured by titration with acid to a fixed end point or to the inflection point on the titration curve, alkalinity is a measure of the acid neutralising capacity. In groundwaters, and most other natural waters, the carbonate equilibria predominate and the end point at pH 4.5 essentially measures the HCO_3^- concentration (Figure 3). In some waters other weak acids - borate, silicate and organic acids - contribute to the alkalinity and must be subtracted for more accurate interpretation of the carbonate system. Computer programmes are available to perform this function. It should be noted that alkalinity is not affected by gain or loss of CO_2 gas. However, the pH will change; it will increase with CO_2 (gas) loss and decrease with CO_2 (gas) gain.

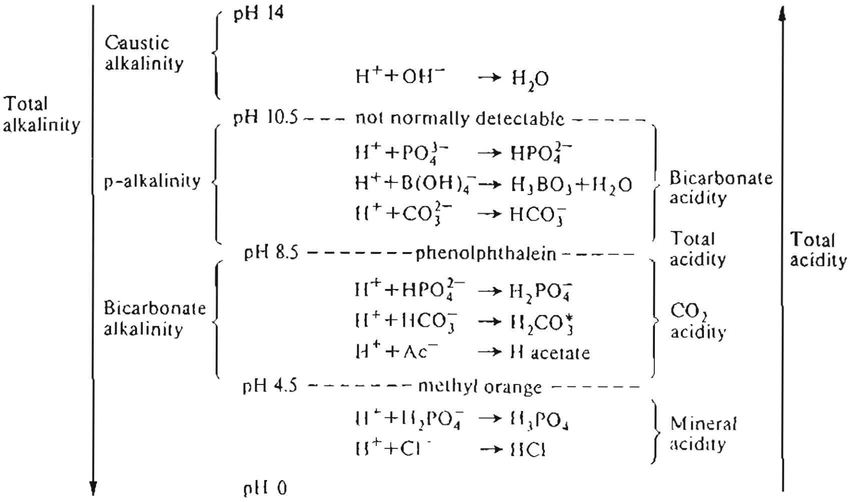


Fig. 3. Reactions during alkalinity titrations (Lloyd and Heathcote, 1975).

Temperature

All geochemical reactions are temperature dependant so that measurements are essential for assessing the extent to which equilibrium has been attained. Measurements can be made with a standard mercury thermometer or using the probe attached to the pH meter for temperature compensation.

Redox Potential

E_H, the oxidation-reduction potential is a qualitative measurement and does not provide information on the specific behaviour of a single redox couple or on the oxidising or reducing condition of the water. However, it provides a convenient theoretical means of studying the limiting conditions of element speciation and mobility in the natural environment.

It is measured using a platinum redox electrode and a reference electrode usually sealed in one unit. As the oxidation-reduction state of a water will change on exposure to the atmosphere E_H measurements have to be carried out in a sealed cell (Figure 4). The cell is attached to a "bleed line" on the discharge pipe from the well pump and water is allowed to flow through the cell until stable conditions are attained.

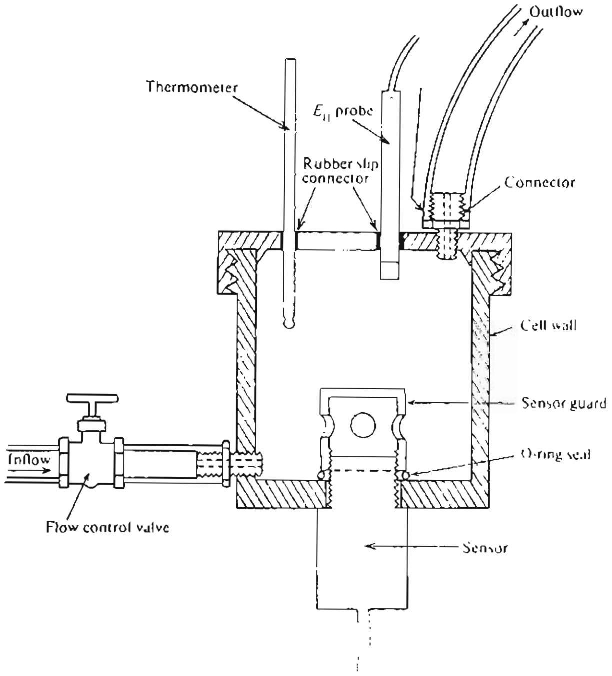


Fig. 4. Isolation cell for the measurement of E_H, temperature and dissolved oxygen (Lloyd and Heathcote, 1985.

Dissolved Oxygen

As with E_H dependable measurements can only be obtained from a sample which is sealed from the atmosphere in the type of cell shown in Figure 4 using an oxygen electrode.

(2) Laboratory measurements

Samples for laboratory analysis need to be collected with the species to be determined firmly in mind. Figure 2 shows 3 distinct routes as follows:

- filtration through a 0.45 µm filter into a plastic container filled to the brim for anion analysis

- collection in a gas bottle of an unfiltered sample for the analysis of organic species.

Filtration is necessary as most groundwaters contain particulate and colloidal material. Samples for cation analysis are normally filtered through a 0.45µm filter and acidified with 1% nitric acid. This should remove particulate matter and maintain the main cations in stable solution. Samples for anion analysis require filtration but not acidification. Samples for organic analysis need to be collected in glass bottles without filtration. Special techniques may be required to collect such samples because of loss of volatile components. Consideration of these is outside the scope of this chapter.

Cation analysis

Atomic absorption spectrometry has been the mainstay of elemental analysis for waters for many years now, and it is quite probably still the most widely used method of analysis (Walsh, 1993). Despite the introduction of more versatile techniques it retains many advantages. It is above all simple to calibrate and operate, it is comparatively inexpensive, rapid and generally free from interferences. Flame atomic absorption is still one of the best methods for routine analysis of some of the easily volatilised metals such as Zn, Cd, Cu, Ag, Pb, Rb etc. For many of these elements it will give excellent detection limits during routine analysis of water samples. However there are numerous occasions when the detection limits attainable through flame atomic absorption analysis are inadequate. This applies even to the 'good' flame AA elements, noted above. Several options are available to improve detection limits. Flame AA can be replaced by flameless AA (electrothermal atomisation), the samples could be preconcentrated or some other analytical instrumental method can be used.

Flameless AA has been a popular method in water analysis, it certainly offers dramatic improvements in detection limits in comparison with flame AA. The other advantage that flameless AA has is the relatively small increase in cost of analysis. Adding the flameless attachment is far less expensive than replacing it with another technique. However flameless AA does have several disadvantages. It is comparatively slow although automation of sample input will reduce this problem. A more serious disadvantage is the matrix interferences that are found. This is especially troublesome where the analyte element is at trace levels and the matrix elements are high levels of refractory elements. This has severely limited the use of flameless AA in analysis of refractory samples (notably geological and ceramic materials). For water analysis the problem is not so severe but high (and very variable) concentrations of calcium and sodium have restricted the application of the technique. Flameless atomic absorption is likely to retain its popularity for selected trace element determinations in water analysis. The determination of Pb is one example of such elements, where the detection limit required is beyond the capabilities of most alternative techniques. However in cases where an alternative analytical technique is available this is likely to be used in preference to flameless AA.

Preconcentration of water samples prior to analysis is an obvious method for enhancing detection limits and it is appropriate for many water samples. The total dissolved content of some waters is low (less than 5000 µm/ml) and preconcentration by simple evaporation can be usefully adopted. A typical level of preconcentration will be 10 or 20 fold. This technique cannot be used for waters that already contain high dissolved solids and it is a somewhat tedious procedure to implement for routine analysis of large numbers of samples.

A technique that has gained widespread acceptance in recent years for elemental analysis of water samples is inductively-coupled-plasma atomic emission spectrometry (ICP-AES). This, together with ICP-mass spectrometry, offers many advantages and its use is bound to grow in the next few years. The main features of merit can be summarised here. ICP-AES has the capability to measure a very wide range of elemental concentrations. The method gives good results for a far greater range of elements than flame AA. It is especially useful for the determination of many of the more refractory elements, notable examples are Si, Al, Ti, etc.

It is also capable of measuring elements simultaneously and this has had substantial implications for the costs of routine analysis. Although the cost of instrumentation is high the costs of analysis per sample where a range of elements are sought, can be dramatically reduced. ICP-AES is capable of measuring element concentrations at both high levels and low levels. Thus for a water analysis programme the traditional major cations Ca, Na and Mg, K, Si, can be determined directly at the 5-500 µm (ppm) level. In addition trace element concentration can be determined at the same time at the parts per billion level.

It is important to also consider the limitations of ICP-AES. It is not entirely without interferences and these have to be carefully documented and detection limits, which are low, may still not be low enough for ground waters. This has lead to the introduction of another ICP based technique - ICP-MS. The linking of the inductively coupled plasma source unit to a conventional quadrupole mass spectrometer gives an analytical technique capable of measuring elemental abundance levels 100 fold lower than ICP-AES (or the best elements in flame AA). ICP-MS is not as rapid as ICP-AES in routine analysis - sample throughput is significantly slower. It is also not really suitable for the analysis of the major cation constituents of waters. A water analysis laboratory would almost certainly require both ICP-AES and ICP-MS to cover the range of elements required. However it does offer outstanding performance at the ultra-trace level. Many elements of environmental concern can be measured down to their natural abundance levels in groundwaters with ICP-MS.

Anion analysis

Although determination of the common anions can be made by conventional colorimetric or titrimetic techniques there is one technique, ion chromatography (IC), which provides a single instrumental method for their rapid, sequential measurement. IC eliminates the need to use hazardous reagents and it also distinguishes between the various halides and the oxides (e.g. between NO_2^- and NO_3^-).

In ion chromotography a water sample is injected into a stream of carbonate/bicarbonate eluant and passed through a series of ion exchanges. The anions of interest are separated on the basis of their relative affinities for a strongly basic anion exchanger. The separated anions in their acid forms are measured by conductivity and identified on the basis of retention times as compared to standards. Any substance that has a retention time coinciding with that of any anion to be determined will interfere. For example high concentrations of low-molecular-weight organic acids interfere with the determination of chloride. Sample dilution overcomes many interferences. A typical inorganic anion separation is shown in Figure 5.

Another technique commonly used for rapid estimation of anion concentrations is the ion selective electrode. These are subject to interference to varying degrees from other ions but are very valuable for monitoring activities.

Organic analysis

Analyses for organic matter in water can be classified into two general types of measurements (Clesceri, Greenberg and Trussell, 1989), those that quantify an aggregate amount of organic matter with a common characteristic and those that quantify individual organic compounds.

In groundwaters the aggregate amount of organic matter present is normally measured in terms of the total organic carbon (TOC). This distinghishes them from wastewaters where organic matter is usually measured in terms of biochemical oxygen demand (BOD) and chemical oxygen demand (COD). Unlike BOD or COD, TOC is independent of the oxidation state of the organic matter and does not measure other organically based elements such as nitrogen and hydrogen which can contribute to the oxygen demand measured by BOD and state of the organic matter and does not measure other organically based elements such as nitrogen and hydrogen which can contribute to the oxygen demand measured by BOD and COD. The carbon molecules are broken down into single carbon units and oxidised to CO_2. The CO_2 is then measured directly in an infra red analyser.

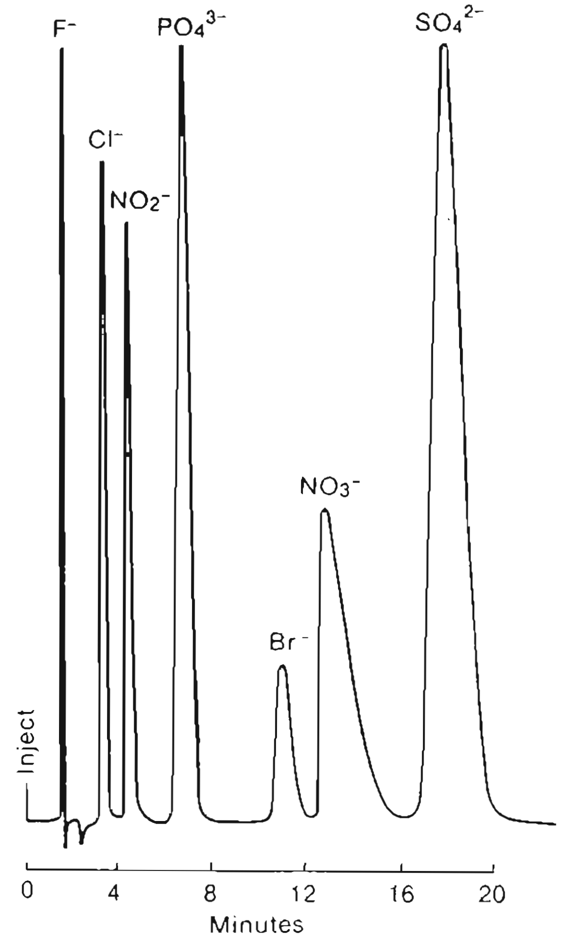


Fig. 5. Ion chromatography - typical inorganic anion separation (Clesceri, Greenberg and Trussell, 1989).

The identification and quantification of trace organic constituents, such as pesticides generally involves isolation and concentration of the organics from a sample by solvent or gas extraction, separation of the components and identification and quantification of the compounds with a detector. Gas chromatographic methods are commonly used in which a carrier gas and a stationary phase in a column are used to separate individual compounds. When the sample solution is introduced into the column, the organic compounds are vaporised and moved through the column by the carrier gas. They travel at different rates, depending on differences in partition coefficients between the mobile and stationary phases.

Various detectors can be used one of the most common being the quadrupole mass spectrometer. The GC/MS technique has become widely used to identify unknown organic compounds occurring in groundwaters.

Another technique is High-Performance Liquid Chromatography (HPLC) in which a liquid mobile phase transports a sample through a column containing a stationary liquid phase. The interaction of the sample with the stationary phase selectively retains individual compounds and permits separation of sample components. Detection is by absorbance detectors - where the detector measures the absorbance of a sample from an incident light source.

Environmental isotope analysis

A detailed discussion of the various techniques is outside the scope of this chapter. A summary of the methods is given in Table 1 (after Lloyd and Heathcote, 1985).

Isotope	Preparation	Analysis
Deuterium	Conversion of 10 μl water to hydrogen using zinc	$^2H/^1H$ ratio is measured by mass spectrometry.
Tritium	Sample is distilled to dryness to remove solids. Tritium is concentrated by electrolysing 100 ml to a volume of 10 ml.	Liquid scintillation counting or conversion to ethane and gas proportional counting.
^{13}C	Carbonate is precipitated in the field by addition of 10 ml of a solution of 400 g $SrCl_2.6H_2O\ l^{-1}$ in 0.880 NH_3 solution to sufficient water to contain 0.5 meq dissolved carbon. After 24 h precipitate is filtered off and washed. In the laboratory CO_2 gas is prepared by reaction of carbonate precipitate with 100% H_3PO_4.	$^{13}C/^{12}C$ ratio is measured by mass spectrometry.
^{14}C	Carbonate is precipitated in the field by addition of 100 ml carbonate free NaOH solution (200 g $NaOH\ l^{-1}$) and 1 litre saturated $BaCl_2$ (500 g $BaCl_2\ l^{-1}$) to 100 litres water. The amount of $BaCO_3$ precipitate required is at least 80 g. If the total alkalinity is less than 250 ppm then carbonate is precipitated from a second 100 litres water sample. In the laboratory the precipitate is acidified producing CO_2 which can be converted to benzene or methane.	Liquid scintillation counting of benzene or gas proportional counting of methane.
^{18}O	10 ml water is equilibrated with CO_2 gas permitting exchange of ^{18}O between water and gas.	$^{18}O/^{16}O$ ratio is measured by mass spectrometry.

Table 1. Summary of analytical techniques for environmental isotopes (after Lloyd and Heathcote, 1985).

(3) Pore water analysis

So far this discussion has assumed that samples are derived from pumping wells and springs. Unfortunately hydrochemical stratification is a common feature of aquifers so that the integrated samples taken from a borehole pump will inevitably have mixed chemistries. Mixing of oxygenated and reducing waters may precipitate iron and if pumping is not continuous stagnation of water in the well will result in chemical changes.

The large volumes of stored intergranular fluids will exert a major influence on water quality over a long period and it is desirable to sample and analyse such matrix fluids. Various drilling techniques can be used to recover uncontaminated core material for investigation, including augering, percussion drilling and airflush or waterflush rotary. If waterflush drilling is used a suitable tracer, such as Li^+, may be added to quantify formation invasion; otherwise air flush (dry) sampling is to be preferred (Edmunds, 1986).

Extraction of interstitial fluids may be carried out by centrifugation, squeezing or fluid displacement. For detailed investigation such interstitial fluid sampling is far the best method of examining the water profile with depth.

REPRESENTATION AND INTERPRETATION OF HYDROCHEMICAL DATA

The units chosen for hydrochemical species are generally mg/l, equivalent in dilute waters to parts per million (ppm). Above concentrations of about 7000 mg/l, however, solution density must be taken into account and values converted according to the equation:

$$\text{mg/kg} = \frac{\text{mg/l}}{\text{solution density}}$$

Trace constituents are often reported as µg/l or parts per billion (ppb).

Whilst the analysis in mg/l is most readily understood, analyses expressed in millimoles per litre (m mol/l) or in milliequivalents per litre (meq/l) are generally more useful in geochemical interpretation where chemical equivalent quantities involved in reactions must be compared.

$$\text{m.mol/l} = \frac{\text{mg/l}}{\text{molecular weight}}$$

$$\text{meq/l} = \text{m mol/l} \times \text{ionic charge}$$

Conversion Factors are given in Table 2 (after Edmunds, 1986)

Element and species	Conversion Factors (a)	Conversion Factors (b)	Element and species	Conversion Factors (a)	Conversion Factors (b)
Aluminium (Al^{3+})	0.11119	0.03715	Lead (Pb)	—	0.00483
Ammonium (NH_4^+)	0.05544	0.05544	Lithium (Li^+)	0.14411	0.14411
Barium (Ba^{2+})	0.01456	0.00728	Magnesium (Mg^{2+})	0.08226	0.04113
Beryllium (Be^{3+})	0.33288	0.11096	Manganese (Mn^{2+})	0.03640	0.01820
Bicarbonate (HCO_3^-)	0.01639	0.01639	Molybdenum (Mo)	—	0.01042
Boron (B)	—	0.09250	Nickel (Ni^{2+})	—	0.01703
Bromide (Br^-)	0.01251	0.01251	Nitrate (NO_3^-)	0.01613	0.01613
Cadmium (Cd^{2+})	0.01779	0.00890	Nitrite (NO_2^-)	0.02174	0.02174
Calcium (Ca^{2+})	0.04990	0.02495	Phosphate (HPO_4^{2-})	0.02084	0.01042
Carbonate (CO_3^{2-})	0.03333	0.01666	Potassium (K^+)	0.02557	0.02557
Caesium (Cs^+)	0.00752	0.00752	Rubidium (Rb^+)	0.01170	0.01170
Chloride (Cl^-)	0.02821	0.02821	Silica (SiO_2)	—	0.01664
Cobalt (Co^{2+})	0.03394	0.01697	Silver (Ag^{2+})	—	0.00927
Fluoride (F^-)	0.05264	0.05264	Sodium (Na^+)	0.04350	0.04350
Hydrogen (H^+)	0.99209	0.99209	Strontium (Sr^{2+})	0.02283	0.01141
Hydroxide (OH^-)	0.05880	0.05880	Sulphate (SO_4^{2-})	0.02082	0.01041
Iodide (I^-)	0.00788	0.00788	Sulphide (S^{2-})	0.06238	0.03119
Iron (Fe^{2+})	0.03581	0.01791	Zinc (Zn^{2+})	0.03060	0.01530
Iron (Fe^{3+})	0.05372	0.01791			

Note: Multiply concentration in mg/l by conversion factor

Table 2. Conversion factors from mg/l to (a) meq/l and (b) mmol/l (Edmunds, 1986).

Individual analyses can then be treated to reveal a large amount of additional information using a variety of computer software. This includes information on speciation, ionic balance, and the extent of saturation with respect to various minerals. The problem with all such modelling of groundwater chemistry is that it is based on the assumption that inorganic chemical equilibrium exists in solution. Unfortunately, such equilibrium is often not attained by redox-sensitive species, such as NO_3^-/N_2 and H_2S/SO_4^{2-} and the models need to be interpreted with caution.

Describing the concentration or relative abundance of major and minor constituents and the pattern of variability is part of nearly every groundwater investigation. Over the years a number of different graphical and statistical techniques have been developed, particularly for the presentation of major ion patterns. Examples are:

1. Bar (or Collins) diagrams
2. Pie diagrams
3. Pattern (or Stiff) diagrams
4. Vertical scale (or Schoeller) diagrams
5. Bivariate or scatter plots
6. Trilinear (or Piper) diagrams
7. Durov diagrams (development of Piper)

Examples of some of these graphical representations are shown in Figure 6.

The most useful plot for comparison of large numbers of groundwater analyses is the Piper diagram. By classifying samples on such a diagram one can identify geological units with chemically similar water and define the evolution of groundwater chemistry along a flow system. Drawbacks are that because concentrations are expressed as percentages variations in absolute concentrations are not shown clearly and analyses of similar mix but different total concentrations are not differentiated.

WATER QUALITY STANDARDS

Finally in this chapter on the chemical analysis of water I would like to briefly consider water quality standards. For use as drinking water and for irrigation, quality standards have been set by many national and intergovernmental agencies.

Water supplied for drinking must be wholesome; in other words it must meet high standards of physical, chemical and biological purity. It must be free from undesirable physical properties such as colour, turbidity, and objectionable odours and tastes. The latter properties are rather subjective and many undesirable substances, e.g. free Cl_2H_2S and phenols, are detectable at or below levels at which they might become toxic. Other substances, eg Cl^-, cannot be distinguished readily until they reach several hundred mg/l. There are now statutory requirements in Europe for the quality of water intended for human consumption which largely supercede WHO standards. A version of the CEC requirements used in the UK is given in Appendix 1.

Water should also be bacteriologically pure and for the purposes of the directive the count (per 100 ml) of coliform bacteria (as indicators of faecal pollution) should be zero. Although coliform bacteria are not themselves harmful they are likely to indicate the presence of pathogenic organisms and viruses.

The most extensive use of groundwater in the world is for the irrigation of crops. The main difference between irrigation waters and domestic supplies is the absence of pretreatment and anaerobic waters, waters high in iron and polluted supplies (eg nitrate-rich) may be used for direct irrigation.

For livestock, many of the quality standards for domestic supplies also apply. However, animals are generally able to tolerate higher total solids concentrations than humans (up to 10,000 mg/l in the case of beef cattle, for example) although levels well below this are clearly desirable.

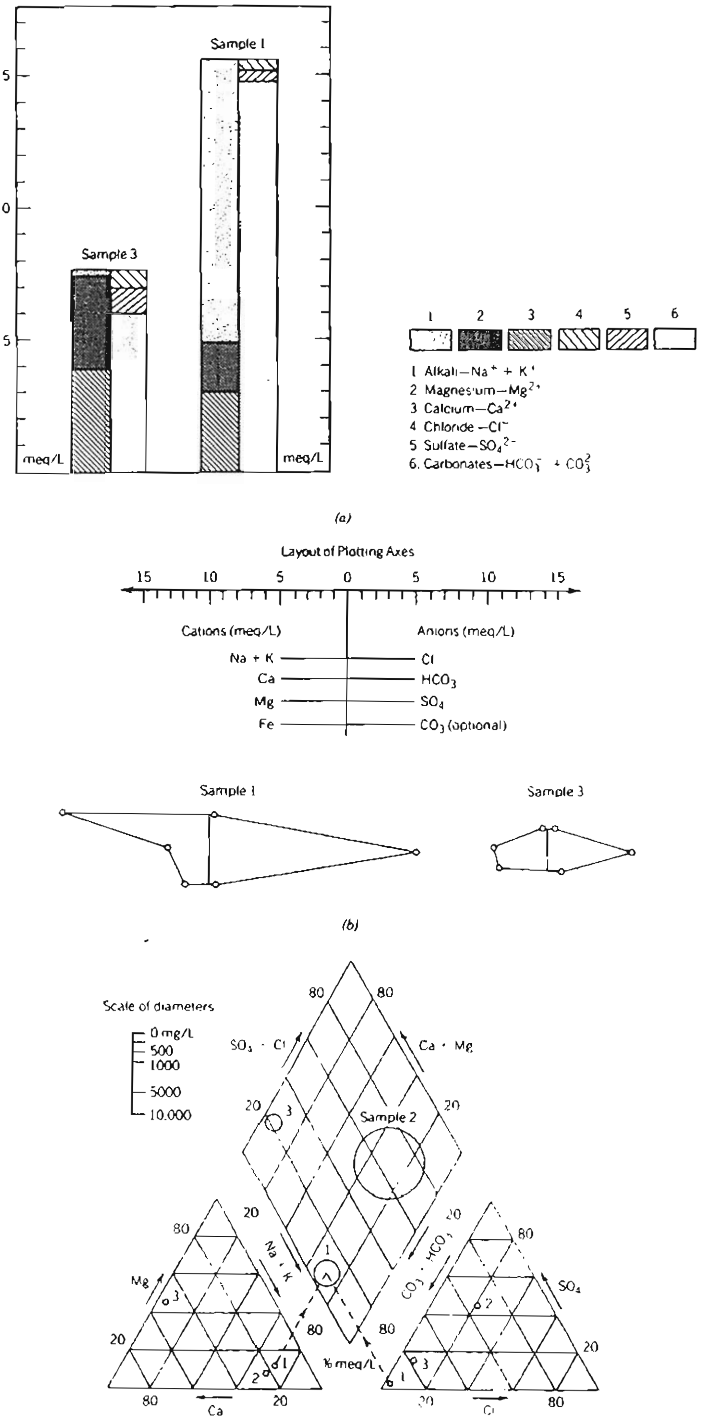


Fig. 6. Examples of plots of major ion data.
(a) Bar diagram
(b) Stiff diagram
(c) Piper diagram
(from Domenico and Schwartz, 1990)

For irrigation, the main concerns are the build-up of salinity levels and presence of elements toxic to plants. In the former case geochemical evaluation can be useful in predicting minerals likely to precipitate following irrigation. The drainage characteristics of the soil are equally as important as the water chemistry. It may be possible to irrigate with quite saline water on well drained sandy soils but on clayey soils saline build-up and the formation of gypsum or other minerals may occur during irrigation with reasonably fresh groundwater. Although some plants, like cotton and sugar beet, are relatively resistant to Cl^-, some fruits are sensitive to as little as 350 mg/l Cl^-.

An excess of sodium in irrigation water may be damaging to soil structure due to the replacement by sodium of calcium and magnesium. The sodium adsorption ratio (SAR) is commonly used to express the tendency for these cation exchange reactions to occur:

$$SAR = \frac{Na^+}{\sqrt{(Ca^{2+} + Mg2^+)/2}}$$

REFERENCES

Clesceri, L.S., Greenberg, A.E. and Trussell, R.R. (editors), 1989. Standard methods for the examination of water and wastewater, 17th edition. American Public Health Association Washington USA.

Domenico, P.A. and Schwartz, F.W., 1990. Physical and chemical hydrogeology. John Wiley, New York.

Edmunds, W.M., 1986. Groundwater chemistry. In: Brandon, T.W. (editor) Groundwater: occurrence, development and protection. Inst. Water Engineers and Scientists, London. 49-107

Lloyd, J.W. and Heathcote, J.A., 1985. Natural inorganic hydrochemistry in relation to groundwater, an introduction. Clarendon Press, Oxford.

Walsh, J.N., 1993. Modern analytical methods applied to groundwater monitoring. In: Conference Documentation E0136 Conference on Groundwater Pollution 16/17 March 1993, London IBC Technical Services Ltd.

Colour	mg/1 Pt/Co scale	20
Turbidity	Formazin Turbidity Units FTU	4
Odour	Dilution number	3 at 25°C
Taste	Dilution number	3 at 25°C
Temperature	°C	25
Hydrogen ion (pH)	pH value	9.5 max. 5.5 min.
Sulphate	mg SO_4/l	250
Magnesium	mg Mg/l	50
Sodium	mg Na/l	150
Potassium	mg K/l	12
Dry residues	mg/l	1500
Nitrate	mg NO_3/l	50
Nitrite	mg NO_2/l	0.1
Ammonium	mg NH_4/l	0.5
Kjeldahl nitrogen	mg N/l	1
Oxidizability	mg O_2/l	5
Dissolved or emulsified hydrocarbons	ug/l	10
Phenols	ug C_6H_5OH/l	0.5
Surfactants	ug/1 as lauryl sulphate	200
Aluminium	ug Al/l	200
Iron	ug Fe/l	200
Manganese	ug Mn/l	50
Copper	ug Cu/l	3000
Zinc	ug Zn/l	5000
Phosphorus	ug P/l	2200
Fluoride	ug F/l	1500
Silver	ug Ag/l	10
Arsenic	ug As/l	50
Cadmium	ug Cd/l	5
Cyanide	ug CN/l	50
Chromium	ug Cr/l	50
Mercury	ug Hg/l	1
Nickel	ug Ni/l	50
Lead	ug Pb/l	50
Antimony	ug Sb/l	10
Selenium	ug Se/l	10
Individual pesticides	ug/l	0.1
Total pesticides	ug/l	0.5
Polycyclic aromatic hydrocarbons	ug/l	0.2
Total coliforms	number/100 ml.	0
Faecal coliforms	number/100 ml.	0
Faecal streptococci	number/100 ml.	0
Sulphite-reducing clostridia	number/20 ml.	<1
Conductivity	uS/cm	1500 at 20°C
Chloride	mg Cl/l	400
Calcium	mg Ca/l	250
Substances extractable in chloroform	mg/l dry residues	1
Boron	ug B/l	2000
Barium	ug Ba/l	1000
Benzo 3,4 pyrene	ng/l	10
Tetrachloromethane	ug/l	3
Trichloroethane	ug/l	30
Tetrachloroethane	ug/l	10
Total hardness	mg Ca/l	60 minimum if supply is softened
Alkalinity	mg HCO3/l	30 minimum if supply is softened

APPENDIX 1

Prescribed concentrations or values specified in the UK Water Supply (Water Quality) Regulations 1989.

Tracer studies

Allan Rodhe
Department of Earth Sciences, Uppsala University

INTRODUCTION

The notes for the present lecture on tracer studies consist of the enclosed article:

Rodhe, A. and Saxena, R., 1990. *Tracer studies in Nordic till soils.* In: Haldorsen, S. (ed) «Hydrogeological Properties of Nordic Tills». Nordic Hydrological Programme, NHP Report, 25: 233-250.

and illustrations shown during the lecture.

TRACER STUDIES IN NORDIC TILL SOILS

Allan Rodhe and Rajinder Saxena
Division of Hydrology, University of Uppsala, V. Ågatan 24
S-752 20 Uppsala, Sweden.

ABSTRACT

Many questions concerning the flow of water and solute transport in soils, such as the velocity of water in different horizons, the fraction of water participating in the flow and transit times of water in a basin can be elucidated by the use of tracers. For studying water flow in till soils, tracers are particularly suitable since the inherent heterogeneity of such soils makes it difficult to asses relevant flow parameters for traditional hydrological flow estimates. A review of tracer studies in till soils in the Nordic countries shows that important general information has been obtained concerning velocities, flow-paths and the process of streamflow generation. The studies are, however, few and mostly descriptive. It is believed that much more quantitative information concerning the flow and dispersion in till soils can be obtained from well-planned artificial tracer experiments.

INTRODUCTION

Though moraine formations are most frequently encountered in Scandinavia yet little is known about the movement of soil water and groundwater in till soils. During the last few decades, interest has been steadily growing on the movement of solutes and pollutants as they are transported in the unsaturated and the saturated zones by soil moisture and groundwater. A clear knowledge of recharge and runoff processes is important to understand the general status of groundwater, its acidification, eventual pollution or other effects of human activity on water. Since chemical processes are involved in solute transport it is essential to know the residence time of water in different soil zones, as well as the origin of recharging water and the processes involved in the streamflow generation. The main problem settles down to a detailed study of the movement of soil water and groundwater.

The movement of soil water and groundwater can not be measured directly. Traditionally, the movement is calculated from head observations using empirically estimated flow parameters. These methods may give good results if the soil is comparatively homogeneous and isotropic with regard to hydraulic conductivity. But these conditions are seldom fulfilled in reality. Tracer techniques offer more direct methods to describe water movement in the unsaturated and saturated zones. The methods are suitable also for heterogeneous soils such as till soils.

Chemical agents and isotopes (radioactive and stable) have been used as tracers of water. In this paper, a summarized review of tracer studies, performed in the Nordic countries to investigate solute and water transport in till soils is presented.

TRACERS USED IN SOIL MOISTURE AND GROUNDWATER STUDIES

The tracers in hydrology can be divided into two groups, artificial or environmental, depending upon whether the tracer is added intentionally to the system or is added by natural processes (which in turn may be strongly affected by human activity). Artificial tracers are normally injected as a slug, and the subsequent areal or temporal variation of tracer concentration in water is interpreted in terms of water flow. Such studies may give direct information of particle velocities of water (speed and direction) and indirectly of flow rates, hydraulic conductivity, dispersion parameters and flow pattern (i.e., matrix/macropore flow). Studies using environmental isotopes may also highlight the above points, but often integrate the various processes involved in water flow through a system, giving information on fractions of event/pre-event water, transit times and reservoir volumes of water. They have also been helpful in the development of the structure of hydrochemical models.

An ideal tracer molecule should normally move with the same velocity as the ordinary water molecule. This implies that the tracer should not be retarded or adsorbed in the soil matrix. Other important considerations in the choice of tracers are: the detectibility at low concentrations (preferably in the field), little influence on the physical properties of water, no health hazards and low cost. Radioactive isotopes which are not adsorbed by the soil are more advantageous since they are detectable in very low concentrations, but health hazard and cost may restrict their use. Flourescent tracers (Rhodamine and flourescein) are also detectable in low concentrations, but their disposition for adsorption in the soil matrix limits their use in soil water and groundwater studies. In this regard inorganic compounds have a slight edge over organic compounds. Tracers so far used in hydrological studies in Nordic till soils are listed in table (1).

Table 1- Tracers used in hydrological studies in Nordic till soils.

Artificial	Organic	Rhodamine-B
	Inorganic	Cl, I, Br, Li and NH_4SCN
	Radioactive	HTO, ^{51}Cr-EDTA and $^{60}Co(CN)_6$
Environmental	Stable isotopes	D and ^{18}O
	Radioactive	HTO
	Water chemistry	Major ions and electrical conductivity

Isotopes as tracers

Isotopic tracers could be either radioactive or stable. The radioactive hydrogen isotope ^3H (tritium, T) is a perfect tracer for many hydrological purposes. Tritium shows almost no adsorption in most soils and a weak solution of HTO in water has almost the same density and viscosity as H_2O. It can be easily detected in low concentrations, but the disadvantage is that samples from the field must be taken to the laboratory for analysis. Due to thermonuclear explosions in the atmosphere during 1952-1963, bomb released ^3H has been extensively used to obtain soil moisture movement, groundwater recharge, its age and origin, and transit times of water in basins.

Apart from HTO, organo metallic compounds like ^{51}Cr-EDTA and $^{60}Co(CN)_6$ (both gamma emitters) have been used in Scandinavia in soil moisture and groundwater movement studies. In the complex form as given above, Cr and Co are generally little or not adsorbed in the soil matrix. Metallic organo compounds containing stable and most abundant isotopes of Cr and Co have also been used in Europe for groundwater flow measurements. For example Cr-EDTA or $Co(CN)_6$ can be injected in the field and the water samples are later irradiated with neutrons to make them radioactive (gamma emitters), allowing detection of very minute quantities of the metals accurately. This reduces the radiation hazard in the field, but the cost can be prohibitive.

Stable isotopes like ^{18}O and ^{2}H (deuterium or D) are some of the most useful and widely used environmental stable isotopes in the hydrological cycle. At high latitudes, the isotopic differences between average summer and winter precipitation is considerable. Such seasonal variations of these isotopes can be traced in the soil moisture and the movement of the depleted or enriched moisture layers can be monitored in time, giving a direct measure of water particle velocity, fluxes and transit time in the unsaturated zone. In a similar manner the variation in the isotope content of precipitation during single precipitation events has been used for component separation of stream discharge and estimation of transit times of water in a basin.

Injection of tracers

In case of laboratory studies, either in the form of columns or lysimeters or scale models of basins, the tracer can be applied in water solution simulating artificial rain and then displaced with ordinary water. For soil moisture studies, the tracer can be also injected below the soil surface in a uniform pattern or as a ring around a small borehole, specially when gamma emitters are used. In field studies of the unsaturated zone, the tracer generally used is tritium, which is applied below the root zone in a uniform pattern mostly in a plain with a minimum possible thickness to achieve a sharp boundary between the old and the fresh incoming water.

In groundwater movement studies, the tracer is normally injected in a borehole and attempt is made to make it well mixed in the injected section of the borehole. Care is taken to isolate the injected section with packers to avoid vertical mixing within the borehole and to limit the study to a particular horizon. In the point dilution technique (Halevy et al., 1967) the rate of tracer dilution with time in the borehole is monitored, enabling estimates of the groundwater flow rate. In other studies, the arrival of the tracer in downstream boreholes is observed to yield estimates of the velocity and the direction of flow.

In basin studies, the tracer could be spread or the ground or it could be injected at certain depths in the soil as point injections. Injection below the root zone may be suitable in order to reduce losses of the tracer due to evapotranspiration.

Detection of isotope tracers

Gamma emitters are measured by crystal scintillaters, beta emitters by liquid scintillation technique whereas the stable isotopes are measured by mass spectrometry. Direct measurement in field is possible in some studies using strong gamma emitters and strong beta emitters, but weak beta emitters like ^{3}H as well as stable isotopes have to be measured in the laboratory.

Sampling of groundwater, or water at or above the atmospheric pressure in general, is straightforward, but sampling in the unsaturated zones requires extraction procedures. In moist soils, soil water can be extracted in field by suction lysimeters. This has been reported in some Scandinavian studies. In other studies, soil samples are taken and the water extracted in the laboratory. In soil column experiments, the soil is sliced into thin plates while in field studies, soil samples are taken either by coring or by digging a pit. Moisture from the soil samples is generally extracted either by simple distillation or by vacuum distillation. Elsewhere in Europe centrifugation has also been advocated. When the stable isotopes ^{18}O and D are involved, special attention has to be given to a complete recovery of the moisture or else fractionation effects will cause a depletion of the extracted water relative to the original sample. The use of suction lysimeters or centrifuges assumes that all soil water at a certain point has the same isotope content, regardless of its binding to the soil particles and in the pores. This question needs further investigation.

When strong gamma emitters, such as ^{51}Cr-EDTA, are used in the unsaturated zone, then the measurements are done in situ with scintillation probes using access pipes, around which the tracer had been earlier injected in circular pattern. These access pipes are also used for soil moisture measurements by neutron meters.

OVERVIEW OF TRACER STUDIES IN NORDIC TILL SOILS

We have tried to find all the reports and publications dealing with tracer studies in till soils in the Nordic countries. Rather few studies were found in the literature. Their main results and conclusions are presented below.

Soil column experiments

Laboratory soil columns have been extremely useful in several aspects of tracer techniques. Foremost, it has been possible to check the physical and chemical behaviours of tracers, i.e adsorption, diffusion, dispersion, exchange with the soil matrix and retardation in a particular type of soil. Knutsson and Forsberg (1971), reported a comparative study on the movement of HTO and ^{51}Cr-EDTA in till soils in laboratory soil columns. ^{51}Cr was obviously delayed and somewhat adsorbed in the sandy till from Krysselboda but only slightly affected in the sandy till from Knappsmåla (both sites are in south-eastern Sweden). HTO did not seem to be adsorbed in the monitored soils. The behaviour of ^{51}Cr and HTO in the above till soils was similar to that found in column studies with illite and illite bearing shale (Knutsson and Forsberg, 1967). Only in montmorillonite ^{51}Cr moved faster than HTO. The retardation of HTO was attributed to isotopic exchange with the inter-layer water.

Vertical flow in the unsaturated zone

One early study on soil moisture movement in till soils was conducted in Sweden by Tamm and Troedsson (1957). They applied a weak solution of thiocyanate (NH_4SCN) on the soil surface. Soil profiles were later dug up and sprayed with a strong solution of $FeCl_3$. All parts of the profile where the tracer had penetrated turned deep red. By this method they could monitor the penetration of the tracer front. The shape of the tracer front was irregular, implying preferential flow-paths of water. No attempts were made to determine the velocity of tracer movement.

In an investigation in the upper reaches of the Suså river catchment in Denmark, Sevel et al. (1981) used bomb tritium to estimate infiltration, groundwater recharge and age of groundwater in the greensand limestone acquifer overlain partly by clayey till and sandy-clay till. They estimated that the vertical velocity ranged between 0.6 to 0.9 m/y, the annual percolation was about 100 to 150 mm and the age of groundwater was around 25 years. Two models for infiltration were developed, a piston flow with dispersion and a box model. The agreement between simulated and observed tritium profiles was moderately good. However, the simulated ^{3}H concentrations were always higher than the observed. This was attributed to the uncertainty in the measurements of ^{3}H inputs.

Seasonal variations of ^{18}O in precipitation have been traced in soil moisture, with the aim to determine rate of soil moisture movement and groundwater recharge in till soils in southern Sweden, Saxena (1987). The ^{18}O depleted soil moisture contributed by snowmelt and the enriched moisture derived from the summer rains were detected in the unsaturated zone. In the till soils at Kloten the mean rate of soil moisture movement from March to September 1983 was found to be about 5.8 mm/d. The annual recharge at this site was estimated to be about 270 mm between summers 1982 and 1983. Another moraine formation was studied on a small ridge at Örbyhus, 42 km north of Uppsala, in October 1984. The isotopic profiles at this site were not well defined, but estimated annual recharge between 1983 and 1984 was about 260 mm. At some sites with glaciofluvial deposits in Uppsala, where water table was below 4 m depth, infiltrated meltwater from two consecutive years snowmelt periods could be traced in the soil profile giving estimate of eventual annual recharge. The annual recharge in Uppsala varied between 200 to 300 mm, which agreed fairly well with estimates done by water balance studies.

In a detailed study in agricultural fields in Denmark, Butts et al. (1988), used ^{60}Co and ^{51}Cr as tracers in the unsaturated zone. The radio tracer labelled solution was irrigated round a borehole and its movement monitored by gamma logging down the borehole. The field studies were carried out in the agricultural fields at Taastrup (till) and Jyndevad (alluvial outwash). Simultaneous application of HTO and ^{60}Co(CN)$_6$ were carried out to evaluate their relative behaviour in comparison to ^{51}Cr-EDTA. It was found that HTO and ^{60}Co moved almost in a similar manner, though HTO appeared to move a bit faster than ^{60}Co. Spatial variability in solute concentration was observed in the same area both by soil coring and gamma logging techniques. These variations were attributed to local differences in soil hydraulic properties, porosity and soil moisture content etc. The water flow and solute transport in the soil profiles in the above experiment was discussed by Jensen and Refsgaard (1988 a, b, c) who used the movement of $^{60}Co(CN)_6$ to model unsteady vertical flow and solute transport. The model simulations were quite successful. However for sandy loam soils at Taastrup Moraine, hysteresis and macropore flow required further modifications of the model parameters.

Johansson (1987) tried to estimate groundwater recharge in till soils in south eastern Sweden by Cl balance method. Estimating the highest deposition of Cl, the groundwater recharge was about 113 to 162 mm annually. Johansson, however, concluded that the uncertainties in the estimation of chloride deposition were large and that it was difficult to make accurate estimates.

Lateral flow

The credit for one of the earliest Nordic field measurements of groundwater flow in till soils using tracers goes to Tamm (1931). In a sloping terrain, holes were dug down to the groundwater table. NaCl was poured in the upstream hole and the arrival of Cl was monitored in the downstream holes. The groundwater velocities in the sloping hard packed deep moraines ranged from 0.01 to 0.03 m/d.

Knutsson (1971) has given a report on groundwater in moraines in south eastern Sweden. Groundwater velocities were determined using HTO, ^{51}Cr-EDTA and Rhodamine B. The field experiments were carried out in a series of filter tubes and the tracers were monitored in groundwater samples and from the ground surface directly. The average groundwater velocity observed was about 0.6 m/d. The flow in the relatively more compact till was slower, 0.15 to 0.25 m/d. The recovery curve of tritium indicated strong dispersion effects. The point dilution technique was also attempted, but due to low lateral velocities at Knappsmåla, which were perhaps comparable to molecular diffusion, this method could not provide information about groundwater flow rates. In Krysselboda, using surface monitoring of the tracer, the flow pattern and the velocity of flow between two small pools at different levels were determined using HTO and ^{51}Cr-EDTA. The average velocity of flow varied between 0.2 to 0.5 m/d. In one of the experiments the groundwater could be traced as a winding band with a maximum velocity of 0.9 m/d and an average velocity of 0.45 m/d. From these investigations Knutsson concluded that groundwater in the drumlin-terrains flows faster than in basal tills.

Lundin (1982) used similar methodology in several experiments to investigate groundwater movement in till soils in Kloten in southern Sweden. NaCl and HTO were added to cavities and boreholes upstream and the arrival of the tracers was monitored in the downstream boreholes. In other experiments porous cups were used to collect the water samples. Percolation studies were also conducted in the same area by applying HTO over a horizontal plain with subsequent monitoring of the tracer in horizontal and vertical directions. Generally the average velocity of groundwater decreased drastically with depth. The experiments with HTO brought out three salient features of water flow:

* Vertical velocity in the unsaturated zone varied between 0.01 to 0.1 m/d. during initial stages of percolation at different sites in the same area.
* Lateral flow was also encountered in the unsaturated zone.
* At near surface levels, which were saturated only temporarily during high water levels, intermittent lateral displacement of the tracer was observed.
* High velocity conduit flow developed in saturated conditions

Möller et al. (1986) injected KBr and NaI, and at a later stage LiCl, to study the relative movement of water (as given by the tracers) and CO_3 in a $CaCO_3$ treated till slope in southern Sweden. In the initial phase, NaI and KBr were applied on the soil surface about 10 to 40 m upstream from three wells at different locations. Iodide could not be detected in the groundwater, but bromide appeared in two of the wells. In one of these, the tracer arrived in two peaks, corresponding to lateral velocities of a few m/d and 0.12 m/d respectively. The first peak was attributed to shallow groundwater flow following snowmelt and the second was assumed to represent the flow through deeper groundwater. The tracer arrival in the other well corresponded to a mean velocity around 0.08 m/d. In a later experiment KI, KBr and LiCl were applied on the surface and injected in boreholes at shallow (0.4-1.0 m) and deep (1.0-1.6 m) depths. None of the tracers applied at the surface or at the shallow depths appeared in the observation wells within the period investigated so far. Li injected at the deeper depths, on the other hand, arrived quickly in a well. It seemed that the lateral velocity at the border between rock and till at 1.5 - 2.5 m depth was large, whereas the vertical velocity through the till soil was small.

Particle velocities of water observed in the reviewed tracer experiments were about 0.002 to 0.01 m/d for the vertical flow in the unsaturated zone (mean values over several months), and the lateral velocities in the saturated zone ranged from 0.1 to 30 m/d (Table 2). Apart from the velocities, the results in the studies have mainly been interpreted in a qualitative way. For instance, no values of dispersion parameters for groundwater flow have been reported from the field experiments.

Table 2 - Vertical velocity of soil water and lateral velocity of groundwater (v_p), groundwater recharge (R), and tracers used in field studies conducted in till soils in Nordic countries.

	Tracer	v_p (m/d)	R (mm/y)	References
Unsaturated	^{3}H	0.0016	100	Sevel et al. (1981)
	^{3}H	0.0025	150	"
	^{18}O	0.0058	266	Saxena (1987)
Saturated zone	Cl	10-30		Tamm (1931)
	^{3}H, ^{51}Cr	0.15-0.9		Knutsson (1971)
	^{3}H, Cl	0.1-29		Lundin (1982)
	Br, I, Li	0.1		Möller et al. (1986)

Basin studies

Basin studies with isotopes or water chemistry have played an important role in the development of the modern concepts of streamflow generation. In such studies, the tracer concentration of the runoff from a basin is interpreted in terms of contributions to the runoff by waters of different transit times in the basin (expressed as transit time distribution, mean residence time or fractions of event/pre-event water) or by waters of different origin (fresh rain/meltwater, soil water, shallow or deep groundwater) and indirectly information can be obtained on the flow-paths of water. The studies show the integrated result of the processes involved in streamflow generation, but little information is obtained on the various processes acting along the flow-paths of water through a basin.

The use of water chemistry as a tracer in runoff studies often relies on the non-conservative property of water chemistry during water flow through a basin. The streamwater is considered to be a mixture of water that has undergone various chemical changes, depending on the flow-paths to the stream. The use of environmental isotopes of the atoms of the water molecules, on the other hand, is based on the conservative behaviour of the isotope content of water during the process of runoff. In this case, hydrograph separation and estimates of transit times are possible due to variations in the isotopic composition of precipitation and the subsequent damping of this variation during water flow through various reservoirs within a basin.

Event studies

A method for quantitative estimates of the various flow components in streams was derived by Ivanov (1948) from continuity conditions for water and solutes. The method was developed by, among others, Pinder and Jones (1969) using water chemistry and Dincer et al. (1970) using isotopes. The now widely used formula for two-component hydrograph separation by water chemistry or isotopes gives the fraction of groundwater or pre-event water in streamflow, X, as

$$X = \frac{c_s - c_p}{c_g - c_p} \qquad (1)$$

where c_s c_p and c_g are the tracer concentrations of streamwater, precipitated water (fresh rain or meltwater) and groundwater (or pre-event water) respectively. With this simplistic model, the water reaching the stream is assumed to originate from two reservoirs, each of which to some degree being homogeneous with respect to the tracer. The tracer concentration of streamwater is measured directly throughout the event, but the application of the model is restricted by difficulties in assessing realistic values of c_p and, particularly, c_g (see Rodhe 1987 for a detailed discussion). In most applications, c_g is given a constant value during the event, equal to the pre-event value of c_g.

The meaning of X, as well as of c_p and c_g, depends on the tracer used. Eq(1) gives the fraction of pre-event water in the stream if the following conditions are met:

* The water in the basin prior to the event is homogeneous with respect to the tracer and this tracer concentration is used as a constant value of c_g during the event.
* The tracer is conservative during water flow through the basin.
* The tracer concentration of the rain or meltwater is constant over the basin and with time during the event.

Areal and temporal variations of c_p may be treated as uncertainties in the input value and are not treated here, i.e., it is assumed that the third condition is fulfilled.

When using isotopes, the above conditions may be more or less fulfilled. The pre-event fraction is an underestimate of the groundwater fraction of streamflow, since some of the discharged fresh rainwater particles may have been discharged as groundwater. If the tracer concentration of the discharging groundwater remains constant, or if its time development during the event can be measured or calculated, then Eq(1) gives the groundwater fraction of streamflow.

When water chemistry is used as a tracer, condition no 2 is usually not fulfilled. Since the infiltrated water changes chemically to groundwater, Eq(1) gives the groundwater fraction of streamflow rather than the pre-event fraction, even if some new water particles are discharged. Further c_p should be the tracer concentration of overland-flowing water, which probably differs from that of rain or meltwater due to solute uptake. Neglecting this would overestimate the groundwater fraction of streamflow.

Hydrograph separation using ^{18}O was performed in several Swedish till basins by Rodhe (1981, 1987) and in the Teeressuonoja basin in southern Finland by Bengtsson et al. (1988). A study of ^{18}O during a snowmelt runoff

464

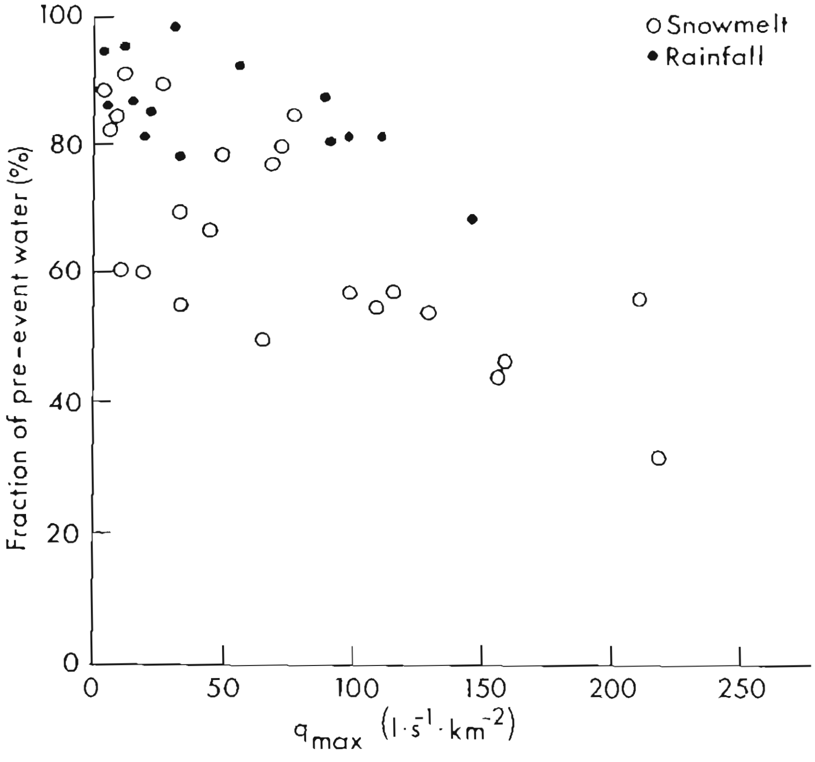


Figure 1. Total volume fraction of pre-event water, as estimated by oxygen-18, versus maximum flow. Snowmelt and rainfall events in 10 Swedish basins. (From Rodhe 1987).

event in a Central Swedish basin was also reported by Jacks et al. (1986), although no quantitative hydrograph separation was performed.

The studies by Rodhe (rainfall and snowmelt events) and Bengtsson et al. (snowmelt events) show that runoff events in till basins often are dominated by pre-event water, see the summarized results from Rodhe (1987) in Figure 1. The calculations underlying the figure are based on the use of a constant value of c_g, equal to the pre-event tracer concentration in the streamwater. According to the discussion above, the pre-event fractions are thus conservative estimates of the groundwater fractions, which are larger than the fractions shown in the figure. Similar results, stressing the dominant role played by groundwater in streamflow generation, has been obtained in a large number of isotope studies in Europe and North America. In contrast to the results of most other isotope studies, Jacks et al: (1986) found that the streamflow during a spring snowmelt was almost entirely new meltwater. A suggested explanation for the dominance of meltwater was the contribution of saturation overland flow from bogs, which covered about half of the basin.

Long time input-output of tracer

Transit times of water in basins, and associated reservoir volumes, have been estimated from comparisons between the temporal variation of the isotope content of precipitation input and streamwater output.

Eriksson (1974) estimated a transit time distribution for the river Råne älv in northern Sweden from the modification of the tritium signal from precipitation to streamflow. Median transit time around 2 years was obtained for the 3800 km^2 drainage basin (which was dominated by till soils, with considerable areas of mires and 3.8 % lake area).

Rodhe (1987) used a simple one-box mixing model to simulate the ^{18}O content of streamwater from that of precipitation on a daily basis. Reservoir volumes around 200-300 mm, corresponding to mean transit times varying from 0.5 to 1.5 year, was obtained for the 7 basins investigated. The model was, however, considered very rough and the results were only indicative. In a more refined approach, Lindström and Rodhe (1986) modelled the ^{18}O-flow through two of the basins investigated by Rodhe (Buskbäcken and Gårdsjön F1 in southern Sweden) by the HBV-PULSE model. With this model, which reproduced the observed ^{18}O-values in the stream considerably better than the simple one box model, reservoir volumes around 480 and 440 mm were obtained for Buskbäcken and Gårdsjön F1 respectively.

Christophersen et al. (1984) attempted to simulate the ^{18}O-flow through the Birkenes basin in southern Norway by the Birkenes model. The reproduction of the -values of streamwater was not very successful, but the ^{18}O -simulations stressed the need for a considerable mixing volume in this hydrochemical model.

Information about groundwater flow-paths has been obtained from water chemistry and environmental isotopes in several studies. Calles (1985) interpreted the gradual increase in electrical conductivity downstream a first order stream, in a till basin in central Sweden, as the result of an increasing relative contribution of deep groundwater with long transit time in the ground. The temporal variation in electrical conductivity at a fixed location in the stream further indicated that the outflow of deep groundwater was dynamic, increasing during periods of high flow in the stream. Espeby (in press) observed major ions and ^{18}O of water sampled at springs, seepage faces and the stream during a spring snowmelt period in a hillslope in southern Sweden. He concluded that preferential flow through macropores played an important role for the saturated flow during periods of high groundwater level, and also for unsaturated flow as long as the soil remained frozen. From Cl-concentrations of soil water and streamwater, observed in connection with a rainstorm exceptionally rich in seasalts, Mulder et al. (in press) concluded that the Birkenes stream in southern Norway to a large degree was fed by translatory lateral flow in the surface near organic layers.

Basin-wide tracer experiments

Nyström (1985) injected tritiated water all over two small till basins with thin soils in the lake Gårdsjön area in south-western Sweden. The injections were made in a 25 m distance grid pattern at 0.4 m depth in the ground. From the observed tritium concentrations in the streams it was concluded that about 40% of the water below the injection depths was replaced in one year. Due to methodological problems, the longer transit times could not be assessed. Reservoir volumes below the injection depth were estimated to 140 and 190 mm in the two basins respectively. Adding the measured water content in the upper 0.4 m, reservoir volumes of 260 and 310 mm were obtained.

In order to investigate the mobility of sodium in the soil, Christophersen et al. (1982) applied a pulse of radioactive ^{24}Na and HTO to the on-going irrigation (about 200 mm water per week) in a 98 m^2 mini-basin in southern Norway. The basin was 61 % barren granitic bedrock with the remaining area covered by till soils with an average depth of about 0.4 m. Comparisons between the relative concentrations of HTO and ^{24}Na showed that ^{24}Na was fairly mobile, but was delayed as compared to HTO. After 6 days, 84% of the applied HTO and about 45% of the ^{24}Na were recovered at the outlet.

CONCLUSIONS

Tracer studies can highlight important problems at different scales, from a detailed description of the flow regime in a specific soil profile to a general view of the water flow through a basin. They are well suited for investigating problems about which we have limited knowledge today and where conventional methods are insufficient, such as the significance of preferential flowpaths during unsaturated and saturated flow and questions concerning transit times and flow-paths of water through a basin. The studies performed in Nordic till soils have given some basic information on particle velocities of water in the unsaturated and saturated zones and on the process of streamflow generation. The reported studies are, however, very few and more studies are needed if we are to understand the nature of the water flow in till soils.

Theoretically, many applications of artificial tracer technique are very simple: the tracer is added at one site and the movement of water is determined from observations of the tracer concentration around the injection area. The practical problems may, however, be large. For instance, the tracer was only partially found in many studies and some times it was not found at all. In order to give quantitative and new information on the nature of water flow in the ground, artificial tracer studies have to be extremely well planned and supported by conventional hydrological measurements.

REFERENCES

Bengtsson, L., Lepistö, A., Saxena, R., and Seuna, P. (1989) Mixing of acid meltwater with groundwater in a forested basin in Finland. Atmospheric deposition, Proc. Symp. Baltimore, May 1989, IAHS Publ. No 179, 251-258.

Butts, M.B., Genders, S. and Sevel, T. (1988) Radioisotope tracer studies of solute transport in the unsaturated zone: field measurement and analysis. Danish Isotope Centre, Copenhagen. 58 pp + appendix.

Calles, U.M. (1985) Deep groundwater contribution to a small stream. *Nord. Hydrol., 16*, 45-54.

Christophersen, N., Seip, H.M., Gvenlid, C. and Tollan, O. (1982) Sodium mobility in a mini-catchment studied with radioactive tracers and artificial precipitation. *Nord. Hydrol., 13*, 105-114.

Christophersen, N., Kjernsröd, S. and Rodhe A. (1984) Preliminary evaluation of flow patterns in the Birkenes catchment using natural ^{18}O as a tracer. In: Johansson, I. (ed.): Hydrological and hydrogeochemical mechanisms and model approaches to the acidification of ecological systems. Rep. 10, Nord. Hydrol. Progr., Oslo, 29-40.

Dincer, T., Payne, B.R., Florkowski, T., Martinec, J., and Tongiorgi, E. (1970) Snowmelt runoff from measurements of tritium and oxygen-18. *Water Resour. Res., 6*, 110-124.

Espeby, B. (in press) Tracing the origin of natural waters in a glacial till slope during snowmelt., *J. Hydrol.*, accepted for publication Sept 1989.

Eriksson, E. (1974) Vattnet kemikaliebäraren. Forskning och framsteg (Stockholm), 5, 41-45.

Genders, S., Sevel, T. and Butts, M.B. (1988) Investigation of solute movement in agricultural fields using isotope techniques. Proc. Nordic Hydrological Conference, Rovaniemi, Aug. 1988, 169-179.

Ivanov, A.T., (1948) Determination of the subsurface component of streamflow by use of the hydrochemical method. Transactions of the Laboratory of Hydrogeological Problems of Acad. Sci. USSR, 3, 243-246.

Halevy, E, Moser, H., Zellhofer, O. and Zuber, A. (1967) Borehole dilution techniques: a critical review. Isotopes in Hydrology, IAEA, Vienna, 531-564.

Jacks, G., Olofsson, E. and Werme, G. (1986) An acid surge in a well-buffered stream. *Ambio, 15*, 282-285.

Jensen, K.,H. and Refsgaard, J.C. (1988) Spatial variability of physical parameters and processes in field soils. Part I: Water flow and solute transport in soil profiles. Proc. Nordic Hydrological Conference, Rovaniemi, Aug. 1988, Rovaniemi, 121-138.

Jensen, K.,H. and Refsgaard, J.C. (1988) Spatial variability of physical parameters and processes in field soils. Part II: Water flow at field scale. Proc. Nordic Hydrological Conference, Rovaniemi, Aug. 1988, 139-154.

Jensen, K.,H. and Refsgaard, J.C. (1988) Spatial variability of physical parameters and processes in field soils. Part III: Solute transport at field scale. Proc. Nordic Hydrological Conference, Rovaniemi, Aug. 1988, 155-168.

Johansson, P.-O. (1987) Methods for estimation of direct natural groundwater recharge in humid climates. Meddelande Trita-Kut 1045, KTH, Stockholm.

Knutsson, G. (1971) Studies of groundwater flow in till soils. Geologiska Föreningens i Stockholm Förhandlingar, 93, 1-22.

Knutsson, G., and Forsberg, H.G., (1967) Laboratory evaluation of ^{51}Cr-EDTA as a tracer for groundwater flow. Isotopes in Hydrology, IAEA, Vienna.

Lindström, G. and Rodhe, A. (1986) Modelling water exchange and transit times in till basins using oxygen-18. *Nord. Hydrol., 17*, 325-334.

Lundin, L. (1982) Mark- och grundvatten i moränmark och marktypens betydelse för avrinningen. UNGI Report 56, Uppsala Univ., Dept Phys. Geogr., 216 pp.

Mulder, J., Christophersen, N., Hauhs, M., Vogt, R.D., Andersen, S., Andersen, D.O. (In press) Hydrochemical controls in the Birkenes catchment as inferred from a rainstorm high in seasalts. *Water Resour. Res.*, In press November 1989.

Möller, Å., von Brömssen, U., Sundlöf, B., Jacks, G., Norrström, A.-C., Sverdrup H., Warfvinge, P. (1986) Markkalkningsförsök vid grävda brunnar i moränmark. Delrapport 1. Statens Naturvårdsverk, Stockholm, Rapport 3100. 50 pp + appendix.

Nyström, U. (1985) Transit time distributions in two small forested catchments. *Ecol. Bull. (Stockholm)* . *37*, 97-100.

Pinder, G.F. and Jones, J.F. (1969) Determination of the groundwater component of peak discharge from the chemistry of total runoff. *Water Resour. Res.*, 5, 438-445.

Rodhe, A. (1981) Spring flood - meltwater or groundwater? *Nord. Hydrol.*, *12*, 21-30.

Rodhe, A. (1987) The origin of streamwater traced by oxygen-18. Uppsala Univ., Dept Phys. Geogr., Div. Hydrol., Report Series A 41, 260 pp + Appendix 73 pp.

Saxena, R.K. (1987) Oxygen-18 fractionation in nature and estimation of groundwater recharge. Uppsala Univ., Dept Phys. Geogr., Div. Hydrol., Report Series A 40, 152 pp.

Sevel, T., Kelstrup, N. and Binzer, K. (1981) Nedsivning. Rapport No Suså-H 6. Dansk komite for hydrologi, Copenhagen. 59 pp.

Tamm, O. (1931) Studier av jordmånstyper och deras förhållande till markens hydrologi i nordsvenska skogsterränger. Medd. fr. Statens skogsförsöksanstalt 26, Stockholm. 248-252.

Tamm, C.O. and Troedsson, T. (1957) A new method for the study of water movement in soil. *Geologiska föreningens i Stockholm Förhandlingar 79*, 581-587.

HYDROGEOLOGICAL PROPERTIES OF NORDIC TILLS

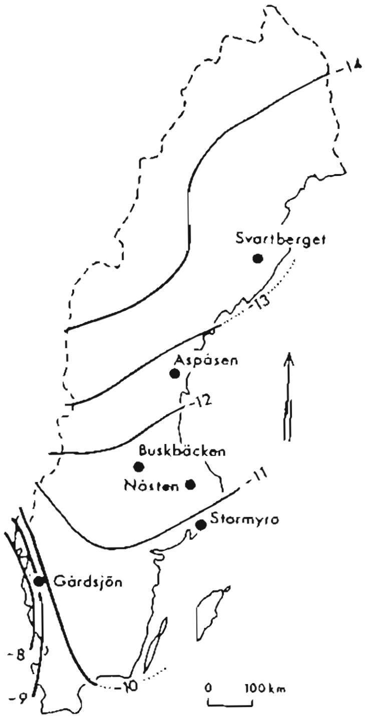
Edited by Sylvi Haldorsen

NORDIC HYDROLOGICAL PROGRAMME
NHP Report No 25

The Coordinating
Commitee for Hydrology
in Norden (KOHYNO)
1990

(

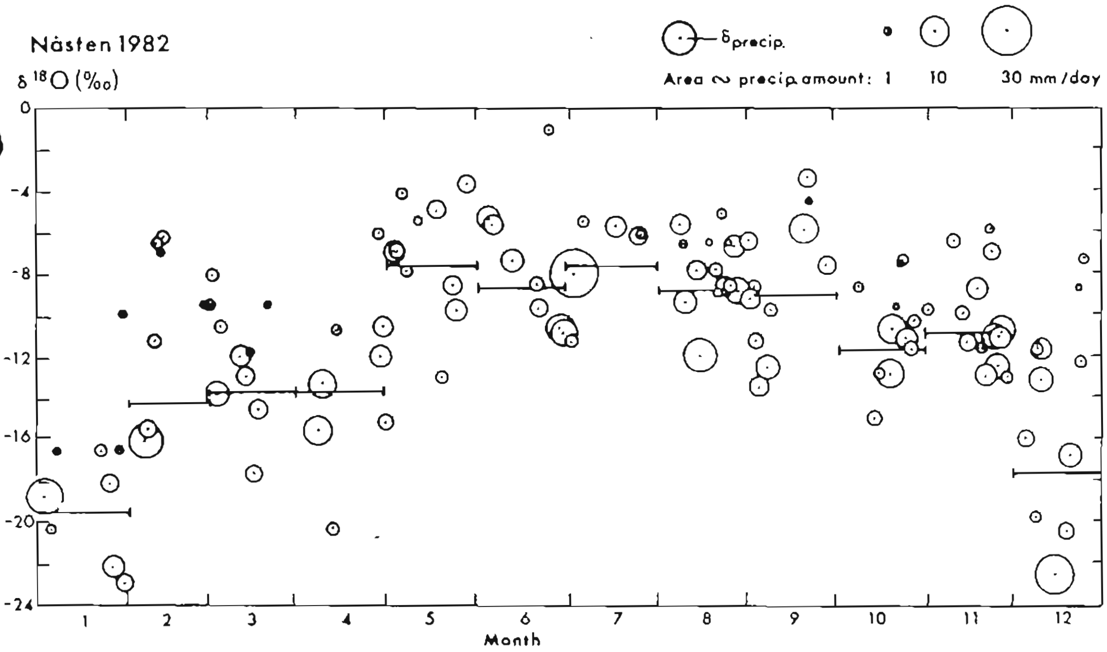
1.



Mean δ18O of precipitation (‰)

Rodhe, 1987

δ18O



MONTH

J J D

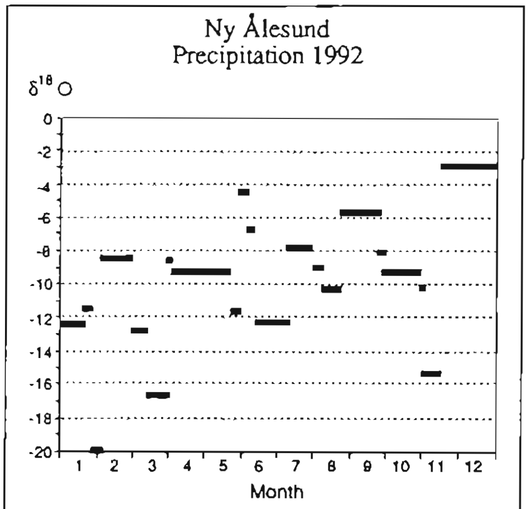


Fig. 4. Gjennomsnittlige månedsverdier for $\delta^{18}O$ i nedbør i Ny-Ålesund, Svalbard, 1992. Beliggenhet, se figur 2.
Monthly average values of $\delta^{18}O$ in precipitation in Ny-Ålesund, Svalbard, during 1992. For location, see Figure 2.

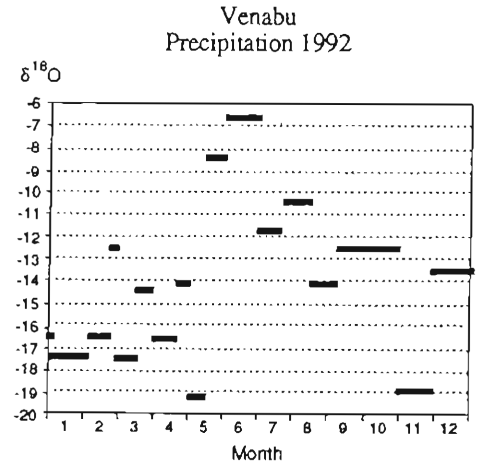


Fig. 5. Gjennomsnittlige månedsverdier for $\delta^{18}O$ i nedbør på Venabu, Hedmark, 1992. Beliggenhet, se figur 2.
Monthly average values of $\delta^{18}O$ in precipitation in Venabu, Hedmark, during 1992. For location, see Figure 2.

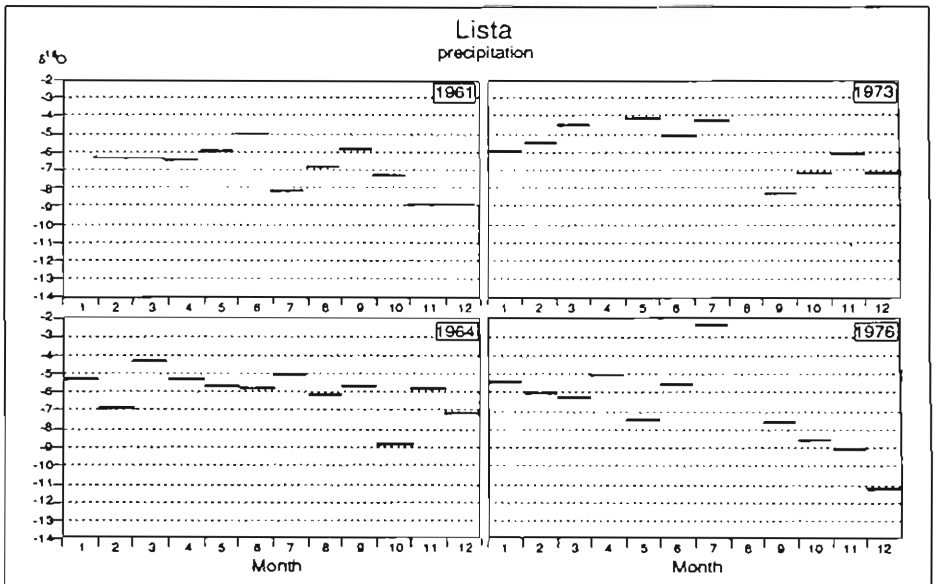


Fig. 3. Gjennomsnittlige månedsverdier for $\delta^{18}O$ i nedbør på Lista, Vest-Agder, 1961, 1964, 1973 og 1976. Data fra IAEA (International Atomic Energy Agency, Wien). Beliggenhet, se figur 2.
Monthly average values of $\delta^{18}O$ in precipitation in Lista, Vest-Agder during 1961, 1964, 1973 and 1976. Data from IAEA. For location, see Figure 2.

from Haldorsen, 1994

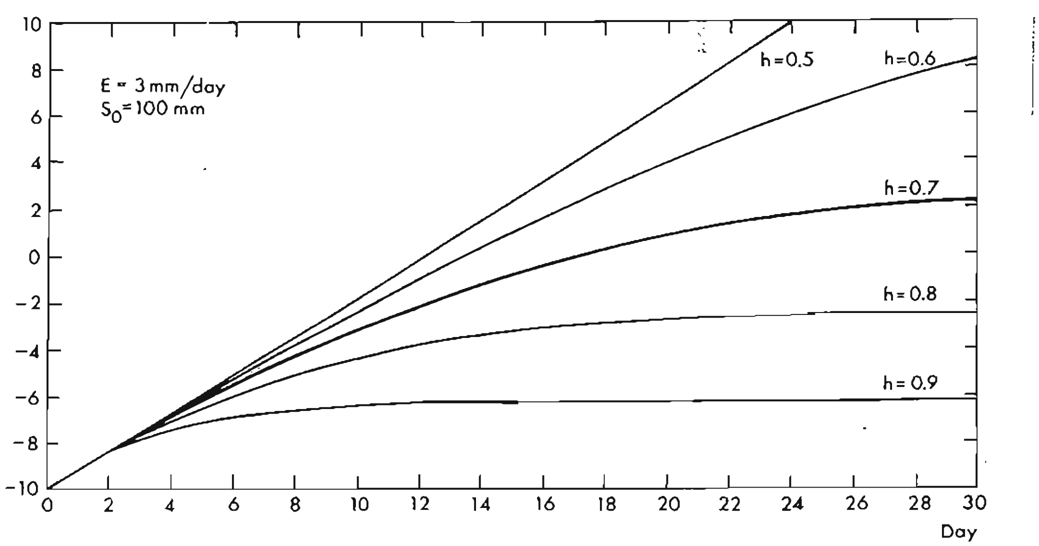
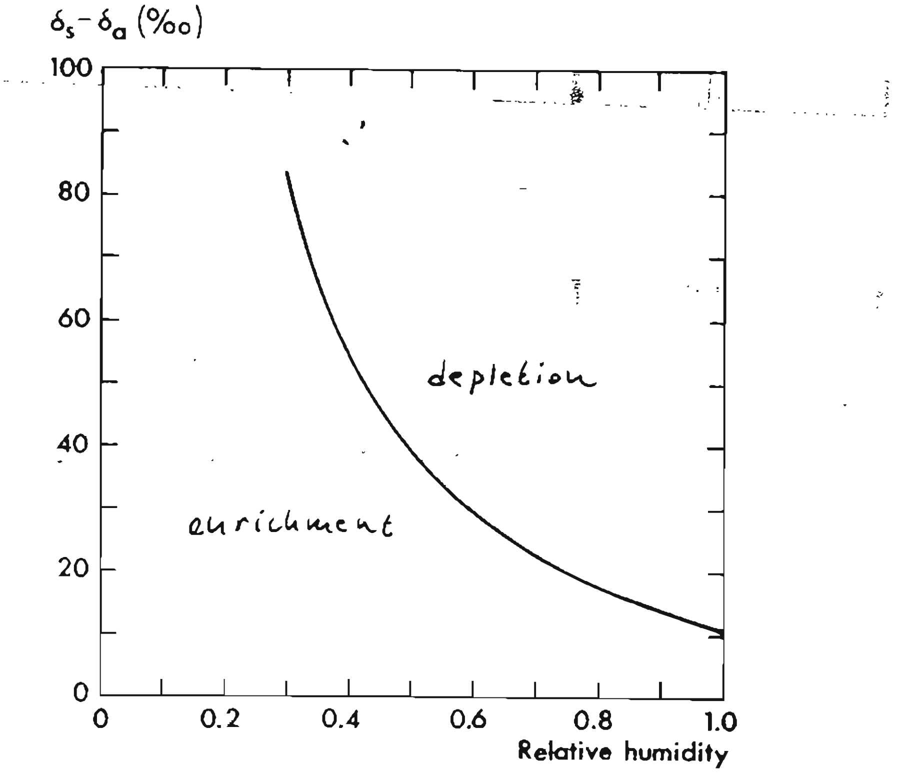


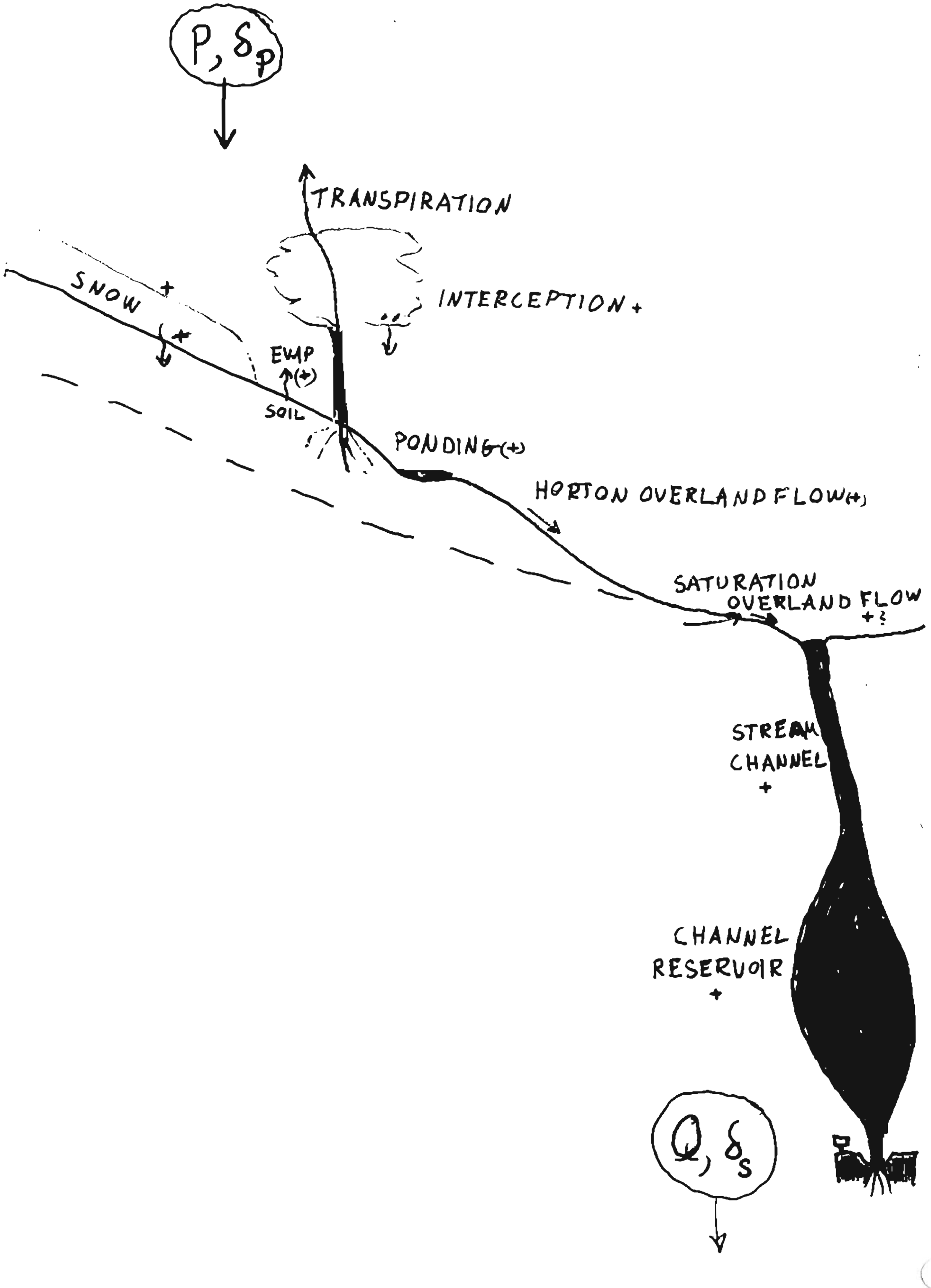
Fig.6.2 Simulation of the $\delta^{18}O$ of water under reference enrichment conditions (the relative humidity h = 0.7) and sensitivity to variations in the relative humidity.



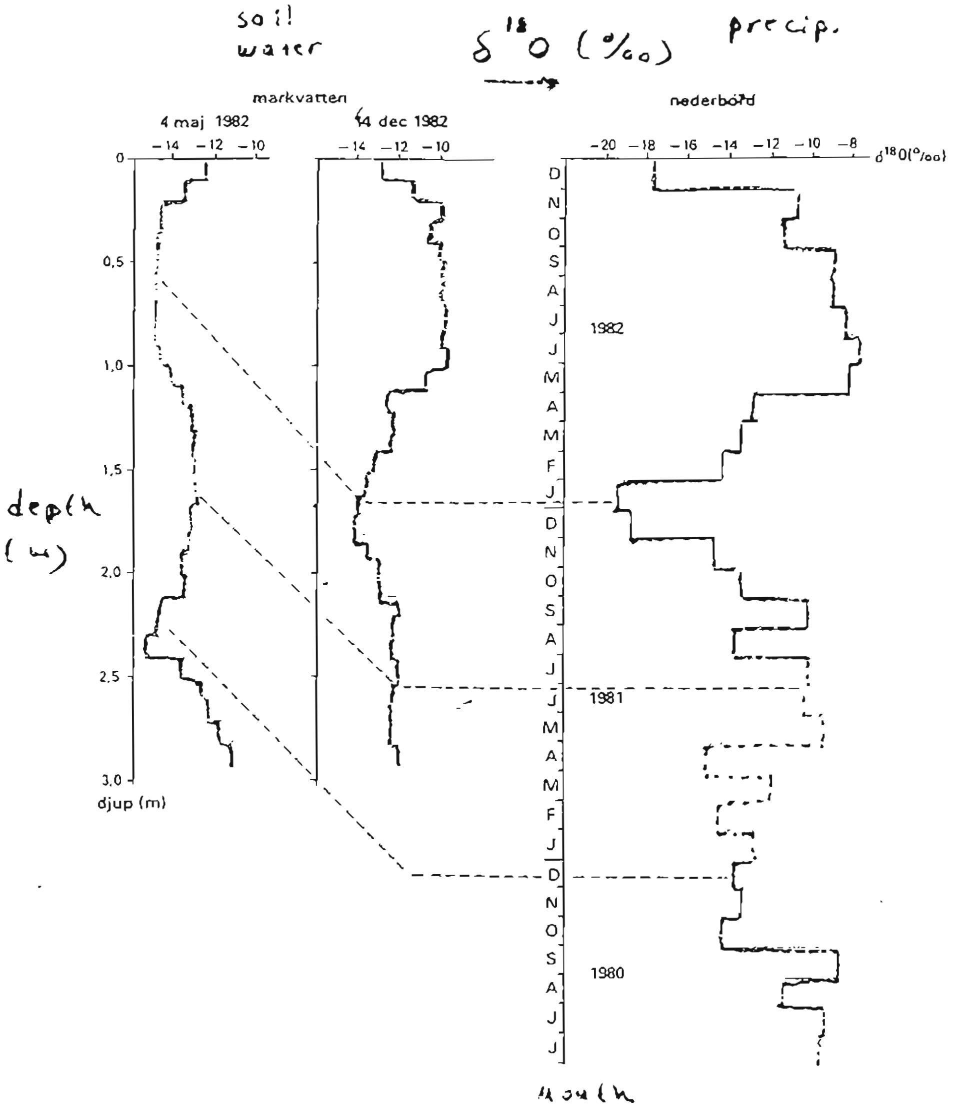
"Steady state value" for $\delta^{18}O$

Rodhe, 1987

^{18}O enrichment possibilities within a basin



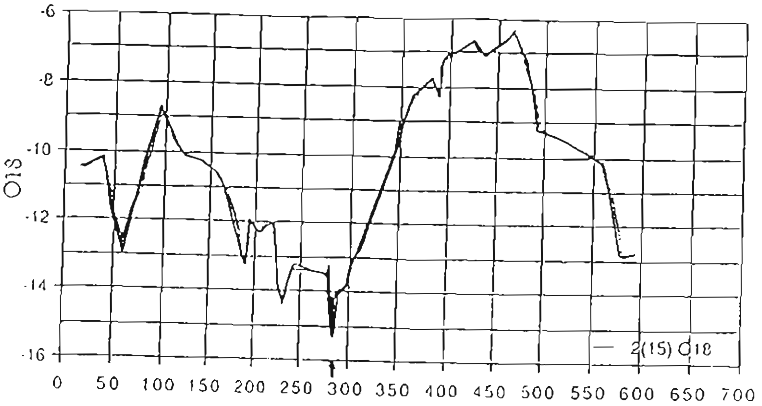
5

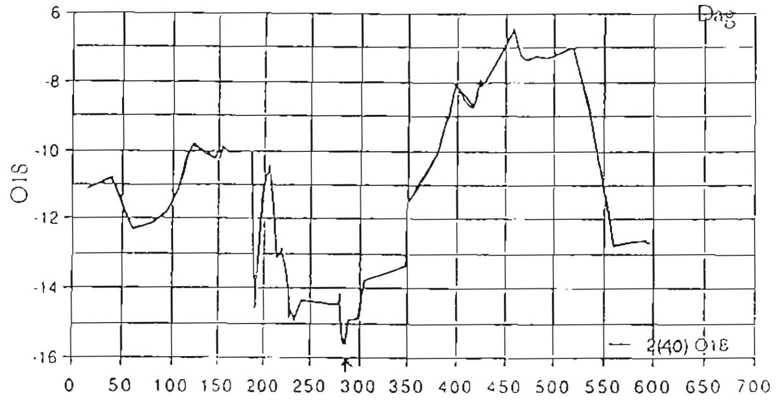


Uppsala 1982

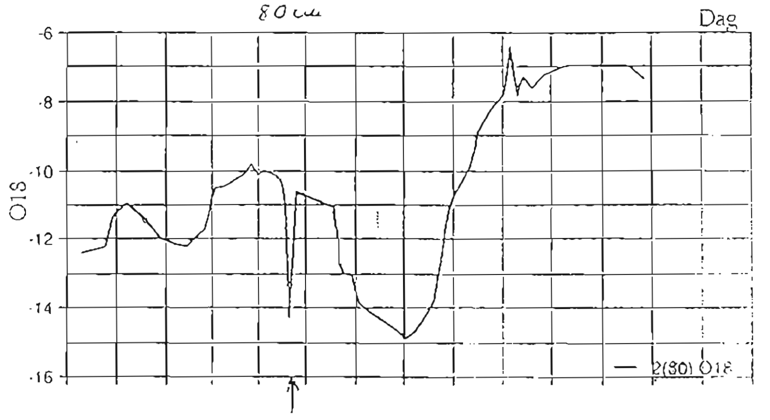
Saxena, 1987

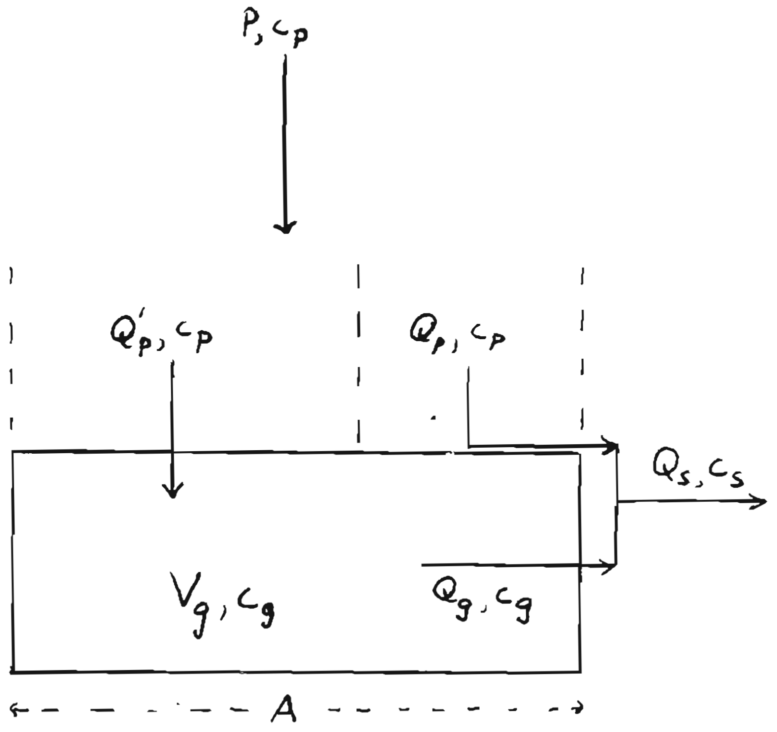
15cm lysimeter Stubbetorp





80 cm





Fraction of groundwater $X = \frac{c_s - c_p}{c_g - c_p}$

Fraction of discharge area of the basin $Y = \frac{Q_s}{A \cdot P}(1-x)$

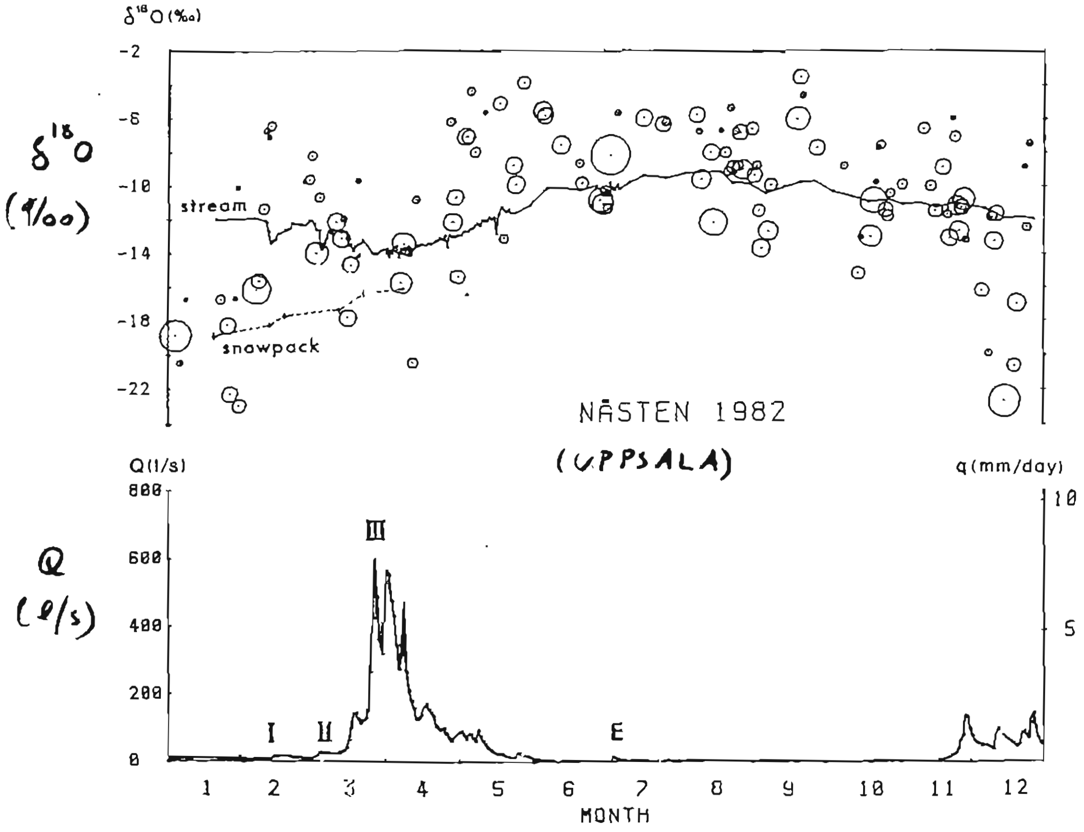
PROBLEM: c_g ?

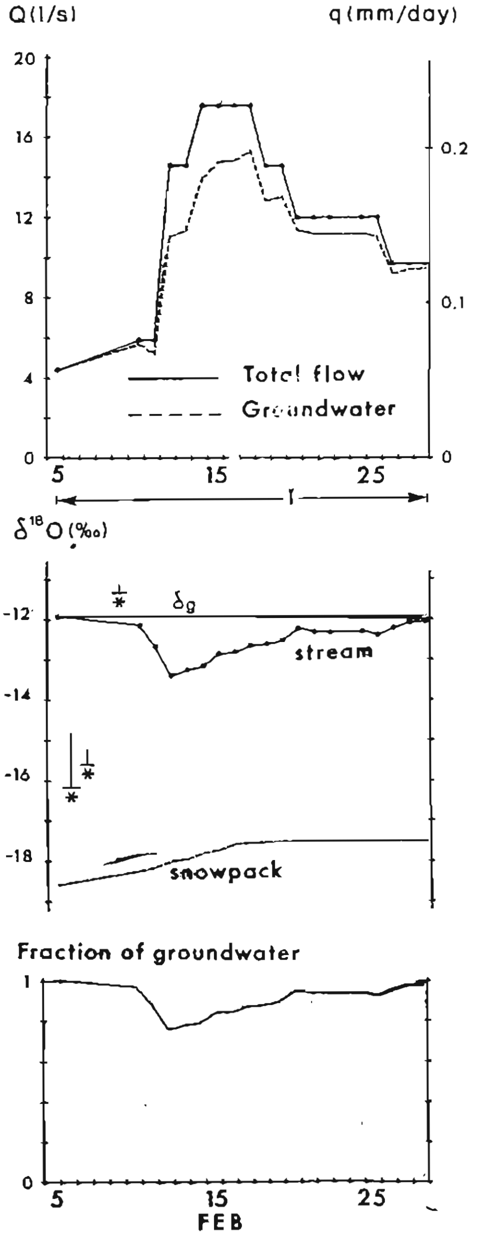
a) c_g constant $\Rightarrow$ ≈ pre-event water

b) c_g variable $\Rightarrow$ actual gr.w. fraction

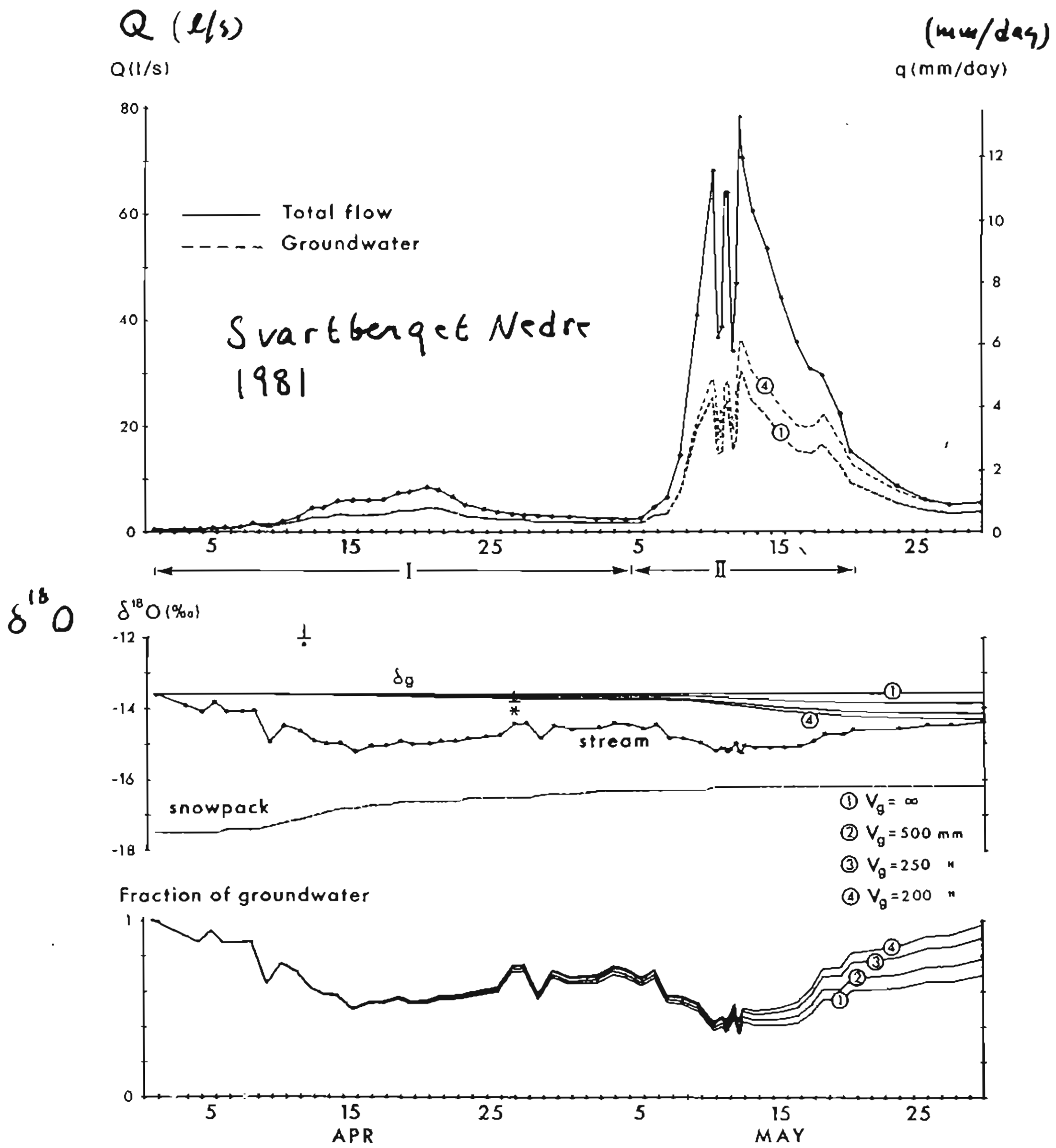
influence of pre-event soil water?

8

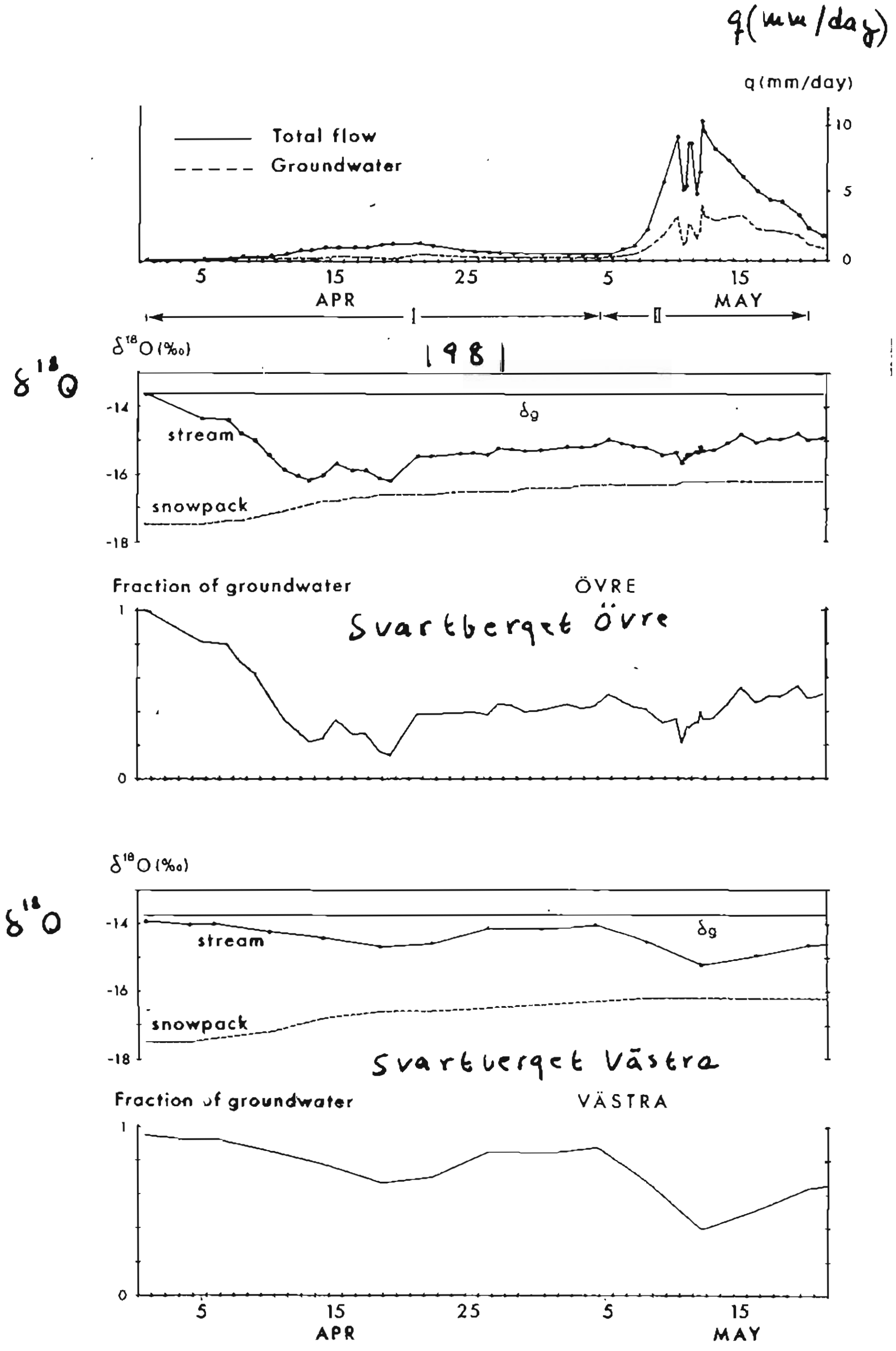

δ18O (‰)
$\delta^{18}O$
(‰)
-2
-6
-10
-14
-18
-22
stream
snowpack
NÄSTEN 1982
(UPPSALA)
Q(l/s)
q(mm/day)
Q
(l/s)
800
600
400
200
0
10
5
III
I
II
E
1
2
3
4
5
6
7
8
9
10
11
12
MONTH


Q(l/s)
q(mm/day)
20
16
12
8
4
0
0.2
0.1
0
Total flow
Groundwater
5
15
25
I
δ18O (‰)
δg
-12
-14
-16
-18
stream
snowpack
Fraction of groundwater
1
0
5
15
25
FEB

Nåsten, 1982


Q (l/s)
(mm/day)
Q(l/s)
q(mm/day)
80
60
40
20
0
12
10
8
6
4
2
0
Total flow
Groundwater
Svartberget Nedre
1981
④
①
5
15
25
5
15
25
I
II
δ18O
δ18O (‰)
-12
-14
-16
-18
δg
stream
snowpack
① Vg = ∞
② Vg = 500 mm
③ Vg = 250 "
④ Vg = 200 "
Fraction of groundwater
1
0
①
②
③
④
5
15
25
5
15
25
APR
MAY

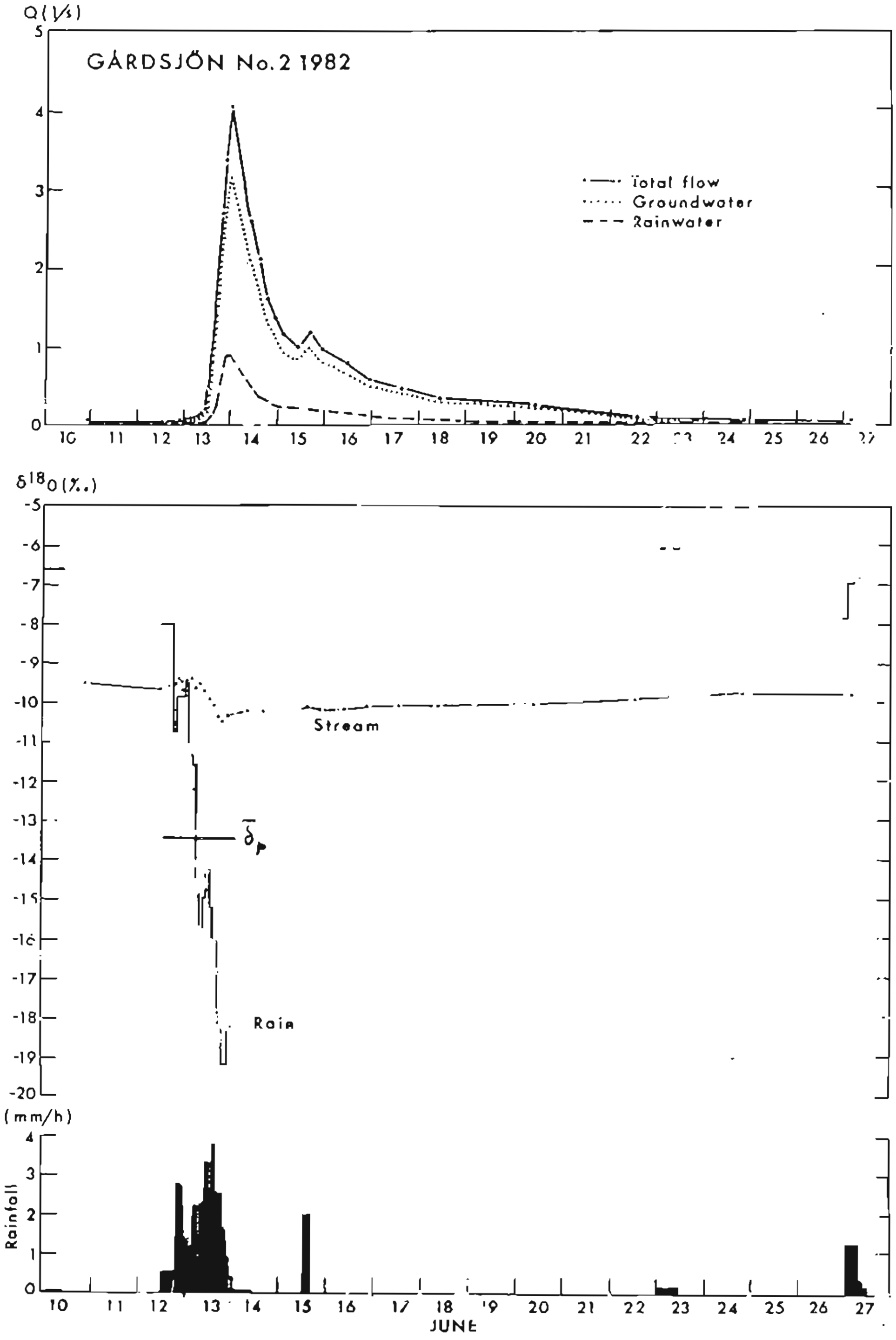
11

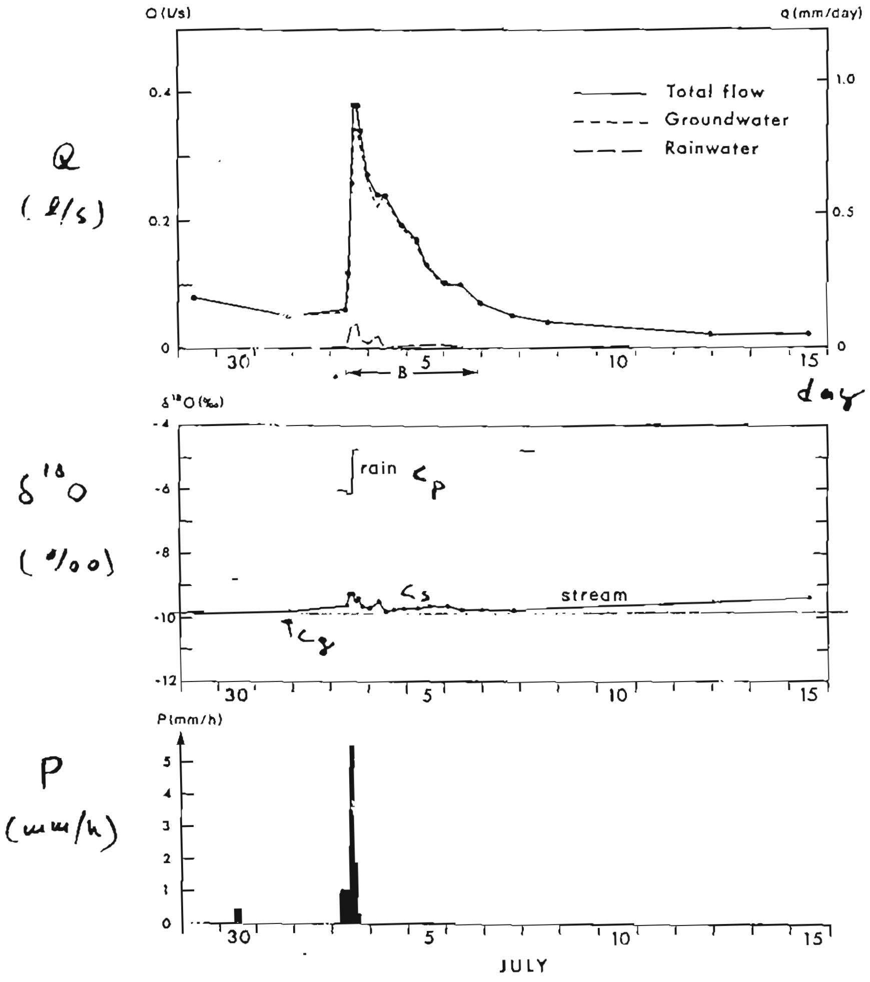

q(mm/day)
q(mm/day)
Total flow
Groundwater
10
5
0
5
15
APR
25
5
15
MAY
I
II
δ18O
δ18O (‰)
1981
-14
-16
-18
δg
stream
snowpack
Fraction of groundwater
ÖVRE
Svartberget Övre
1
0
δ18O
δ18O (‰)
-14
-16
-18
δg
stream
snowpack
Svartberget Västra
Fraction of groundwater
VÄSTRA
1
0
5
15
APR
25
5
15
MAY

Q

$\delta^{18}O$

rainfall intensity





Gårdsjön F1 1982

Fraction of groundwater
(pre-event water)

$$X = \frac{c_s - c_p}{c_g - c_p}$$

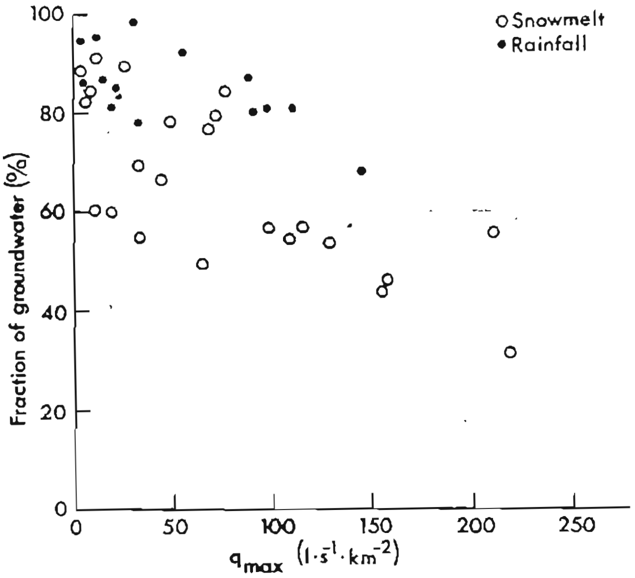


Fig.10.1 Total volume fraction of groundwater versus maximum streamflow. Snowmelt and rainfall events.

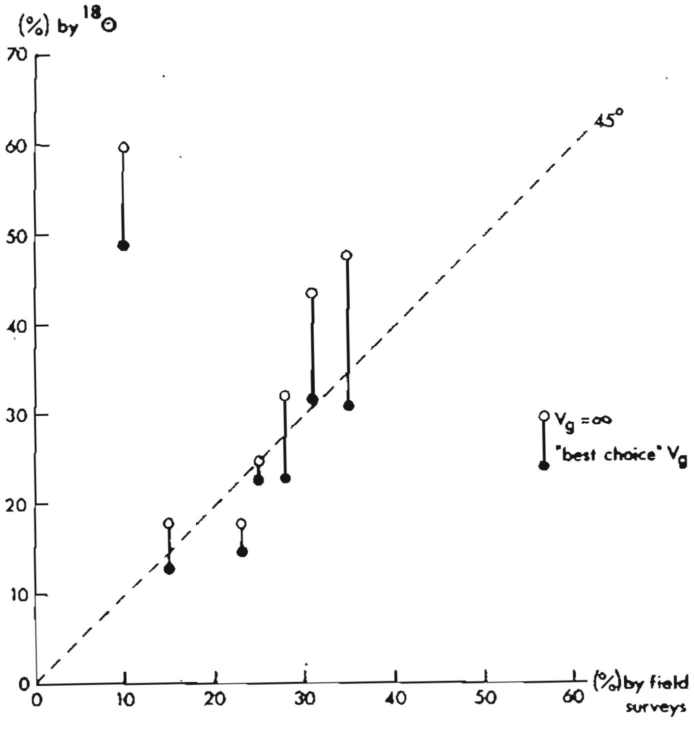
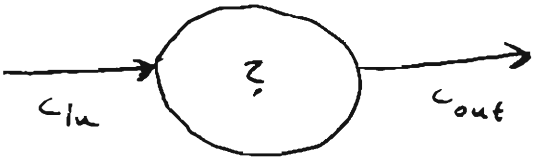
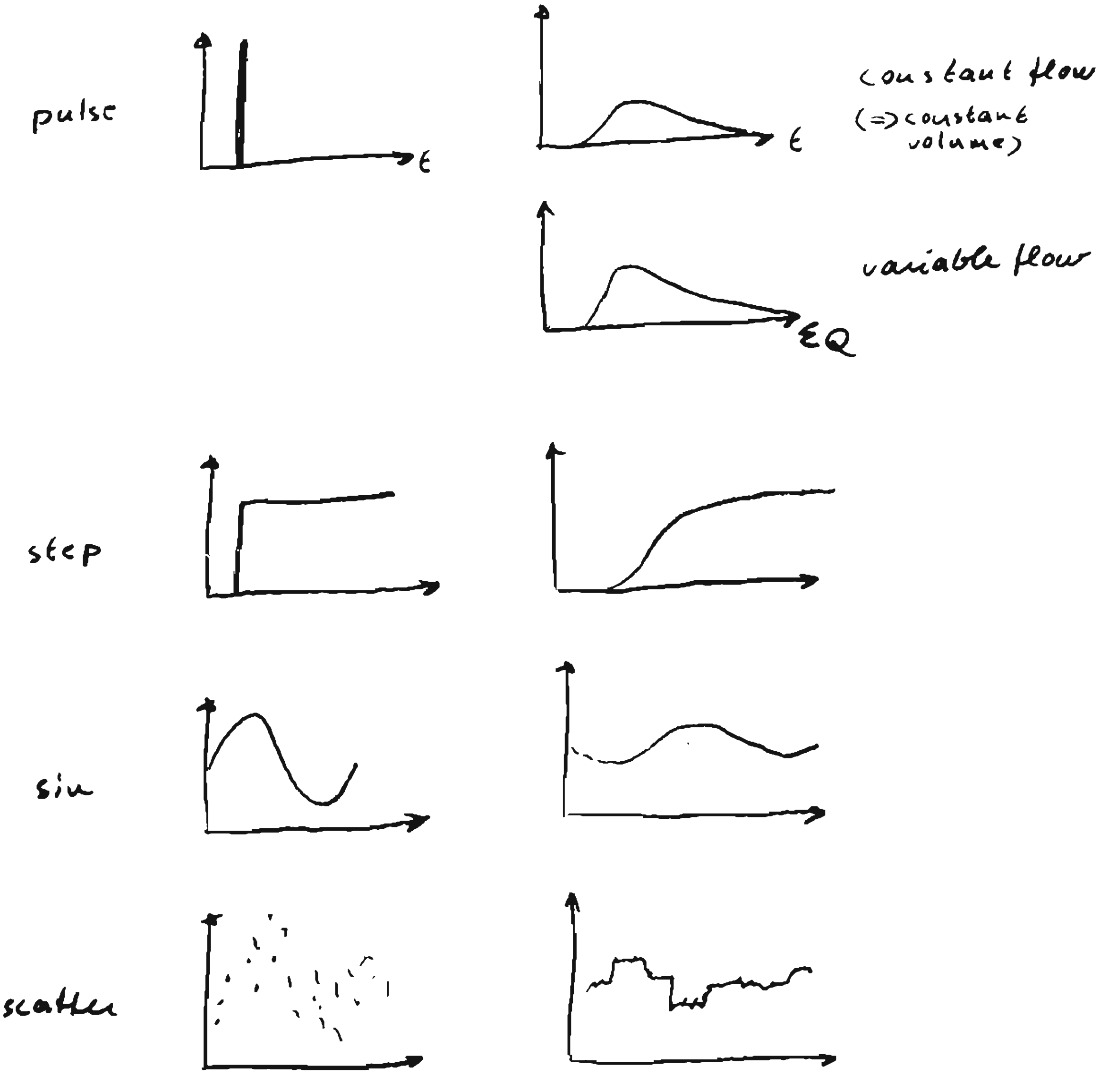
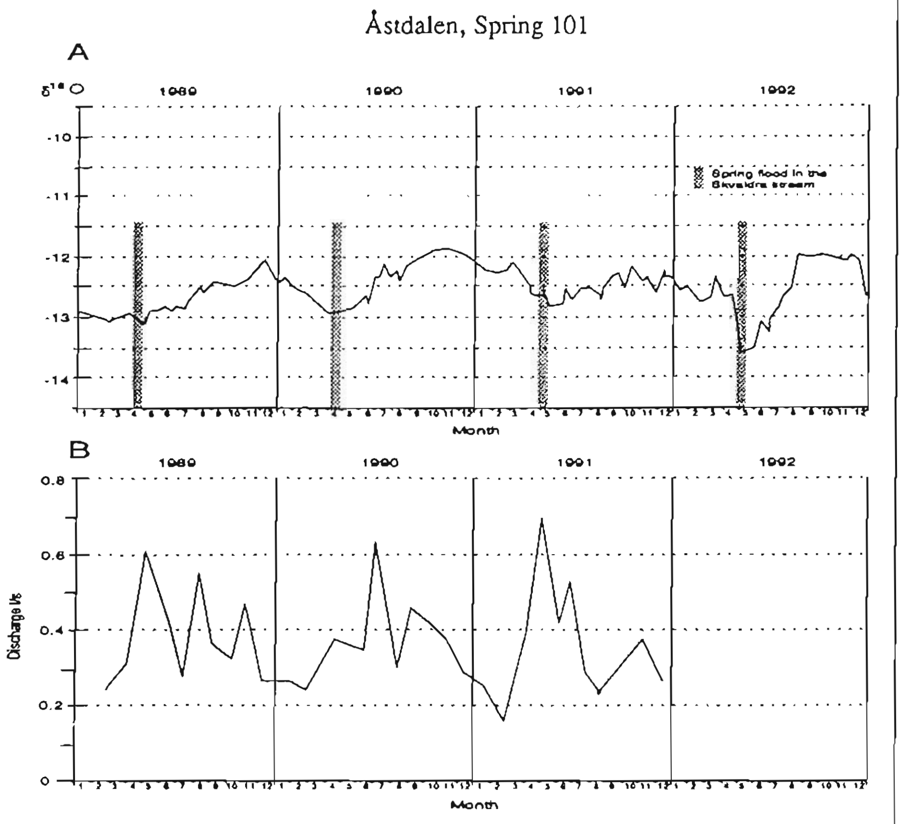


Fig.8.8 Basin fraction of discharge area estimated by ^{18}O and by field surveys. The "best-choice" value is calculated by using the best fitting model reservoir volume of 250 or 500 mm (see text).

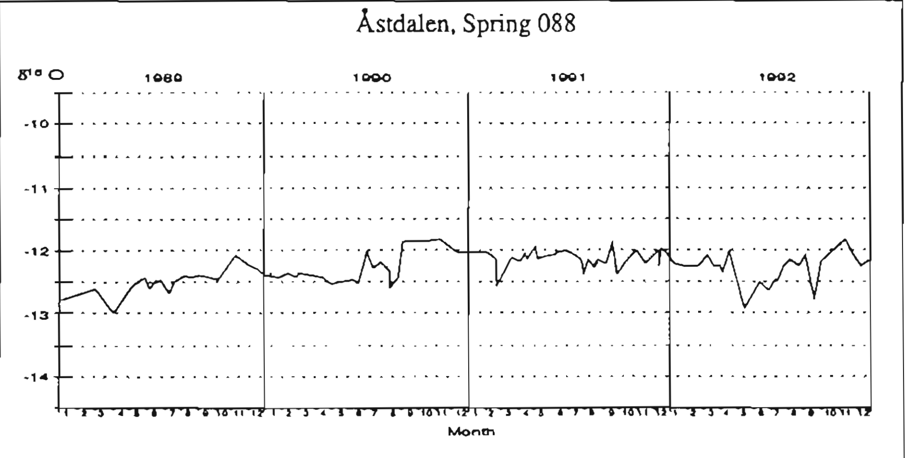

c_in
?
c_out


pulse
t
t
constant flow
(=> constant volume)
variable flow
tQ
step
sin
scatter



Åstdalen, Spring 101

Fig. 9. A. Variasjon i $\delta^{18}O$-verdier. B. Vannføring for kilde 101. Vårflommen i elva Skvaldra i Åstdalen er angitt på figur A. Beliggenhet, se figur 7.
Variation in $\delta^{18}O$-values and discharge of groundwater from spring 101. The time of the spring flood in the Skvaldra stream in Åstdalen is indicated. For location, see Figure 7.



Åstdalen, Spring 088

Fig. 10. Variasjon i $\delta^{18}O$-verdier for kilde 088. Beliggenhet, se figur 7.
Variation in $\delta^{18}O$-values of groundwater from spring 088. For location, see Figure 7.

from Haldorsen, 1994

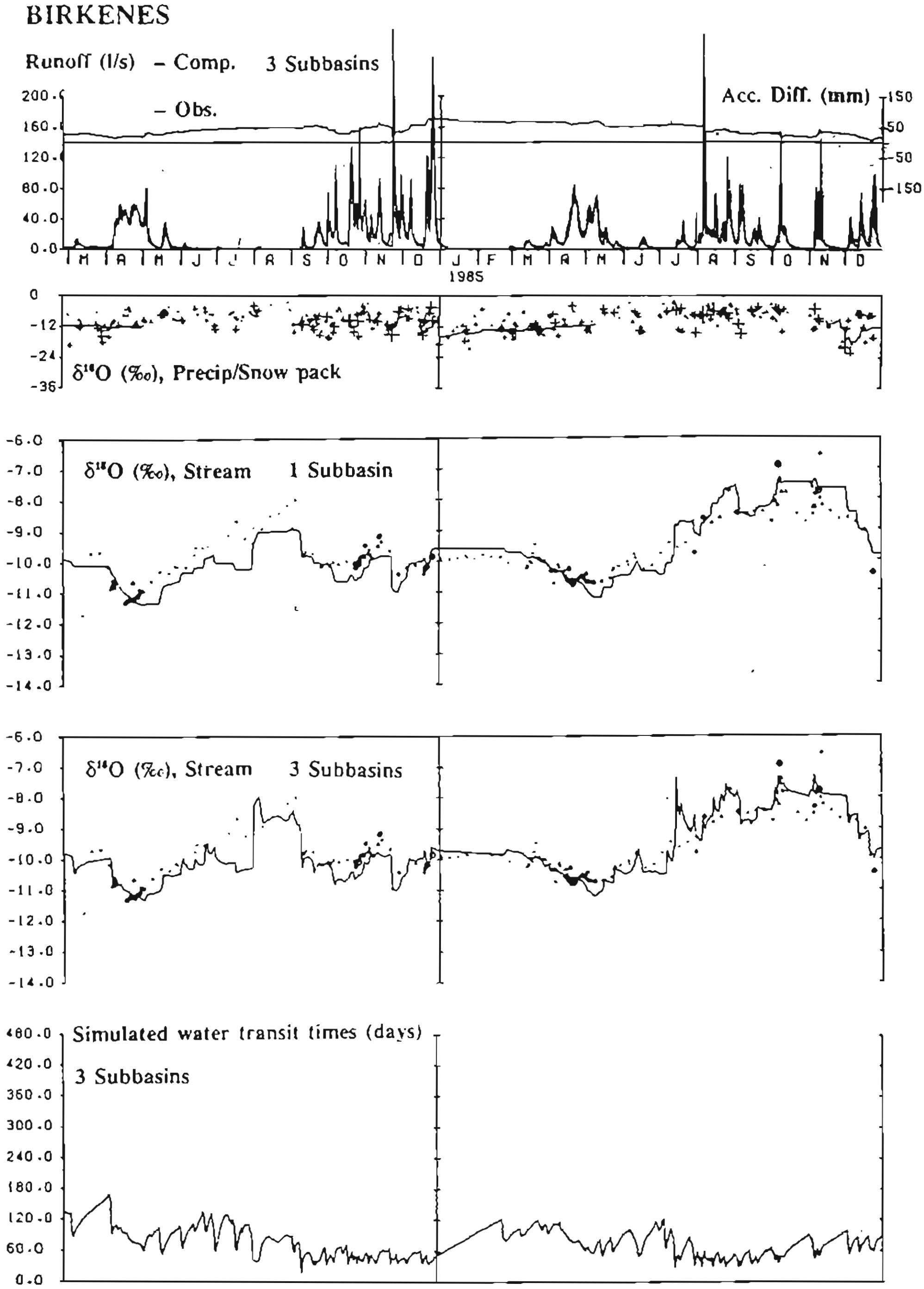
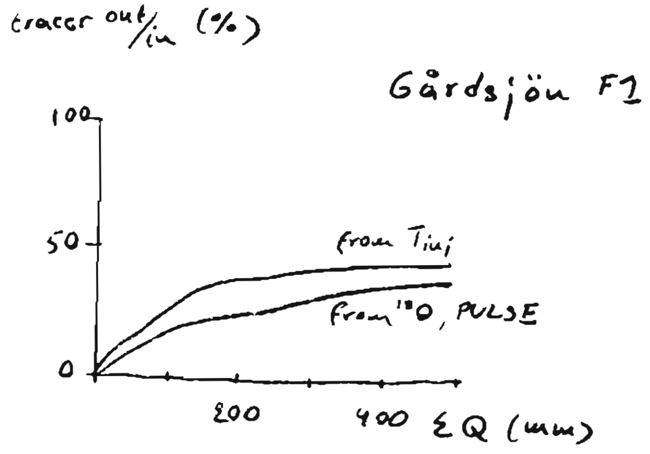
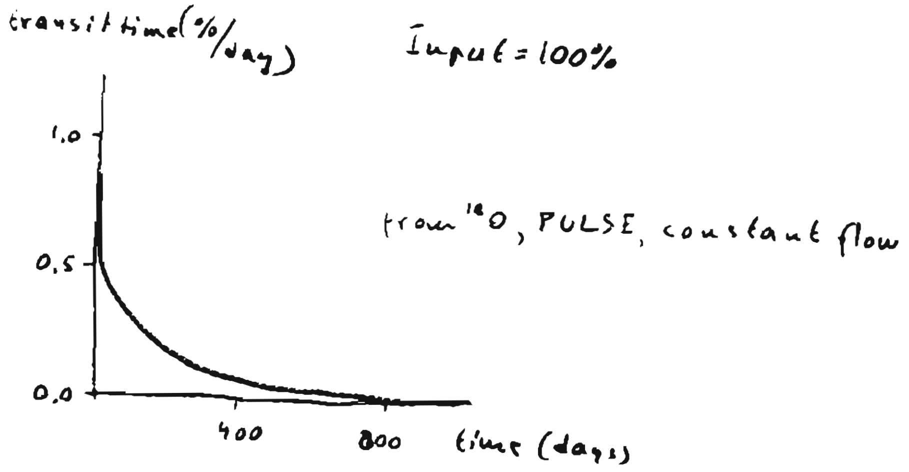
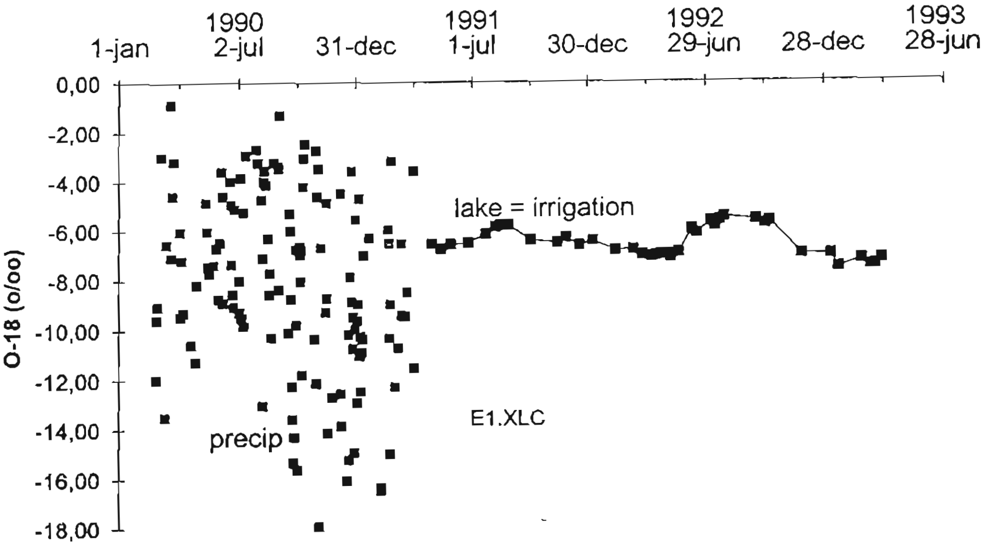
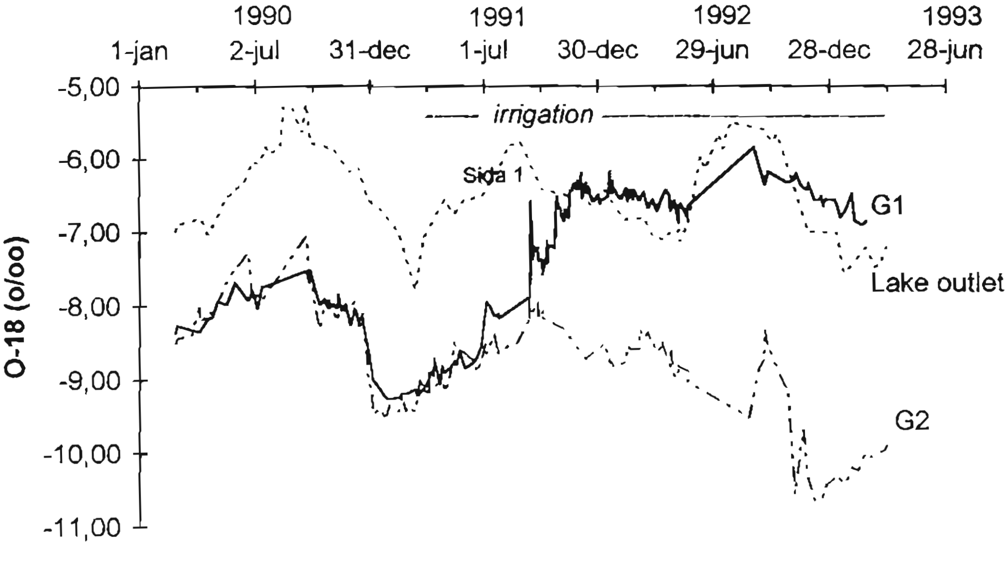


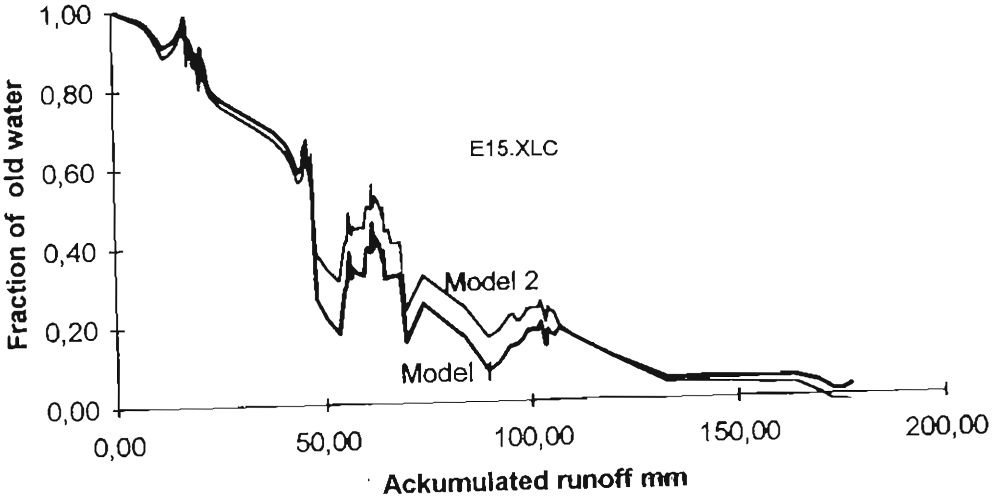
Figure 3. Simulation of runoff, ^{18}O and mean transit time for Birkenes, Norway (Observations received from SI, Oslo).

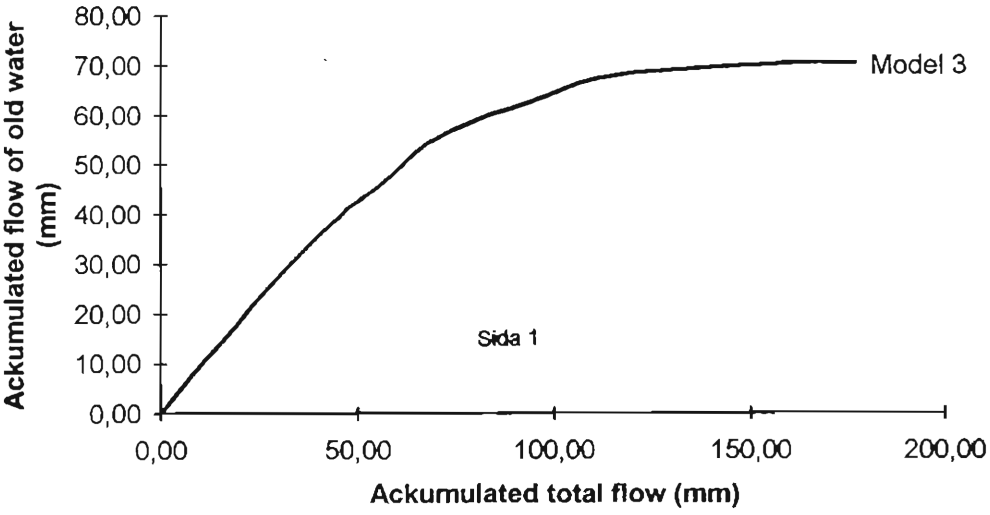

tracer out/in (%)
Gårdsjön F1
100
50
0
from Tin;
from 18O, PULSE
200
400
ΣQ (mm)


transit time (%/day)
Input = 100%
1,0
0,5
0,0
from 18O, PULSE, constant flow
400
800
time (days)


1990
1991
1992
1993
1-jan
2-jul
31-dec
1-jul
30-dec
29-jun
28-dec
28-jun
0,00
-2,00
-4,00
-6,00
-8,00
-10,00
-12,00
-14,00
-16,00
-18,00
O-18 (o/oo)
lake = irrigation
precip
E1.XLC


1990
1991
1992
1993
1-jan
2-jul
31-dec
1-jul
30-dec
29-jun
28-dec
28-jun
-5,00
-6,00
-7,00
-8,00
-9,00
-10,00
-11,00
O-18 (o/oo)
irrigation
Sida 1
G1
Lake outlet
G2


E15.XLC
Fraction of old water
1,00
0,80
0,60
0,40
0,20
0,00
0,00
50,00
100,00
150,00
200,00
Model 2
Model 1
Ackumulated runoff mm


Ackumulated flow of old water (mm)
80,00
70,00
60,00
50,00
40,00
30,00
20,00
10,00
0,00
0,00
50,00
100,00
150,00
200,00
Model 3
Sida 1
Ackumulated total flow (mm)

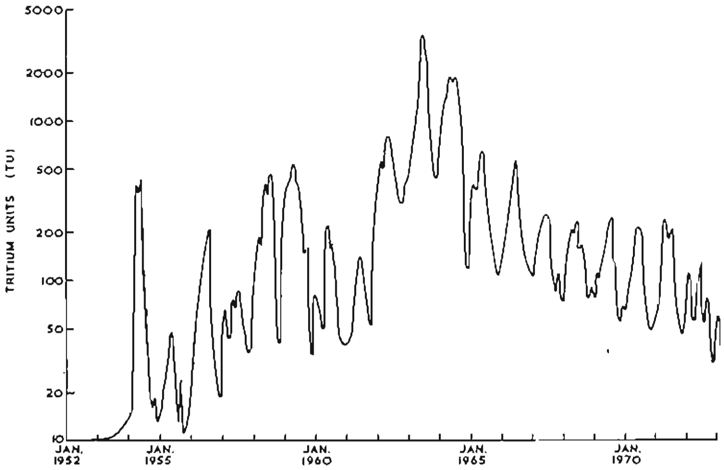


Figure 3.2 Tritium concentration in UK rainfall. Observed or computed values 1952–72 (not corrected for radioactive decay)

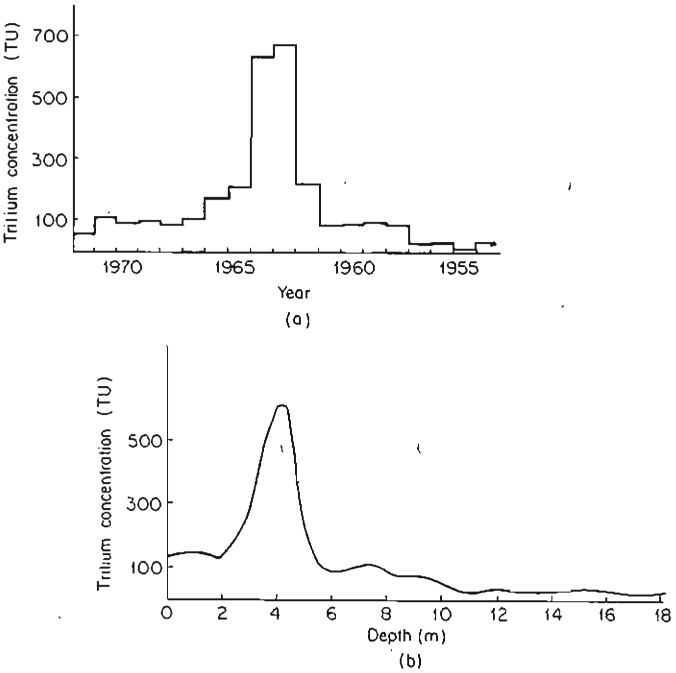


Figure 3.3 (a) Tritium in rainfall (UK). Weighted means corrected to January 1973. (b) Chalk profile. Tritium content of water in core, October 1968

Manual measurements and automatic instrumentation

Ånund Killingtveit
Division of Hydraulic and Environmental Engineering
Norwegian Institute of Technology

Knut Sand
SINTEF Norwegian Hydrotechnical Laboratory

1. STREAMFLOW MEASUREMENTS

1.1 Introduction

Measurement of stage and discharge is a very central part of operational hydrology. This chapter covers measurement of stage, discharge measurements and the establishing of rating curves, and methods for direct measurement of stage. Only the main principles are covered, for details the reader is referred to specialized literature as the WMO manuals (WMO, 1980, and Tilrem, 1986a & 1986b, based on experience from Norway and developing countries), and the authoritative works by Herschy on streamflow measurements (Herschy, 1978, 1985).

1.2 Measurement of stage

Stage is usually measured either to be used as an index for discharge estimation, or to give levels in reservoirs. Although the fundamental measuring techniques are the same in both cases, the practical considerations are usually different, and the demands on the installations and operation procedures differ. The measurement range is usually larger in reservoirs than in rivers, up to more than a hundred meters, while the rate of change with time is lower, except in small intake reservoirs in peaking systems. This section will focus on stage used as discharge index.

1.2.1 Non-recording stations

The classical instrument for stage measurements is the staff gauge, as shown in figure 1.1. The gauge will usually have marks for every centimeter, often with alternating black/white one centimeter wide marks, as in 1.1. Such alternating markings make the gauge easier to read, but may lead to individual preferences of choosing even or odd values. Most investigations also show overrepresentation of stages with modulo 5 or 10.

Waves and turbulence will often make it difficult to read stage with 1 centimeter accuracy. An easy way to overcome this difficulty is to attach a transparent tube with a narrow opening at the bottom end along the gauge, illustrated in figure 1.1. The water level in this tube will be strongly damped compared to the fluctuations of the free surface, but still representative.

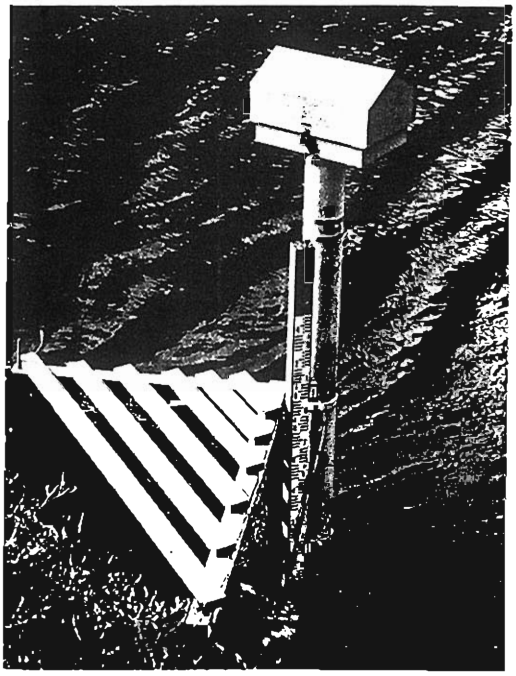


Figure 1.1. Staff gauge

The staff gauge must have a hard and durable surface and markings, resistant to corrosion, rot, mechanical damage (for instance floating ice), UV decay and algae growth. The Norwegian gauge has the scale printed on a plastic film with silk screen print technique, covered by two mm thick transparent polycarbonate. This makes a light but strong gauge. It is produced in lengths of one meter.

1.2.2 Crest gauges

As non-recording stations usually only will be read once or twice a day, flood crests will usually not be logged. To register crest level, stations can be equipped by a simple crest gauge, consisting of a vertical tube with small opening at the bottom end, with a removable central measuring stick and floating granular cork inside. Some cork will adhere to the measurement stick, and show highest water level since last measurement.

1.2.3 Recording stations

Non-recording gauges are labour consuming (at least in relation to the volume of information they collect), unfit in uninhabitated areas, and give inadequate time resolution in rapidly varying streams. The solution to this is to use recording gauges. The classical recording station is a stilling well with a chart recording instrument, measuring time with a mechanical clock and stage with a counterweighted float. The stilling well is connected to the stream with a tube, narrow enough to damp out rapid water level fluctuations, but large enough to avoid being easily clogged. See figure 1.2. The design of recording stations varies from prefabricated instrument shelters with float wells made from metal pipes, to individually designed huts with shelter for observer and even possibilities for overnight accommodation.

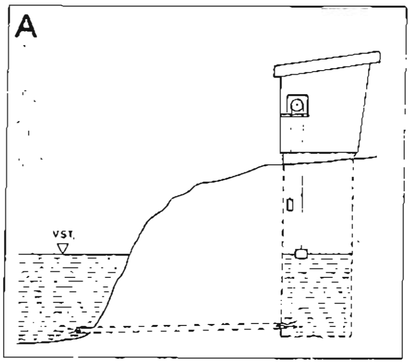


Figure 1.2 Recording gauge

The main problems with mechanical stage recorders arise from negligence and improper maintenance. The most common errors are

- poor trace due to faulty pen or pencil
- missing data due to outrunning paper capacity or clock running down

Traditionally, recorder charts have been manually read. This is time consuming, even if only daily values are read. It takes even more labour if timing errors have to be corrected. Today, charts will usually be converted to digital form by systems based on digitizing boards.

Precision mechanical instruments like stage recorders are becoming increasingly expensive, while prices on electronic equipment are falling. This, together with the vulnerability of mechanical recorders to improper maintenance, has led to increased interest in electronic stage recorders. The transitional generation of loggers was the punching instruments; mechanical instruments as the graphical recorders, but using punched paper tape as recording media. The next generation was electrical equipment, registering on audio tape. Today the most common types are fully electronic instruments, recording on solid state memory, either in form of exchangable cassettes or cards, or in internal memory. The storage units can be tapped locally or by telemetry, usually over standard serial interfaces. The most advanced stations are telemetering stations, transmitting either in real time or at fixed intervals. See section 2 for more information on telemetering systems.

Electronic recording systems are presently quite reliable, and offer several advantages over mechanical equipment. The most important are probably:

- more precise timing
- exact synchronization of several variables if necessary
- less vulnerable to improper maintenance
- can combine high time resolution with long observation period
- easy postprocessing, as data are stored in digital form

The last point is usually considered very important, but the actual manpower savings compared with a good digitizing system is probably often exaggerated.

One drawback has been the need for electrical power, but the power consumption of electronic devices and memory is going down, while the battery capacities are increasing, so this is hardly a problem anymore. For many systems it has been difficult for the observers to detect malfunctioning of the instrument, but now most systems have displays that can show measured stage, battery state, system time etc.

On an electronic recording instrument, stage from a float water level sensor must be converted to a signal that can be electronically read and stored. The most common translating systems convert the angle of the float wire pulley to resistance by means of a potentiometer, and further to digital form through an AD converter. An alternative is to convert the pulley position directly to digital form by a mechanical or photoelectric decoding device.

The use of electronic recording instruments opens for other sensor types for water level recording than float types. The most commonly used is the pressure transducer, which registers water pressure on a membrane. To avoid the influenced of varying atmospheric pressure (the normal variations of atmospheric pressure corresponds to one meter water head), the other side of the membrane must be exposed to air at atmospheric pressure, or water pressure measurements must be combined with air pressure measurements. In reservoirs with large water heads, the thermal expansion of water with varying temperature will significantly influence the water pressure/water stage relation.

The pressure transducer with air tube is usually protected by an anchored steel or PEL pipe.

The main advantage of the pressure transducer over float type equipment is that stilling well is not needed, reducing the cost of installation dramatically. The absence of the stilling well also solves the problem with ice in the well. There has, however, also been serious operational problems with pressure transducers, such as drifting away from the calibration curve, condensing water in air tube etc.

1.3 Discharge measurement methods

The rating curve must be established and controlled by discharge measurements. In principle there is a wealth of methods for discharge measurements, but the most commonly used, both in Norway and most of the world, is dilution gauging and the use of current meters.

1.3.1 Volumetric measurement

This is the most straightforward method to measure discharge, but unfortunately seldom applicable, as it requires

1) Low discharge, less than 100 l/s

2) Site where it is possible to divert all water to a measuring vessel.

The measuring principle is simply to measure the time needed for filling a measuring vessel of known capacity. As the uncertainty of the timing and of the capacity of the measuring vessel should be easy to determine, it is easy to make uncertainty analysis of the method, and as it is swift, it is easy to make repeated measurements.

1.3.2 Dilution gauging

The principle of dilution gauging is that if the amount of a conservative tracer in the river is known, and the concentration can be measured, the discharge can be calculated. Two techniques are used, constant injection and bulk injection. In the case of constant injection, the method for calculating the discharge is simply

$$c = \frac{i}{q+i}$$

where q is discharge and i is the injection rate (in the same measurementunits), and c is the measured concentration. Solving for q yields

$$q = i(\frac{1}{c} - 1)$$

The bulk injection method is based upon injecting a known quantity of tracer in the river. Provided all of the tracer passes the measuring point, the following relation holds

$$\int_0^T i(t)\,dt = I$$

where I is the volume of tracer, T is the time necessary for all of the tracer to pass the measuring point, and $i(t)$ the time varying tracer "discharge" at the measuring point. If i is negligible compared to q, it is easily found that

$$q = \frac{I}{\int_0^T c(t)\,dt}$$

where $c(t)$ is the measured concentration.

The method can be used where the turbulence of the river provides full mixing of the tracer in the water (i.e., homogeneous concentration in a river cross-section) within a practical distance from the injection points. This makes the method suited for highly turbulent streams, where the conditions may be difficult for current meter measurements. Thus, the two methods can be considered complementary.

The tracer must be easily detectable, not chemically reactive with water or material in the river reach, and environmentally harmless. Any natural content of the tracer must be low compared to the injected amount, and not rapidly varying. Fluorescent tracers like rhodamine are quite common, but are difficult to used in turbid water.

In Norway sodium chloride (NaCl, ordinary table salt) is the most used tracer. Concentration is measured by reading changes in river water conductivity by a standard conductivity meter, preferably logging, with adjustable range. The tracer is cheap and easy available, but the method requires care when dissolving the salt, and the calibration procedure is somewhat time-consuming. The use of bulk injection of sodium chloride was pioneered by Norwegian hydrologists Aastad and Søgnen (Østrem, 1964). Sodium chloride has been used for discharges up to 3-400 m3/s (Otnes & Ræstad, 1978).

Radioactive tracers have been used, and are accurate in small doses, but the restrictions on the handling and application of such tracers limit their use.

1.3.3 Use of current meters

The use of current meters is the most widespread methd for discharge measurements. The principle is to measure the flux of water through a river cross-section, by determining the velocity of water in a large number of points. The most commonly used current meters are propeller type water velocity meters.

The cross-section is defined by a wire with distance markings stretched across the water or by a permanent or removable cableway. The current meter is operated from cableway, bridge, boat or by wading, depending on the size of the river.

Usually the discharge is calculated from mean velocity in cross-section verticals, usually fifteen to thirty evenly spaced verticals. In each vertical, the estimate of the mean velocity is based on one or several measurements. In a "standard" turbulent vertical velocity profile, the velocity 0.4 times the water depth above the bottom corresponds to the mean profile velocity, so this point would be the obvious choice if only one measurement is taken. If two are used, they will normally be 0.2 times the water depth and 0.8 times the water depth, as the average of these velocities is close to the vertical mean. See figure 1.3.

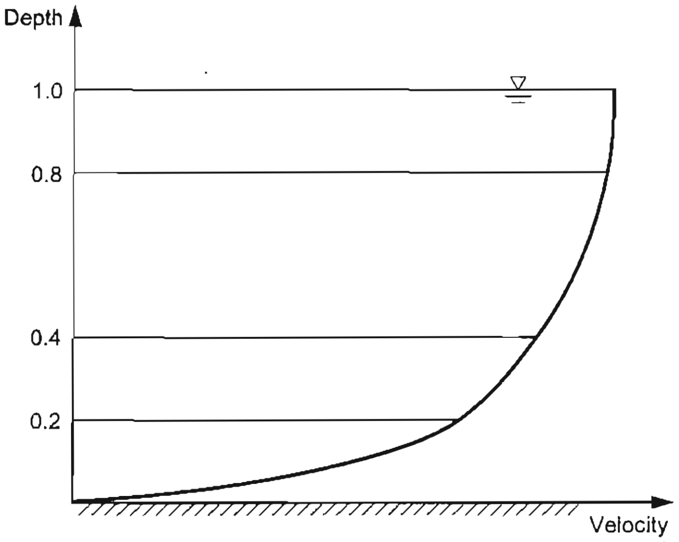


Figure 1.3 Vertical velocity profile

For the main current region of the section usually more points are measured - one close to the surface and one close to the bottom, and two to five between. The mean velocity is obtained by graphical or numerical integration.

The recommended practice in Norway is to measure in five points; near bottom, 0.2, 0.4, 0.8 times water depth, and near surface (Tilrem & Viken, 1988).

When the current meter is suspended in a wire, it will drift off somewhat if the current is strong, and this must be corrected for when the depth of the measuring point is calculated.

The water velocity can display quite large fluctuations due to moving eddies, and it is important that a sufficiently long averaging time to even out such fluctuations. On the other hand, too long averaging time will increase the total discharge measurement time unnecessarily. A common recommendation for averaging time is minimum 45 seconds.

Measurements from ice cover is done through holes drilled in the ice, one hole at each vertical. If the current meter is held by rod, it is hinged to allow the meter to be swivelled up along the rod by a wire when it is lowered and lifted through the hole. Frazil ice disturbs the measurements, and should if possible be avoided.

When the mean velocity in the verticals has been determined, the total discharge is calculated by weighting each vertical velocity by a representative area. Three methods are commonly used:

- mean section method
- mid section method area
- specific discharge method (Harlachers method)

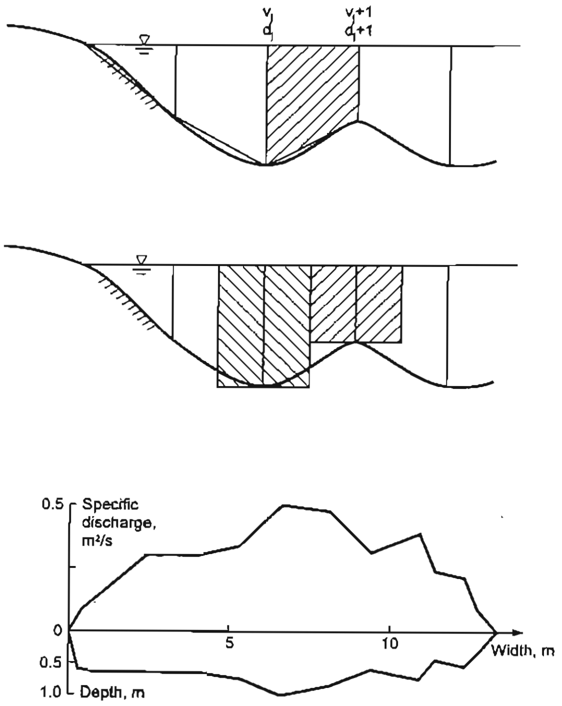


Figure 1.4 Three methods for calculating total discharge from mean vertical velocities

The principles of the three methods are shown in figure 1.4. In the mean section method discharge is calculated for the trapezoidal segments between verticals, using the mean of the velocities in the confining verticals as representative velocity. The mid section method calculates the discharge in a rectangular segment embracing a vertical, using the velocity in that vertical as representative velocity. Harlachers method calculates the discharge by integrating the specific discharge (velocity times depth) along the cross section. In Norway Harlachers method is the standard method.

The calculations can be made manually, graphically or by programmed calculators. For field use programmable calculators or notebook type battery powered personal computers are convenient. Two examples on available commercial or freeware programs are FLYGEL (NVE) and CUMEC (The Danish Land Development Service). Nordic comparisons have shown that the programs can give somewhat differing results. The differences normally stem from how the velocity profiles near the bottom are treated, and how velocities are interpolated towards the shore.

Measurements in ice covered rivers present special problems. The current meter has to be lowered through holes drilled in the ice, and the thickness of ice in the verticals must be measured. The velocity profiles will be different under ice, as the water velocity is zero at the water/ice boundary. Calculations of mean profile velocity must take this in account, and the programs must have an option for handling ice covered sections.

1.3.4 Float method

The simplest way to measure water velocity is to time floating objects over known distances. This will give the surface velocity. Mean vertical velocity must be obtained by empirical corrections factors - typically 0.7 to 0.8 (the higher value for larger depth). An alternative is to use floats with sunk drifting anchors - for instance rods. To calculate the discharge, the velocity has to be measured in 15 to 25 verticals spread evenly across the river.

The method requires parallel current, but is otherwise often suitable under difficult conditions. The method is not very accurate, and is mainly used for rough estimates of discharge.

Sargent (1981) developed an interesting variation of the float method - the integrating float technique, where rising air bubbles from an air tube anchored to the bottom are used as floats. The bubbles will drift with the current, and when the depth of the nozzles and the buoyant speed is known, the crossectional velocity distribution can be calculated from the trace of air bubbles breaking the surface. The trace is registered by photographing it.

1.3.5 Moving boat

The moving boat technique is a fast method to make discharge measurements in large rivers. A fast reacting current meter with a wane is attached to the boat with a rod, and measures the instant velocity one meter below the surface. The boat traverses the stream at right angles. to the current, with known velocity. The relative water speed and the direction are measured at typically twenty to thirty points across the river, while the depth is measured with an echo sounder. The true water velocity is calculated by resolving the apparent velocity vector in the transverse direction (boat speed) and the current direction (water speed). As the water speed is measured in one fixed level, a correction factor must be applied to get the mean velocity in verticals. About 0.9 is common.

Usually several traverses are made, to reduce the uncertainty, and an accuracy of approximately 5 per cent is claimed at the 95% confidence level.

1.3.6 Ultrasonic and electromagnetic measurements

While the discharge measurement methods mentioned so far are sampling methods, the ultrasonic (acoustic) method is capable of measuring mean velocity continuously, and the electromagnetic methods provide continuous measurement of discharge.

If a sound signal is transmitted through water having a velocity component parallel with the sound path, the travel time of the signal will depend on water velocity. In the ultrasonic gauges, this principle is used to measure the mean water velocity. The velocity will be averaged only along the path defined by the line between the transmitter and the receiver, and depth variations are not considered. To cover the depth variations of the velocity, several gauges must be used, or the instruments must travel up and down.

Ultrasonic gauges are usually permanent installations, and are mainly used where it is difficult to establish rating curves. The method is best suited for approximately rectangular sections. The signal path is usually about 45 degrees on the current.

While the ultrasonic method measures the mean water velocity along a path, the electromagnetic method measures the mean velocity and total discharge by determining changes in the electromotive force induced in the water by the moving water in the earth magnetic field. The method requires large installations with probes buried under the river bed. It is particularily suited for rivers with unstable profiles or vegetation growth.

1.3.7 Hydraulic structures

Under given conditions, hydraulic structures have known stage/discharge relationships. They also provide a stable hydraulic control. Their use is limited by the installation cost, which increases rapidly with maximum discharge capacity. There is a wealth of designs of hydraulic control structures, and the criteria for choice of design are

- rating curve characteristics
- capacity/cost relationship
- available head difference, possibility forraising the upstream water level
- sediment load
- fish migration through structure

The two main categories are raised structures with full hydraulic jump, and submerged structures with partial jump (subcritical to critical/supercritical flow and back). In Norway the standard sharp edged 90 or 120 degree V-notch weir is most commonly used structure of the first category, while the Crump weir is used where submerged structures are needed. In most Norwegian streams, the range of discharge is very large. To obtain a satisfactory resolution on low flow, while maintaining high total capacity, composite weir profiles are sometimes used..

To avoid ice disturbing the rating curve during winter, V-notch weirs can be covered with removable lids to prevent icing in theprofile (Tveit, 1974).

For comprehensive descriptions of the different types of control structures, the reader is referred to the specialized literature, for example Bos (1976) and Ackers *et al* (1978).

1.4 Stage-discharge relation

When stage is used as index on discharge, it is necessary to establish a stage-discharge relationship, the rating curve. In the general case, the discharge is dependent on stage and head slope, as illustrated for instance by the Manning formula

$$Q = MAR^{\frac{2}{3}}S^{\frac{1}{2}} \qquad [2.1]$$

where the head slope S can vary independently of stage, while the friction parameter M, the wetted area A and the hydraulic radius R is dependent on stage h; $M(h)$, $A(h)$, $R(h)$. If the slope can be considered constant, the discharge is a function of stage only (as long as the friction-stage and geometry-stage relations are unchanged).

The stage-discharge relationship is unique upstream a hydraulic control, as long as the distance is not long enough to allow the slope to vary. Otherwise it will be unique if the stage gauging station is not affected by backwater effects, and the discharge does not vary so quickly that the floodwave significantly alters the slope. The last condition may arise in the case of artificial releases, or flash floods producing flood waves.

Even if the discharge is a function of stage only, this relationship may change with time. As illustrated by the Manning formula, this may be caused by changes in the friction-stage relationship, *M(h)*, or the geometry-stage relationship; *A(h)*, *R(h)*. Changes in friction Are usually caused by vegetation changes or ice cover, while changes in geometry is typically caused by ice cover formation, channel bed erosion or sedimentation, or hydraulic constructions or works. Ice and vegetation changes are seasonal, while most other changes are permanent. Changes may arise abruptly, usually caused by floods or construction works, but will often evolve over time.

1.4.1 Rating curve

For natural channels, the rating curve has to be established by discharge measurements over as wide range as possible. The curve must also be checked regularly by follow up measurements, more frequently for unstable channels and hydraulic controls.

The most commonly used mathematical formulation of the rating curve is

$$q = k(h+c)^n \qquad [2.2]$$

where k, c and n are constants. Most weirs will conform to this formula, but most composite weirs and natural controls will need two or more segments, each of the type above, to have an acceptable fit over all stage ranges.

$$q = k_j(h+c_j)^{n_j} \qquad for \quad h_{j-1} \leq h < h_j \qquad [2.3]$$

For the lowest segment, c_1 will be equal to $-h_0$, the lowest point in the weir or saddle point in the control. A rectangular weir has n close to 1.5, and a triangular weir 2.5. Most natural controls will have n between these two values.

Usually the power formula is fitted by logarithmic regression, in the form

$$\log q = k + n \log(h + c)$$

This requires c to be preset, as in case of the lowest segment, or to be optimized to produce the best fit. Standard least square logarithmic regression presumes errors to be proportional to the variable, and this is not an unlikely assumption - discharge measurement errors are probably near proportional to discharge. Fitting the curve may be done numerically or graphically, both procedures are somewhat cumbersome on multisegment curves, as c_j usually has to be determined by trial and error, and h_{j-1} has to be determined in a separate step - by finding the intersection between segments. The segments do not necessarily intersect.

Although there is som physical rationale behind the power curve, fitting it to a natural control is primarily an exercise in curve fitting. The curve gets many parameters as the number of segments increase (four extra for each new segment), and the resulting curve is not smooth. Other mathematical formulations have been proposed, usually polynomials (WMO....). The problems with these are mainly that they cannot be used for rating curve extrapolation - the extrapolation has to be done prior to the curve fitting. This is - by the way - a healthier procedure anyhow.

1.4.2 Extrapolation of the rating curve

A rating curve cannot be calibrated to the possible extremes, neither at the low nor the high end, and the curve has to be extrapolated to cover the range of stages. If the stage of zero discharge is known, extrapolating at the lower end will not pose any problems. Extrapolation at the high end will easily introduce large errors, as the control section often will exhibit large changes and even change position on flood stages. The value of flood discharge measurements cannot be emphasised too much.

Extrapolation the upper power curve segment [2.3] is quite common practice, but not advisable for large extrapolations. Better methods are based on description of the channel or control geometry. Most methods require at least the cross section area - stage relationship to be known. The simplest method is based on the assumption that it is easier to extrapolate the mean velocity than the discharge. Discharge Q is then calculated by the relationship

$$Q = A(h)\,V(h)$$

where $A(h)$ is known and $V(h)$ is extrapolated.

Stevens method and similar methods are based on the Manning or Chezy formulas, which state that

$$Q \;\; vs \;\; A(h)R(h)^{\frac{2}{3}} \qquad \textit{(Manning formula)}$$

$$Q \;\; vs \;\; A(h)R(h)^{\frac{1}{2}} \qquad \textit{(Chezy formula)}$$

should be close to linear relationships, as

$$\frac{Q}{A(h)R(h)^{\frac{2}{3}}} = MS^{\frac{1}{2}} \qquad \textit{(Manning formula)}$$

$$\frac{Q}{A(h)R(h)^{\frac{1}{2}}} = CS^{\frac{1}{2}} \qquad \textit{(Chezy formula)}$$

and C, M and S usually are relatively independent of stage. Stevens method is illustrated in figure 1.5.

Where few discharge measurements are available, physical hydraulic scale models, or mathematical models for stationary flow like HEC-2 (US Army Corps of Engineers, Hydrologic Engineering Center) may be used for rating curve extrapolation.

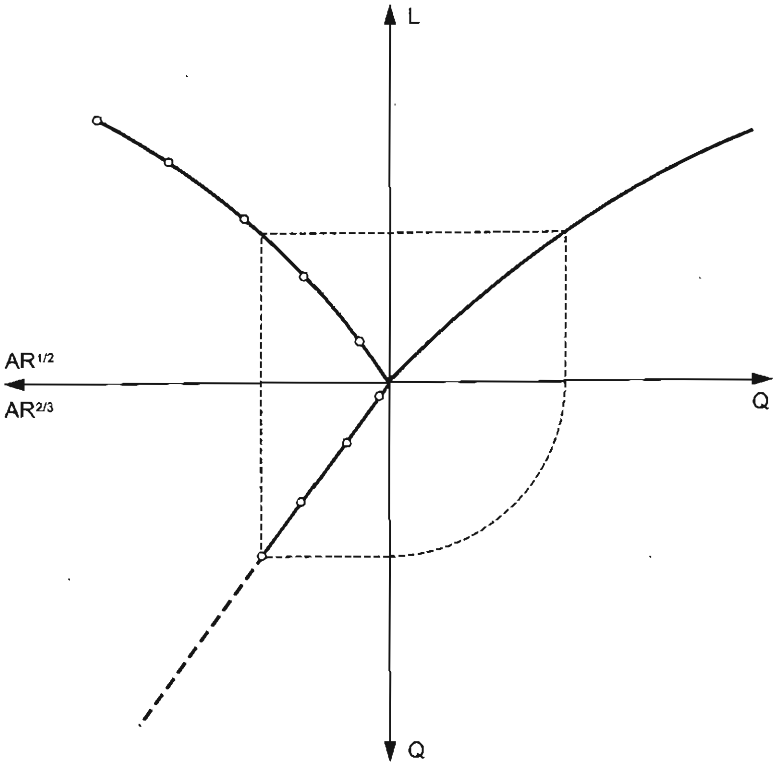


Figure 1.5 Graphical rating curve extension, using the Stevens method

1.4.3 Discharge-stage-slope relationships

As mentioned previously, backwater effects may cause the water head slope to vary for a given stage. As indicated by the Manning formula, the result is that discharge is not a unique function of stage, but of stage and slope. To determine the slope, it is necessary to use two stage gauges. The distance between the gauges must be large enough to determine the slope with sufficient certainty. The rating curve is replaced by a surface or two-dimensional table, describing discharge as function of stage at one gauge and slope, or - more practical - stage at both gauges.

As calibrating a two-dimensional surface requires more data than calibrating a one-dimensional rating curve, it will be efficient to use the Manning formula or hydraulic models (physical or mathematical) to establish the relationship.

Backwater effects may be caused by tidal influence, downstream reservoirs, sluices etc. In principle such conditions affects the water level and slope a long distance upstream under subcritical flow conditions, the effect sinking asymptotically to zero, only to be broken by supercritical flow reaches.

1.4.4 Effects of river ice

River ice reduces the discharge at given stage, compared to ice free conditions. This is caused by reduced active water conveyance area due to ice cover and at times bottom ice formation. In addition the velocity profile is altered, due to increased friction at the water/ice boundary.

The problem cannot be solved by changing to another rating curve, as the ice conditions and their effects on stage changes with time. An example on ice influence is given in figure 1.6.

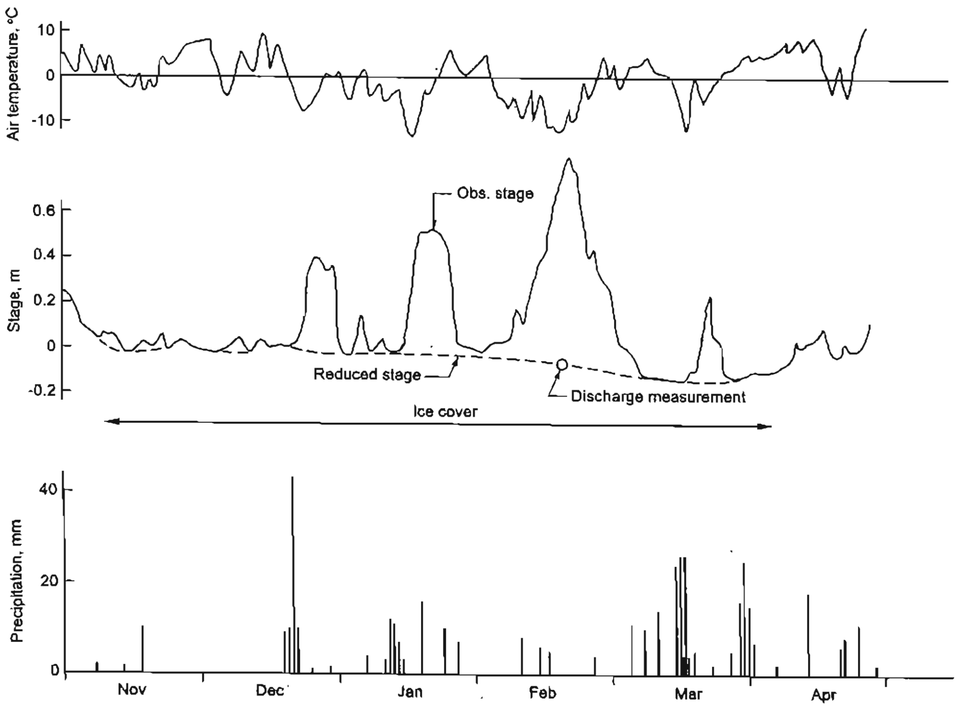


Figure 1.6 Effects of ice on river stage.

Correction for ice influence must be based on discharge measurements during winter. This is time consuming and expensive, and the number of measurements accordingly low. The interpolation between measurement points is a subjective procedure, based on knowledge about temperature and precipitation, and comparison with ice free gauging stations. An alternative is to interpolate between measurement points with a hydrological model.

Usually the ice correction is established by finding a finding an "equivalent" stage, allowing the standard rating curve to be used, also for winter conditions. The procedure is in principle the same as adjusting for vegetation growth (Stout's method), but the influence of ice is usually greater and more fluctuating through the season. Cold spells can give dramatic freezeups with stage increases that can resemble floods, figure 1.6.

1.5 Practical considerations

1.5.1 Choosing a gauging station location

A stable hydraulic control is the most important criterion for choosing a gauging stage location. Sites with hydraulic control where flow changes from subcritical to supercritical flow are very valuable, because they tend to be stable (are often on rock ledges), are easy to inspect for changes and break potential backwater effects. Narrow control sections give better discharge resolution than wide ones.

If river ice is a problem, gauging stations in lakes are often less influences, as the outlet often is ice free due to upwelling deep water.

Other important criteria are absence of backwater effects, accessibility, suitability for gauge mounting and station construction, access to electrical power supply.

1.5.2 Stage datum

It is imperative that a fixed datum is established and maintained within easy access from the stage station. This is usually a bolt in rock, and is used for controlling the vertical alignment of the staff gauge(es). The bolt must be clearly marked, to avoid it being inadvertently being removed or made inaccessible by road works, land fills etc.

1.5.3 Measuring range

Obviously, the equipment must cover the range of possible stage variation. On graphical recorders, there is a tradeoff between stage resolution and measuring range. This is often overcome by using a reversing pen guide spindle. On electronic equipment there may also be a similar tradeoff, if the data word size is small. 8 bit dataword can only cover 2.5

meters if 1 cm resolution is wanted. 16 bit dataword will have such limitation.

The possible levels of extreme floods are often underestimated, and many installations have been proved inadequate by floods causing the float to roof in the stilling well, recording equipment to be drowned or complete installation to be washed away. A lamentable situation, as very valuable data is lost in addition to the direct damages, not to speak about how embarrassing such situations are to a hydrologic service!

1.5.4 Observation time

Manually read staff gauges are usually read once a day at a fixed time, and this reading is often regarded as an unbiased estimate of daily mean stage. This may not be true. Many rivers display a diurnal variation, which may bias the estimation of daily mean stage. Such diurnal variations are caused by snow or glacier melt, afternoon convective showers, or peaking releases from hydropower plants. The remedy is either to have two or more observations a day, to find a representative observation time, or to switch to recording equipment. A special problem arises on manually operated reservoirs, where stage often is read in connection with gate operation. It will then be difficult to know which time period the stage reading is representative for.

1.5.5 Controlling the stage

The recorder must be checked against observed stage regularly. A recording station must therefore also be equipped with a staff gauge. If the stilling well is large enough, the staff gauge is often located inside the well. If the gauge is outside the well, it is important to be aware that the recording gauge will measure the actual head where the connecting tube emerges, while the staff gauge usually is located in still water. This will introduce a difference between the recorded level and the still water level equal to the velocity head

$$\Delta h_v = \frac{v^2}{2g}$$

where v is the mean water velocity in the vertical section over the connecting tube. The velocity head is not negligible, at 1 m/s velocity it amounts to 5 cm, at 2 m/s to 20 cm. As the rating curve will usually be established on the manually read stages, the velocity head water level reduction will lead to underestimation of flood stages and discharge. The differences between registered and observed stage can also be misinterpreted as instrument drift when start and end stage (or other control stages) differ.

1.5.6 Winter maintenance

Measuring stage in winter conditions with river ice cover introduces problems. For manual measurements an ice well has to be kept open at the gauge. In stilling wells the well has to be kept free of ice. If electric supply is available, this can be done by lowering a light bulb in the stilling well. Without electrical supply, the stilling well can be kept ice free by filling and maintaining a paraffin cover in the stilling well. This will give a higher level in the stilling well than in the river, due to the different densities of water and paraffin. Stilling wells dug in ground will be less vulnerable to ice formation than pipe type wells exposed to air.

1.5.7 Clogged pipe

A common cause of operation errors is clogged connection pipe, giving constant or highly damped measured stage. It can take time before this is noticed, especially if the staff gauge used for control is inside the stilling well. A clogged pipe can often be opened by flushing.

1.5.8 Regular inspection

A stage station should be inspected at regular intervals by trained field technicians, to check gauge alignment, maintenance etc. In addition, recording stations will be visited by local personnel for change of charts, data cassettes etc., and to control stage. With the availability of electronic loggers with large storage capacity, it may be tempting to increase the inspection interval. All operation experience indicates that also this kind of equipment should be inspected regularly, and that data cassettes should be changed at least every six months to avoid large data losses in case of malfunction.

1.5.9 Sheltering of instruments

Instruments, both mechanical and electronic, will have differing demands on operation environment. It is important to consider these demands, especially on temperature and humidity, when instrument mounts and shelters are designed. In addition it is also wise to take into account the risk for vandalism and theft, and secure the instruments accordingly. Usually information signs about the instruments and their purpose may reduce the risk of damage.

1.5.10 Choosing a section for current meter discharge measurements

The main criteria for the selection of a section for current meter measurements are (Tilrem & Viken, 1988):

- the difference in discharge between the stage measurement station and the discharge measurement must be considerably less than the measurement uncertainty

- the flow must be parallel, and normal to the section
- no significant backwater flow should be present
- no boulders, giving irregular velocity profiles and nonnormal flow components, should be present
- eddys of any kind should be avoided, if possible
- the site should be easily accessible and safe, even in high floods
- water velocity should not be below 0.15 m/s, and not above 2.5 m/s

At some stations, it may be practical to have separate measuring sections for high and low discharges.

2 Automatic data acquisition systems in hydrology

2.1 Introduction

The use of hydrological models for runoff forecasting has been increasing rapidly during the last 10 years. To be able to use the models efficiently for real time flow forecasting in hydro power systems, observations must be made available for model input and updating as quickly as possible. This is even more important for flood forecasting, where efficient collection, transmission and quality control is mandatory to be able to issue reliable forecasts. For small catchments it may be nessesary to collect meteorological and hydrological data with hourly timesteps or less. In fact, in most data collection systems built in Norway during the last years, hourly timesteps has become the standard.

The increasing need for automatic data acquisition systems has led to considerable efforts in research and development to find optimal system design and improve reliability. This work has led to development of sensors, electronics, software and complete systems which have been tested for extended periods both in experimental and operational networks.

2.2 Main structure and system components

The main components in an automatic data collection system for hydrological and meteorological data are shown in Figure 2.1. The components are explained below:

Sensor	Only a few meterological and hydrological phenomena can be measured directly. In most cases one observes the phenomena indirectly by observing the effects in a suitable way. For measurements of wind for example one can observe how clouds, smoke or trees are affected. More practically, one can use a vane anemometer and observe how this is rotated by the wind. This rotation can be observed directly, mechanically or electrically, and by a suitable transformation give the wind speed. Similarely one can observe the wind direction by a wind vane, temperature by changes in electical resistance, precipitation as weight of rain and snow and so on. The combination of what is influenced and the equipment to present the effect is called a **sensor**.
Field station	This is the outermost link in the data collection system. It is usually equipped with one or several sensors and is designed to operate in localities with no mains. Power is normally supplied by batteries or solar cells. Data transmission to the next link (the outstation) is usually done by simplex radio but can also be by line. Measurements and transmissions are triggered by commands from an internal clock on predetermined time steps, usually once an hour. The field station has to be able to withstand extreme climatic

512

variations. To conserve power at field stations, they are switched on only for the period from the measurement sequence is started until the data has been transmitted to the outstation.

Outstation An outstation is primarily a connecting link between one or several field stations, and the master station. It can, however, also have its own sensors similar to the field station. Since an outstation often will serve several field stations, it is sometimes also called a **concentration station**. It will normally be localized on places with mains and an established communication network, for example telephone or power line transmission. Normally an outstation is required to receive and store data from the fieldstation(s) for at least one week without loosing any data. Transmission of data to the master station is normally done by duplex communication via public telephone lines or through other communication channels. The data transmission from outstation to central station is initiated and controlled from the master station.

Master station The master station (sometimes also called *Central station* or *Base station*) is the final destination for data collected at fieldstations and outstations. The master station can establish communication with several ,outstations, either on regular time steps or upon user request. The master station will also normally have software for data quality control, error correction and a database for data storage and retrieval. The master station can also send data further to other computers, for example for use in runoff forecasting models.

Repeater station The repeater station is a station used to establish radio communication between a field stations and an outstation where a direct radio path is not possible. Data are received from the field station and sent on to the outstation without beeing changed. The signal strength and frequency may be changed, however. The repeater station is usually placed on hills or mountain tops. One repeater station may serve several field stations.

Modem Modem is an abbreviation for **mod**ulate/**dem**odulate. A modem is a device which allows computers to communicate along telephone lines. The modulator converts digital signals into audio frequency tones, and the demodulator transforms the audio tones back into digital signals in the other end.

2.3 Hydrological data for hydropower design and operation

2.3.1 Purpose of data collection

The purpose of the hydrological data collection can vary, but at least three different purposes can be identified:

* Runoff data for design of new plants
* Measurements for control and verification of concession rules given by the government
* Data for economical optimization of operation

The three different uses of data have different requirements for time resolution in data and speed in data acquisition. Data for operational use must generally be made available much quicker than for the two other types of use. Data which are beeing used for load or inflow forecasting must be collected in real-time or near real-time. This can only be accomplished by using automatic systems. For the two other uses manually data collection or off-line collection by the use of data-loggers or analog recorders may be cheaper and more reliable.

2.3.2 Local recording or remote transmission

Traditionally most hydrometric data has been recorded locally, and transmitted by mail for later processing. Remote transmission has mainly been used for monitoring water levels in reservoirs and other types of data needed for the control of gates and valves in hydropower systems. Local recording can be done by:

* Manual readings
* Analog recording
* Data loggers

Even if data are not needed for real-time use there may still be several advantages using an automatic system:

* Quicker access to data
* Easier processing of data
* Errors can be identified early
* Observers are hard to find and getting more expensive
* Increasing time resolution and number of parameters makes manual reading impractical

Due to these and other reasons the use of automatic systems tend to be increasing, even for data types that are not beeing used immediately. In many cases it can also be beneficial to extend an automatic system to include all types of measurements.

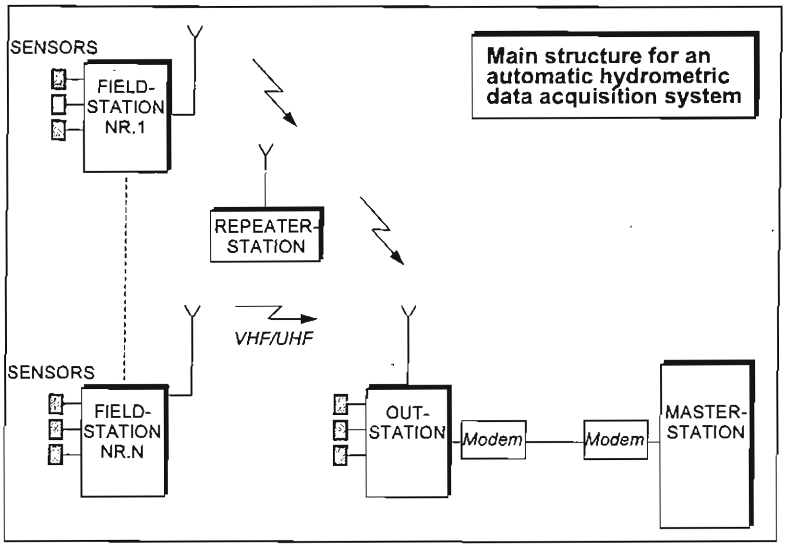


Figure 2.1 Main components in an automatic data acquisition system for hydrometric data.

2.3.3 Types of data needed

A number of different hydrological and meteorological parameters can be of interest, depending on the planned use. The most important types of parameters for use in runoff forecasting are underlined.

Meteorological parameters:
* <u>Precipitation</u>
* <u>Air temperature</u>
* Wind speed
* Solar radiation
* Air humidity

Hydrological parameters:
* <u>Stage</u>
* <u>Streamflow</u>
* Water temperature

- Snow depth
- Snow water equivalent
- Snowmelt
- Evaporation
- Ground water level

Precipitation, air temperature and wind speed (for correction of precipitation data) are commonly used as input to runoff forecasting models. Snow water equivalent and streamflow are used for model updating (Section 7.4). Radiation and air humidity can be useful for computation of evaporation and snowmelt.

For water quality monitoring and chemical mass balance studies a number of different water quality parameters can be of interest. Sensors for automatic measurements are available for parameters as:

Water quality parameters:

- pH
- Turbidity
- Conductivity
- Colour
- Light penetration
- Ammonium
- Nitrogen
- Phosphorus

2.4 Data transmission

Hydrometric data are usually collected at remote sites, far from the operational center where they are going to be used. Thus, transmission links have to be established to transmit data in an economical and reliable way. The user requirements specification will have to determine the degree of reliability needed for a given system. This may, in next turn, determine the most economical transmission type. A number of different transmission channels may be used for hydrometric data transmission:

- Hard wire links
    - Private land lines
    - Private (rented) telephone wires
    - Public switched telephone network
    - Power lines
- Radio
    - Radio line transmission
    - VHF/UHF-transmission
- Satellite telemetry

- * Meteor-burst telemetry

2.4.2 Hard-wire links

Private land lines requires the user to lay his own cables and maintain these. Such lines are only feasible over short distances or for control systems where a very high degree of reliability can be economically justified. For hydropower utilities the use of power line communication may be an alternative, especially if hydrometric data can be sent on existing communication channels via power lines.

Private (or rented) telephone wires provide a high standard and high capacity, but they are also expensive. For hydrometric data such lines are only feasible where the user requirements for speed and reliability are very high.

Public switched telephone network lines usually meets the transmission needs for most hydrometric users. By using automatic dialling equipment at the master station, a large number of outstations may be called sequentially from one telephone. Since the volume of data transferred normally is low (typically 1-10 sensors with 1 measurement each hour for each station) the data transfer from each outstation is very fast. A modem is needed at each outstation and at the master station. Today modems with transmission speed of 2400 and up to 9600 baud are readily available at a low price, and even higher speeds are becoming possible by special modems.

2.4.3 Radio transmission

Radio communication provides a good alternative to hard wire links, provided that frequency allocations are available. Both VHF and UHF equipment is used for data transmission. All radio transmission equipment must meet government performance specifications for parameters such as transmitter power, frequency stability, spurious emissions and signal distortion. It may be difficult to obtain a frequency allocation, especially if alternate communication links exists, for example telephone lines. In planning a radio-based system it is very important to do detailed studies of radio paths between the planned station sites. Much of this work can be done using topographical maps, but in marginal cases direct measurements of signal strength should be carried out. Radio equipment is vulnerable to ligthning strike, particularely where long arials are put on hill tops to obtain good line of sight. Radio transmission can be of simplex, semi-duplex or full duplex type:

Simplex	is data transmission where signals can be passed one way only.
Half duplex	is a two-way transmission where data can be sent both ways but not simultaneously.
Full Duplex	or two-way transmission of data. A full duplex link allows to transmit and

receive information simultaneously in both directions.

Simplex transmission is the simplest type, and is commonly used for radio transmission between a field station and the outstation. Half or full duplex transmission is used between master station and outstations, where the master station requests data and the outstation responds by sending back observed data. Full duplex transmission is most reliable, but also most expensive. With duplex communication two stations can exchange control information to verify that a message has been correctly transmitted. In case errors are detected, the receiving station can ask for retransmittal until the transmission is ok. For simplex transmission a good strategy is to retransmit a number of identical messages and let the receiving station compare them to each other. At least three messages should be used. Tests performed in mountainous terrain in Norway under difficult conditions has shown that this may improve communication reliability substantially.

2.4.4 Satellite telemetry

Hydrologists can use satellite telemetry to collect data from networks of remote hydrometric stations. Such systems are both reliable and cost effective, and there are almost no restrictions on the geograpic area that can be covered. Three operational satellite networks are now readily available for environmental data transmission, the ARGOS-system, the GOES-system and the METEOSAT-system.

The polar-orbiting ARGOS-system consists of three main components: The terrestrial radios that transmit environmental data, the radio receiver/transmitter on the satellite and the radio reception, processing and distribution system at receiving ground stations. The transmitting radios in the ARGOS-system are known as Platform Terminal Transmitters (PTT's). They are small, battery operated radio transmitters that cost approximately $2000 each (in 1993). They transmit short data messages in less than 1 s at random time intervals (100-200 s) at the frequency 401.650 MHz. When a message is received by the satellite it is immediately retransmitted at a different frequency, and can then be picked up by a Direct Readout Ground Station (DRGS). The message is also stored aboard and retransmitted later when one of the main receiving ground stations are within radio view. From here data can be sent by mail, telex or by modem to the user. The ARGOS-system has proved to be both reliable and economical, and has been used for hydrometric data transmission from many remote areas.

The GOES-system (Geostationary Operational Environmental Satellite) is operated by the United States. The data transmission operates in the same way as described for the ARGOS-system. Radio transmitters operating in the GOES-system are known as Data Collection Platforms (DCP's) and they can be operated in three modes: self-timed (simplex), random adaptive (simplex) and interrogated (duplex). The GOES-system is most widely used in North- and South-America.

The METEOSAT-system is operated by the European Space Agency (ESA). This system is functionally equivalent to the GOES-system, and offers the same type of services. By using plug-in modules data from up to 120 sensors can be collected, processed, formatted and transmitted by METEOSAT. The data can be received directly by the user by installing a 1.5 m dish antenna, or from the main groundstation in Darmstadt, Germany.

2.4.5 Meteor burst communication

This communication type utilized ionizes meteorite trails in the upper parts of the stratosphere (80-120 km) to establish a radio link. The ionized meteor trails are reflecting radio waves back to earth, making communications between stations up to 2000 km apart possible. This communication system is primarily used in remote areas where other communication types are not easily available. A large snow measurement program in USA (SNOTEL) use meteor burst communication to collect data from over 500 stations in the western part of USA.

2.4.6 Communication protocols

A communication protocol is a set of computing rules for ordering, transmission and error control for data communication. Both hardware and software protocols can be defined. The EDC-protocol (see below) is an example of a software protocol. The lack of standard communication protocols have long been a serious problem for environmental data collection systems. Different suppliers have developed their own proprietary protocols which usually can only communicate with thir own type of hardware. This can lead to difficulties where equipment from several suppliers are being used together, or when the user wants to start buying equipment from new suppliers. In Norway the hydropower industry has initiated the specification of two important software protocols for data communication to minimize such problems, the EDC-protocol and the ELCOM-protocol. Both are considered very important for the standardization of data-acquisition systems for hydropower utilities in Norway.

The EDC-protocol
The Environmental Data Communication (EDC) protocol is a standardized communication protocol for transmission of hydrological and meteorological data between outstations and the master station. The EDC-protocol is a layered protocol built on the OSI-model for open systems communication. The EDC-protocol has been tested in several projects and is now implemented in equipment from many suppliers of environmental data collection instruments in Norway. The hydropower industry in Norway usually specifies that the EDC-protocol must be included in all equipment for hydrometric data acquisition.

The ELCOM protocol is a standardized communication protocol built on the OSI-model for transmission of data between two computers, usually between a process control computer and a background computer. Many types of hydrometric data (eg. reservoir levels) are collected into the process control computer and then transferred to a background computer or master station by the ELCOM protocol. Examples of such systems are described in Section 4.5.

2.5 System design and operation

The main components in an automatic data collection system for hydrometric data were briefly described in Section 4.1. In this section the systems operational characteristics and the integration with other measurements in the power station and the remote control system will be discussed in more detail. It is assumed that all hydrometric data are going to be transferred to a master station for processing, quality control and use. The term *background computer* is used synonymously with master station in this context. The background computers in hydropower stations are used for short and long term operational planning, and for administrative tasks such as the preparation of reports and statistics. Data from the process control computer(s) are usually transferred to the background computer under the ELCOM protocol. Figure 2.1 gives an overwiev of a typical system.

A field station will usually be equipped with more than one sensor. At meteorological stations the standard sensors used are precipitation, air temperature and wind speed. To conserve power the field stations are usually only activated at regular intervals by an internal clock which triggers a measurement sequence. The measurements are usually performed once an hour. The field station scans each of its sensors in sequence, and performs the nessecary analog to digital conversion for analog sensors. The results are stored temporarily in the field stations memory. When all sensors have been scanned, data transmission is activated. The field stations are usually located at remote sites, where telephone connections can not be established, so radio transmission is the best alternative. Data transmission by radio from field stations to outstation is usually by simplex communication, and the field station can not know if data transfer were successfull. One simple method to improve data communication is to repeat the transmission several times. This will increase the probability of success if communication is hindered by temporal problems. The use of duplex communication will improve the transmission, but also lead to more expensive field station and higher power consumption. Data transmission between field stations and outstations is not standardized by a specified protocol yet, but in Norway a de facto standard has been established by the firm Aanderaa for this communication link.

When the transmission is finished the field station turn off all power consuming activities, and goes "into sleep" where only the internal clock is active.

The outstation will normally be operating continuously, waiting for data transmission from field stations or requests for data from the master station. If the outstation is equipped with sensors it will also activate a measurement cycle at regular intervals, similar to the fieldstation.

520

When a data transmission from a field station is detected, the outstation will analyze the message and verify if it is correct. If it is accepted a date and time stamp is added, and the message is stored in a rolling store memory. This memory usually has a capacity to store all incoming data from 7-10 days. The data are stored in anticipation of a request call from the master station. Data transmission from outstation to master station is normally initiated and controlled from the master station, but it is also possible to let the outstation establish communication. This can be beneficial to alert the master station in case of unexpected floods, unusual events or if functional errors are detected.

Communication between outstation and master station is usually by done by telephone, and here the EDC-protocol is used. Even if the communication lines or the master station is not functioning, no data is lost until the rolling store memory is full. Normally this happens first after 1-2 weeks. From then on the oldest data are lost by being overwritten by new data.

The master station is the final destination for all measurements. The master station can request data from all outstations sequentially, at fixed time intervals or when data are needed for special purposes. The master station can be a background computer in the hydropower station, a PC or a dedicated database computer.

Different system designs - three examples

The integration of a hydrometric data acqusition system in an existing hydropower system can be done in different ways, depending on user requirements and available communication channels. In Figure 2.2 three alternative and commonly used system designs are shown.

SYSTEM X:

In this system hydrometric data are beeing collected both by a special hydrometric data collection system with outstations and field stations, and by the process control system for the hydropower system. The background computer here uses the EDC-protocol to transmit data from outstations. Hydrometric data are also beeing collected by the process control system with three different solutions:

- Some sensors are connected directly to a remote control terminal. Only sensors with standard interface (0-5V/4-20mA/BCD code) can be used in this way
- One outstation is connected to a remote control terminal by BCD-code interface. The outstation can have its own sensors or receive data from fieldstations in a normal way.
- One outstation is connected directly to the process control computer. This requires that the process control supplier is willing to release specifications for his proprietary communication protocol (K1) between control terminals and process control computer, or vica verca.

Data are beeing transmitted from the process control computer to the background computer by the ELCOM protocol. Data from the two remote control terminals are transmitted to the process control computer by using the internal protocol K1.

SYSTEM Y:
In this system data from some hydrometric stations are beeing collected directly by the background computer, other by a separate master station. The background computer uses the EDC-protocol for data transfer from the outstation(s), but uses a special proprietary protocol (K2) for communication with the master station. By using the EDC-protocol data from many different outstation makes can be accessed without knowing details about the outstations.

SYSTEM Z:
In this system the master station can only communicate with its "own" outstations using the proprietary protocol K4. Data can be sent to a background computer by protocol K5 or displayed on a terminal or PC.

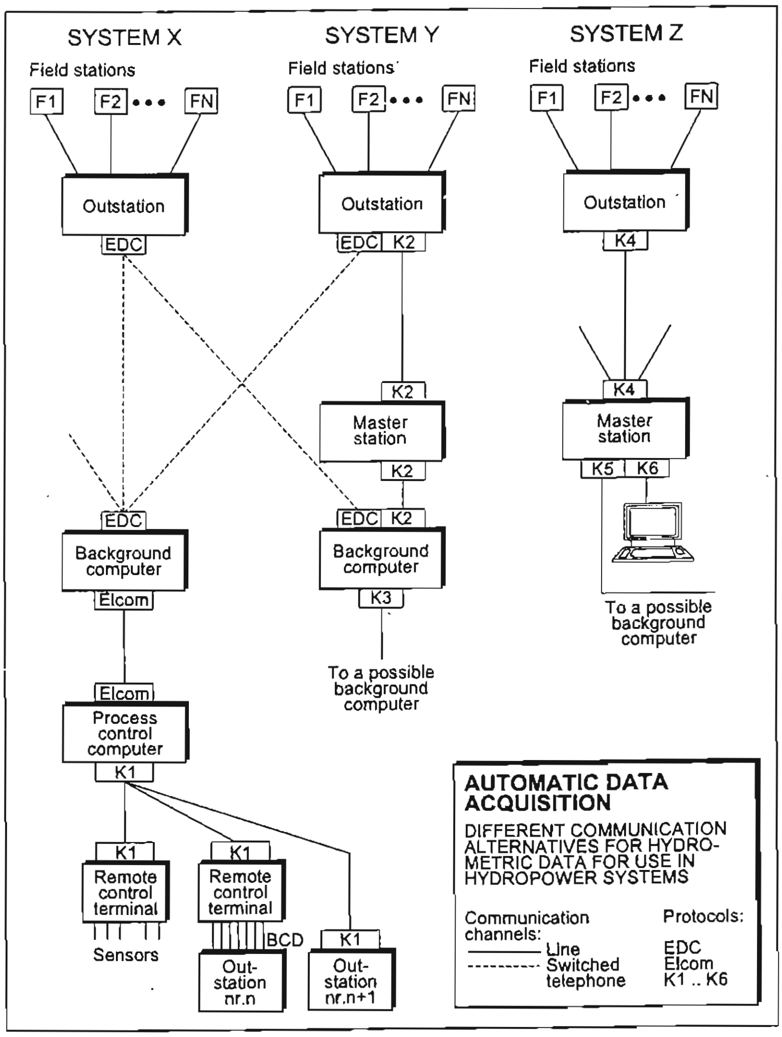


Figure 2.2 *Hydrometric data acquisition in a hydropower system. Three alternative solutions*

2.6 Hydrometric data for hydropower systems operation - 2 examples

The first test systems in Norway for automatic transfer of data to hydropower stations were put in operation around 1980. The first systems have since then been replaced by more modern equipment, some old systems have been closed down and a large number of new systems are under planning or has been put in operation. At present at least 20 automatic hydrometric systems are beeing operated by Norwegian hydropower companies, and the number is increasing steadily. Here two different systems that both have been in operation for several years are presented.

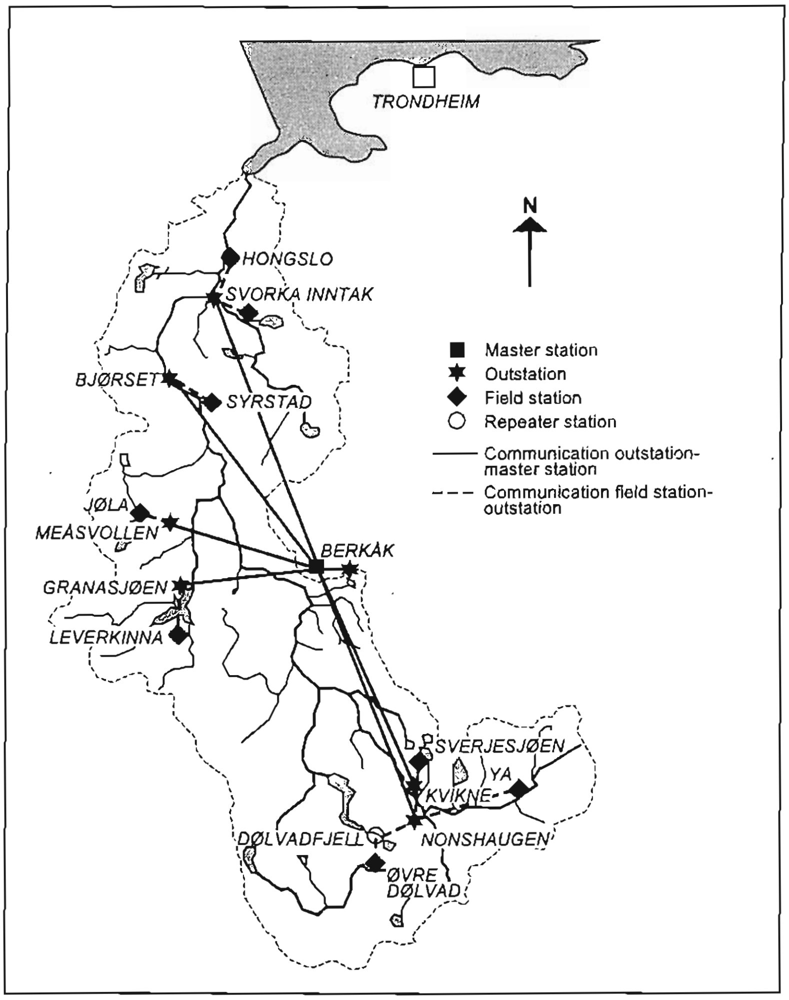
2.6.1 The Orkla hydropower system (KVO)

One of the first and most extensive automatic hydrometric data collection systems was installed in the Orkla river in Mid-Norway when the Orkla development with its 5 hydropower stations were finished around 1980. The hydrometric data system was planned to supply data for a number of runoff forecasting models, integrated with short and long term planning software. All data are transmitted automatically to the main control center at Berkåk, from where all the power stations in the river are controlled. The system consists of the following components (pr. 1/1-1988), (Figure 2.3).

- 1 master station in the control center at Berkåk
- 7 outstations, 3 with their own sensors connected
- 8 field stations
- 1 repeater station
- 48 sensors

The first version of the electronic equipment were installed back in 1981, and proved not to be reliable enough for operational use. It has later been replaced by more modern equipment. The fieldstations and outstations in the present system are made by the Norwegian firm, A/S Scan-matic. This firm has also delivered similar equipment to a substantial number of other power companies.

Data transmission from fieldstations to outstations is by simplex VHF-radio. One repeater station was nessecary to obtain radio communication from station ØVRE DØLVAD to the outstation at NONSHAUGEN. From outstations to master station public switched telephone is used, with automatic dialling equipment. Power supply is from solar cells and batteries at the field stations, and mains at the outstations and the master station. Most of the sensors are made by the Norwegian firm Aanderaa, except for the weighing type precipitation gauges which are made by Belfort (USA) and Geonor (Norway).



Figur 2.3 Automatic data acquisition system in the Orkla River

The Alta hydropower system

The Alta hydropower scheme is built and operated by Statkraft. From 1992 a new automatic hydro optimization system were put into operation, partly replacing an older system. It includes a large and very complex data acquisition system. (Hjertenes & al., 1992).

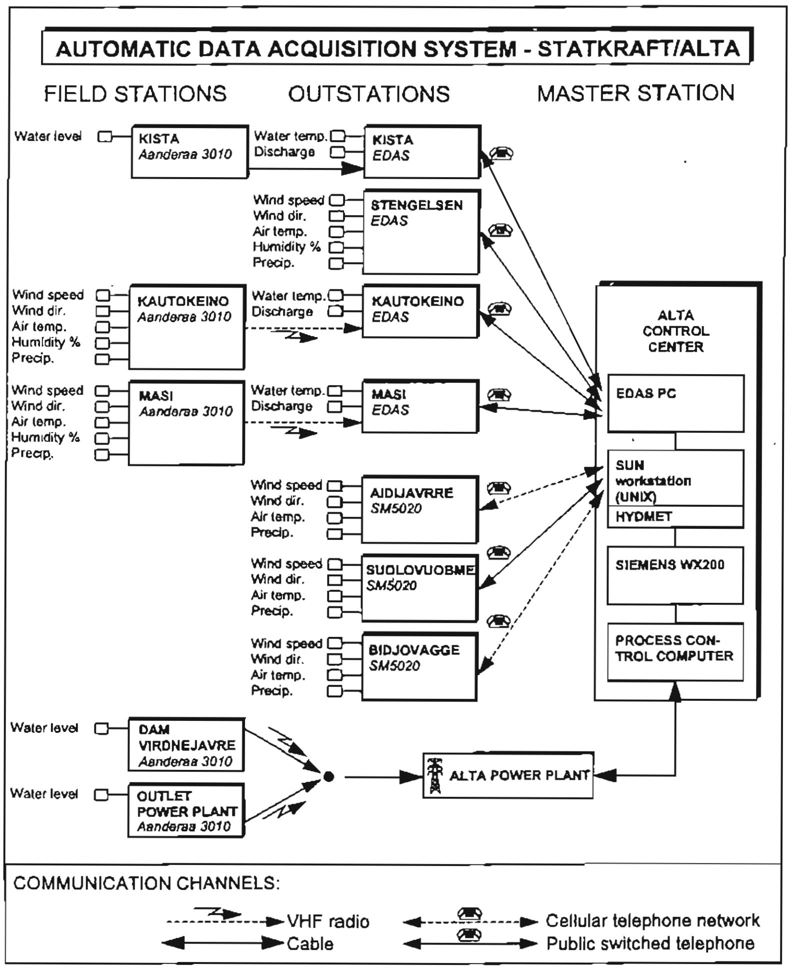


Figure 2.4 Main components in the Alta automatic data acquisition system

The system today consists of the following main components (Figure 2.4):

- The master station with HYDMET software and database (See Section 4.7) installed on a Sun workstation, an IBM-compatible PC and a process control computer from Siemens, all collecting data automatically from outstations and field stations
- 7 outstations (4 from EDAS, 3 from Scan-matic)
- 5 fieldstations (all from Aanderaa)
- 36 different sensors

The older EDAS-equipment was decided used even if the data communication became more complex. The master station is located in the main operational center for Altaverkene. All data are sent to the HYDMET database on a Sun workstation. Data from the old EDAS-system is transferred to HYDMET/Sun by two RT-programs, one putting data out on file and another reading data into HYDMET on Sun. Data from the power station and reservoir is transferred by the process control system from Siemens, and from the Siemens WX200 computer into HYDMET on Sun by ELCOM.

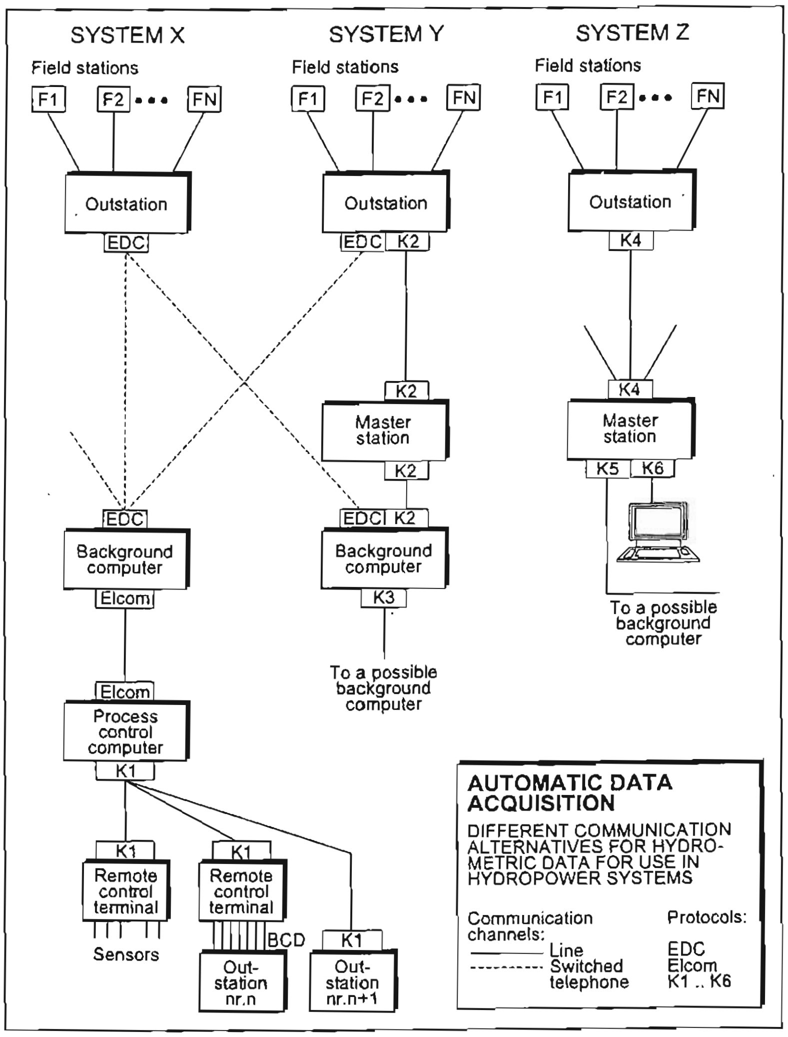
Data transmission

The data transmission paths are complex and a number of different transmission channels are used (Figure 2.4) to transmit data to the 3 different receiving computers at the master station in Alta control center.

The field stations in Kautokeino and Masi transmit data by VHF-radio to outstations located in a populated area. The field station at Kista (in the Alta river) transfer data to its outstation by line. The outstations at Kista, Stengelsen, Kautokeino and Masi all communicate with the old EDAS Master station by public telephone.

The outstations at Bidijavrre and Bidjovagge use cellular telephone for data transmission, while the Suolovuobme outstation send data by ordinary telephone. These three stations have no field stations and all are equipped with same type of sensors: Air temperature, Precipitation, Wind speed and Wind direction.

The field stations at dam Virdnijavre and in the river below tailrace outlet both are connected to the process control system in Alta power station. Data are transmitted together with the process control signals to the process control computer. From here data are sent by ELCOM to the Sun workstation.



Figur 2.4 Main components in the Alta automatic data acquisition system

2.7 Software

The hydropower industry in Norway has financed development of software for collection, transmission, storage, retrieval and presentation of hydrometric data during the last 10 years. In addition to the above components, software for automatic quality control and data correction has been developed. The complete software system is called HYDMET and has mainly been developed by EFI. The software has been tailored for hydropower users, with real time data collection for use of data in forecasting models. The data storage is based on relational database technology. From 1992 on implementations for SYBASE and ORACLE will be available, using standard SQL query language.

2.7.1 Software components

The HYDMET system consists of a number of independant but interconnected modules as shown in Figure 2.6. The HYDMET system can collect data using either by the EDC-protocol or the ELCOM-protocol (See Section 2.4.6). The main features in the system can be grouped into the following 4 sections:

The database
This is where all data are beeing stored. Five different types of data can be stored in the database today:

- Sensor calibration data
- Time series of observed and/or corrected data
- Data for the hydraulic system (reservoir, gate capacity, etc.)
- List of users
- Report definitions

Time series can be stored as hourly, daily or weekly values, or as irregular timeseries where time of observation is stored for each value.

Data collection
HYDINN is the main data collection program. By using the EDC-protocol data can be collected by public switched telephone or other communication lines either automatically or on demand. The HYDINI/HYDRES programs are used for data collection from process control computers by using the ELCOM-protocol.

Preprocessing/Quality control
In order to be able to use data for forecasting in real time, an automatic method for quality control and data correction had to be implemented. These functions are performed by the HYDAKO module. A number of tests can be made automatically to control data collected into the HYDMET system:

- Test for permitted maximum and minimum values
- Test for permitted changes in data values
- Test for repeated constant values, for example from a frozen wind speed sensor or water level sensor

If missing or erronous values are detected, several methods for data correction or filling in missing data are available:

- Use average value between previous and next value
- Use a predefined constant value
- Use previous accepted value
- Use value computed by multiple regression from observations at several nearby stations
- Use value computed from autocorrelation in observed values

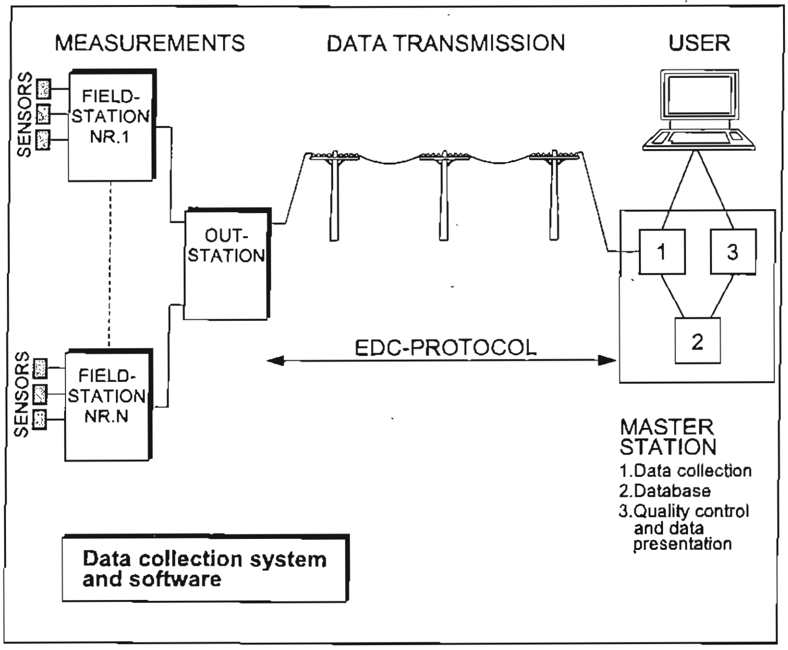
Other types of computations can be done by the module HYDBER. This can be any type of routine computations, for example water level to runoff or inflow to a reservoir based on water level and reservoir release.

Reports and data maintainance

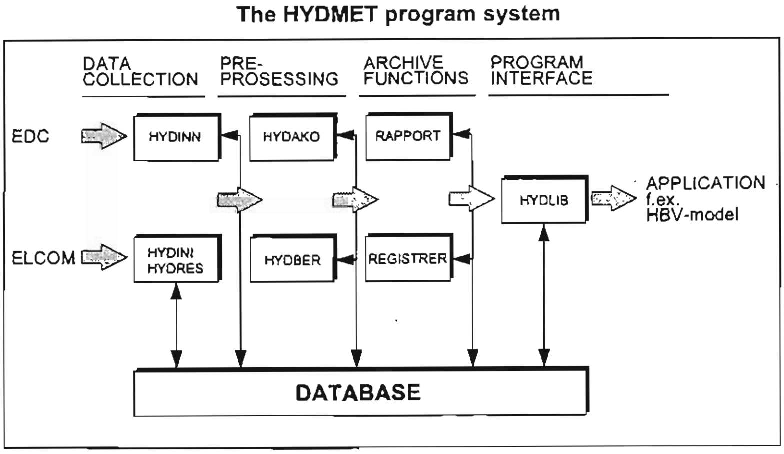
Standard reports can be defined by a report generator and stored in the database. Such reports, as text or graphically, can be produced at predetermined intervals or on demand. Data can also be inspected and edited manually by a separate module (REGISTER). This module can also be used to enter data manually into the database.

Subroutine interface

A subroutine library (HYDLIB) has been developed for use by programs that needs to interface the HYDMET-database directly. This is used for example by the HBV-model for model data input and updating. This library can also be used to develop special reports, not avaliable in the standard report generator, or in any other program needing data directly from HYDMET.



Figur 2.5 Data collection system and software components



Figur 2.6 Main components in the HYDMET software system

2.8 Recommendations

When an automatic data acquisition system is beeing planned, some of the most important things to consider are found to be:

- Try to coordinate measurements for different purposes (historic and operational) within the same system
- Coordinate meteorological and hydrological measurements
- Work out a detailed user specification for the system *before* detailed planning is started
- Make plans for system operation, maintainance and sensor recalibration *in advance*

User requirements specifications
The preparation of a detailed user specification is an important first step in the design of an automatic data acquisition system. The users of the system has to identify the essential goals for the system, and decide on standards for reliability in operation. A user specification should identify what the user wants from the system, not details about the equipment. Possible future needs should be taken into consideration, as for example the need for increased number of sensors or increased time resolution. Some of the main steps in the user specification are commented below.

Accuracy requirements
In Table 4.1 a selected list of hydrological and meteorological parameters are presented, together with recommended units and accuracy requirements as defined for standard meteorological and hydrological observations in Norway. Generally it is recommended to adhere to these standard requirements, even if the accuracy is not needed at present for the planned use. The reason is that data series may in the next run be used in other types of applications, new model developments or in combination with other standard meteorological observations. In these applications the accuracy requirements may be higher then in the first applications.

Time resolution
For most purposes a time step of 1 hour is adequate. Only in very rapid rivers and small catchments (eg. urban areas) measurements more often than this may be needed. Increasing the time resolution will increase the total volume of data, and put more strain on database and quality control systems. If integrated or averaged values within the timestep are needed, the sensors may have to perform continuous measurements and process them to compute such values. Examples are global radiation, precipitation and snow melt where hourly sums are needed, or wind speed and air temperature where hourly averages are needed.

Environmental considerations
The operating environment for electronic equipment has to be considered carefully. Some

of the most important environmental parameters to be considered are:

- Temperature (low/high/cycling/shock)
- Air humidity (condensating)
- Shock/Vibrations (during transport, installation and operation)
- Biological effects (insects, rodents, fungus)
- Chemical effects (Salt spray, acid precipitation,..)
- Electromagnetic interference

Power supply

Field stations are always required to operate from local power supply, using internal batteries combined with solar cells or wind generators. For field stations with low energy consumption the use of internal batteries can be the most economical solution. With higher energy consumption the use of solar cells has proved to be reliable and economical. Some types of field stations, eg. from Aanderaa can be equipped with solar cells integrated in the station hardware, making the installation and operation very simple. When using solar cells at high latitudes the available solar radiation during the winter periode has to be considered carefully. Snow and freezing rain may clog the solar cells and reduce power capacity. The solar cell output and/or battery voltage can be monitored and transmitted together with sensor data for control.

At outstations and in the master station a mains power supply is usually available. Still, a backup power supply may be needed. Often one of the most frequent failure rates is due to electricity mains failure. A battery backup with automatic switching and trickle recharging may be needed to ensure reliable operations.

Maintainance and sensor calibration

Both manual and automatic systems needs regular control and maintainance to ensure a high quality in data. It is important to make such plans in advance, to make sure that manpower and money are available to perform regular inspections and preventive maintainance. Most types of sensors needs regular maintainance and recalibration, especially sensors with mechanical parts like a wind cup anemometer. The Norwegian Meteorological Institute (DNMI) has given recommended intervals between recalibration for most types of sensors mentioned here. These recommadations are listed in Table 4.1 together with other specifications for different sensor types.

Documentation

A good documentation of the system is vital for maintainance and repair of the system. Make sure that all details about sensors, cables, electronics, software and power supply are documented and kept available for later use. Equally important, update the documentation when the system has been changed!

Table 2.1 Recommended units and standards for some hydrometeorological parameters. Adapted from Colderup Jensen & al (1992) and WMO (1970))

Parameter	**Unit**	**Range**	**Accuracy**	**Resolution**	**Recalibration interval (Years)**
Water level (stage)	cm	-	± 1 cm	1 cm	-
Stream discharge	m^3/sec	-	-	-	-
Snow depth	cm	-	± 1 cm	1 cm	-
Snow water eqivalent	mm	-	± 5 mm	1 mm	-
Evaporation	mm	0-100 mm/d	±0.1 mm < 10 mm/d ±2% > 10 mm/d	±0.1 mm	-
Air temperature	°C	-60-+60 °C	±0.1 °C	±0.1 °C	2 years
Precipitation	mm	0-300 mm	±0.1 mm < 10 mm/d ±2% > 10 mm/d	±0.1 mm	0.5 year
Radiation	watt/cm^2	300-3000nm	-	-	1 year
Wind speed	m/s	0-75 m/s	± 0.5 m/s< 5m/s	0.1 m/s	Group test
Wind direction	°	0-360°	± 5°	5°	Group test
Relative humidity	%Rh	10-100%Rh	±5% (<50%) ±2% (>50%)	1%Rh	2 years
Atmospheric pressure	hPa	Average ±100 hPa	± 0.1 hPa	0.1 hPa	3 years

Excursion 1

Thursday 20 October

a) Sagelva

Leader: Aanund Killingtveit, NTH

Excursion 1
Thursday 20 October

a) Kaldvella

Leader: Gaute Storrø, NGU

Description of the Kaldvella catchment

Gaute Storrø
Geological Survey of Norway

TOPOGRAPHY

The Kaldvelldalen valley has a NE-SW orientation and is connected to the Selbusjøen and Gauldalen valleys near Ler. The valley is filled with large sand and gravel bodies. Langvatnet lake lies on the large valley plain and has no visible outlet. The Kaldvella river originates from the marsh areas on the south side of the valley. The river has eroded a large ravine in the valley sediments. The edge of the plain lies at an elevation of 175m above sea level (a.s.l.) and projects 50-75m above the underlying clay sediments.

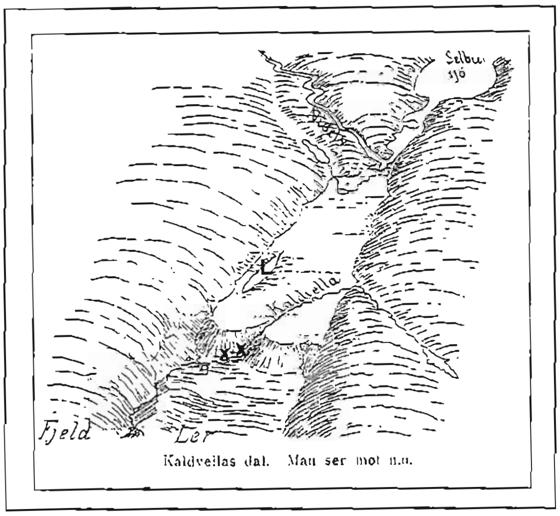


Figure 1: "*The most remarkable feature in theKaldvell valley is the lake which lies on the terrace surface, Langvatnet (L); it resembles the Dead Sea in that it has no outlet.*." *X=springs* Quote and drawing from Hans Reusch, 1919.

BEDROCK GEOLOGY

The bedrock in the Kaldvella catchment consists of a cover of Cambro-Silurian rocks thrust over Precambrian rocks. Between the upper allocthonous rocks and the underlying autochthonous rocks is a thin layer of sedimentary rocks of local origin.

The allochthonous rocks are divided into an upper and a lower unit. The upper unit is designated the Trondheim nappe. In many areas, this unit is cut by intrusive rocks such as Trondhjemites and gabbroes (possibly of Caledonian age).

These rocks belong to the Gula-nappe and are dominated by hornblende/biotite schists, migmatite gneisses and phyllites/greywackes. In the east are rocks belonging to the Støren/Meråcker nappe, together with rocks from the lower nappe unit. In Kaldvelldalen itself, shales and greywackes dominate.

The principal minerals for the above named rocks are quartz, biotite, muscovite, plagioclase feldspar, potassium feldspar, proxene and chlorite. A characteristic common to most of the rock types in the area is that they have a high calcite content. Pyrite is also present in many of the rocks. All the rocks in the area, apart from the migmatite gneisses and the mylonites, are soft and easily eroded.

QUATERNARY GEOLOGY

The extent of the Scandinavian ice sheet during the last Ice Age (the Weichselian) is shown in Figure 2. This figure shows the location of the Trondheim-Gauldalen region relative to the ice sheet. It can be seen that ice-margin deposits from the Younger Dryas Chronozone are a major feature of the regions Quaternary geology.

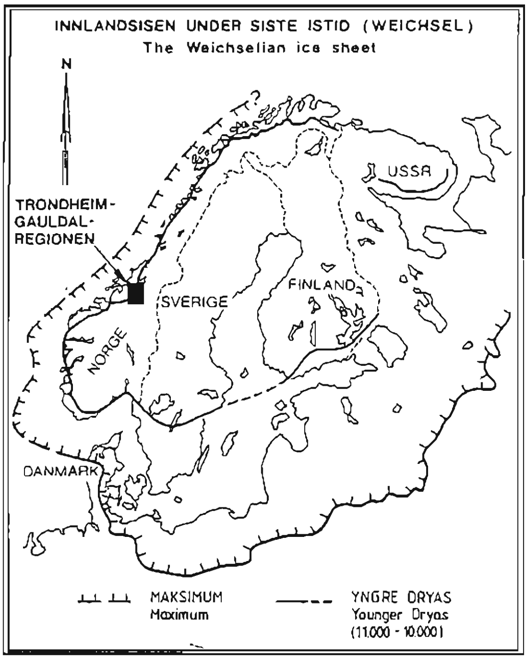


Figure 2: *The inland ice sheet during the last ice age*

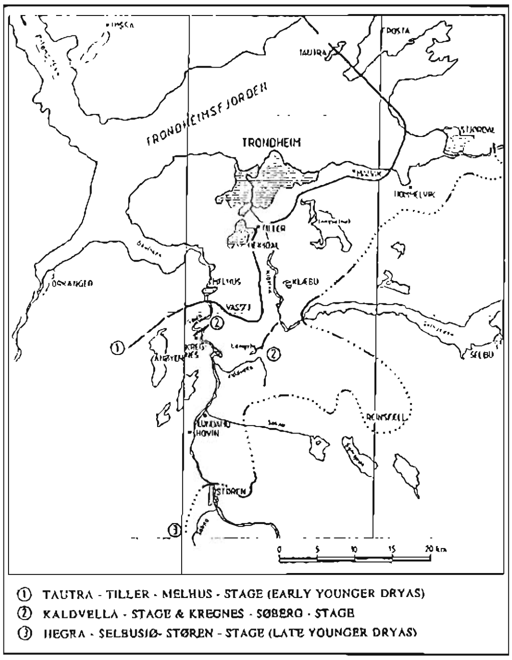


Figure 3: *Location map, the Trondheim region*

A series of distinct ice-margin deposits from the Ra Chronozone are present in the Trondheim region. Such deposits are also found in Gauldalen; the Melhus Ridge is interpreted as an ice-push moraine from the early Younger Dryas. The glacier calved further southwards in Gauldalen to Kregnes-Søberg, where the large fluvio-glacial deposit indicates a relatively long residence time. As the thickness and gradient of the ice decreased, the glacier calved further south to Hovin where fluvio-glacial terraces mark a local hiatus in the glacier retreat.

Marginal moraines at different locations within the region indicate that there were still significant glacier remains west of Gauldalen and probably also in the east in Kaldvelldalen and the Selbusjøen basin. At the time when these ice-margin deposits were being formed, the sea level in Gauldalen lay at approximately 175-180m a.s.l.. Figure 3 shows an overview of the Trondheim region.

FORMATION OF THE GLACIAL DEPOSITS IN KALDVELLDALEN

The main features of Kaldvelldalens Quaternary geology are shown in Figure 4. The superficial deposits can be divided into three units:

1) The terminal ridge northeast of Langvatnet
This ridge rises 30-40m above the surrounding landscape and has a moraine-like surface in terms of lithology and geomorphology. Under this moraine-like upper layer is a sorted fluvio-glacial deposit with a distinct southwest-dipping cross-stratification This fluvio-glacial deposit is clearly influenced by ice movements and consolidation at the ice-front. The ridge has an elevation of 225m a.s.l., i.e. sedimentation occurred at least 40-50m above the marine limit.

2) Sandur deposits from the esker to the southern end of Langvatnet
The surface of the sandur lies at an elevation of approximately 200m a.s.l. near the terminal ridge and falls towards the southwest to a level of 180m a.s.l. by the shore of Langvatnet. The surface has numerous kettle holes and meltwater and run-off channels. The biggest kettlehole, Langvatnet itself, is more than 50m deep.

3) The fluvio-glacial delta
The delta is built out from the southern end of Langvatnet and extends approximately 2km southwest. The deltas upper surface coallesces with the marine limit (175-180m a.s.l.).

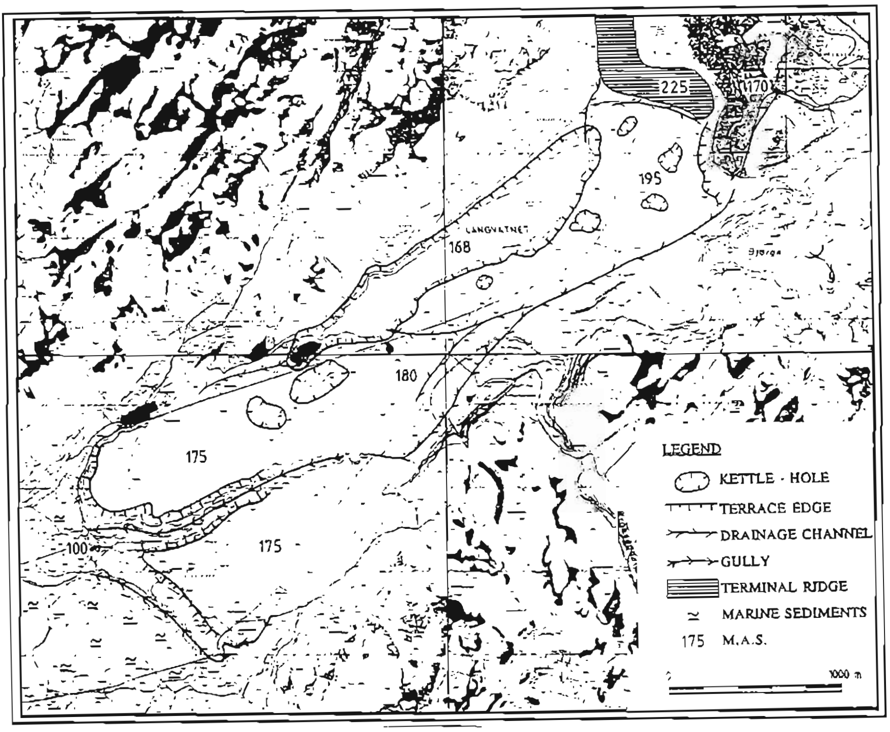


Figure 4: *The principal geomorphological and Quarternary geological features in the Kaldvelldalen valley.*

DEGLACIATION

It is not possible to determine the exact location of the early Younger Dryas step in Gauldalen. This makes the reconstruction of the deglaciation in the area difficult. Two possible deglaciation models are presented here, Models I and II (Figures 5 and 6 respectively). The ice-front locations shown are approximate. Lines mark the junction between the marginal deposits in Gauldalen and Kaldvelldalen.

Model I

A) During the early Younger Dryas ice-advance, meltwater drained northwards from the Nea valley and the Selbusjøen basin depositing the large fluvio-glacial sediment bodies near Tiller (Figure 3). The fluvio-glacial deposits and drainage channels on the eastern side of Vassfjellet indicate that this drainage direction was maintained after the glacier thickness had decreased. The central Vassfjellet massif area is assumed to have been ice-free by the time of this ice-advance.

B) After the ice- advance in the early Younger Dryas, the glacier calved in Gauldalen to Kregnes where a large marginal step formed. The ice-front was located in the Selbusjøen basin near the northern end of Kaldvelldalen. An ice-dammed lake formed between the glaciers in Gauldalen and the Selbusjøen basin. Marginal deposits built up at the Selbusjøen glacier front in Kaldvelldalen. Drainage of the ice-dammed lake must have occurred towards Gauldalen, i.e. northward drainage was prevented by ice. The local topography supports such a conclusion.

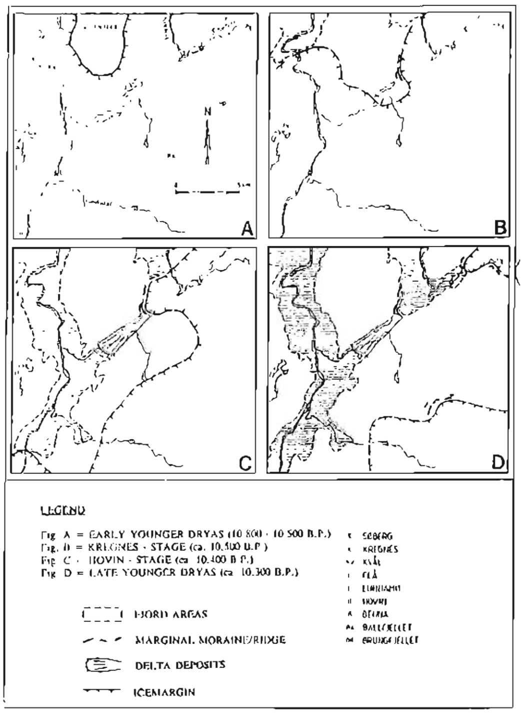


Figure 5: *Assumed course of the ice-fronts retreat, Model I*

C) Favourable conditions for glacier calving (i.e. a deep fjord) existed from Kregnes to Hovin. This enabled a relatively fast ice-retreat from Gauldalen. An open fjord arm was formed leading to Kaldvelldalen. The potential for the glacier calving and retreat was significantly less in the shallow water area at the western end of the Selbussjøen basin. The ice-front east of Vassfjellet is therefore interpreted to have essentially been stagnant during this phase of the deglaciation. In the area of the ice-front in Kaldvelldalen, sediments were deposited between and on the ice remains. A sandur containing dead-ice blocks and a delta were formed. The elevation at which deposition took place was controlled by the marine limit in Gauldalen

D) During the ice-advance in the late Younger Dryas (approximately 10 300years BP), the ice-front in Gauldalen advanced to Kvasshylla, north of Støren. The long and relatively deep fjord arm in Gauldalen enabled the deposition of thick marine sediments. In the Nea valley and the Selbusjøen basin, the ice advanced to the western end of Selbusjøen where terminal moraines were deposited.

Model II:

A) The marginal moraines in the Kregnes-Søberg area belong to the early Younger Dryas. Otherwise the conditions are as described for Model I A).

B) The Gauldalen glacier calved southwards towards Flå and probably up to the marginal deposit at Hovin before or during the formation of the terminal ridge in Kaldvelldalen. The glacier lake in Kaldvelldalen was caused by remaining ice blocks south of Vassfjellet. The melting of these ice blocks and the sand scree/delta deposit took place with sediment transport towards Gauldalensfjord. This model also assumes the prevention of drainage to the north and consequent drainage of the glacier lake towards Gauldalen.

C) The last phase in the deglaciation is as described for Model I D).

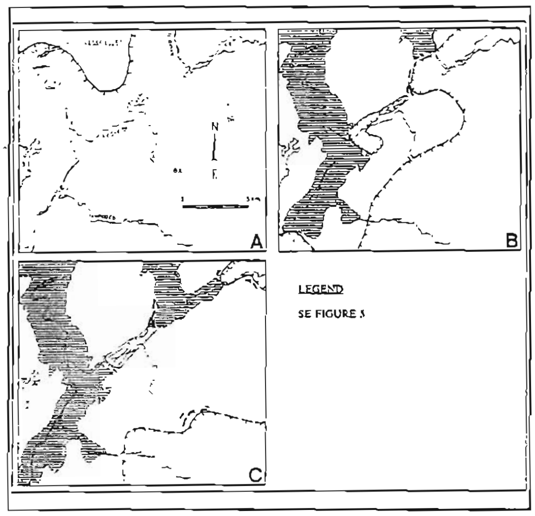


Figure 6: *Assumed course of the ice fronts retreat, Model II.*

HYDROLOGY

The total area of the surface water catchment draining towards the Kaldvell deposit is approximately 17,5 km^2. The actual catchment boundaries are uncertain in the area along the moraine ridges in the northeast and along the delta front in the southwest. This uncertainty is caused by the large sediment thicknesses and the potential this gives for hidden run-off through groundwater flow. Near the deltas steep edge in the southwest, groundwater emerges from springs spread over a large area.

The catchment area is characterized by a shallow moraine cover and a thin heather/peat vegetation. The delta surface and the marsh area each form approximately 20% of the total area. Langvatnet is the only large lake in the catchment area and forms 3% of the total area. The average elevation of the area is 330m a.s.l.. The Kaldvellbekken stream is the only large watercourse in the area with a total length of 7km. River water infiltrates into the groundwater as the stream crosses the delta.

Calculations show that the infiltration of precipitation and snowmelt on the delta surface contributes approximately 35% of the total groundwater recharge whilst infiltration from Langvatnet and Moatjern (a small shallow tarn) together contribute approximately 60%. The remaining 5% comes from infiltration along the Kaldvellbekken stream.

The mean annual rainfall measured at Løksmyr DNMI-station which lies approximately 1,5km northwest of Langvatnet is 1349 mm. The mean runoff for the valley sides in Kaldvella is 845mm. Water balance studies in the Sagelva area gives a long-term mean annual evapo-transpiration of 438 mm. It is assumed that this number also represents a mean value for the higher areas of the Kaldvella area.

HYDROGEOLOGY

The hydrogeological mapping is concentrated on the Kaldvell delta aquifer. Monitoring of the groundwater levels within the aquifer was carried out during the period November 1984 to February 1989. The aquifer shows extremely regular seasonal variations with the lowest annual water level occurring in the middle of April, the highest level occurring after the Spring thaw in the middle of June followed by another decrease before the Autumn rains in the middle of September. The timing of the highest water level after the Autumn rains varies to some extent, but typically occurs in the middle of November.

The groundwater level fluctuations through the winter season are approximately the same from year to year (2,1-2,3m). Assuming an effective porosity of 30%, this represents a reduction in the groundwater volume of 630-690 l/m^2 (mm) through the winter season. The aquifers seasonal variations show good agreement with type curves for the Lowland/Østland area (see Figure 7).

The temporal variations of the groundwater levels at Kaldvella are extremely similar to those of a moraine aquifer in the Sagelv area measured in the period 1973-89, despite the differences in aquifer type. It is likely therefore that the Kaldvell aquifer has followed the same long term trends as the aquifer in Sagelva. Using this assumption as a basis, the long-term mean groundwater level for the aquifer at the measuring point on the deltas surface was calculated as -23m, with minimum and maximum values at the same point of -26m and -19m.

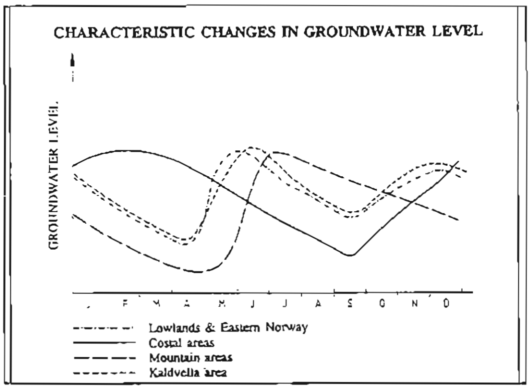


Figure 7. *Type curves for annual variaton in groundwater level.*

Figures 8 shows a vertical profile through the Kaldvell delta whilst Figure 9 shows a contour elevation map for the groundwater table. The groundwater elevation map shows the average groundwater level for 1988 which corresponds to the long-term mean groundwater level for the aquifer. The groundwater gradient increases gradually through the flow system from 0,4 mm/m in the northeast to 45 mm/m towards the delta front in the southwest. The reason that such a high gradient can be maintained over an extensive area can be seen in Figure 8.

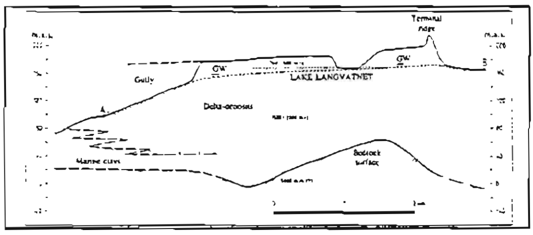


Figure 8: *Vertical profile A-B through the Kaldvell deposit. The location of the profile is shown in Figure 9.*

The ravine was formed at a time when the groundwater was at its original level. It acted as a drainage channel. Groundwater flow takes place towards the ravine, but drainage is hindered by a low permeability upper layer in the ravine together with fine-grained sediments which lie in front of and partly under the shelf of the delta. This results in a build up of excess pressure in the aquifer, which results in high groundwater gradients towards the ravine and the formation of springs and spring horizons at a level 3-4m over the level of the Kaldvellbekken stream.

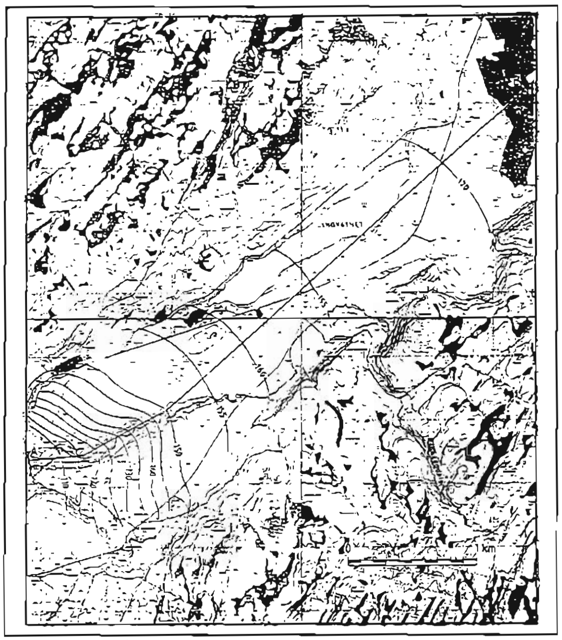
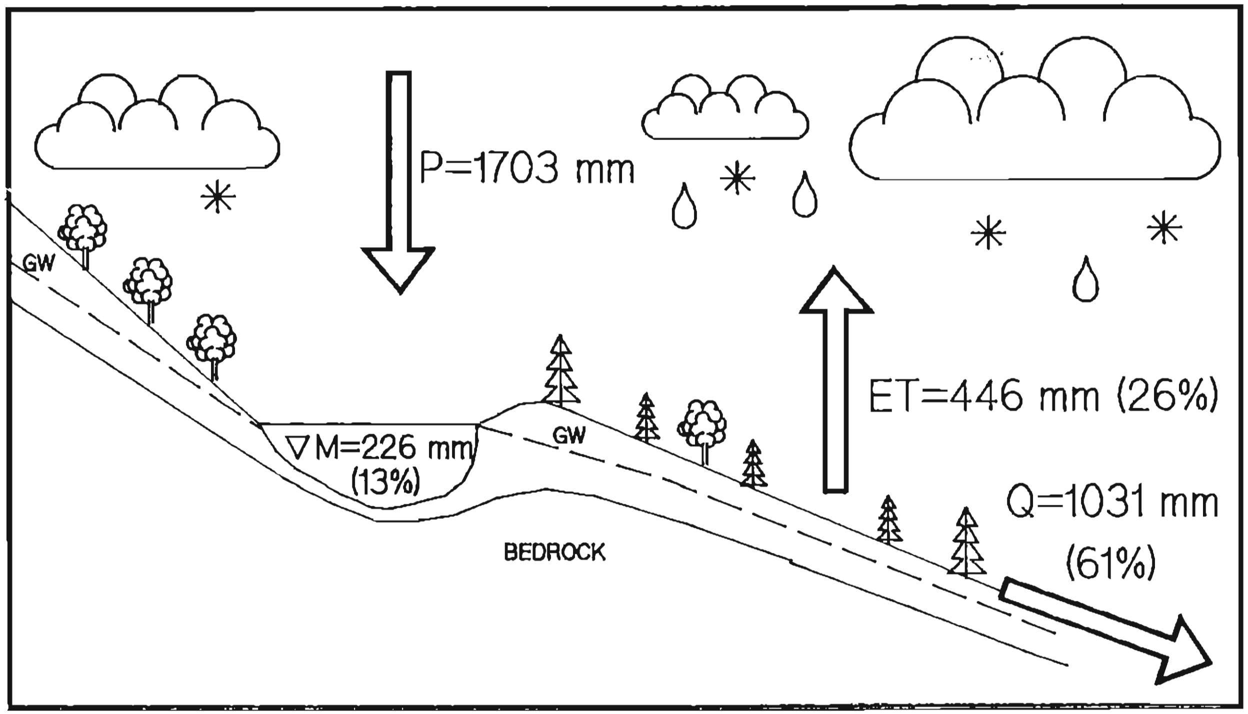


Figure 9: *Contour elevation map of the groundwater table in the Kaldvell delta. The contour lines give groundwater elevation in m a.s.l..*

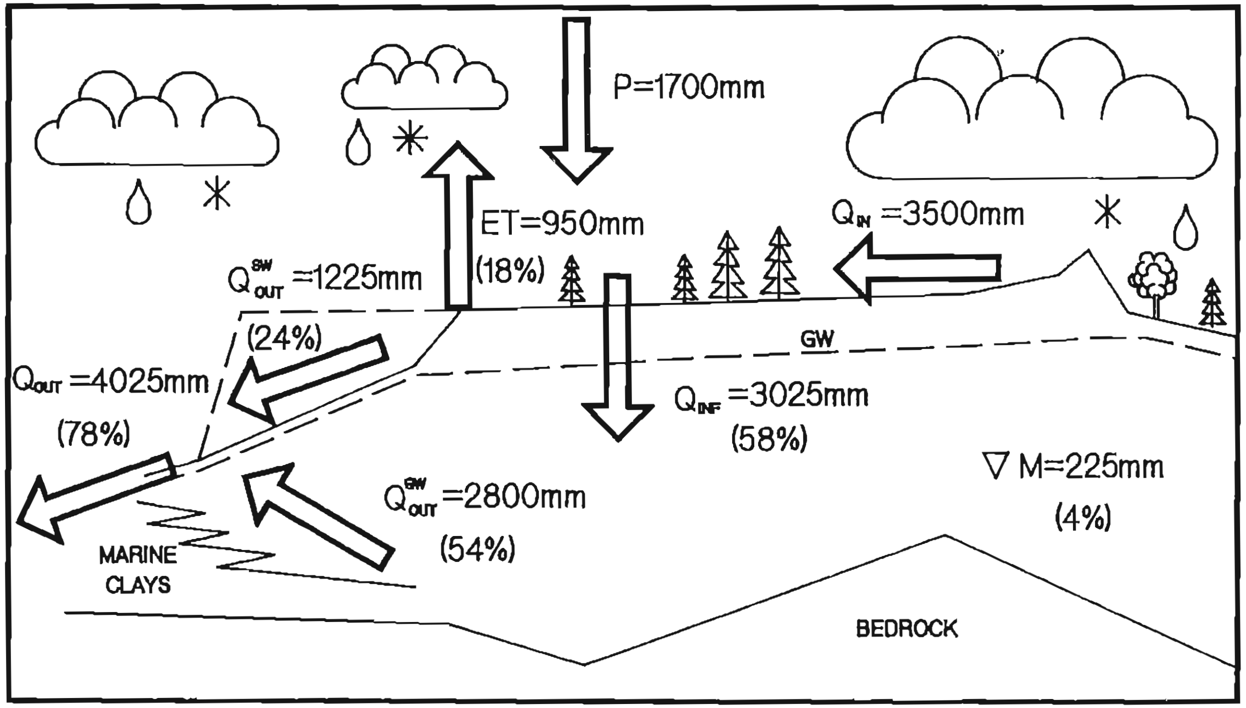
WATER BUDGET OF VALLEY SIDE AREAS,
KALDVELLA CATCHMENT 1988.



$$P = Q + ET + \nabla M$$

Total area = 13,5 km^2

WATER BUDGET OF DELTA AREA, KALDVELLA CATCHMENT 1988.



$P + Q_{IN} = Q_{OUT} + ET + \nabla M$

$Q_{OUT} = Q^{GW}_{OUT} + Q^{SW}_{OUT}$ (GW=GROUND WATER, SW=SURFACE WATER)

$Q_{INF} = Q^{SW}_{OUT} + \nabla M$ (INF=INFILTRATION)

Total area = 4,0 km^2

CHEMICAL BUDGET OF VALLEY SIDE AREA, KALDVELLA CATCHMENT 1988.

	NaCl 10^3kg	SEASALT SO4 10^3kg	1M H2SO4 m^3	1M HNO3 m^3	MgCO3 10^3kg	CaCO3 10^3kg
FROM PRECIPIT.	72.3	7.8	145.2	141.8	12.2	4.0
TO MAGAZINE	9.4	1.0	18.9	18.4	1.6	0.5
TO RUNOFF	67.2	8.5	158.1	39.2	24.1	125.1
CHANGES (+/–)	+ 4.3	+ 1.7	+ 31.8	- 84.2	+13.5	+ 125.1

WEATHERING:
3 X 10^3 kg S as FeS2 = 0.2 g/m^2
14 X 10^3 kg Mg as MgCO3 = 1.0 g/m^2
122 X 10^3 kg Ca as CaCO3 = 9.0 g/m^2

DRY FALLOUTS: 4.3 X 10^3 kg NaCl = 0.3 g/m^2

NITROGEN CONSUMPTION: 1.2 X 10^3 kg N as N = 0.1 g/m^2

Total area 13.5 km^2

CHEMICAL BUDGET OF DELTA DEPOSITS, KALDVELLA CATCHMENT 1988.

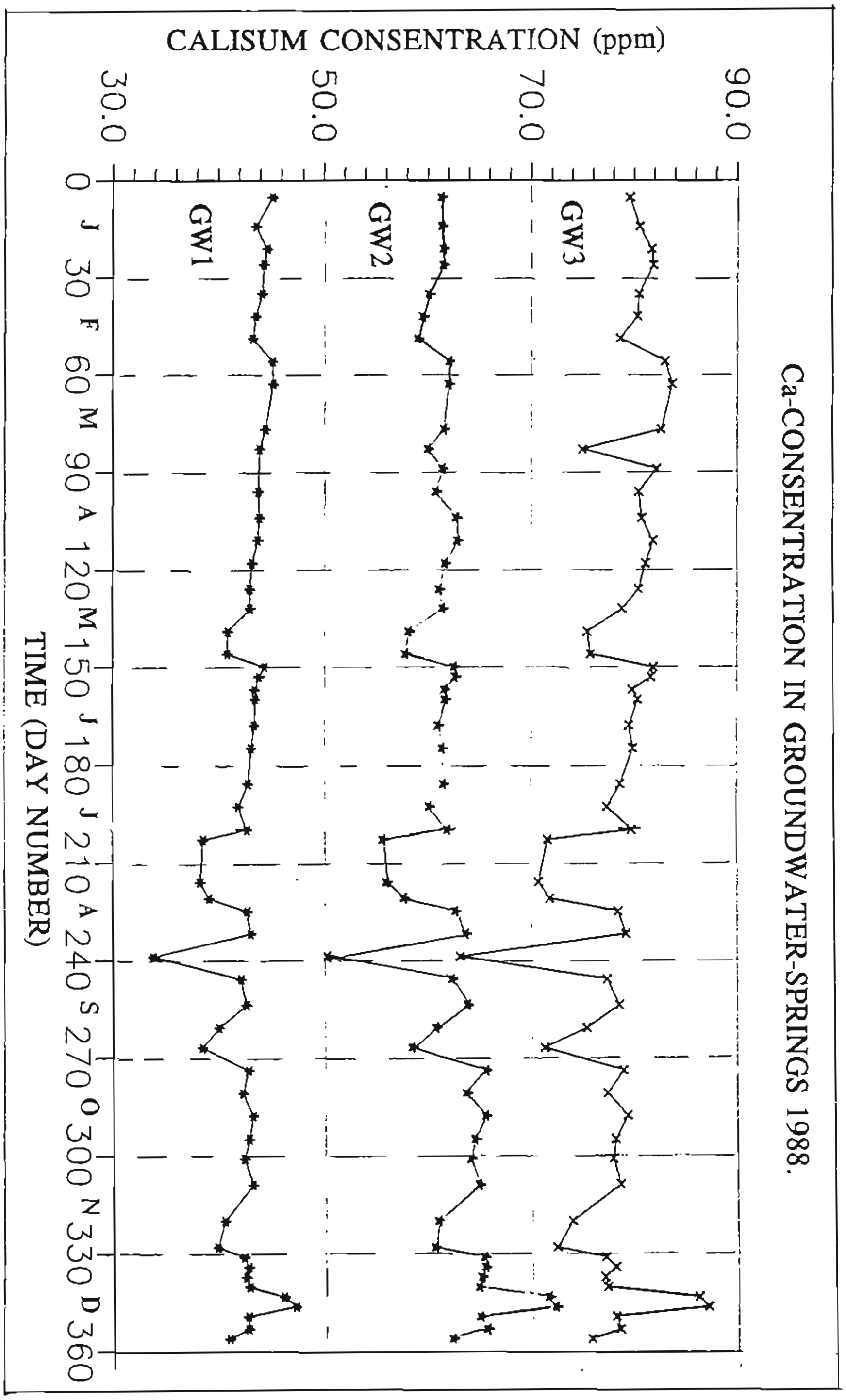
		NaCl 10^3kg	1M H2SO4 m^3	1M HNO3 m^3	MgCO3 10^3kg	CaCO3 10^3kg
INFILTRATION	PRECIPITATION	18.6	58.5	36.6	3.1	1.0
	SURFACE WATERS	43.7	161.8	26.0	15.8	81.8
GW RUNOFF		105.2	1376.4	175.9	160.8	1267.0
CHANGES (+/–)		+42.9	+1156.1	+113.3	+141.9	+1184.2

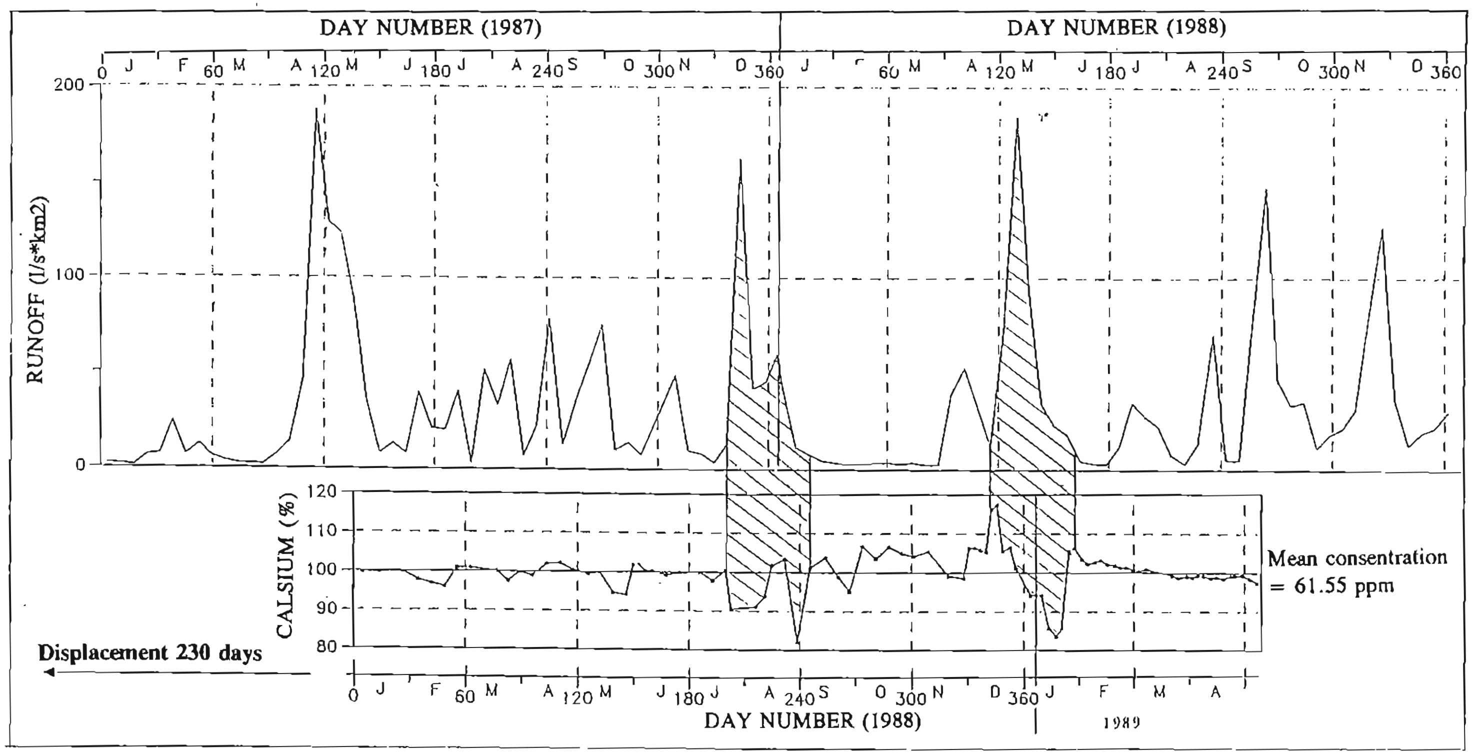
WEATHERING:
69 X 10^3 kg S as FeS2 = 0.7 g/m^3
142 X 10^3 kg Mg as MgCO3 = 1.4 g/m^3
1184 X 10^3 kg Ca as CaCO3 = 11.6 g/m^3

SEASALT OUTWASH: 43 X 10^3 kg NaCl = 0.1 g/m^3

NITROGEN POLLUTION: 1.6 X 10^3 kg N as N = 0.4 g/m^3

Totale volume
unsaturated zone
= 102 mill. m^3


Ca-CONSENTRATION IN GROUNDWATER-SPRINGS 1988.
CALISUM CONSENTRATION (ppm)
90.0
70.0
50.0
30.0
GW3
GW2
GW1
0 J 30 F 60 M 90 A 120 M 150 J 180 J 210 A 240 S 270 O 300 N 330 D 360
TIME (DAY NUMBER)


DAY NUMBER (1987)
DAY NUMBER (1988)
RUNOFF (l/s*km2)
0
100
200
0 J F 60 M A 120 M J 180 J A 240 S O 300 N D 360
J F 60 M A 120 M J 180 J A 240 S O 300 N D 360
CALSIUM (%)
80
90
100
110
120
Mean consentration
= 61.55 ppm
Displacement 230 days
0 J F 60 M A 120 M J 180 J A 240 S O 300 N D 360 J F M A
DAY NUMBER (1988)
1989

Excursion 2

Sunday 23 October

Selbu-Røros r/t

Leader: Ola Magne Sæther, NGU

NorFa-kurs Excursion 2: Selbu-Røros r/t Sunday Oct. 23, 1994

Ola Magne Sæther
Geological Survey of Norway, Trondheim

ROUTE

Selbu-Flora-Ås-Nesjødammen-Stugudal-Rien-Brekken-Røros-Glåmos-Brekken-retur

PURPOSE

Present geology/soil/vegetation in Central Norway from sea-level up to the Norwegian peneplain (c. 1000 m a.s.l.). A visit to the mining museum and a tour of the town of Røros is planned at 1200-1430 hours. Total driving distance is c. 200 km.

BACKGROUND ON BEDROCK AND QUATERNARY GEOLOGY

Structural geology and rocktypes

The Trondheim Region constitutes the central part of the Norwegian Caledonides and contains all the major geological elements of the orogen. It embraces a wide range of metasedimentary and igneous rocks from Middle Proterozoic to Silurian age (Fig. 1).
The tectonostratigraphic arrangement of the rocks within the Trondheim region comprise a succession of thrust nappes translated onto the Fennoscandian shield in an east to southeasterly direction during an orogenic history that was both polyphasal and diachronous (Roberts and Wolff, 1981). Two main phases have been defined: The Finnmarkian of late Cambrian to early Ordovician age and The Scandinavian of middle to late Silurian age (Sturt, 1978 and Roberts and Sturt, 1980). The stratigraphy of the succession or piles of nappes as shown in Tab. 1 and simplyfied in Fig. 2, is separated into five allochtonous units or Nappe-complexes.

The fossil-bearing, low-grade metamorphic, Cambrosilurian (570-400 Ma) volcanosediments of the Trondheim Region overlie the Precambrian (c. 1500 Ma) gneisses of the Western Gneiss Region (Strand, 1961; Gee and Sturt, 1985; Gee, Roberts and Wolff, 1985). The Trondheim Region belt is fairly calcareous and contain large amounts of dark minerals with a high cation exchange capacity. The metamorphosed sediments are arranged in a symmetric broad-banded pattern around several fold axes which strike SW-NE at the west end of Lake Selbusjøen. The «Father of geochemistry» V. M. Goldschmidt, who was a professor at the University of Oslo and later director of the Mineral Resource Laboratory, mapped the metamorphic zones of this area in a treatise in 1915. This was two decades after Barrow published his classic work on metamorphic facies around the Great Glenn Fault in Scotland (1893). This means that paralell bands with similar mineral assembleages occur on each side of a central zone where rocks that were buried deepest are exposed. The mineral assemblages will be indicative of lower and lower temperature as one moves progressively away from this fold axis. Within this metamorphic belt, which is thrusted and folded and actually is the core of the Caledonide orogen running through all of Norway and the western part of Sweden, there are different bodies of dioritic (i.e. quartz-rich) and gabbroic (i.e. dark minerals) rocks. Examples of the former are Trondhjemite, a term introduced by Goldschmidt, and exported globally as Silver or White granite, which is quarried at Folstad quarry, Støren (in Gauldalen) and Olaberget, Tolga (in Østerdalen), and the latter examplified by the Fongen-Hyllingen intrusion in Tydal.

Our trip starts within the graphite-bearing phyllite (labelled 63 on the Geologic Map of Norway 1:1mill., 1985) surrounding the eastern part of Lake Selbusjøen and proceeds through the calc-silicate schists and gneisses (86) and pass the mica-schist (65), which in some areas contain rare crystals (i.e. porphyroblasts) of garnet ($(\mathrm{(Fe,Mg,Ca)_3Al_2(SiO_4)_3})$) and staurolite ($(\mathrm{FeAl_4Si_2O_{10}(OH)_2})$), both minerals with hardness 7-7.5 on Mohs' scale. This has been exploited as raw material for millstones over many centuries (since 1100's). The presence of these minerals in addition to kyanite ($\mathrm{Al_2SiO_5}$) indicates minimum P and thus depth of burial of 2-3 km or more and a maximum T of 300-500 degrees. At Ås, the community center of Tydal county (called Norway's Kuwait because of its large revenues from selling hydroelectric power to Trondheim) the Fongen-Hyllingen gabbro will be seen to the north and south of the valley, respectively. This is a 160 km^2 layered intrusion with predominantly mafic (i.e. dark minerals) and somewhat similar to the Bushveld Complex in South Africa containing chromium and platinum, which has been studied in depth(!), i.e. in two dimensions by Wilson and his students at Aarhus University (DK), over the past twenty years (Wilson & Brink Larsen, 1985). It was intruded into the metamorphosed volcanosediments of the Trondheim Region during the Silurian. It is believed to have crystallized from a basaltic parent magma at 5-6 kbars. The minerals are mainly Fe-olivine ($\mathrm{Fe_2SiO_4}$) and Ca-pyroxene ($\mathrm{CaMgSi_2O_6}$) with plagioclase ($\mathrm{CaAl_2O_8}$).

The study of the distribution of these minerals reveals at what temperature and pressure the minerals were formed, suggesting a depth of about 15 km, and by dating radiogenic isotopes (U-Pb or Rb-Sr) it is possible to frame its emplacement in time at 430 Ma, and thus put some constraints on the formation of the Caledonide structures in this area.

At Nesjødammen we will (hopefully no fog!) have a view towards the Sylane mountains, which are rich in amphibole minerals, resistant towards erosion, and part of the Seve and Essandsjø nappes.

Quaternary geology

During the Weichselian glaciation, central Norway was completely covered by the Scandinavian inland ice. Drumlins, flutings and glacial striae indicate that the inland ice moved in a WNW direction from an ice culmination zone situated to the south and east of the water divide (Fig. 3). The coastal areas became deglaciated about 12,500 yBP. During the Younger Dryas cooling, the inland ice readvanced and distinct terminal moraines were formed (Fig. 4). The final deglaciation is characterised by a fast thinning of the inland ice, leaving the last remnants of ice in depressions in the terrain. During the deglaciation, an ice-directed glaciofluvial drainage took place across the water-divides, and locally lakes dammed by glaciers were formed. The weight of the inland ice led to an isostatic depression of the landmass that increased towards the east. In Selbu, areas that lie below 200 m a.s.l. were submerged by the sea during the deglaciation, and marine sediments are found nearly up to this level (Fig. 5).

At Kråkstad, c. 10 km to the east of Selbusjøen, a large glaciofluvial delta can be observed from the bus. It was built up to the sea-level during the deglaciation, and was later influenced by fluvial erosion during the shoreline displacement.
In Tydal there are many examples of till and glaciofluvial sediments. In the mountains, the till cover is often thin and discontinuous, and the highest mountains generally consist of exposed bedrock.

South of Stugudal we will pass over the mountain on a road that was kept open all winter for the first time last season. Here we will pass above the tree-line and see a part of the undulating mountainous landscape, which is so favorable for cross-country skiing. Only a thin veneer of till covers the bedrock, -erosion by the Quaternary ice has dominated over deposition-, before we get down to the lakes Rien and Aursunden at Brekken. These lakes are the headwaters for the Glomma River, Norway's longest. In this area we will deposits of Quaternary sediments.
At Brekken there is a large glaciofluvial delta, with several kettle holes made when huge ice blocks were buried in the sediments and later melted. The delta was built up to the lake level during the deposition, but later oblique uplift caused by the isostatic recovery explains how this delta is now situated c. 16 m above the lake level.
At Hittersjøen, there is an esker formed by a subglacial drainage through the valley. Just to the north of Røros, a very marked esker with several kettle holes, can be observed from the bus. At Kvitsanden this esker, which can be traced continuously for 40 km, is partly covered by eolian sand.
Along the road towards Glåmos, road sections reveal glaciolacustrine sediments, deposited in a glacier-dammed lake.
At Molinga there is another large esker, deposited by a glaciofluvialdrainage across a water divide to Haltdalen.

Along the northern side of Aursunden, there are sections in the till which consists of large amounts of fine material (clay and silt). These tills are derived from the Cambrosilurian schists found in this area. Tills in the southernmost part of the Røros area are coarser than these.

The massive sulphide deposits of the Røros-vidda

We have now passed the water divide into the Røros-vidda area. Geomorphologically this is part of a large peneplain, where mining of copper from the metalliferous «greenstones» and gabbros have been carried out for three-four hundred years. The Røros Kobberverk A/S closed down in 1976 after 300 years of continues operation. We will get some statistics on the size of the mining activity and production during and between different wars(!) at the Røros Mining Museum. The massive sulfide deposits are epigenetic, have pyrite and chalcopyrite as the main minerals and contain about 2% Cu and 2% Zn. The deposits occur as bodies of compact ore of varying shape and dimensions; the bodies are often lens- or ruler-shaped, with the length greatly exceeding the other dimensions. The ore-body at Kongens grube (the King's mine) with a length of about 3 km, a width between 50 and 120 m and a thickness between 1 and 4 km with a maximum of 8 m is often used as an example of the epigenetic deposits.
Within the Trondheim Region metasediments there are also some uneconomical syngenetic, sedimentary sulphide deposits containing pyrite and pyrrhotite (within the Gula nappe complex) which are distinctly layered and impregnated with a bituminous substance. These are distinct from the epigenetic depostis and practically free of copper. They are believed to have been deposited under reducing conditions in sea-water rich in hydrogen sulfide from a volcanic exhalative source. Similar processes are occurring today along the Mid-Atlantic Ridge, East Pacific Rise and in the Red Sea.

At Røros we will visit the old miners cathedral on a short walk around town. The town of Røros is on UNESCO's world heritagelist of sites to behold. The story of the life of old days Røros has been told by the Norwegian author Johan Falkberget («The Fourth Nightguard», «Christian Sextus» a.o.). The mining required vast amounts of wood, which was exploited over large areas around Røros. This has not been completely reforested with birch and spruce, although it was started in 1730 by a dane Langen and in this harsh climate reforesting is a challenge in itself (average temperature is 0.5 C and precipitation is less than 300 mm per year). The mining was managed primarily by Germans who had learned the trade at some of the old European mining schools like Freiberg, Clausthal or even Kongsberg (Norway) and had little understanding of the sub-arctic conditions in this area.

REFERENCES

Asklund, 1938: Hauptzüge der Tektonik und Stratigraphie der mittleren Kaledoniden in Schweden. Sveriges geologiska undersøkning, C417, 1-99.

Gale, G. and Roberts, D., 1974: Trace element geochemistry of Norwegian Lower Paleozoic basic volcanics and its tectonic implications. Earth Planet Sci. Letters 22, 380-390.

Gee, D.G. 1975: A Geotraverse through the Scandinatian Caledonides from Østersund to Trondheim. Sveriges geologiska undersøkening, 717, 1-66.

Gee, D.G. 1980: Basement - cover relationships in the central Scandinavian Caledonides. Geol. Føren. Forha, v. 102, 455-473.

Goldschmidt, V.M., 1915: Übersicht der Eruptivegesteine in Kaledonischen Gebrige Zwischen Stavanger und Trondhjem, Vidensk. Skrifter I. Mat.Naturv, no. 10.

Nilsen, O. 1978: Caledonian Sulphide Deposits and Minor Iron-formations from the Southern Trondheim Region Norways. NGU Bull. 340, p. 35-85.

Oftedahl, Chr., 1980: Geology of Norway, NGU Bull. 356, p.1-114.

Roberts, D. 1967: Structural observations from the Kopperå-Riksgrensen area and discussion of the techtonics of Stjørdalen and the NE Trondheim region, NGU Bull 245, p. 64-122.

Roberts, D. 1978: Caledonides of south central Norway in Caledonian - Appalachian orogen or the North Atlantic region. Geol. Surv. Canada, Paper 78-12, 31-37.

Roberts D. and Sturt, B.A., 1980: Caledonian deformation in Norway. Journ. Geol. Soc. London, v. 137/ 3,2,4, 1-250.

Roberts, D. and Wolff, F.C., 1981: Technostratigraphic development of the Trondheim region Caledonides, Central Norway. Journ. of Struct. Geol, v.3/4, p. 487-494.

Size, W.V. 1979: Petrology, geochemistry and genesis of the type-area Trondhjemite Central Caledonides, Norway. NGU Bull. 351, 51-76.

Sollid, J.L. and Reite, A.J. 1983: The last glaciation and deglaciation of central Norway. In Glacial Deposits in North-Western Europe, Eklers, J. (ed). Geol. Surv. Hamburg, 41-59.

Sturt, B.A. (see references this volume).

Vokes, F.M. 1962: Mineral parageneses of the massive sulphide orebodies of the Caledonides of Norway. Econ.Geol. 57, 890-903.

Vokes, F.M., 1968: Regional metamorphism of the Paleozoic geosynclinal sulphide ore deposits of Norway. Inst. Min.Met. Trans. 77, B53-B59.

Wolff, F.C., 1976: Geologisk kart over Norge, berggrunnskart Trondheim 1:250 000, NGU.

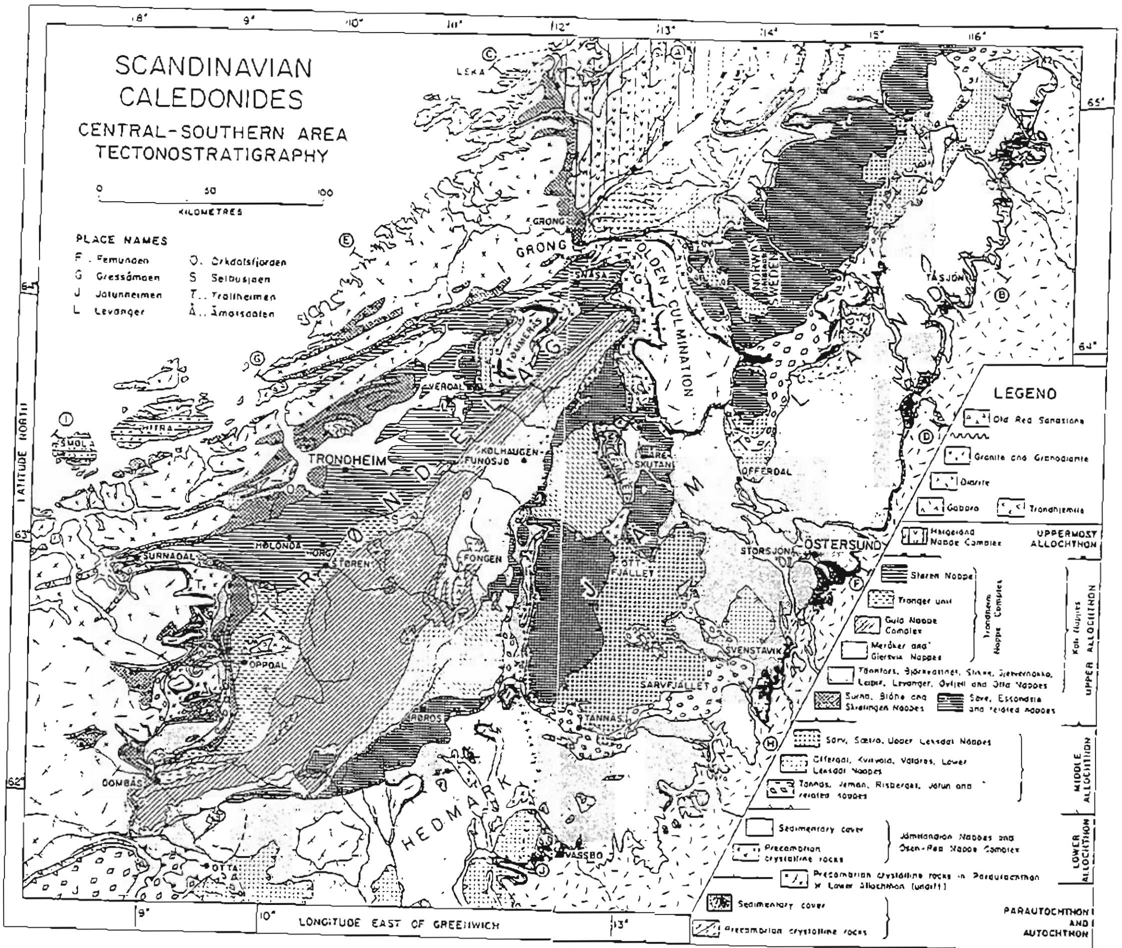


Fig. 1A Tectonostratigraphy of the central southern part of the Scandinavian Caledonides. A–B, C–D, E–F, G–H, and I–J mark the lines of profiles illustrating the structure shown in Fig 1B (From Gee & Sturt, 1985)

Table 1 Comparison of the Trondheim Supergroup stratigraphic units in the Støren and Meraker Nappes. Hardenby's (1980) interpretation of a younger Liafjellet Group above the Slagån Group in the Meraker Nappe is included in this table, although this interpretation is controversial (From Gee & Sturt, 1985)

Trondheim Supergroup Stratigraphy			
Western Trondelag–Støren Nappe		Eastern Trondelag–Meraker Nappe	
Horg Group	Greywackes and conglomerates (age uncertain).	Liafjellet Group (?)	Greywackes and conglomerates. (probably Llandovery age).
		Slågan Group	Black phyllites. (Llandovery age).
Upper Hovin Group	Greywackes and conglomerates. Subordinate limestone and rhyolitic tuffs. (Mid Ordovician and possibly younger)	Kjølhaugen Group	Greywackes and calcareous phyllites (intruded by concordant gabbros).
Lower Hovin Group	Black shales, greywackes and conglomerates. Limestones. Andesites and rhyolites. Massive and pillow basalts present locally. Volcanoclastic turbidites. (Mid Arenig to Caradoc)	Sulåmo Group	Grey and calcareous phyllites. Subordinate greenstones and greenschists. Basal volcanoclastic conglomerates.
Støren Group	Pillow lavas, cherts and very subordinate phyllites. (pre-Mid Arenig)	Fundsjø Group	Greenschists (locally pillow lavas) and keratophyres, extensively intruded by gabbros albite granites and dolerites. (In part of Tremadoc age?)
Gula Nappe Complex			

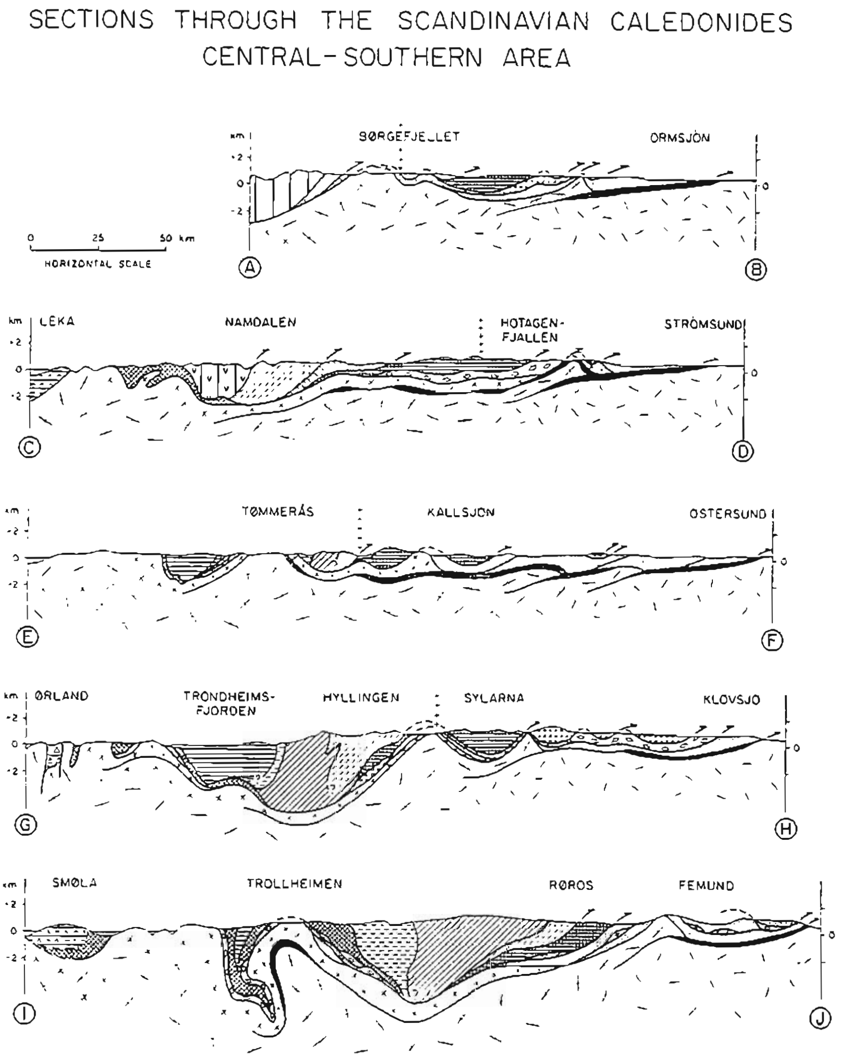


Fig. 12. Schematic profiles through the central southern part of the Scandinavian Caledonides. The lines of the profiles are shown in Fig. 1

6

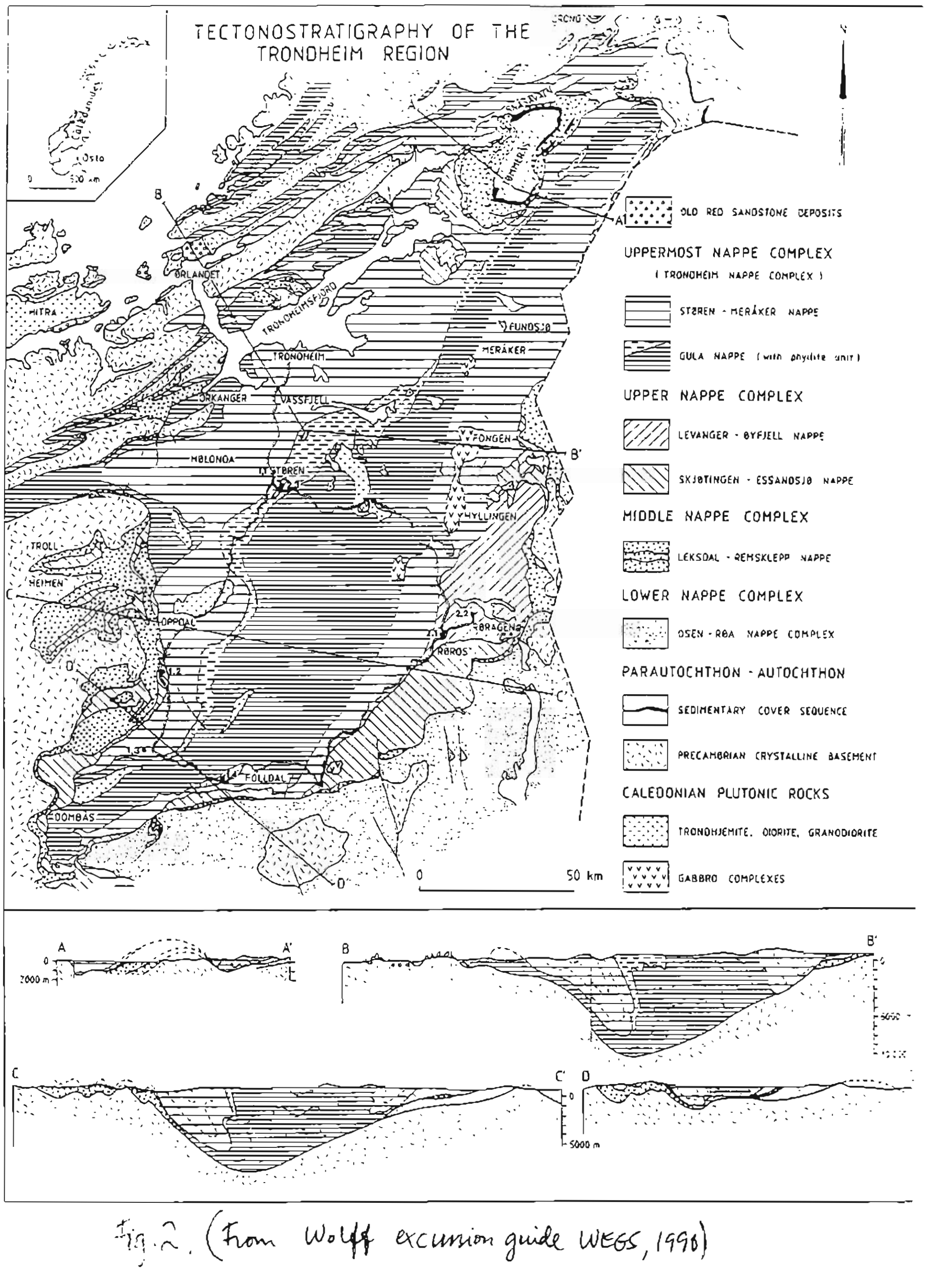


Fig. 2. (From Wolff excursion guide WEGS, 1990)

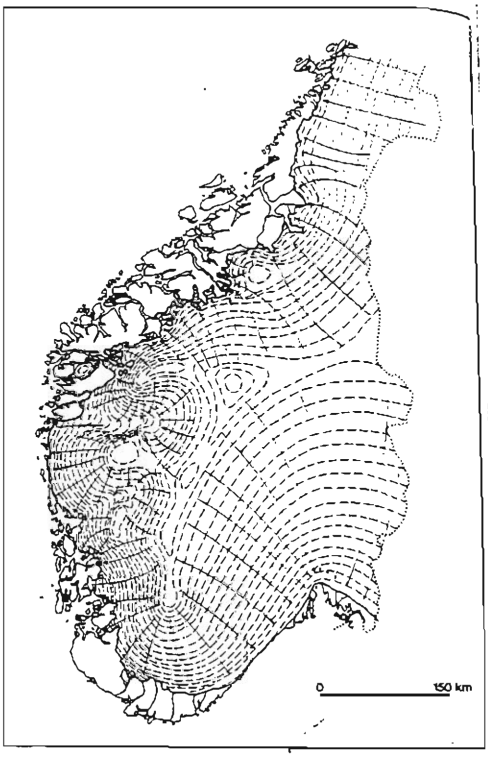


Figure 57. Reconstruction of the continental ice sheet during the Younger Dryas. The arrows indicate the main flow lines. Sea level of today. After Sollid & Sørbel (1982b).

Fig. 5 (NorFa) (from Sollid & Reite, 1983)

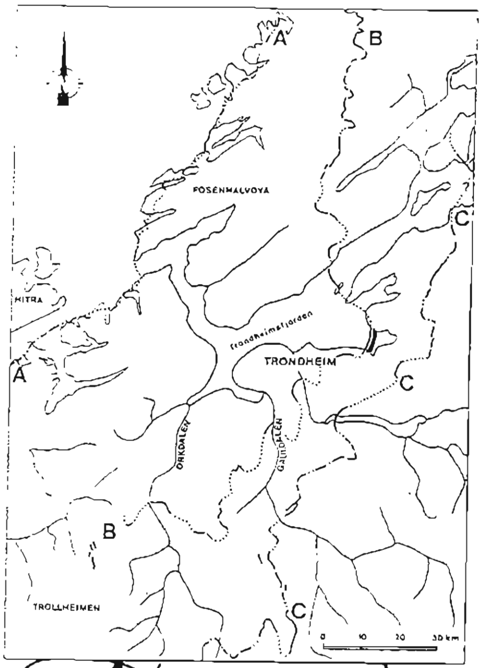


Fig. 4 (by A. J. Reite)

Randmorener i Trøndelag (etter Arne J. Reite)
A-A Randmorener dannet ved oppkalving til ytre kyststrøk (12.500 år gamle)
B-B Randmorener fra hovedframstøtet i tidlig Yngre Dryas (10.800-10.500 år gamle).
C-C Randmorener fra breframstøt i sen Yngre Dryas (10.300 år gamle).

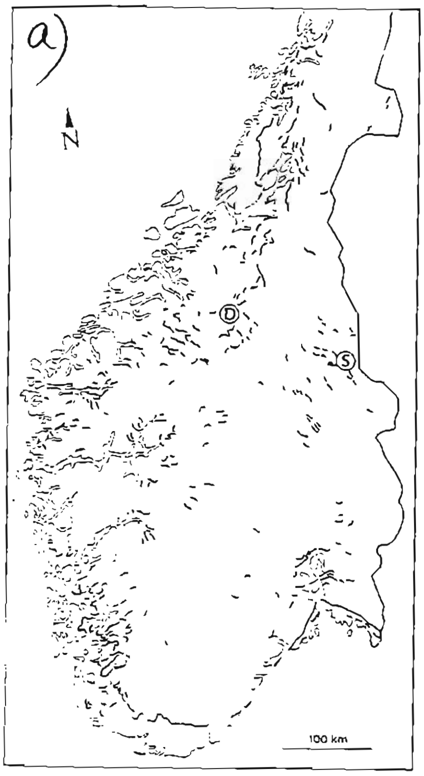


Fig. 2. Distribution of ice-marginal deposits (end moraines and ice-~~marginal deltas)~~ in southern Norway. D = Dovrefjell, S = Søln.

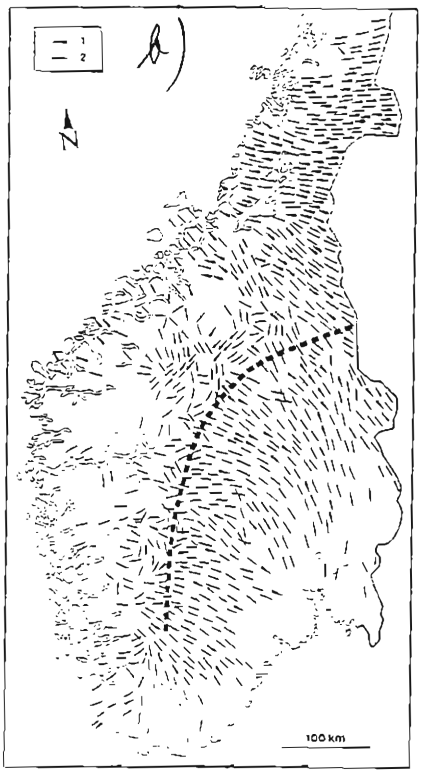


Fig. 3. Distribution of drumlinoide forms in southern Norway. 1 = drumlins, 2 = flutes. In addition, the main position of the ice divide is marked on the figure.

Fig. 3 (from Sollid og Sørbel, 1994, Norsk Geografisk Tidsskrift

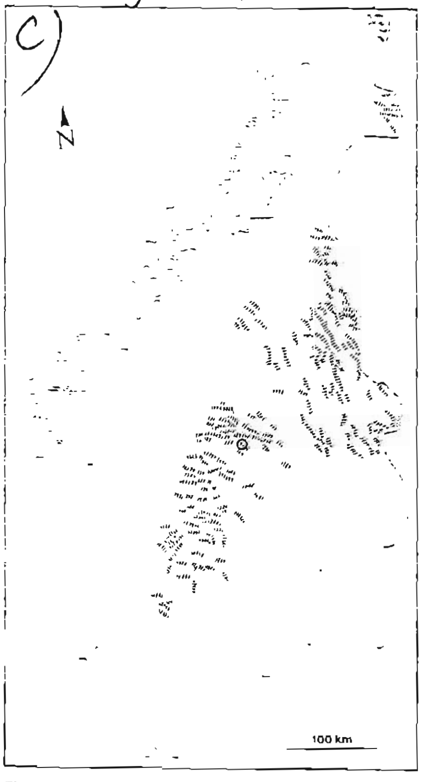


Fig. 5. Distribution of Rogen moraines in southern Norway. The circle marks the location of Fig. 4.

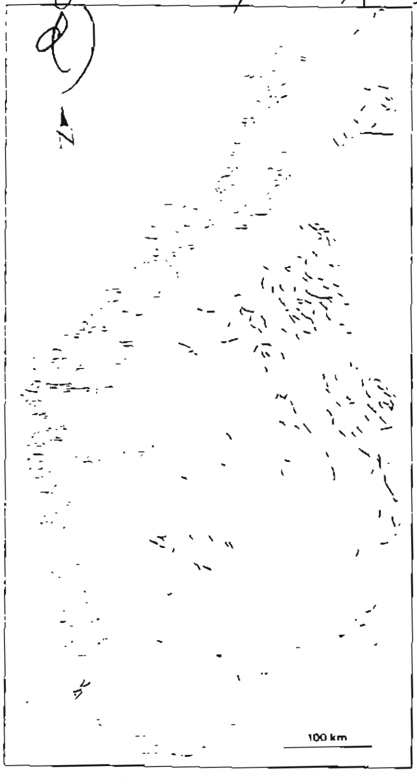


Fig. 6. Distribution of eskers in southern Norway.

9

Temperaturen styrer CO_2-nivået – ikke omvendt

I en fersk, vitenskapelig artikkel hevdes det at temperaturendringer kontrollerer den dramatiske økningen i CO_2-innholdet i atmosfæren. Dette er stikk i strid med rådende teori.

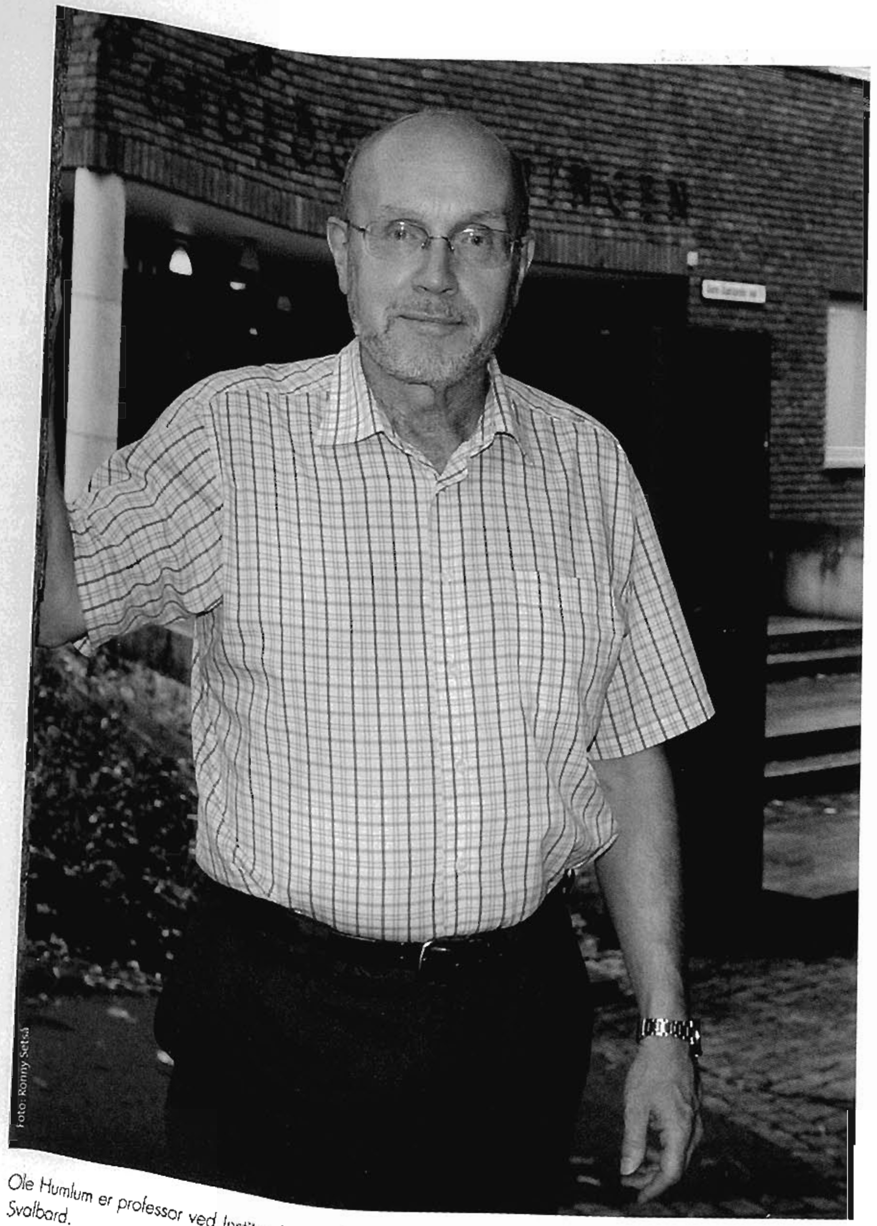


Foto: Ronny Setså

Ole Humlum er professor ved Institutt for geofag ved Universitetet i Oslo og professor II ved Universitetet på Svalbard.

– Data fra iskjerner har tidligere demonstrert at atmosfæriske endringer i CO_2 har etterfulgt temperaturendringer de siste hundretusener av år. Nå har vi funnet bevis for at dette også gjelder i moderne tid, forteller Ole Humlum, professor ved Institutt for geofag ved Universitetet i Oslo og professor II ved Universitetet på Svalbard.

«Atmospheric CO_2 is therefore not initiating the large glacial-interglacial climate changes, and presumably these are controlled by orbital Milankovitch cycles.

Artikkelen, som snart publiseres i tidsskriftet Global and Planetary Change, peker i retning av at det er temperaturendringene på Jorda som har styrt økningen i konsentrasjonen av karbondioksid i atmosfæren de siste tiårene.

Forskerne bak artikkelen er, foruten Ole Humlum, Kjell Stordahl, statistiker ved Telenor og Jan-Erik Solheim, professor emeritus ved Institutt for fysikk og teknologi ved Universitetet i Tromsø.

Undersøkelsene baserer seg på åtte ulike datasett i tidsperioden fra januar 1980 og fram til desember 2011. Datasettene inkluderer målinger av overflatetemperaturen i havet, temperaturen på landområdenes overflate og temperaturen i troposfæren, samt målinger av CO_2 i atmosfæren og beregninger av hvor store CO_2-utslipp menneskene står bak.

KLASSISK FORESTILLING BØR REVURDERES

Forskerne viser i artikkelen til at det har vært antatt at den globale oppvarmingen siden 1975 har vært et resultat av økningen av konsentrasjonen av CO_2 i atmosfæren. Men, hevdes det, resultatene fra deres forskning viser det motsatte.

– Det som formidles til beslutningstakerne og allmennheten er at økning av CO_2-nivået leder til global oppvarming. Denne klassiske forestillingen bør revurderes i lys av de nye resultatene, framholder Humlum.

Siden 1980 har økningen av atmosfærisk CO_2 ligget elleve til tolv måneder etter økningen i global overflatetemperatur i havene, nesten ti måneder etter økningen i den globale lufttemperaturen, og omtrent ni måneder etter økningen i global temperatur i troposfæren.

«Changes in ocean temperatures appear to explain a substantial part of the observed changes in atmospheric CO_2 since January 1980

Tilsvarende sammenheng er også påvist fra iskjerner over de siste 420 000 år. Tiden mellom årsak og virkning har imidlertid vært lengre enn i moderne tid. Det er snakk om århundrer til årtusener.

Forskerne skriver samtidig at koblingen mellom økt karbondioksid i atmosfæren og økte temperaturer er svært svak den motsatte veien. I enkelte tilfeller har de faktisk sett at en økning i konsentrasjonen av klimagassen har gitt lavere globale temperaturer.

HAVENE VIKTIGERE ENN MENNESKELIGE UTSLIPP

Humlum og medforfatterne påpeker videre at korrelasjonen mellom menneskelige (antropogene) utslipp av karbondioksid og endringer av CO_2 i atmosfæren er ustabil og viser liten korrelasjon.

De mener derfor at det først og fremst er temperaturen i overflatevannet i havene som kontrollerer endringer i atmosfærisk CO_2.

Forskerne har også undersøkt forholdet mellom antropogene utslipp av karbondioksid og atmosfærisk CO_2-nivå ved å se på et pol-til-pol-tverrsnitt av jordkloden. Tverrsnittet antyder at den største kilden til atmosfærisk CO_2 befinner seg like sør for ekvator. Den største mengden av menneskelige utslipp av klimagassen stammer derimot fra midtre breddegrader på den nordlige halvkule, altså fra de vestlige industrialiserte landene.

Artikkelforfatterne konkluderer derfor at det ikke er menneskelige utslipp av CO_2, men muligens havene, som har kontrollert den atmosfæriske endringen av CO_2 siden 1980.

«CO_2 released from use of fossil fuels have little influence on the observed changes in the amount of atmospheric CO_2

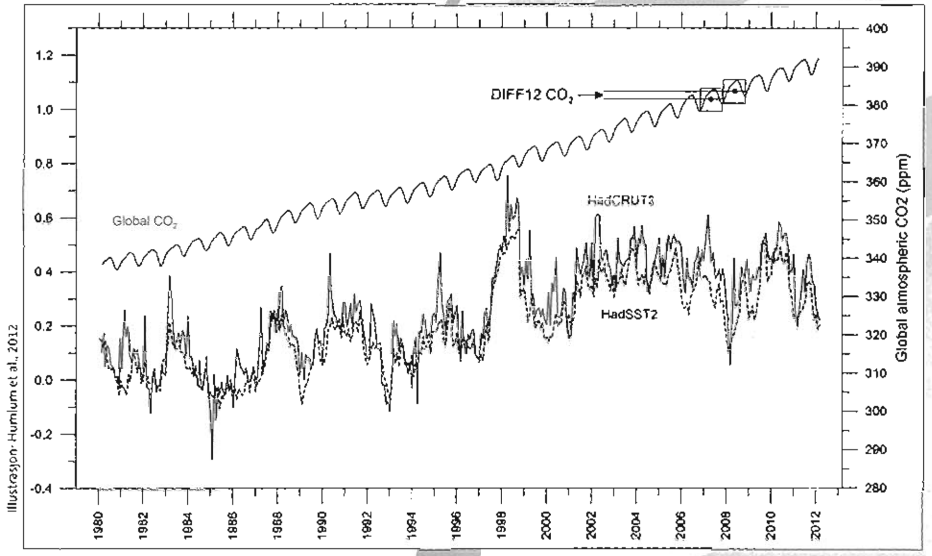
ØNSKER MER ÅPEN HOLDNING

Humlum mener resultatene antyder at detaljene rundt CO_2-kretsløpet ikke er godt nok kjent, og at menneskelige utslipp av klimagassen kan være overvurdert.

– Jeg tror CO_2 spiller en for stor rolle i dagens klimamodeller, og at de ikke på langt nær beskriver de naturlige klimavariasjonene på en god måte.

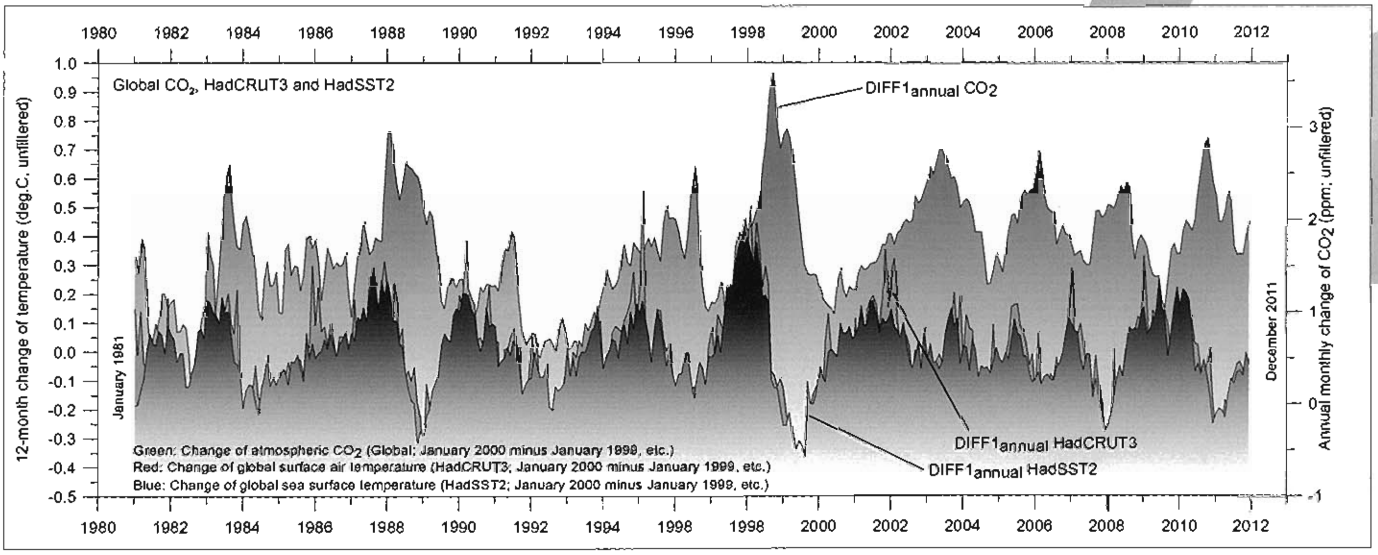
– Forskningsmessig håper jeg funnene våre kan bidra til en mer åpen holdning rundt årsakene til nåtidens klimaendringer, avslutter Ole Humlum.

RONNY SETSÅ



Illustrasjon: Humlum et al., 2012

Figuren viser konsentrasjonen av CO. i atmosfæren siden 1980 (grønt), jevnt stigende opp mot 390 ppm i dag. Blå kurve viser hvordan overflatetemperaturen i havene har endret seg, mens rød kurve viser endringer i den globale lufttemperaturen. Legg merke til hvordan CO_2-kurven undulerer i takt med årstidene, forårsaket av endringer i havtemperatur og fotosyntesen i den terrestriske biosfæren.



Illustrasjon: Humlum et al., 2012

Figuren viser globale endringer av CO i atmosfæren siden 1980 (grønt), global overflatetemperatur i havene (blå) og global lufttemperatur (rød). Legg merke til at grønn kurve ligger etter både blå og rød kurve.

REFERANSE:

Ole Humlum, Kjell Stordahl og Jan-Erik Solheim: The phase relation between atmospheric carbon dioxide and global temperature; Global and Planetary Change, 2012