

Diss. ETH ex. B

Diss. ETH No. 12868

Identification of Runoff Processes in Catchments with a Small Scale Topography

A dissertation submitted to the
SWISS FEDERAL INSTITUTE OF TECHNOLOGY ZÜRICH
for the degree of
DOCTOR OF NATURAL SCIENCES

presented by

HANS MARIA JOZEF FEYEN
Ir., Faculty of Agricultural Sciences, Catholic University of Leuven (BE)
born November 14, 1970
Belgium

accepted on the recommendation of

Prof. Dr. Hannes Flübler, examiner
Prof. Dr. Paolo Burlando, co-examiner
Dr. Andreas Papritz, co-examiner
Prof. Dr. Andrea Rinaldo, co-examiner

1998

Table of Contents

Table of Contents	i
List of Figures	iv
List of Tables	vi
Preface	vii
Abstract	xi
Kurzfassung	xiii
1. Introduction	1
1.1 Scientific Background	2
1.2 Runoff Generating Processes	7
1.2.1 Flow Processes on the Soil Surface	7
1.2.2 Flow Processes in the Sub-surface	9
1.3 Time Series Modelling in Hydrology	10
2. The Modelling of Hydrological Time Series	13
2.1 Introduction	13
2.2 Time Series Models in the Observation Space	15
2.2.1 Sources of Stochastic Disturbances	16
2.3 Estimation of Transfer Function Models	22
2.3.1 Recursive Least Squares	22
2.3.2 Biased Least Squares Estimation	26
2.4 Instrumental Variables	29
2.4.1 Ordinary Instrumental Variables	29
2.4.2 Refined and Simplified Instrumental Variables	31
2.5 On Model Identification and Evaluation	32
2.5.1 Criteria based on the Goodness of Fit	33

2.5.1.1 The Mean Squared Error	33
2.5.1.2 The Nash/Scutcliffe Criterion (Coefficient of Determination)	34
2.5.2 Criteria based on the Parameter Efficiency	34
2.5.2.1 EVN (Error Variance Norm)	34
2.5.3 Information Criteria	35
2.5.3.1 Young Information Criterion (YIC)	35
2.5.3.2 Akaike Information Criterion (AIC)	35
2.6 Time Series in the State Space	36
2.6.1 The Kalman Filter	36
2.6.2 Smoothing	39
2.7 Data Based Modelling	42
2.7.1 Definition	42
2.7.2 Data Based Modelling of Nonlinear Time Series	44
2.7.3 Derivation of a Time Variable Parameter Model	44
2.7.4 Estimation of Linear Transfer Functions with Time Variable Parameters	46
3. A Tracer Experiment to Detect Flow Paths in Forest Soils	53
3.1 Introduction	54
3.2 Material and Methods	57
3.2.1 Experimental Site	57
3.2.2 Soil Plots	58
3.2.3 Sampling and Analyses of the Runoff	60
3.2.4 Sprinkling Device	60
3.2.5 Tracer Application	61
3.3 Results	64
3.3.1 Results of the Steady State Irrigation	64
3.3.2 Tracer Breakthrough on the Muck Humus Plot	67
3.3.3 Tracer Breakthrough on the Mor Humus Plot	72
3.4 Discussion	75
4. Runoff Processes at Three Spatial Scales	81
4.1 Introduction	82
4.2 Theory	84
4.2.1 Linear Reservoir Model	84
4.2.2 Linear Time Series Model	84
4.2.3 Estimation of the Transfer Function Model	85
4.2.4 Nonlinear Time Series	86
4.3 Material and Methods	87
4.3.1 Erlenbach	88
4.3.2 Artificial Sub-catchments	88
4.3.3 Soil Plots	91
4.3.4 Event and Pre-event water	92
4.4 Results	93
4.4.1 Irrigation of the Soil Plots	93
4.4.1.1 Muck Humus Plot	93
4.4.1.2 Mor Humus Plot	97

4.4.2 Runoff from Natural Rainstorms	99
4.4.3 Modelling the Nonlinear Properties of the Runoff Generation	101
4.4.4 Runoff from the Soil Plots	103
4.4.5 Runoff from the Sub-catchments	112
4.4.6 Runoff from the Erlenbach headwater catchment	117
4.5 Comparison of Scales	120
4.6 Summary and Conclusions	124
5. Concluding Remarks	129
<u>Appendix A</u>	133
A.1 Recursive Formulation of the Least Squares Solution	133
A.2 Derivation of the Kalman Filter	135
A.3 Gauss-Markov sequences	139
References	141
Curriculum Vitae	147

List of Figures

1. Introduction	1
Figure 1.1: Situation of the Alptal valley in Switzerland.	3
Figure 1.2: Map of the sub-catchments N1, N2 and N3 in Alptal (P. Schleppe, pers. comm.).	4
Figure 1.3: An overview of runoff generating processes.	8
2. The Modelling of Hydrological Time Series	13
Figure 2.1: Stochastic and deterministic components of time series models.	17
Figure 2.2: The Instrumental Variable procedure (Section 2.4.1).	31
Figure 2.3: Example of the filtering and smoothing procedures.	49
Figure 2.4: Overview of the nonlinear transfer function estimation procedure.	52
3. A Tracer Experiment to Detect Flow Paths in Forest Soils	53
Figure 3.1: Side-view of the soil plots.	59
Figure 3.2: Timing of the tracer application on the soil plots.	61
Figure 3.3: Tracer application areas on the soil plots (to scale).	62
Figure 3.4: Tracer application in the sub-surface.	63
Figure 3.5: Results of the steady state irrigation ($2.3 \cdot 10^{-6} \text{ m}^3 \text{ m}^{-2} \text{ s}^{-1}$) on the muck humus plot.	66
Figure 3.6: Results of the steady state irrigation ($2.3 \cdot 10^{-6} \text{ m}^3 \text{ m}^{-2} \text{ s}^{-1}$) on the mor humus plot.	68
Figure 3.7: Tracer breakthrough on the muck humus plot.	71
Figure 3.8: Tracer breakthrough in the mor humus plot. Depth runoff only.	74
Figure 3.9: Summary of the tracer experiments on the soil plots.	76
4. Runoff Processes at Three Spatial Scales	81
Figure 4.1: Overview of the sub-catchments N_1 , N_2 and N_3	89
Figure 4.2: Soil type distribution on the sub-catchments N_1 (right) and N_2 (left).	90
Figure 4.3: Results of the irrigation experiments on the muck humus plot.	94
Figure 4.4: Results of the irrigation experiment on the mor humus plot.	98
Figure 4.5: Result of the power law optimization for the depth runoff from the muck humus plot.	102
Figure 4.6: Runoff from the soil plots for S1. S1a and S1b denote the subsets studied in detail.	104
Figure 4.7: Runoff from the soil plots for S2. S2a and S2b denote the subsets studied in detail.	105
Figure 4.8: Electrical conductivity (EC) for S2a and S2b at all scales.	107

LIST OF FIGURES

Figure 4.9: Nonlinear transfer functions for the runoff from the soil plots.108
Figure 4.10: Runoff from the sub-catchments for S2.114
Figure 4.11: Runoff from the sub-catchments during S1.115
Figure 4.12: Nonlinear transfer functions for the runoff from the sub-catchments116
Figure 4.13: Runoff from the Erlenbach headwater catchments for S1 and S2.119
Figure 4.14: Nonlinear transfer functions for the runoff from the Erlenbach headwater catchment. . .120
Figure 4.15: Contribution of event and pre-event water to the runoff at all scales.121
Figure 4.16: Overview of the runoff processes at all scales.125

List of Tables

3. A Tracer Experiment to Detect Flow Paths in Forest Soils	53
Table 3.1: Properties of the mor humus soil.	57
Table 3.2: Properties of the muck humus soil.	58
Table 3.3: Summary of the tracer application on both plots.	64
4. Runoff Processes at Three Spatial Scales	81
Table 4.4: Properties of the sub-catchments.	90
Table 4.5: Nonlinear transfer function models for the irrigation experiments.	95
Table 4.6: Water balances during S1 and S2.	100
Table 4.7: Power law exponent γ estimated for both time series and all studied areas.	103
Table 4.8: Nonlinear transfer functions for the runoff from the soil plots during natural rainfall. . .	109
Table 4.9: Origin of the runoff water during S2a and S2b at all scales.	110
Table 4.10: Nonlinear transfer functions for the runoff from the sub-catchments.	112
Table 4.11: Nonlinear transfer function models for the runoff from the Erlenbach.	118

Preface

I carried out this dissertation in the framework of two scientific projects which were funded by the Swiss Federal Government

'Nitrogen Saturation Experiments' (NITREX), part of the European Research Programme 'Environment'. The Swiss contribution to this project was funded by the 'Bundesamt für Bildung und Wissenschaft', BBW Nr. 94.0139.

'Modelling runoff from forested hillslopes with a small scale topography', financially supported by the Swiss National Science Foundation, NF 21-46927.96.

Structure of the Dissertation

In this section, I will briefly describe how the different chapters of this dissertation merge together. This thesis reports on the research on runoff generating processes at three spatial scales in the northern Swiss prealps. It supports a study on the cycling of nitrogen (NITREX), which is performed on artificially delineated sub-catchments. Typically, runoff is generated somewhere in a catchment by the generating process and flows towards the catchment outlet along possibly varying flowpaths. A complete characterization of runoff processes thus needs a description of the dynamics of the underlying generating process and a characterization of the flowpaths of the runoff process.

The first chapter presents a short overview of the important runoff processes and of some applications of time series analysis in hydrology. To characterize the dynamics of the runoff generating processes, I used methods of time series analysis. The techniques I applied are rather uncommon in soil physics and hydrology. For the reader to become more accustomed

to the time series analysis methods, I dedicated a chapter to them (Chapter 2). I mainly focus on two topics: i) how to deal with the inherent nonlinearity of the rainfall-runoff relationship, and ii) how to estimate linear transfer functions from observed data possibly corrupted by measurement errors. The next chapter (Chapter 3) describes a tracer experiments performed on two soil plots. The purpose of this experiment was to detect the transport regime and the flowpaths in top- and subsoil layers of the two soil plots. In Chapter 4, the algorithms derived in Chapter 2 are applied to characterize the nonlinear dynamics of the runoff generating processes at three spatial scales. In addition, measurements of the water quality serve to determine the residence time of the runoff water in the vadose zone. Combining this information with the results of the time series modelling is the key to understanding the dominating runoff processes. In the last Chapter, the results of the study at three spatial scales are combined with tracer experiment at the smallest scale and I discuss the fate of rainwater and dissolved nitrogen components in the NITREX sub-catchments.

Acknowledgements

During my dissertational work, many people supported me in one way or another. First of all, I want to express my gratitude to *Andreas Papritz*, for the open way of discussing, for the unstinted help with the statistical theory on time series analysis and for 'repairing' my writings. It was a great pleasure to me to work in the group of Prof. *Hannes Flühler*. He gave me the freedom to decide by myself on the line of my work, but always supported me. I am also very grateful to Prof. *Paolo Burlando* for his constructive comments and advice on my work. Without the experience and commitment of the technical staff of the soil physics group, this dissertation would not have happened the way it did. Thanks to *Jörg Leuenberger* for the many hours of digging and carpentry, to *Hanspi Läser* for his workshop work and to *Hannes Wydler* and *Hans Wunderli* for their outstanding and only superficially chaotic technical installations. Also as 'Informatikbetreuer', I often experienced the energy and the dash of *Hans Wunderli*, which very often encouraged me. The availability of simultaneous data at three spatial scales was a unique situation, which was only possible by the cooperation of *Felix Forster* and *Patrick Schleppei* from the Swiss Federal Institute of Snow, Forest and

Landscape Research. The fieldwork in Alptal was without doubt the most fascinating part of the research. During the experiments (summer '97) and on many other occasions, I was pleased by so many helping hands: *Philipp Aeby*, *Nicole Augustin*, *Maya Bundt*, *Paul Gähwiler*, *Michael Gysi*, *Frank Hagedorn*, *Jürg Hutter*, *Roy Kasteel*, *Daniel Stadler*, *Christian Stamm* and *Bjørn Studer*. If I've left someone out, I owe you a pint, or a coffee. I enjoyed sharing the office with *Philipp Aeby*, always having time for some jokes. A big thank you to *Jeannette Hollinger* for doing so much paperwork. Finally, I am very grateful to my friends, my family, my wife *Nicole* and my son *Maurus* for their patience and for supporting me during these years.

Leer - Vide - Empty

Abstract

This thesis presents a study on the generation of runoff from prealpine catchments at three spatial scales. The study was part of the Nitrogen Saturation Experiments (NITREX), a European project on the cycling of nitrogen in forest ecosystems. On the Swiss NITREX site in the Erlenbach catchment (Alptal valley, canton of Schwyz), the nitrogen deposition onto forested sub-catchments (1500 m^2) delineated by trenches was experimentally increased, and its transfer into the soil and vegetation and the losses by runoff and denitrification were investigated. Preliminary results of the nitrogen budget strongly suggested that the nitrate losses of the sub-catchment were hydrologically controlled. To obtain detailed information on the effect of runoff processes on nitrogen turnover, we monitored the fluxes of water and dissolved solutes from the sub-catchments, from the Erlenbach headwater catchment (0.7 km^2) and from two isolated soil plots (13 m^2). The Erlenbach is a steep (20 percent) and narrow catchment, with a network of drainage trenches connected to permanent brooks in deeply incised gorges. Two sub-catchments are forested, the third is a wetland. In the forested sub-catchments, closely linked to the topography, two soil types predominate: a wet muck humus soil in small depressions and a better drained mor humus soil on mounds and ridges. On the wetland sub-catchment only the muck humus occurs. The soil plots, which represent these two soil types, were instrumented to monitor the outflow at three depths and the moisture status of the different soil layers. On the soil plots, we carried out controlled irrigation and tracer experiments as well as observations under natural boundary conditions. By comparing the dominating runoff processes at all scales, we studied the effect of vegetation, soil type distribution and drainage area on the runoff generation.

To identify the runoff dynamics, we estimated linear discrete transfer functions from water input and outflow measurements. Before estimating the transfer functions, the data series were linearized by means of the following approach. First, we fitted a linear transfer function with parameters which were allowed to vary in time to the observed data. The fitting procedure we applied based on Kalman filtering and fixed interval smoothing techniques. Then, the parameter which describes the hydrological gain (i.e. the amount of runoff caused by one unit of water input) was compared to additional data characterizing the water regime of the soils. Since the level of the groundwater correlated best with the time variable gain, this variable served to transform the measured precipitation nonlinearly into an effective precipitation. From the linearized rainfall-runoff data we estimated linear transfer functions with constant parameters, which could be given a physical interpretation. To obtain unbiased parameter estimates, we employed a recursive and iterative instrumental variable method. To

avoid overparameterization, we selected the model structure (i.e. number of parameters of the transfer functions) based on an information criterion. This procedure was applied to data from the irrigation experiments on the soil plots and to rainfall-runoff measurements on the soil plots (13 m²), the sub-catchments (1500 m²) and the Erlenbach headwater catchment (0.7 km²). To locate the pathways of the runoff processes in the soil plots, we performed tracer experiments under a controlled water regime. To unravel the effect of the various soil horizons on the runoff dynamics, Cl⁻ and Br⁻ tracers were applied onto the soil surface and also injected below ground at the interface between the upper humus layer and the gleyic subsoil of both plots. Water sampled at the outflow of the plot, sub-catchments and headwater catchment allowed to quantify the ratio of event or pre-event water in the runoff.

On the soil plots, runoff reacted generally very fast to rainfall or irrigation water inputs. On the wet muck humus plot, few but large macropores in the gleyic subsoil contributed substantially to the runoff generation. On the better drained mor humus plot, the tracer recovery and the runoff ratio were higher, but the runoff dynamics of the flow processes were clearly slower. We attributed this to the effect of the better structured subsoil of this plot. On both soil plots, water entering through the subsoil contributed considerably to the runoff generation during rainfall events. At the same time, large quantities of water drained from the soil plots as deep seepage without directly contributing to the outflow.

Despite the different vegetation and soil type distribution, the runoff dynamics were very similar in all sub-catchments. Thus, it was not the abundance of the mor and muck humus soil which determined the characteristics of the runoff, but rather the spatial arrangement of the muck humus zones and the presence of waterlogged areas. All sub-catchments reacted faster to rainfall than the smaller soil plots, which was likely due to the presence of waterlogged areas. The electrical conductivity indicated, particularly during peak runoff, that the residence time of rainwater in the sub-catchments was very limited. Furthermore, the contribution of the mor humus areas to the sub-catchment runoff was very limited.

Regardless of the larger area of the headwater catchment, the dynamics of the runoff processes resembled those of the sub-catchments closely. The transfer functions and the electrical conductivity measurements indicated that two flow components added to the runoff. The slow component had an electrical conductivity comparable to that of the runoff of the mor humus soil plot. It passed through the macropores of the subsoil which contributed directly to the main brook of the Erlenbach catchment. The fast component originated from wet or waterlogged muck humus areas and was rapidly routed to the brooks by the drainage trenches.

The identified flow processes in the sub-catchments agree well with the preliminary nitrogen budget. Because of the short residence time of water and solutes, nitrate may not be removed from the rainwater by the slow microbial activity. Still, the contact time may be sufficient to retain the ammonium by the much faster adsorption process.

Kurzfassung

Diese Arbeit präsentiert eine Untersuchung zur Abflussbildung in voralpinen Einzugsgebieten auf drei räumlichen Skalen. Sie wurde im Rahmen des NITREX Projektes (Nitrogen Saturation Experiments) durchgeführt, ein Europäisches Projekt, das den Stickstoffkreislauf in Waldökosystemen untersucht. Im Erlenbach Einzugsgebiet (Alptal, Kanton Schwyz) wurde der Stickstoffeintrag in bewaldeten Teileinzugsgebieten künstlich erhöht, um die Stickstoffumsetzung im Boden, die Aufnahme durch die Vegetation sowie die Austräge durch Abfluss und als Denitrifikation zu untersuchen. Erste Ergebnisse zeigten, dass die Stickstoffausträge aus den Teileinzugsgebieten durch hydrologische Prozesse gesteuert werden. Um detailliertere Informationen über den Einfluss der Abflussprozesse auf den Stickstoffkreislauf zu erhalten, beobachteten wir in den Teileinzugsgebieten (1500 m^2), im Erlenbach Einzugsgebiet (0.7 km^2) und auf zwei isolierten Flächen (13 m^2) die Abflüsse und die darin gelösten Stoffe. Das Erlenbach Einzugsgebiet umfasst einen schmalen, steilen Hang (20 Prozent) und wird von Abflussgräben durchzogen. Diese münden in tief eingeschnittene Bachläufe. Die Teileinzugsgebiete wurden durch 80 cm tiefe Gräben abgetrennt. Zwei Teileinzugsgebiete sind bewaldet, das dritte ist eine feuchte Wiese. In den bewaldeten Teileinzugsgebieten treten in Abhängigkeit von der Topographie zwei Humusformen auf: ein nasser Anmoor Humus in Senken und ein gut drainierter Rohhumus auf Kuppen. In der feuchten Wiese kommt nur Anmoor Humus vor. Die 13 m^2 grossen isolierten Flächen wurden in je einer dieser Humusformen eingerichtet um den Abfluss und die Bodenfeuchte in verschiedenen Tiefen zu erfassen. Auf diesen Flächen führten wir unter kontrollierten Randbedingungen Bewässerungs- und Tracerexperimente durch und beobachteten das Verhalten des Systems bei natürlichen Niederschlagsereignissen. Um die Auswirkungen der Vegetation, der Humusform und der Gebietsgrösse auf die Abflussbildung bestimmen zu können, verglichen wir die Abflüsse auf allen drei Skalen.

Zur Identifikation der Abflussdynamik schätzten wir lineare diskrete Transferfunktionen von den Wasserinput- und Abflussmessungen. Bevor wir die Transferfunktionen schätzen konnten, wurden die Daten folgendermassen linearisiert: zuerst schätzten wir eine lineare Transferfunktion mit zeitabhängigen Parametern ausgehend von den erhobenen Daten. Dieses Schätzverfahren basiert auf dem Kalman Filter und auf Glättungsalgorithmen. Danach verglichen wir die zeitabhängigen Parameter mit zusätzlichen Messungen zum Wasserhaushalt. Weil die Grundwassertiefe am besten mit dem hydrologischen Gewinn (i.e. die Abflusszunahme, die durch eine Einheit Wasserzufuhr verursacht wird) korrelierte, verwendeten wir diese Grösse, um die gemessenen Niederschläge in effektive Niederschläge umzuwandeln.

Von den so linearisierten Niederschlags- und Abflussdaten schätzten wir erneut lineare Transferfunktionen, die wir physikalisch interpretieren können. Damit systematische Abweichungen der geschätzten Parameter vermieden werden, benutzten wir einen auf Hilfsvariablen basierenden rekursiven und iterativen Schätzalgorithmus. Die Struktur der Transferfunktion (d.h. die Anzahl Parameter) wurde anhand eines statistischen Kriteriums bestimmt. Dieses Verfahren wendeten wir auf die Daten der Bewässerungsexperimente und auf die Daten zu natürlichen Abflussereignissen auf allen drei Skalen an. Zusätzlich führten wir Tracerexperimente auf den isolierten Flächen durch, um die Fliesswege des Wassers zu lokalisieren. Wir applizierten Cl^- und Br^- an der Bodenoberfläche und in der Tiefe, um die Auswirkung unterschiedlicher Bodenhorizonte auf die Abflussdynamik zu charakterisieren. Über die elektrische Leitfähigkeit der Abflüsse auf allen drei Skalen quantifizierten wir den direkten Beitrag des Niederschlags zu den Abflüssen.

Der Abfluss der isolierten Flächen reagierte schnell auf Niederschlag oder Bewässerung. Im nassen Anmoor Humus trugen vereinzelte Makroporen wesentlich zum Abfluss bei. Die Wiederfindungsrate für Wasser und Tracer war zwar höher im Rohhumus, die Abflussbildung war jedoch wesentlich langsamer. Wir führten dies auf den besser strukturierten Unterboden der Rohhumusfläche zurück. Auf beiden Flächen flossen während natürlichen Niederschlagsereignissen nicht unerhebliche Wassermengen unterirdisch aus angrenzende Gebieten zu; gleichzeitig aber sickerten grössere Wassermengen in tiefere Bodenschichten, ohne zum Abfluss beizutragen.

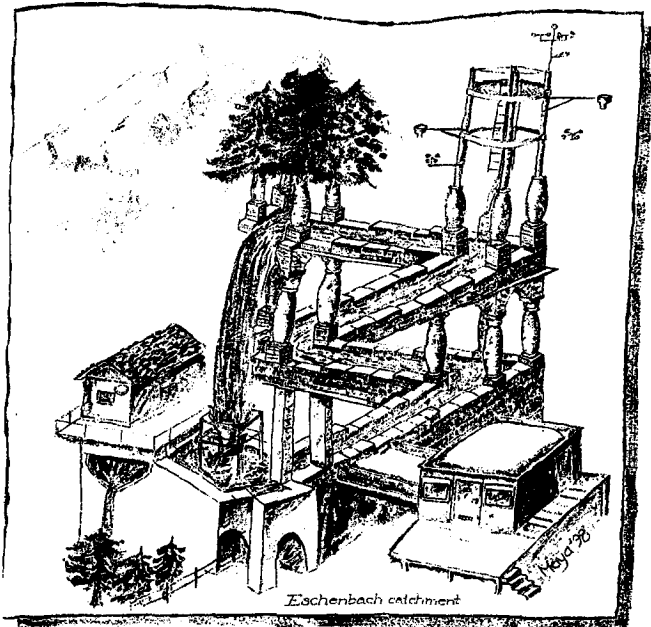
Trotz der unterschiedlichen Verteilung von Vegetation und Humusformen war die Abflussdynamik der Teileinzugsgebiete ähnlich. Daraus folgt, dass nicht der Flächenanteil an Anmoor oder Rohhumus sondern deren räumliche Anordnung und die Anwesenheit wassergesättigter Gebiete die Abflussbildung bestimmte. Die Reaktion der Teileinzugsgebiete auf Regenereignisse war schneller als diejenige der isolierten Flächen, was wir ebenfalls auf die Präsenz wassergesättigter Gebiete zurückzuführen. Die elektrische Leitfähigkeit der Abflüsse zeigte zudem, dass die Verweilzeit des Niederschlagswassers in den Teileinzugsgebieten sehr gering und der Beitrag der Rohhumusgebiete zum Abfluss limitiert war.

Trotz der grösseren Fläche des Erlenbacheinzugsgebietes war dessen Abflussverhalten ähnlich demjenigen der Teileinzugsgebiete. Die Transferfunktionen und die elektrische Leitfähigkeit des Abflusswassers zeigten, dass eine langsame und eine schnelle Komponente zum Abfluss beitrugen. Die langsame Komponente hatte eine Leitfähigkeit vergleichbar mit derjenigen der Rohhumusfläche. Das Wasser floss durch Makroporen im Unterboden, die direkte Verbindung zum Vorfluter hatten. Die schnelle Abflusskomponente entstand auf nassem und wassergesättigtem Anmoor Humus und wurde durch die Drainagegräben schnell zum Vorfluter geführt.

Die identifizierten Abflussprozesse stimmen mit den ersten Ergebnissen der Stickstoffhaushaltsstudie gut überein. Wegen des relativ grossen Zeitbedarfs zum Abbau des Nitrats durch Mikroorganismen bei gleichzeitig kurzer Verweilzeit des Wassers und der gelösten Stoffen im Boden, wird Nitrat nicht aus dem Regenwasser entfernt. Die Verweilzeit reicht aber für die schnellere Adsorption von Ammonium aus.

Chapter 1

Introduction



The hydrological cycle according to M. Bundt

1.1 Scientific Background

The Swiss contribution to the European NITREX project (Wright and Rasmussen, 1998) started in 1993. The NITREX project was set up to study the consequences of an increased atmospheric deposition of nitrogen onto forested ecosystems. The project had two main objectives: i) to detect the critical load of nitrogen which does ultimately not result in a nitrogen saturation of the ecosystem, and ii) to determine whether the nitrogen saturation of a forest ecosystem is reversible. An ecosystem is saturated with nitrogen if the input of this element into the ecosystem exceeds the uptake by plants and soil micro-organisms and the internal storage capacity for N-components. The excess nitrogen leaves the ecosystem in the runoff, in the deep seepage that percolates through the root zone towards deeper groundwater layers or by an increased denitrification. It may also be accumulated in the organic soil material. The nitrogen exports by runoff and as denitrification can be monitored and are therefore used as indicators for the nitrogen state of the ecosystem. Results from NITREX studies in Sogndal (Wright *et al.*, 1995) and Gardsjön (Moldan *et al.*, 1995) indicate that ecosystems do not respond gradually to nitrogen inputs. For instance, if the export of nitrogen is hydrologically controlled, export even occurs from ecosystems not (yet) saturated with nitrogen. In such cases, the residence time of nitrogen in the ecosystem is too short to be entirely taken up by plants and micro-organisms. Thus, any study on the cycling of nitrogen in natural ecosystems requires detailed information on the hydrology of the studied area. Given such information, it is possible to distinguish between hydrologically controlled leaching of nitrogen and leaching of nitrogen which is caused by a saturation of the ecosystem.

In the framework of NITREX, two types of experiments have been performed. In areas which suffered for a long time from large atmospheric nitrogen deposition, "clean roof" experiments were set-up to study the reversibility of the nitrogen saturation of the forested ecosystem. In this type of experiments, the rainwater is collected on a roof between the tree canopy and the ground and purified before re-application onto the soil surface. In areas with small deposition of nitrogen, selected ecosystems are treated with well-defined amounts of

nitrogen. In both types of experiments, the reaction of several compartments of the ecosystem to the altered nitrogen load were monitored.

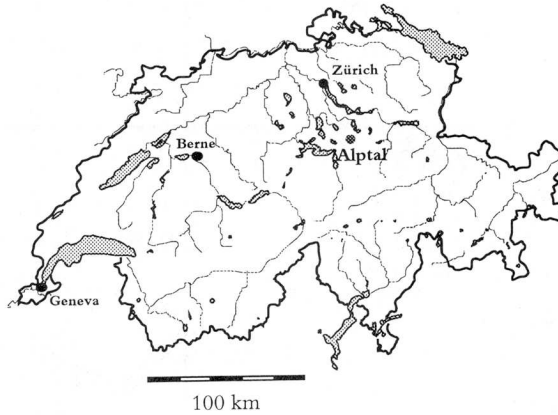


Figure 1.1: *Situation of the Alptal valley in Switzerland.*

The Swiss NITREX site is situated in the valley of Alptal, in the northern Swiss prealps (Fig. 1.1). Since more than 30 years, hydrological data records of several headwater catchments within the Alptal valley have been collected by the Forest Hydrology group of the Swiss Federal Institute of Forest, Snow, and Landscape Research (Burch, 1994). One of the best documented headwater catchments in this area is the Erlenbach, a long, narrow and steep catchment with a total surface area of 0.7 km^2 . Within the Erlenbach catchment, three areas were selected for the Swiss NITREX project. Two of them are forested, the third one is a wetland, grown with grasses only (Fig. 1.2).

The selected areas were delineated by means of trenches to obtain the so-called sub-catchments, in the following referred to as N_1 , N_2 and N_3 . The subsoils in the area are clayey and gleyic, which suggests that they have small hydraulic conductivities. Based on this pedologic information we considered the sub-catchments as hydrologically tight systems. To establish a

balance for water and nitrogen, the sub-catchments were instrumented to quantify the major fluxes of water and nitrogen entering or leaving the sub-catchments. To detect changes in the nitrogen content of the understory vegetation and the tree layer, a detailed monitoring program and long term experiments were accomplished (Müller, 1997). In addition, an initial inventory of the soil nitrogen pools was performed based on a geostatistical approach.

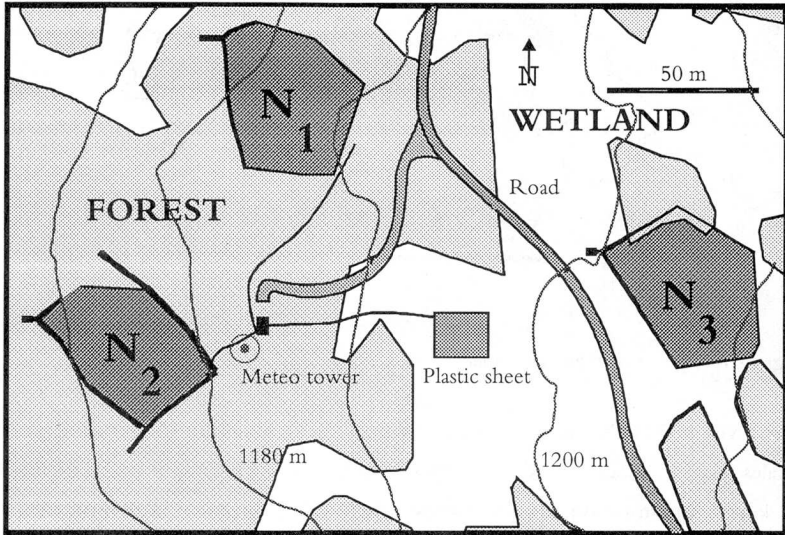


Figure 1.2: Map of the sub-catchments N_1 , N_2 and N_3 in Alptal (P. Schleppei, pers. comm.).

No detailed information on the hydrology of the sub-catchments was available before the NITREX experiment started. However, some information was available from a site nearby where Stadler (1996) investigated the influence of soil freezing on the runoff generation. This study was performed on two soil plots of about 10 m^2 on a steep slope. Initially, the plots were equipped to measure the total runoff but the instrumentation in one of the plots was later extended to measure flow parallel to the surface in the gleyic subsoil, in the humic A layers and on the surface. Furthermore, the plot was delineated by trenches to a depth of 80

cm to avoid lateral inflow of water. In spring and early summer 1995, runoff generation during natural rainstorms was studied by Gysi (1996). Surprisingly, he observed a large flow of water in the gleyic subsoil. Since the soil matrix is virtually impermeable, we hypothesized there was a network of large pores, channels or fissures in the subsoil (Feyen *et al.*, 1997). This, however, conflicted with the assumption that the sub-catchments had an impermeable lower boundary. This contradiction stimulated further efforts to study the hydrology of the sub-catchments. I focused my work on the processes of runoff generation within the sub-catchments. Two distinct soil types occur in the sub-catchments and their distribution is closely linked to the micro-topography. Therefore, I studied water flow and runoff generation for the two soil types by controlled experiments and by observing the system without any intervention.

The main goals of the detailed observations on the soil plots were:

- to detect along which flowpaths the water flows through these soil plots, to study the dynamics of the flow processes, to compare these processes for both plots under natural rainfall and under controlled irrigation and to draw conclusions on the origin of the water leaving the soil plots at the downslope draining ditch.

Initial hypothesis

Given the distinct vegetation and soil moisture status of both plots, I expected much larger runoff volumes from the muck humus as compared to the mor humus plot. In addition, because of the different soil structure in the subsoil of both plots, I assumed that runoff generation is much slower in the mor humus plot.

For the NITREX project the runoff of the sub-catchments N_1 , N_2 and N_3 was and still is continuously observed. Thus, it is possible to directly compare the runoff processes from the sub-catchments and from the soil plots. This led to two additional objectives:

- to determine the relative role of the two soil types for the runoff generation in the sub-catchments.

Initial hypothesis

Given the topographic distribution of both soils, i.e. the muck humus in the depressions and the mor humus on the hillocks, and the fairly wet soil moisture status of the muck humus areas, I assumed that the main part of the runoff is generated in the muck humus areas.

- to determine the influence of the soil type distribution in the three sub-catchments on the generation of runoff

Initial hypothesis

Since only muck humus occurs on the N₃ sub-catchment, I expected larger amounts of runoff and faster response to rainfall in this sub-catchment.

The Forest Hydrology group of the Swiss Federal Institute of Forest, Snow, and Landscape Research kindly provided the runoff data of the Erlenbach headwater catchment. Thus, simultaneous data records measured at three different scales were available. This widened the possibilities of studying the effect of soil, vegetation and scale on the runoff generation and led to a fourth objective of this thesis

- to combine the data on the water fluxes at three spatial scales to examine the effect of catchment size on the dynamics of the runoff processes and to reveal the sources of the runoff water and its flow paths throughout the Erlenbach catchment.

Initial hypothesis

After observing the runoff during a high intensity rainstorm (July, 1995) I assumed that the runoff generation at all scales is very fast and that most of the runoff processes can be located on the soil surface or in the upper soil layers.

To achieve these objectives, I applied two independent strategies:

- First, I modelled the hydrological data by time series analysis and transfer function estimation. After linearization the rainfall-runoff data series of all study areas based on their soil moisture status, discrete transfer functions were estimated. These transfer functions, representing a class of black box and scale independent models, could be interpreted in terms of physically meaningful processes.
- Second, I used water tracers to explore routing and residence times of soil water and solutes. Water quality data of rainfall and runoff were used to classify the origin of runoff water as event or pre-event water. In addition, controlled irrigation experiments with artificial tracers were carried out on the soil plots to detect the flow paths of water and solutes.

The sequel of the introduction is a short overview of the most common hydrological processes which have been reported on. Afterwards, I discuss some applications of time series modelling for identifying the dynamics of runoff generating processes.

1.2 Runoff Generating Processes

Runoff generation in natural catchments is driven by many processes. Parallel processes at different locations or successive processes at the same location are excited by the rainfall and influenced by vegetation, topography, structural properties of the soil, moisture state and scale of the catchment. Since the pioneering work by Horton (1933), who distinguished between infiltration and lateral flow of water on the soil surface, many other hydrological processes have been reported on (e.g. Anderson and Burt, 1990; Bonell, 1993). Hydrological processes span about eight orders of magnitude in space and time (Blöschl and Sivapalan, 1995). Ideally, processes should be observed at the scale they occur: processes larger than the observation scale appear as trends in the data, processes smaller than the observation scale appear as noise (Cushman, 1987).

Despite the growing knowledge about water flow in saturated and unsaturated media, it is not yet possible to investigate the hydrology of a catchment without any field measurements. I briefly describe the processes of runoff generation and discuss the most common modelling approaches. Figure 1.3 schematically shows the most common runoff processes.

1.2.1 Flow Processes on the Soil Surface

According to Horton (1933), it is the infiltration capacity of the upper soil layer which determines how the precipitation is partitioned into surface runoff (“Hortonian overland flow”, **g**) and infiltration (**a**). Under certain conditions, e.g. if a dry hydrophobic litter layer is present (Burch *et al.*, 1988), a temporally reduced infiltration capacity may lead to a “temporary Hortonian overland flow”. This flow process, however, disappears after wetting of the water repellent upper soil layer. If the infiltration continues until the soil matrix is saturated, water will start to flow on the soil surface. This process, which is called “saturation overland flow” (**d**), will then occur independently of the rainfall intensity (Kirkby, 1988). Especially in lower basin areas or near the stream channel, where large sub-surface fluxes of water converge, the ability of the soil to transport water in the lateral direction may be insufficient. This can lead to “return flow” (**i**), the exfiltration of water from the soil surface. Since return

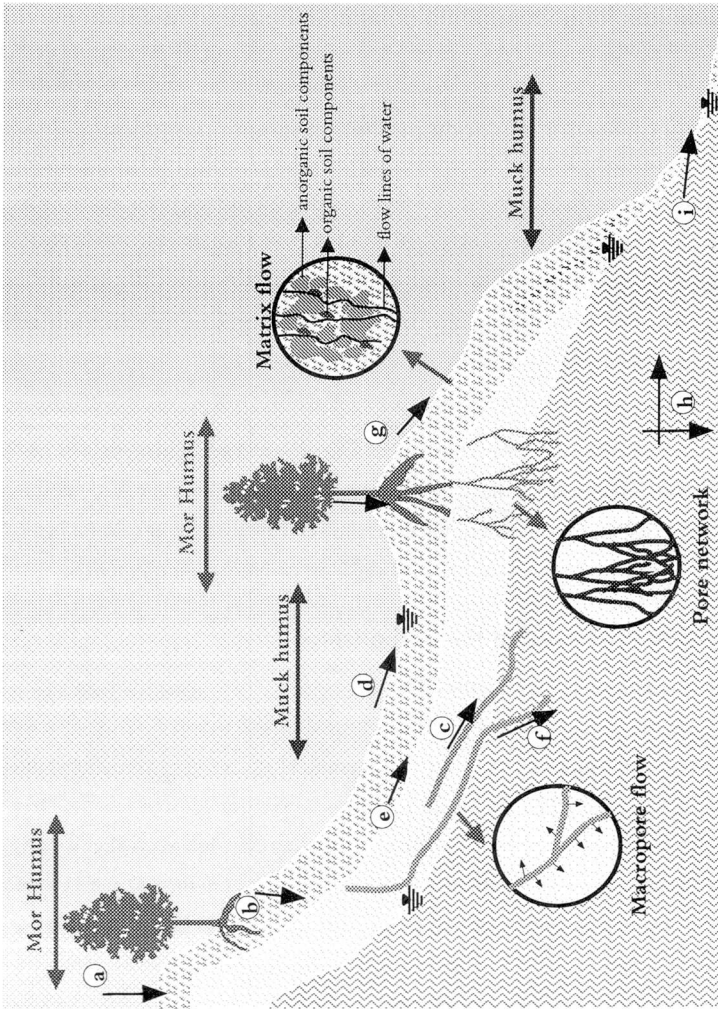


Figure 1.3: An overview of runoff generating processes. Infiltration (a), matrix flow (b), lateral macropore flow (c), saturation overland flow (d), sub-surface lateral matrix flow (e), vertical macropore flow (f), Hortonian overland flow (g), groundwater flow (h) and return flow (i).

flow may originate from the soil matrix or from macropores, it is closely linked to flow processes in the subsoil. According to Dunne and Black (1970), return flow is one of the components contributing to saturation overland flow.

1.2.2 Flow Processes in the Sub-surface

Since most forest soils in humid climate areas have a permeable upper layer, the main portion of the precipitation can be taken up by the soil matrix (**b**) (Bonell, 1993). The water regime of forests is therefore dominated by flow processes in the soil. They can be subdivided into processes within the unsaturated (throughflow, interflow) and in the saturated zone (groundwater flow, deep sub-surface flow) (Ward and Robinson, 1990). In the unsaturated soil layers, sub-surface lateral matrix flow (**e**) is induced when the hydraulic conductivity is larger in the lateral than in the vertical direction. This is the case when more permeable upper soil layers overlay subsoil layers with a small permeability. The less permeable soil layers are natural boundaries for the water flow which leads to local ponding (perched water table). Since the hydraulic conductivity decreases with depth in most soils, this situation is quite common. The importance of sub-surface lateral matrix flow has been shown in numerical calculations by Freeze (1972) and has since been confirmed by many experiments (e.g. Beven, 1982). The streamlines of the lateral sub-surface matrix flow are parallel to the impermeable layers, but they can be diverted by lateral or vertical macro- or mesopores (**c**). When the infiltrating water reaches the capillary fringe of the groundwater, it quickly raises the level of the fully saturated zone. The resulting downslope flow process is governed by the lateral saturated hydraulic conductivity and the presence of preferential flowpaths such as macropores or pipes. Numerous experimental studies (McDonnell, 1990; Turton *et al.*, 1995) have demonstrated the importance of macropores for lateral flow and transport. Sub-surface flow in macropores (**c**, **f**) can contribute substantially to the fast runoff. It has been shown by Mosley (1982), that even flow in highly permeable matrix layers can cause identically fast runoff processes. If the subsoil and the bedrock are highly permeable, or if cracks or fissures into deeper geological layers exists, a substantial part of the infiltrated water percolates towards deeper groundwater layers (**h**).

1.3 Time Series Modelling in Hydrology

Time series analysis is the most common technique to explore, describe and forecast environmental time series. Most time series models have in common that they are a lumped description of natural processes with a small number of known or adjustable parameters. Apart from providing a good description of the data series, time series modelling should lead to models which are physically meaningful and statistically well defined. I will focus on some applications for exploration and forecasting of hydrological time series to illustrate the main features of this model category. For a comprehensive overview of modelling rainfall, flow and mass transport in hydrological systems, see Conell and Todini (1996).

Probably the best known linear model in hydrology is the Instantaneous Unit Hydrograph (IUH), which is mainly used for prediction and forecasting purposes (e.g. Bras, 1990). The IUH can be presented as a time series model in observation space (Jakeman *et al.*, 1990) or in the state space (Rodríguez-Iturbe *et al.*, 1978; Ihringer, 1985). The IUH is a transfer function which transforms the effective water input into the hydrological response of a basin. The central hypothesis of this approach is that the watershed response is linear. For this reason, it can only be applied successfully if the effective, runoff generating water input is known. It is, however, not clear how to make reliable a priori estimates of the effective input. A further disadvantage, besides the constraints of proportionality of rainfall and runoff, is the time invariance of the ordinates of the IUH. Since in general the response of a catchment depends on antecedent moisture conditions, the application of the IUH approach to predict or forecast the nonlinear response of a catchment is cumbersome.

The nonlinear properties of the runoff generation strongly depend on the soil moisture state of the drainage basin. This perception of a natural drainage basin is often used to deal with the nonlinearity. Since direct measurements on soil moisture state variables are rarely available, they are often replaced by a surrogate variable. Since the runoff itself reflects the actual and past moisture conditions of a catchment, is often used as a surrogate variable. Although this is physically not opportune (it implies a feedback between the runoff and the precipitation), it has been successfully applied to make predictions of the runoff. Kachroo and

Natale (1992) subdivided the rainfall-runoff data series into smaller sections with a comparable hydrological status. They classified the runoff as into low, medium, and high flow and estimated the appropriate transfer functions for each classification by means of the method of weighted least squares. A comparable approach are the NeTAR (Nested Threshold Autoregressive) models of Astatkie *et al.* (1997), which were applied to predict the discharge from a stream in northwest Iceland. The authors identified the air temperature during periods of snow- or glacier melt and the state of the basin storage as the main sources of nonlinearity. Daily measurements on the air temperature were available and as a surrogate for the moisture state of the basin, daily discharge values were used. First, they estimated a threshold discharge to distinguish between high and low basin storage. Afterwards, for periods with low basin storage, three temperature classes were distinguished which accorded to low snow or ice melt, high snow or ice melt and glacier ice melt only. For each of the so-defined four regions linear time series models with discharge, precipitation and temperature were determined. This class of threshold models are an extension of the SETAR (Self Exiting Threshold Autoregressive) and TARSO (Open-loop Threshold Autoregressive System) models of Tong (1990) and Tong (1983). A similar approach are the constrained linear systems (CLS), a simple piecewise linear black-box model (Natale and Todini, 1977). Still, this type of threshold models are only one possibility to deal with the nonlinearity of runoff generation. An overview of several techniques to model nonlinear time series is presented by De Gooijer (1992) and by Priestley (1988).

In the state dependent modelling approach of Priestley (1988) the parameters of the time series models vary with time. The parameter variation is linearly related to changes in the precipitation and the runoff. An alternative to the state dependent models is proposed by Young (1989). He proposed to use algorithms of times series analysis to estimate discrete linear transfer functions from observed data. To avoid an overparameterization, the number of parameters of the transfer functions is determined by information criteria based on common statistical inference. These information criteria find a trade-off between the goodness of fit and the precision of the estimated parameters. As a result, parsimonious linear transfer functions, which can often be interpreted in terms of physical processes, can be esti-

mated from observed data. Since processes underlying hydrological data are mainly nonlinear, linear transfer functions may be inappropriate descriptors. Still, time series analysis techniques may be applied to model the nonlinear properties of the runoff generation. This approach uses information on the hydrological state of the catchment to transform the measured precipitation into effective, runoff generating precipitation. Again, if no information on the soil moisture state of the basin is available, the runoff can be used as a surrogate variable.

Also for modelling the transport of solutes in porous media, time series approaches have been applied. An approach described by Beven and Young (1988) treats the porous medium as an Aggregated Mixing Zone (AMZ): two parallel reservoirs, one accounting for the slow component of the solute transport, the other for the fast component. In a way this is similar to the concept of mobile and immobile water of van Genuchten and Wierenga (1976). The slow component accords to that part of the solute which is exchanged between the mobile and the immobile water, the fast component equals the solute which remains in the mobile water. From tracer experiments under steady flow conditions, linear transfer functions may be estimated based on time series of solute input and breakthrough. From these transfer function, the volume of the mobile and immobile pore water can be computed. In fact, the AMZ approach is the discrete counterpart of the continuous transfer functions of Jury (1982). Comparable to the AMZ is the ADZ (Aggregated Dead Zone) model of Young and Wallis (1986) for the dispersion of solutes in rivers.

Often time series models can be given a physically meaningful interpretation. Parlange *et al.* (1992) calculated the mean diffusivity of the soil from the parameters of an AR(1) model. The AR(1) model was fitted to the depth profile of the soil moisture distribution. Young and Beven (1994) showed how the deterministical part of a transfer function model can be interpreted directly in terms of dynamic conservation equations associated with environmental flow processes. The authors divided the total runoff of a catchment into a fast and a slow flow component. A direct extrapolation of both flow components to stormflow and baseflow is, however, not possible. For baseflow separation purposes additional geochemical information is a prerequisite. Moreover, the uncertainty of the parameters of the time series model may lead to associated uncertainty in the derived 'physically meaningful' processes.

Chapter 2

The Modelling of Hydrological Time Series

2.1 Introduction

In time series analysis much of the research concentrates on defining time series models for natural processes. In contrast to many physically based deterministic models, time series models which describe natural systems have a fairly simple model structure. Deterministic models are often based on the researcher's subjective interpretation of reality. In addition, such models tend to be complex and highly overparameterized. The purpose of time series modelling is not the development of complex and subjective physically based simulation models, but rather the development of algorithms which allow to estimate the underlying structure of the observed time series and to capture their main characteristics. The estimation algorithms are objective in the sense that they are based on methods of statistical inference. Although time series models do not rely on a deterministic representation of a natural system, the fitted time series model may in many cases be interpreted in terms of physical processes.

The structure and parameters of a time series model can be determined by a wide variety of estimation, filtering and smoothing algorithms, both in the frequency and the time domain. In the time domain, where the interpretation of time series models is more obvious, dynamical processes can be represented in the 'observation space' and the 'state space'. In the observation space, the input and the output of a system are directly linked by a transfer function. In state space setting, information on the input and output of a system are used to model the temporal variation of a (mostly) unobserved internal system state. The system output is linearly related to the system state. In both representations parsimonious models are preferred and selected by statistical inference.

Parsimonious and physically interpretable time series models are a valuable alternative to the classical deterministic approaches in hydrology and soil physics. However, most time series applications in hydrology focus on *forecasting* of streamflow or climatological phenomena. This thesis rather focuses on *detecting* (soil) physical and hydrological processes from nonlinear time series of water input and runoff. Although very well known in system engineering, most of the estimation techniques applied in this work are not widespread in soil science and hydrology. Therefore, I include a detailed description of estimation procedures and algorithms used in the following chapters. This introduction to the basic theory elements mainly relies on the following textbooks:

- Box, G.E.P., and G.M. Jenkins. 1970. *Time Series Analysis, Forecasting and Control*. Revised edn published 1976. San Francisco, Holden-Day.
- Gelb, A. (ed.). 1979. *Applied Optimal Estimation*, The M.I.T. press.
- Norton, J.P. 1986. *An Introduction to Identification*, London: Academic Press.
- Priestley, M.B. 1988. *Non-linear and Non-Stationary Time Series Analysis*, London: Academic Press.
- Young, P.C. 1984. *Recursive Estimation and Time Series Analysis*, Berlin: Springer-Verlag.

All other sources are mentioned explicitly in the text.

This chapter first discusses time series models in the observation space (2.2). A linear discrete transfer function model is extended to take stochastic disturbances due to measurement

errors and inadequacies of the transfer function into account (2.2.1). To estimate the parameters of the transfer function, a recursive least squares procedure can be applied (2.3.1), but may lead to biased estimates (2.3.2). To by-pass this problem I applied alternative (unbiased) estimation procedures based on the application of an instrumental variable (2.4.1 and 2.4.2). To select an optimal model structure (i.e number of parameters of the transfer functions), it is useful to employ statistical criteria which combine the goodness of fit and the parameters uncertainty (2.5).

Since runoff is mostly generated by a nonlinear process, the application of linear transfer functions is often inconvenient. Still, it is possible to approximate a nonlinear process by a linear transfer function with parameters which vary in time (2.7.2). The estimation of such transfer functions benefits from time series analysis methods in the state space (2.6), which are (in this thesis) mainly filtering (2.6.1) and smoothing (2.6.2) procedures. Although these estimation procedures have been developed to estimate the noise free state of a system, they may be applied to estimate time variable parameters of a linear transfer function too. The correlation of the time variable parameters with information on the hydrological state of the system is applied to linearize the initial rainfall and runoff data. Finally, discrete linear transfer functions may be estimated from the linearized data by the estimation methods of section 2.2. In sections 2.7.3 and 2.7.4, an overview of this linearization procedure is given.

2.2 Time Series Models in the Observation Space

If an exogeneous input variable (e.g. rainfall) acts on a system (e.g. soil), it activates the internal system dynamics (e.g. by changing the soil water content or the depth of the groundwater table), which in turn may result in a system output (e.g. runoff). After cessation of the driving inputs, the internal dynamics steadily diminish and tend towards a state of rest again. A single-input, single-output linear time series model in the *observation space* describes these dynamics mathematically as

$$y_k^* = -a_1 y_{k-1}^* + b_0 u_{k-\delta}^* \quad (2.1)$$

In Eq. (2.1), y_k^* denotes the true system output at time k , $u_{k-\delta}^*$ is the true value of the exogeneous input and δ is a time delay between the beginning of input and output. The gain of the system, b_0 , is the change of the output induced by a unit increase of the input. The autoregressive decrease of the output after cessation of the inputs is expressed by a_1 . As can be seen in the observation space, there is a direct link between input and output, without explicitly modelling the changes in the internal state of the system.

Equation (2.1), being a fully deterministic description of a dynamical process, can be extended to take additional autoregressive and exogeneous components into account as

$$y_k^* = - \sum_{i=1}^n a_i y_{k-i}^* + \sum_{j=0}^m b_j u_{k-j-\delta}^* \quad (2.2)$$

with a_1 to a_n being the parameters defining the autoregressive dynamics and b_0 to b_m the parameters associated with the exogeneous input. In observation space, the model structure of a time series model is summarized as $[n (m+1) \delta]$, where n is the number of autoregressive components, $(m+1)$ the number of input terms and δ a time delay as before. Since the variables y_k^* and $u_{k-\delta}^*$, and the model parameters a_i and b_j have no spatial significance, Eq. (2.2) describes a lumped system with linear dynamics.

The output of dynamical systems depends on its history, not just on the present (or delayed by δ) exogeneous inputs. The past of a system influences the future by a number of initial conditions or 'stored energies', which determine the dynamics of the system. This number of stored energies equals n , the autoregressive model order of the time series model. For hydrological systems (e.g. a drainage basin), the number of stored energies is tantamount of the number of parallel or serial reservoirs.

2.2.1 Sources of Stochastic Disturbances

This section describes how stochastic disturbances may be represented in a time series model. However, for estimation purposes a much simplified time series model will be applied. In Eq. (2.2), the variables associated with the unknown parameters a_i and b_j are exactly known quantities. Further, I assume that the $[n (m+1) \delta]$ model structure describes the system

dynamics completely. This does, however, not cohere with reality: several stochastic disturbances may corrupt the purely deterministic model (2.2). We recognize three sources of stochastic disturbances which may corrupt a time series model (Fig.2.1)

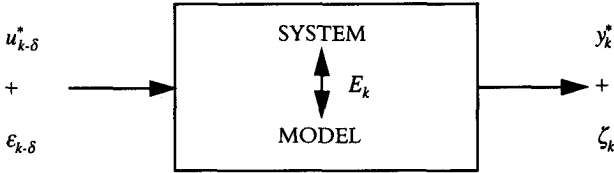


Figure 2.1: Stochastic and deterministic components of time series models.

Because of the invasive character of the measurement process, or because of the specific properties of the measurement devices, the true system inputs $u_{k-\delta}^*$ and system outputs y_k^* are unobservable. They are corrupted by unknown stochastic disturbances. We can define two observable variables y_k and $u_{k-\delta}$ as

$$y_k = y_k^* + \zeta_k \quad (2.3)$$

$$u_{k-\delta} = u_{k-\delta}^* + \varepsilon_{k-\delta}, \quad (2.4)$$

with y_k^* and $u_{k-\delta}^*$ as before the true (but unobservable) system output and input. The measurement errors on the output and the input are stochastic disturbances denoted by $\{\zeta_k\}$ and $\{\varepsilon_{k-\delta}\}$. A third stochastic disturbance may occur if the $[n(m+1)\delta]$ model structure is an imperfect representation of the true system. This can be due to further exogeneous inputs which are not considered in the model or to additional dynamics which are not expressed. Also, since Eq. (2.2) is a linear time series model, possibly nonlinear properties of the true system are not accounted for. All the deviations of the model from the true system are enclosed in the stochastic model error $\{E_k\}$. For the moment, we do not make assumptions on the statistical properties of the noise processes $\{\zeta_k\}$, $\{\varepsilon_{k-\delta}\}$ and $\{E_k\}$.

Taking the stochastic disturbances into account, model (2.2) can be extended to

$$(y_k^* + \zeta_k) = - \sum_{i=1}^n a_i (y_{k-i}^* + \zeta_{k-i}) + \sum_{j=0}^m b_j (u_{k-j-\delta}^* + \varepsilon_{k-j-\delta}) + E_k, \quad (2.5)$$

or

$$y_k = \sum_{i=1}^n a_i y_{k-i} + \sum_{j=0}^m b_j u_{k-j-\delta} + E_k. \quad (2.6)$$

By using the backward shift operator z^{-1} ($z^{-1}y_k = y_{k-1}$ and $z^{-0}y_k = y_k$), the time series model (2.5) can be written as

$$A(z^{-1})(y_k^* + \zeta_k) = B(z^{-1})(u_{k-\delta}^* + \varepsilon_{k-\delta}) + E_k, \quad (2.7)$$

with

$$A(z^{-1}) = z^{-0} + a_1 z^{-1} + \dots + a_n z^{-n} \quad (2.8)$$

and

$$B(z^{-1}) = b_0 z^{-0} + b_1 z^{-1} + \dots + b_m z^{-m}. \quad (2.9)$$

Because of the linearity of the backward shift operator, the right hand-side of Eq. (2.7) takes the form

$$B(z^{-1})(u_{k-\delta}^* + \varepsilon_{k-\delta}) + E_k = B(z^{-1})u_{k-\delta}^* + B(z^{-1})\varepsilon_{k-\delta} + E_k. \quad (2.10)$$

Without loss of generality, the time delay δ can be set equal to zero. Equation (2.10) changes to

$$B(z^{-1})(u_k^* + \varepsilon_k) + E_k = B(z^{-1})u_k^* + b_0 \varepsilon_k + b_1 \varepsilon_{k-1} + \dots + b_m \varepsilon_{k-m} + E_k. \quad (2.11)$$

If the stochastic disturbances $\{\varepsilon_k\}$ and $\{E_k\}$ are white noise (i.e. zero mean, serially uncorrelated random variables), they can be combined into a new stochastic term

$$B(z^{-1})(u_k^* + \varepsilon_k) + E_k = B(z^{-1})u_k^* + (b_0 \varepsilon_k + E_k) + b_1 \varepsilon_{k-1} + \dots + b_m \varepsilon_{k-m}. \quad (2.12)$$

Since a linear combination of two white noise processes again equals a white noise process, e.g. $(b_0\varepsilon_k + E_k) = e_k$, Eq. (2.12) can be reformulated as

$$B(z^{-1})(u_k^* + \varepsilon_k) + E_k = B(z^{-1})u_k^* + e_k + c_1e_{k-1} + \dots + c_p e_{k-p}, \quad (2.13)$$

or, by using the backward shift operator

$$B(z^{-1})(u_k^* + \varepsilon_k) + E_k = B(z^{-1})u_k^* + C(z^{-1})e_k. \quad (2.14)$$

The parameters c_1 to c_p of the polynomial $C(z^{-1})$ are introduced to allow for more flexibility in the description of the stochastic disturbances. The stochastic time series model (2.7) now takes the following form

$$A(z^{-1})(y_k^* + \zeta_k) = B(z^{-1})u_k^* + C(z^{-1})e_k \quad (2.15)$$

or

$$y_k = \frac{B(z^{-1})}{A(z^{-1})}u_k^* + \frac{C(z^{-1})}{A(z^{-1})}e_k. \quad (2.16)$$

This model clearly shows how the observed output variable depends on a purely deterministic input variable and a stochastic noise model. The same result can be obtained if the combined stochastic disturbances $\{\varepsilon_k\}$ and $\{E_k\}$ are described by a moving average process of the form

$$e_k = \sum_{i=0}^r f_i \varepsilon_{k-i} + \sum_{j=0}^{r'} f_j' E_{k,j}. \quad (2.17)$$

In other words, if we assume that the measurement error of the input and the model error do not introduce their own dynamics (i.e. they have no autoregressive components), then the noise model has the same autoregressive behaviour as the observed output variable. Further, it is no longer possible to distinguish in model (2.16) between the measurement errors of the input and the model errors and we shall use the same stochastic process $\{\varepsilon_k\}$ to denote both processes.

It is possible to describe the combined noise processes $\{\varepsilon_k\}$ and $\{E_k\}$ with an autoregressive (AR) model of the form

$$e_k = \sum_{i=1}^s g_i e_{k-i} + \varepsilon_k, \quad (2.18)$$

or an autoregressive moving average (ARMA) model (Eq. 2.19).

$$e_k = \sum_{i=1}^s g_i e_{k-i} + \sum_{i=0}^r f_i \varepsilon_{k-i}. \quad (2.19)$$

This changes the dynamics of the noise model in Eq. (2.16) towards a more general transfer function of the form

$$y_k = \frac{B(z^{-1})}{A(z^{-1})} u_k^* + \frac{C(z^{-1})}{D(z^{-1})} e_k, \quad (2.20)$$

with

$$D(z^{-1}) = z^0 + d_1 z^{-1} + \dots + d_q z^{-q}. \quad (2.21)$$

In model (2.20), $D(z^{-1})$ is a combination of the autoregressive dynamics in $A(z^{-1})$ and of the ARMA model (2.19). This model is a special case of the most general parametric model structure¹

$$A(z^{-1})y_k = \frac{B(z^{-1})}{F(z^{-1})} u_k^* + \frac{C(z^{-1})}{D(z^{-1})} e_k, \quad (2.22)$$

which allows for more flexibility in the noise description and for common or different poles (or dynamics) for the different inputs.

Transfer function models (2.16) and (2.20) express how stochastic disturbances of the input variable and the model error can be dealt with in a time series model. However, the effect

¹. The so-called *Box-Jenkins* model structure (Box and Jenkins, 1970), can be derived from Eq. (2.22) by setting $A(z^{-1})$ equal to one, the *Output-Error* model structure by setting $A(z^{-1})$, $C(z^{-1})$ and $D(z^{-1})$ equal to one.

of the stochastic disturbance on the output measurements $\{\zeta_k\}$ is still not considered. Because disturbances on the output measurements do not pass the system, it can be argued that they have a smaller impact on the estimation procedure than $\{\varepsilon_k\}$ or $\{E_k\}$, and may often be neglected. Nevertheless, as will be shown later, their presence affects the estimation procedures which are used to estimate the model parameters.

The stochastic disturbances $\{\zeta_k\}$, $\{\varepsilon_k\}$ and $\{E_k\}$ are unknown a priori. As a consequence, it is not possible to use the transfer functions (2.16) or (2.20) as a basis for inference. For practical purposes we can assume a fairly general transfer function model of the form

$$y_k = \frac{B(z^{-1})}{A(z^{-1})} u_{k-\delta} + \eta_k, \quad (2.23)$$

with y_k and $u_{k-\delta}$ the observable system output and input, and a model for the stochastic disturbances of the form

$$\eta_k = \sum_{i=1}^q d_i \eta_{k-i} + \sum_{i=1}^p c_i e_{k-i} = \frac{C(z^{-1})}{D(z^{-1})} e_k, \quad (2.24)$$

with $\{e_k\}$ a white noise process. In this model, we treat $\{\eta_k\}$ as a general noise term which accounts for all the stochastic and deterministic effects on y_k which are not due to the effect of the unmeasurable true input signal. Therefore, the parameters of the noise model (2.24) must be interpreted independently of Eq. (2.20) and of the general transfer function model (2.23).

The transfer function model (2.23) can be rewritten in matrix notation by first using the definition of the backward shift operator

$$y_k = -a_1 y_{k-1} - \dots - a_n y_{k-n} + b_0 u_{k-\delta} + \dots + b_m u_{k-m-\delta} + \eta_k \quad (2.25)$$

and then grouping the parameters and the observations into two vectors. This results in

$$y_k = \mathbf{z}_k^T \mathbf{a} + \eta_k, \quad (2.26)$$

with

$$\mathbf{z}_k^T = [y_{k-1}, \dots, y_{k-n}, u_{k-\delta}, \dots, u_{k-m-\delta}] \quad (2.27)$$

and

$$\mathbf{a} = [-a_1, \dots, -a_n, b_0, \dots, b_m]^T. \quad (2.28)$$

In Eq. (2.26), \mathbf{z}_k^T is a $(n + m + 1)$ row vector with the present and the past input and output variables, \mathbf{a} is a $(n + m + 1)$ column vector containing the model parameters and the scalar noise process $\{\eta_k\}$.

2.3 Estimation of Transfer Function Models

2.3.1 Recursive Least Squares

The least squares approach is a simple and concise technique to estimate parameters of regression models. Assume we have observed the input to and the output from a system at a number N of points $t_k = t_0 + k\Delta t$, $k = 1, 2, \dots, N$. Using Eq. (2.26), we model the k -th observation, $k > \max(n, m + 1 + \delta)$, as a linear combination of past output and present and past input values. Without loss of generality we can assume that the time delay δ equals zero. In an initial step, using only the first $K = \max(n, m + 1) + 1$ observations, we obtain an estimate $\hat{\mathbf{a}}^{(K)}$ of the parameter vector, \mathbf{a} , from the well known normal equations

$$\hat{\mathbf{a}}^{(K)} = \left[\sum_{k=1}^K \mathbf{z}_k \mathbf{z}_k^T \right]^{-1} \sum_{k=1}^K \mathbf{z}_k y_k. \quad (2.29)$$

Of course, $\hat{\mathbf{a}}^{(K)}$ minimizes the sum of squared deviations

$$J = \sum_{k=1}^K (y_k - \mathbf{z}_k^T \hat{\mathbf{a}}^{(K)})^2, \quad (2.30)$$

and the estimation error after K observations, $\tilde{\mathbf{a}}^{(K)} = \hat{\mathbf{a}}^{(K)} - \mathbf{a}$, of the parameter vector is given by

$$\tilde{\mathbf{a}}^{(K)} = \left[\sum_{k=1}^K \mathbf{z}_k \mathbf{z}_k^T \right]^{-1} \sum_{k=1}^K \mathbf{z}_k y_k - \mathbf{a} = \left[\sum_{k=1}^K \mathbf{z}_k \mathbf{z}_k^T \right]^{-1} \sum_{k=1}^K \mathbf{z}_k \eta_k. \quad (2.31)$$

To derive this expression we used the model (2.26) and the definition of $\hat{\mathbf{a}}^{(K)}$ (Eq. 2.29). In principle, the same approach can be used to estimate \mathbf{a} from all the available data. However, such an 'en bloc' solution to the estimation problem is computationally inefficient and a recursive algorithm which updates the estimate $\mathbf{a}^{(K-1)}$ at each time instant k , $k > K = \max(n, m + 1) + 1$, is preferable. To clearly distinguish between the 'en bloc' solution and the recursive solution, the recursive estimate of the parameter vector after $k = K$ observations will be denoted as $\hat{\mathbf{a}}_k$, the corresponding estimation error as $\tilde{\mathbf{a}}_k$. The recursive form of the least squares equation is

$$\hat{\mathbf{a}}_k = \hat{\mathbf{a}}_{k-1} - \mathbf{K}_k \left[y_k - \mathbf{z}_k^T \hat{\mathbf{a}}_{k-1} \right], \quad (2.32)$$

with

$$\mathbf{K}_k = \mathbf{P}_{k-1} \mathbf{z}_k \left[\mathbf{1} + \mathbf{z}_k^T \mathbf{P}_{k-1} \mathbf{z}_k \right]^{-1} \quad (2.33)$$

and

$$\mathbf{P}_k = \mathbf{P}_{k-1} - \mathbf{P}_{k-1} \mathbf{z}_k \left[\mathbf{1} + \mathbf{z}_k^T \mathbf{P}_{k-1} \mathbf{z}_k \right]^{-1} \mathbf{z}_k^T \mathbf{P}_{k-1}. \quad (2.34)$$

The derivation of the recursive form of the least squares estimator, a straightforward application of linear algebra, is given in Appendix A.1.

Superficially, the time series model (2.26) equals a multiple regression model. In regression analysis, however, the independent and dependent variable are clearly separated. In time series model (2.26), the observed output y_k serves as both the independent (in \mathbf{z}_k) and dependent variable. However, Mann and Wald (1953, cited in Priestley, 1981, pp. 353) showed that the optimal properties of the least squares estimation are preserved for $N \rightarrow \infty$.

The estimation algorithm (2.32), as it has been derived, is a deterministic procedure in the sense that it ignores the statistical nature of the input and output measurements and of the noise process $\{\eta_k\}$. As a consequence, it does not provide any information on the quality of the estimates. In order to provide statistical information on the nature of the estimates, it is necessary to make some assumptions about the noise process $\{\eta_k\}$.

- a) $E[\eta_k] = 0$, i.e., $\{\eta_k\}$ is a zero mean noise process.
- b) $E[\eta_k \eta_{k+i}] = \sigma^2$ if $i = 0$, else $E[\eta_k \eta_{k+i}] = 0$, which means that $\{\eta_k\}$ is serially uncorrelated.
- c) $E[\eta_k \mathbf{z}_k] = 0$, i.e. the noise $\{\eta_k\}$ is independent of the measurements vector \mathbf{z}_k .

This condition implies the independence of $\{\eta_k\}$ and \mathbf{z}_l for $l \leq k$

Because of these assumptions, the expected value of the estimation error at time instant k , $E[\tilde{\mathbf{a}}_k]$, is equal to zero. Thus, the covariance matrix of the estimation errors at time instant k , \mathbf{P}_k^* , can be defined as

$$\mathbf{P}_k^* = E[\tilde{\mathbf{a}}_k \tilde{\mathbf{a}}_k^T] = \text{Cov}[\tilde{\mathbf{a}}_k, \tilde{\mathbf{a}}_k]. \quad (2.35)$$

With Eq. (2.31) this expression can be written as

$$\mathbf{P}_k^* = E \left[\left[\sum_{l=1}^k \mathbf{z}_l \mathbf{z}_l^T \right]^{-1} \sum_{l=1}^k \mathbf{z}_l \eta_l \sum_{l=1}^k \mathbf{z}_l^T \eta_l \left[\sum_{l=1}^k \mathbf{z}_l \mathbf{z}_l^T \right]^{-1} \right]. \quad (2.36)$$

Taking the expectation of the right-hand side of this expression conditional on $\mathbf{z}_l, l = 1, 2, \dots, k$ and using the second and the third assumption, we find

$$\mathbf{P}_k^* = \left[\sum_{l=1}^k \mathbf{z}_l \mathbf{z}_l^T \right]^{-1} \sum_{l=1}^k \mathbf{z}_l \mathbf{z}_l^T \left[E_{\eta|\mathbf{z}} [\eta_l \eta_l] \right] \left[\sum_{l=1}^k \mathbf{z}_l \mathbf{z}_l^T \right]^{-1}. \quad (2.37)$$

Then taking the expectation with respect to the marginal distribution of \mathbf{z}_i , $i = 1, 2, \dots, k$, and recalling that the conditional expectation is identical to the marginal expectation for independent random variables, we obtain

$$\mathbf{P}_k^* = \sigma^2 \mathbf{P}_k = \sigma^2 \left[\sum_{l=1}^k \mathbf{z}_l \mathbf{z}_l^T \right]^{-1}. \quad (2.38)$$

Thus, the covariance matrix of the estimation errors \mathbf{P}_k^* is the product of the noise variance σ^2 and the \mathbf{P}_k matrix. The covariance matrix may be calculated at any time instant k with either an a priori estimate of the noise variance σ^2 or an estimate recursively obtained from the data as

$$\hat{\sigma}_k^2 = \hat{\sigma}_{k-1}^2 - \frac{1}{k} (\hat{\sigma}_{k-1}^2 - \eta_k^2). \quad (2.39)$$

The recursive equations offer the possibility to evaluate the accuracy of the estimates at each time instant k . In addition, in contrast to the 'en bloc' least squares solution of Eqs. (2.29) and (2.31), there is no need for a matrix inversion. However, the least squares solution has to be used with care. The assumptions made to derive the covariances of the estimation error of the least squares solution conflict with the stochastic properties of Eq. (2.26) and, in addition, the algorithm is based on hypothetical noise free output measurements. Moreover, the elements of \mathbf{P}_k^* in Eq. (2.38) behave like that of \mathbf{P}_k except for the scaling factor σ^2 and are a strictly decreasing function of the sample size. This means that the influence of observational data on the estimate of $\hat{\mathbf{a}}_k$ decreases towards the end of the data series. Although this is statistically consistent, it limits the use of the least squares approach for nonstationary or nonlinear time series.

As a further limitation of the least squares approach, errors in the measurement of the output and the input variable will impair the estimate $\hat{\mathbf{a}}_k$. The next section shows, how the presence of the noise process $\{\zeta_k\}$ associated with the variables in \mathbf{z}_k and the properties of $\{\eta_k\}$ may provoke a bias of the estimated parameters $\hat{\mathbf{a}}_k$, no matter how many data are used.

2.3.2 Biased Least Squares Estimation

The \mathbf{z}_k vector in Eq. (2.26) contains noisy past output and noisy present and past input variables. We can, however, assume that the influence of noise sources associated with the input variables is completely comprised in the noise process $\{\eta_k\}$. Thus, in analogy to Eq. (2.26), we can write

$$y_k = \sum_{i=1}^n a_i y_{k-i} + \sum_{j=0}^m b_j u_{k-j-\delta}^* + \eta_k. \quad (2.40)$$

Since the input variable is supposed to be noise-free, the noise process $\{\eta_k\}$ of (2.40) differs from that of (2.26). By separating the true output variable and the measurement errors

$$y_k^* + \zeta_k = \sum_{i=1}^n a_i (y_{k-i}^* + \zeta_{k-i}) + \sum_{j=0}^m b_j u_{k-j-\delta}^* + \eta_k, \quad (2.41)$$

or

$$y_k^* = \sum_{i=1}^n a_i (y_{k-i}^* + \zeta_{k-i}) + \sum_{j=0}^m b_j u_{k-j-\delta}^* + \eta_k - \zeta_k \quad (2.42)$$

After combining the elements of the noise vectors, i.e. $\eta_k - \zeta_k = E_k'$, and by writing Eq. (2.42) in matrix notation follows

$$y_k^* = \mathbf{z}_k^T \mathbf{a}_k + E_k', \quad (2.43)$$

with

$$\mathbf{z}_k^T = \left[(y_{k-1}^* + \zeta_{k-1}), \dots, (y_{k-n}^* + \zeta_{k-n}), u_{k-\delta}^*, \dots, u_{k-\delta-m}^* \right] \quad (2.44)$$

and \mathbf{a}_k the vector with the parameters as before.

The normal equation which gives the least squares solution to this problem is

$$\hat{\mathbf{a}}_k = \left[\sum_{l=1}^k \mathbf{z}_k \mathbf{z}_k^T \right]^{-1} \sum_{l=1}^k \mathbf{z}_k y_k, \quad (2.45)$$

or, after combining with Eq. (2.43) and rearranging terms

$$\hat{\mathbf{a}}_k = \mathbf{a}_k + \left[\sum_{l=1}^k \mathbf{z}_k \mathbf{z}_k^T \right]^{-1} \sum_{l=1}^k \mathbf{z}_k E'_k. \quad (2.46)$$

To calculate the bias of the estimation, we take the expectation of the above expression

$$E_{E'|\zeta}[\hat{\mathbf{a}}_k] = E_{\zeta} \left[\mathbf{a}_k + \left[\sum_{l=1}^k \mathbf{z}_k \mathbf{z}_k^T \right]^{-1} E_{E'|\zeta} \left[\sum_{l=1}^k \mathbf{z}_k E'_k \right] \right], \quad (2.47)$$

with $E_{E'|\zeta}$ the expected value of the noise process $\{E'\}$ given $\{\zeta\}$ and $E_{\zeta}[\mathbf{a}_k] = \mathbf{a}_k$. Thus, the least squares solution is only unbiased if the second term of the right-hand side of Eq. (2.47) is zero. We can separate the stochastic disturbances from the vector with the observations \mathbf{z}_k as follows

$$\mathbf{z}_k = \mathbf{z}_k^* + \zeta_{zk}, \quad (2.48)$$

with

$$\zeta_{zk} = \left[\zeta_{k-1}, \dots, \zeta_{k-n}, 0, \dots, 0 \right] \quad (2.49)$$

and

$$\mathbf{z}_k^* = \left[y_{k-1}^*, \dots, y_{k-n}^*, u_{k-\delta}^*, \dots, u_{k-\delta-m}^* \right]. \quad (2.50)$$

It is now possible to write the expectation term of Eq. (2.47) as

$$E_{E'|\zeta} \left[\sum_{l=1}^k \mathbf{z}_k E'_k \right] = E_{E'|\zeta} \left[\sum_{l=1}^k \mathbf{z}_k^* E'_k \right] + E_{E'|\zeta} \left[\sum_{l=1}^k \zeta_{zk} E'_k \right], \quad (2.51)$$

or in the following expression for all the time instants in the time series, assuming $n > m + 1$ and by setting δ equal to zero

$$E_{E'|\zeta} \left[\sum_{l=1}^k \mathbf{z}_k E'_k \right] = E_{E'|\zeta} \begin{bmatrix} y_n^* & y_{n+1}^* & \cdots & y_{N-1}^* \\ y_{n-1}^* & y_n^* & \cdots & y_{N-2}^* \\ \cdots & \cdots & \cdots & \cdots \\ y_1^* & y_2^* & \cdots & y_{N-n}^* \\ u_{n+1}^* & u_{n+2}^* & \cdots & u_N^* \\ \cdots & \cdots & \cdots & \cdots \\ u_{n+1-m}^* & u_{n+2-m}^* & \cdots & u_{N-m}^* \end{bmatrix} \cdot \begin{bmatrix} \eta_{n+1} - \zeta_{n+1} \\ \eta_{n+2} - \zeta_{n+2} \\ \cdots \\ \eta_N - \zeta_N \end{bmatrix} \quad (2.52)$$

$$+ E_{E'|\zeta} \begin{bmatrix} \zeta_n & \zeta_{n+1} & \cdots & \zeta_{N-1} \\ \zeta_{n-1} & \zeta_n & \cdots & \zeta_{N-2} \\ \cdots & \cdots & \cdots & \cdots \\ \zeta_1 & \zeta_2 & \cdots & \zeta_{N-n} \\ 0 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 0 \end{bmatrix} \cdot \begin{bmatrix} \eta_{n+1} - \zeta_{n+1} \\ \eta_{n+2} - \zeta_{n+2} \\ \cdots \\ \eta_N - \zeta_N \end{bmatrix}$$

By inspecting Eq. (2.52), it is obvious that the least squares estimation will be unbiased only if both noise processes $\{\zeta_k\}$ and $\{\eta_k\}$ are zero mean serially uncorrelated random variables and are mutually independent. If a priori information on the noise statistics is available, this bias can be corrected for. However, since this is rarely the case, other methods are required to obtain an unbiased estimate.

2.4 Instrumental Variables

2.4.1 Ordinary Instrumental Variables

Instrumental variable (IV) techniques are based on the assumption that an additional variable $\hat{\mathbf{x}}_k^*$ exists with the following properties

- $\hat{\mathbf{x}}_k^*$ is strongly correlated with the unmeasurable noise-free observations vector \mathbf{z}_k^*
- $\hat{\mathbf{x}}_k^*$ is independent with the noise processes $\{\zeta_k\}$ and $\{\eta_k\}$

If an IV is available which fulfils both requirements, it can be easily used to formulate an alternative normal equation of the form

$$\hat{\mathbf{a}}_k = \left[\sum_{l=1}^k \hat{\mathbf{x}}_l^* \mathbf{z}_l^T \right]^{-1} \sum_{l=1}^k \hat{\mathbf{x}}_l^* y_l. \quad (2.53)$$

It can be seen from Eq. (2.52) and from the properties of the IV that the bias of the least squares estimation based on the application of an IV is zero. Estimating the bias of $\hat{\mathbf{a}}$ calculated by means of Eq. (2.53) results in

$$E_{E'|\zeta}[\hat{\mathbf{a}}_k] = E_{\zeta} \left[\mathbf{a}_k + \left[\sum_{l=1}^k \hat{\mathbf{x}}_l^* \mathbf{z}_l^T \right]^{-1} E_{E'|\zeta} \left[\sum_{l=1}^k \hat{\mathbf{x}}_l^* E_k' \right] \right], \quad (2.54)$$

which is analogue to Eq. (2.47). Since the IV is independent of the noise processes $\{\zeta_k\}$ and $\{\eta_k\}$, the expected value

$$E_{E'|\zeta} \left[\sum_{l=1}^k \hat{\mathbf{x}}_l^* E_k' \right] \quad (2.55)$$

equals zero and the estimation is unbiased. Clearly, with the definition of an IV the estimation problem is not solved. Rather it has reduced to finding an adequate IV vector. Recalling section 2.3.2, we showed that the error $\{\zeta_k\}$ on the output measurements can cause a biased

least squares estimation. In most cases, $\{\zeta_k\}$ is not a white noise process. For this reason, $\{E_k^1\}$ will be correlated with present and past values of $\{\zeta_k\}$. One possibility to deal with this problem is the introduction of an IV in the normal equations, as in Eq. (2.53).

A possible IV vector can be a second measurement of y_k , with independent noise characteristics. Another possibility is using 'smoothed' values of the output variable. The smoothing is likely to reduce the effect of the noise process $\{\zeta_k\}$ on the estimation. Along this line of thinking, Young (1985) proposed an auxiliary model to estimate an IV vector. The auxiliary model has the following form in transfer function notation

$$\hat{y}_k^{IV} = \frac{\hat{\beta}(z^{-1})}{\hat{\alpha}(z^{-1})} u_{k-\delta}, \quad (2.56)$$

with

$$\hat{\beta}(z^{-1}) = \beta_0 + \beta_1 z^{-1} + \dots + \beta_m z^{-m} \quad (2.57)$$

and

$$\hat{\alpha}(z^{-1}) = 1 + \alpha_1 z^{-1} + \dots + \alpha_n z^{-n}. \quad (2.58)$$

The polynomials $\hat{\alpha}(z^{-1})$ and $\hat{\beta}(z^{-1})$ are chosen and updated in a recursive or iterative fashion, based on the estimation results. The IV vector $\hat{\mathbf{x}}_k^*$ is then composed by replacing the output measurements in \mathbf{z}_k^T (Eq. 2.44) by the outcome of the auxiliary model (2.56). Proceeding through the iterations or recursions, the estimate $\hat{\mathbf{x}}_k^*$ will progressively become more correlated with the true but unmeasurable observations vector \mathbf{z}_k^* . For a more extensive description of the IV technique, please refer to Sönderström and Stoica (1983).

An iterative scheme, which may be applied to estimate the parameter vector by means of an instrumental variable approach, takes the following form:

1. Calculate an initial estimate of the parameter vector $\hat{\mathbf{a}}_k$ by means of a (biased) least squares procedure. This results in a set of a_i and b_j parameters of the transfer function Eq. (2.23).

2. By means of Eq. (2.56), compute the instrumental variable $\hat{\mathbf{x}}_k^*$. The parameters of the polynomials $\hat{\alpha}(z^{-1})$ and $\hat{\beta}(z^{-1})$ are set equal to those obtained from the least squares estimation at the previous step, or, $\alpha_i = a_i$ and $\beta_j = b_j$.
3. Re-estimate the parameter vector $\hat{\mathbf{a}}_k$ by application of the altered least squares normal equation (2.53). This again results in a set of a_i and b_j parameters, which can be utilized to compute the IV in step 2. The iteration continues until the parameter estimation converges. A schematic overview of the IV procedure is given in Fig. 2.2.

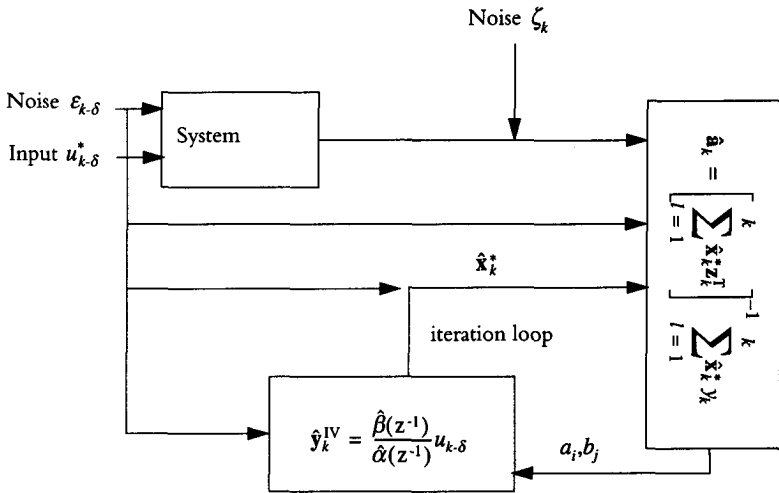


Figure 2.2: The Instrumental Variable procedure (Section 2.4.1).

2.4.2 Refined and Simplified Instrumental Variables

More sophisticated versions of the IV methodology may be formulated. In the previous section (2.4.1), we made no specific assumptions on the nature of the noise process $\{\eta_k\}$. Still, the noise process $\{\eta_k\}$ can be modelled by means of an ARMA process, as in Eq. (2.59).

$$y_k = \frac{B(z^{-1})}{A(z^{-1})}u_{k-\delta} + \frac{C(z^{-1})}{D(z^{-1})}e_k, \quad (2.59)$$

with $\{e_k\}$ being a zero mean serially uncorrelated random variable. The Refined Instrumental Variable (RIV) technique can be used to estimate the model parameter for the complete transfer function model—Eq. (2.59). The details of this estimation procedure, which is both recursive and iterative, are explained in Young (1985). A simplification of the RIV procedure, the SRIV (Simplified RIV), can be obtained by assuming that both $C(z^{-1})$ and $D(z^{-1})$ are equal to one.

2.5 On Model Identification and Evaluation

The time series model (2.23) contains a deterministic part which is associated with the observed input and a stochastic part, associated with a white noise process $\{e_k\}$ as described in Eq. (2.24). The estimation of time series models of this kind, with a deterministic and a stochastic part, typically follows a sequential approach: in a first step, the deterministic part of the time series model is considered to obtain an estimation of the polynomials of Eq. (2.60)

$$y_k = \frac{\hat{B}(z^{-1})}{\hat{A}(z^{-1})}u_{k-\delta} + \hat{\eta}_k. \quad (2.60)$$

For this purpose, instrumental variable techniques as in the previous section (2.4) may be applied. From the difference between the estimated and the measured output, the noise sequence $\{\hat{\eta}_k\}$ can be computed. In a second stage of the estimation procedure, an adequate noise model is determined, most of the time with an autoregressive moving-average model structure of the form

$$\hat{\eta}_k = \frac{\hat{D}(z^{-1})}{\hat{C}(z^{-1})}e_k, \quad (2.61)$$

or as an autoregressive model of the form

$$\hat{\eta}_k = \frac{1}{\hat{C}(z^{-1})} e_k. \quad (2.62)$$

For both steps of the estimation procedure two issues are crucial:

- which is the optimal model structure (order), i.e. the most appropriate order of the model polynomials A , B , C and D and the size of the time delay δ , and
- how well does the selected optimal model structure describe the time series.

Evidently, if the number of model parameters increases, more of the initial variation in the data is accounted for. However, the uncertainty on the parameter estimates will not necessarily decrease. If a model is overparameterized, the importance of a single parameter in the total model structure decreases, which is expressed by a larger variance of the estimation error of the parameter estimation. Thus, it is necessary to apply criteria which define how well a model performs (deviation between measured and estimated outputs) and which reflect the variance of the estimation error of the parameter estimation (thus, the ‘parameter efficiency’) of the model. Despite considerable efforts (see de Gooijer *et al.*, 1985), there is still no unique solution to determine the structure and goodness of fit of a time series model from experimental data.

2.5.1 Criteria based on the Goodness of Fit

2.5.1.1 The Mean Squared Error

$$\text{MSE} = \frac{1}{N} \sum_{k=\mu}^N (y_k - \hat{y}_k)^2 \quad (2.63)$$

with $\mu = \max(n, m + \delta + 1) + 1$. This criterion can be used to compare different model structures estimated from the same input-output data series. Because the expected value of the residual variance changes with the length of the time series, the criterion cannot be used to compare models estimated from different series of data.

2.5.1.2 The Nash/Scutcliffe Criterion (Coefficient of Determination)

$$R_T^2 = 1 - \frac{\sum_{k=\mu}^N (y_k - \hat{y}_k)^2}{\sum_{k=\mu}^N (y_k - \bar{y}_k)^2} \quad (2.64)$$

with \bar{y}_k being the arithmetic mean of the output data series. Using this criterion, it is possible to compare models fitted to different series of data. However, if the available data are split into a calibration and a verification series, then this criterion should be used with care. If, for instance, the verification period is drier than the calibration period, there will be less runoff. As a consequence, the differences between y_k and \bar{y}_k will be smaller, and in turn, smaller R_T^2 values than for the calibration period will result.

2.5.2 Criteria based on the Parameter Efficiency

2.5.2.1 EVN (Error Variance Norm)

The EVN was specifically developed for use with instrumental variable methodologies.

$$\text{EVN} = \frac{\sigma_\eta^2}{n+m+1} \sum_{i=1}^{n+m+1} \hat{p}_{ii}(N) \quad (2.65)$$

with $\hat{p}_{ii}(N)$ the covariance of the estimation error of the i -th parameter. If the parameter estimation is performed by means of a recursive least squares or instrumental variable technique, then the estimation error at the last recursion (time instant N) is used. This corresponds to the i -th diagonal element of the \mathbf{P}_k^* matrix at $k = N$. The EVN is a measure of the overall variance (and thus efficiency) of the model parameters. The more parameters a model contains, the larger the estimation errors on the parameters and thus the larger the EVN will be. It can be altered into a normalized form (NEVN)

$$\text{NEVN} = \frac{\sigma_{\eta}^2}{n+m+1} \sum_{i=1}^{n+m+1} \frac{\hat{p}_i(N)}{\alpha_i^2(N)} \quad (2.66)$$

with $\alpha_i^2(N)$ the recursive estimate of the i -th parameter of the $\hat{\mathbf{a}}_k$ vector at time instant $k = N$ and σ_{η}^2 the variance of the residuals from the modelling. If the estimation error of the parameters small compared to the parameter value, the NEVN will be small, too.

2.5.3 Information Criteria

Methods based on information measures are a mathematical expression of the principle of parsimony. They combine the goodness of fit with the efficiency of the model parameters. The purpose of the application of an information criterion is determining the time series model which describes the observed data best by using the smallest number of parameters.

2.5.3.1 Young Information Criterion (YIC)

The YIC, which has been developed for single-input single-output transfer function models, combines the residual variance of the model with the parameter efficiency into one criterion:

$$\text{YIC} = \ln \left(\frac{\sum_{k=\mu}^N (y_k - \hat{y}_k)^2}{N} \right) + \ln(\text{NEVN}) \quad (2.67)$$

The YIC will be minimal only if the model fit and the parameter efficiency are optimal.

2.5.3.2 Akaike Information Criterion (AIC)

To identify the autoregressive order of the noise model (Eq. 2.62), an alternative identification criterion based on Akaike (1974) can be used

$$\text{AIC} = (N - \mu) \log \hat{\sigma}_{\eta}^2 + 2\nu \quad (2.68)$$

with ν the number of fitted parameters, $\mu = \max(n, m + \delta + 1)$ and

$$\hat{\sigma}_{\eta}^2 = \frac{1}{N - \mu} \sum_{k = \mu + 1}^N (y_k - \hat{y}_k)^2 \quad (2.69)$$

The AIC tries, as the YIC, to reach an optimal balance between reducing the residual variance and increasing the number of fitted model parameters. The AIC criterion can be applied to moving average (MA) and autoregressive moving average (ARMA) models, too.

2.6 Time Series in the State Space

2.6.1 The Kalman Filter

The state space representation of a finite parameter linear model provides a system description which is more closely related to physical reality than observation space representations. It clearly distinguishes between internal dynamics of a system (Eq. 2.70) and the observation process (Eq. 2.71)

$$\mathbf{x}_k = \Phi_{k-1} \mathbf{x}_{k-1} + \mathbf{L}_{k-1} \mathbf{U}_{k-1} + \mathbf{G}_{k-1} \mathbf{w}_{k-1} \quad (2.70)$$

$$\mathbf{Y}_k = \mathbf{H}_k \mathbf{x}_k + \mathbf{v}_k, \quad (2.71)$$

where the time index k takes the values from $k=1, 2, \dots, N$ and

- \mathbf{x}_k is an n -dimensional state vector
- \mathbf{U}_{k-1} is an m -dimensional vector with the actual inputs
- \mathbf{w}_{k-1} is a p -dimensional system noise vector
- \mathbf{v}_k is a l -dimensional measurement noise vector
- \mathbf{Y}_k is a l -dimensional vector of actual outputs
- Φ_{k-1} is the system dynamics matrix of dimension $n \times n$
- \mathbf{L}_{k-1} is the input transition matrix of dimension $n \times m$

- \mathbf{G}_{k-1} is the system noise transition matrix of dimension $n \times p$
- \mathbf{H}_k is the measurement matrix of dimension $l \times n$

The state or model equation of the state space model (2.71) describes how the internal system state at time instant k (e.g. soil water content) depends on the state at the previous time instant $k-1$ and the exogeneous inputs (e.g. rainfall). The system noise $\{\mathbf{w}_{k-1}\}$ takes account for all influences on the system state which are not caused by the exogeneous inputs. By means of the observation equation, the system state is linearly transformed into the system output (e.g. runoff), which is corrupted by the measurement error $\{\mathbf{v}_k\}$.

The state space representation as written above equals a vector system of difference equations. Any high order differential or difference equation, possibly with time variable coefficients, may be represented as a vector system of first order difference equations. Since the elements of the \mathbf{L}_k , ϕ_{k-1} , \mathbf{G}_{k-1} and \mathbf{H}_k matrices may vary in time, the state space representation is not limited to stationary or linear systems. In addition, the individual elements of \mathbf{U}_{k-1} and \mathbf{w}_{k-1} may influence several system states simultaneously. This can be expressed by allowing the elements off the main diagonal of \mathbf{L}_k and \mathbf{G}_{k-1} to be nonzero.

In the observation space, the purpose of the estimation procedures was to find appropriate parameters for the transfer function. In state space, these parameters are supposed to be known from the underlying physical system description or from a fairly general stochastic model. Now the estimation concerns the state vector in the presence of measurement and system noise, but with given model parameters. Of particular interest is the estimation of a state vector which minimizes the estimation error. By means of filtering, i.e. estimating the state vector at the current time based upon all past and present measurements, an unbiased and minimum variance estimate $\hat{\mathbf{x}}_k$ of the state vector can be found. A possible solution to this estimation problem is the Kalman filter (Kalman and Bucy, 1961), a linear recursive estimation which constructs an optimal estimate of the state vector given past and present values of the actual outputs vector and past values of the actual input vector. The complete two stage Kalman solution to the estimation problem takes the following form:

- Extrapolation

$$\hat{\mathbf{x}}_{k|k-1} = \Phi_{k-1} \hat{\mathbf{x}}_{k-1|k-1} + \mathbf{L}_{k-1} \mathbf{U}_{k-1} \quad (2.72)$$

$$\mathbf{P}_{k|k-1} = \Phi_{k-1} \mathbf{P}_{k-1|k-1} \Phi_{k-1}^T + \mathbf{Q}_{k-1} \quad (2.73)$$

- Updating

$$\hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k \left[\mathbf{Y}_k - \mathbf{H}_k \hat{\mathbf{x}}_{k|k-1} \right] \quad (2.74)$$

$$\mathbf{P}_{k|k} = \left[\mathbf{I} - \mathbf{K}_k \mathbf{H}_k \right] \mathbf{P}_{k|k-1}. \quad (2.75)$$

Details on the derivation of the Kalman filter are given in Appendix A.2. The optimality of the Kalman filter algorithm is expressed by the Kalman gain matrix \mathbf{K}_k . The Kalman gain matrix is proportional to the uncertainty in the estimate and inversely proportional to the measurement noise contained in \mathbf{R}_k . The influence of the system noise and the measurement noise on the state estimation can be summarized by normalizing all elements of $\mathbf{P}_{k|k}$ with the appropriate element of \mathbf{R}_k . In the following, this is shown for a system with a state vector of dimension $n = 1$ and with one input and one output variable only. The Kalman filter equations (2.72) to (2.75) can be altered to

- Extrapolation

$$\hat{\mathbf{x}}_{k|k-1} = \Phi_{k-1} \hat{\mathbf{x}}_{k-1|k-1} + \mathbf{L}_{k-1} u_{k-1} \quad (2.76)$$

$$\mathbf{P}_{k|k-1}^* = \Phi_{k-1} \mathbf{P}_{k-1|k-1}^* \Phi_{k-1}^T + \mathbf{NVR}_{k-1} \quad (2.77)$$

- Updating

$$\hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k \left[y_k - \mathbf{H}_k \hat{\mathbf{x}}_{k|k-1} \right] \quad (2.78)$$

$$\mathbf{P}_{k|k}^* = \left[\mathbf{I} - \mathbf{K}_k \mathbf{H}_k \right] \mathbf{P}_{k|k-1}^*, \quad (2.79)$$

with

$$\mathbf{NVR}_{k-1} = \frac{\mathbf{Q}_{k-1}}{\mathbf{R}_k} \text{ and} \quad (2.80)$$

$$\mathbf{P}_{k|k-1}^* = \frac{\mathbf{P}_{k|k-1}}{\mathbf{R}_k}. \quad (2.81)$$

The noise variance ratio \mathbf{NVR}_k , which has to be specified by the analyst, determines the filtering properties of the Kalman filter. If we assume that the noise processes \mathbf{v}_k and \mathbf{w}_k are stationary in the wide sense, the covariance matrices \mathbf{Q}_k and \mathbf{R}_k are constant for all k . The noise variance ratio matrix \mathbf{NVR}_k reduces to \mathbf{NVR} , a matrix with time invariant values. For large \mathbf{NVR} values, the filter will react more on changes in the data, rather than estimating a smooth time series of the state vector. The opposite, a smooth time series of the state vector, which clearly expresses the a priori known internal dynamics, can be obtained by choosing a small \mathbf{NVR} value. In most applications, the \mathbf{NVR} can be defined at a fixed value for all data points. However, if rapid changes in the state vector are expected, it is possible to give the \mathbf{NVR}_k matrix locally larger values. This technique, which is called 'variance intervention' reflects the analyst's interpretation of the system dynamics.

2.6.2 Smoothing

The Kalman filter has been used successfully for a wide range of applications. Ihringer (1985) modelled hydrological time series by means of the Kalman filter, and showed that the non-linear properties of runoff generation can be modelled by letting the matrices \mathbf{L}_{k-1} and ϕ_{k-1} vary in time. The Kalman filter has also been applied for the inverse estimation of parameters characterizing the hydrodynamic behaviour of aquifers (Ferraresi *et al.*, 1996). Especially the on-line forecasting of hydrological series is an intensively used application of the Kalman filter (Chao-Lin Chiu, 1978;). However, if a Kalman filter is applied for off-line modelling and interpretation of hydrological time series, it does not fully explore the information contained in the data. A filter estimates a noise free state vector at time instant k , based on all observations up to k . In off-line analysis, however, data at time instants $k+1$ to N are available

as well. To estimate the system state, $\hat{\mathbf{x}}_k$, based on all available measurements, a smoother can be used. Smoothing is a data processing scheme which uses all data between the time instants 1 and N to estimate the system state at $1 \leq k \leq N$. The smoothed estimate of \mathbf{x}_k is then denoted as $\hat{\mathbf{x}}_{k|N}$, which expresses that the estimate is conditioned on all available data. Three types of smoothing are of interest. In *fixed point* smoothing, the state vector at time instant $k = k'$ is estimated as the total number of observations N increases. In *fixed lag* smoothing, the state vector at time $k' = N - l$, which has a fixed lag l to the total number of observations N , is continuously updated. The former two smoothing procedures can be used on-line, i.e. for applications where the number of available data still increases. The Fixed Interval Smoothing (FIS), which will be discussed in this chapter, is for off-line procedures. It is used to estimate the state vector at all time instants k between 1 and N , exploiting all available data. FIS can be considered as a combination of a forward filtering procedure from $k = 1$ up to $k = N$ (e.g. by means of the Kalman filter), which is followed by a backward filter which revises the forward estimates using data from $k = N-1$ up to $k = 1$. The result of the smoothing process is an optimal estimate $\hat{\mathbf{x}}_{k|N}$ of the state vector at time instant k , based on all data up to $k = N$. The most straightforward derivation of a fixed interval smoother is obtained by combining the results of a forward and a backward filtering process. Let $\hat{\mathbf{x}}_{k|k}$ as before denote the result of the forward filtering and ${}_b\hat{\mathbf{x}}_{k|k}$ the outcome of the backward filter. Because of the Gaussian assumptions on the noise processes \mathbf{v}_k and \mathbf{w}_k , the estimation errors of the backward and the forward filter results are uncorrelated. Thus, an optimal smoother takes the form of a linear combination of $\hat{\mathbf{x}}_{k|k}$ and ${}_b\hat{\mathbf{x}}_{k|k}$ as

$$\hat{\mathbf{x}}_{k|N} = \mathbf{A} \hat{\mathbf{x}}_{k|k} + \mathbf{A}^* {}_b\hat{\mathbf{x}}_{k|k}, \quad (2.82)$$

with \mathbf{A} and \mathbf{A}^* time variable weighting matrices to be determined. As for the derivation of the Kalman filter, the smoother is optimal if it is unbiased and the smoothed estimates have a minimal error variance. By replacing the state estimates in Eq. (2.82) by the true values plus an estimation error, we obtain

$$\tilde{\mathbf{x}}_{k|N} = [\mathbf{A} + \mathbf{A}^* - \mathbf{I}] + \mathbf{A} \tilde{\mathbf{x}}_{k|k} + \mathbf{A}^* {}_b\tilde{\mathbf{x}}_{k|k}. \quad (2.83)$$

For unbiased filtering errors, the smoothing error $\tilde{\mathbf{x}}_{k|N}$ equals zero only if

$$\mathbf{A}^* = \mathbf{I} - \mathbf{A} . \quad (2.84)$$

The smoothing error of Eq. (2.83) can be reformulated as

$$\tilde{\mathbf{x}}_{k|N} = \mathbf{A}\tilde{\mathbf{x}}_{k|k} + (\mathbf{I} - \mathbf{A})_b \tilde{\mathbf{x}}_{k|k} , \quad (2.85)$$

and the covariance of the estimation error of the smoother is

$$\mathbf{P}_{k|N} = \text{Cov}[\tilde{\mathbf{x}}_{k|N}, \tilde{\mathbf{x}}_{k|N}] = \mathbb{E}[\tilde{\mathbf{x}}_{k|N}\tilde{\mathbf{x}}_{k|N}^T] \quad (2.86)$$

$$\mathbf{P}_{k|N} = \mathbf{A}\mathbf{P}_{k|k}\mathbf{A}^T + (\mathbf{I} - \mathbf{A})_b \mathbf{P}_{k|k} (\mathbf{I} - \mathbf{A})^T . \quad (2.87)$$

with ${}_b\mathbf{P}_{k|k}$ the covariance matrix of the estimation error of the optimal backward filter. Differentiating Eq. (2.87) with respect to \mathbf{A} and equating the result to zero results in

$$\mathbf{I} - \mathbf{A} = \mathbf{P}_{k|k} [\mathbf{P}_{k|k} + {}_b\mathbf{P}_{k|k}]^{-1} , \quad (2.88)$$

and

$$\mathbf{P}_{k|N} = \left[(\mathbf{P}_{k|k})^{-1} + ({}_b\mathbf{P}_{k|k})^{-1} \right]^{-1} . \quad (2.89)$$

From Eq. (2.89) it is obvious that the estimation error of the smoothed estimate is always smaller than or equal to the filter estimation error. It is now possible to write the state vector as

$$\hat{\mathbf{x}}_{k|N} = \mathbf{P}_{k|N} \left[(\mathbf{P}_{k|k})^{-1} \hat{\mathbf{x}}_{k|k} + ({}_b\mathbf{P}_{k|k})^{-1} {}_b\hat{\mathbf{x}}_{k|k} \right] . \quad (2.90)$$

Equations (2.90) and (2.89) can be used to calculate the smoothed estimate of the system state if the backward and the forward filtered system state and the corresponding error covariance matrices are available. Evidently, also the normalised error covariance matrices of Eqs. (2.76) to (2.81) can be used to estimate the smoothed state vector. Since the above smoother solution processes the time series three times (backward, forward and their com-

ination), it is computationally rather inefficient. Therefore, several alternatives of the basic smoother equations (2.90) and (2.89) have been developed. An example of an optimal smoother which does not involve backward Kalman filtering are the Rauch-Tung-Striebel (Rauch *et al.*, 1965) algorithms

$$\hat{\mathbf{x}}_{k|N} = \hat{\mathbf{x}}_{k|k} + \mathbf{A}_k \left[\hat{\mathbf{x}}_{k+1|N} - \hat{\mathbf{x}}_{k+1|k} \right] \quad (2.91)$$

$$\mathbf{P}_{k|N} = \mathbf{P}_{k|k} + \mathbf{A}_k \left[\mathbf{P}_{k+1|N} - \mathbf{P}_{k+1|k} \right] \mathbf{A}_k^T, \quad (2.92)$$

and

$$\mathbf{A}_k = \mathbf{P}_{k|k} \Phi_k^T \mathbf{P}_{k+1|k}^{-1}. \quad (2.93)$$

This backward recursion once more filters the initial forward Kalman filter state vector. The combination of a forward and backward filtering process assures that the estimate at time instant k is conditioned on all available data. As can be seen from the smoother equations, at least one complete filter solution is necessary to calculate the smoothed solution. Therefore, fixed interval smoothing can only be performed off-line, if all the observational data are available.

2.7 Data Based Modelling

2.7.1 Definition

Most of the modelling work in soil physics and hydrology relies on deterministic approaches. Deterministic modelling of natural systems often depends on prior scientific knowledge and may lead to far too complex and overparameterized models for natural dynamical systems. This conflicts with the nature of many natural systems, which are often governed by only a few dominant processes. Data Based Modelling (DBM), as proposed by Young and Beven (1994), approaches the modelling of natural systems from an alternative, statistically reliable perspective. The DBM approach is based on objective statistical inference of the model struc-

ture and parameters. Instead of postulating a model structure (e.g. as a set of differential equations), DBM uses time series analysis to characterize the dominating system properties from series of observational data without making assumptions on the nature of the processes. In its purest form, DBM avoids subjective analyst opinions from entering the model identification and parameter estimation process. However, Jakeman *et al.* (1994) have shown how a limited amount of a priori knowledge can substantially improve the DBM technique. Moreover, the modelling of hydrological time series involves a physical interpretation of the identified model structure and of the estimated parameters. As a consequence, only time series models which have a physically plausible meaning are of interest. This puts, however, restrictions on the model identification and parameter estimation procedure.

The number of a_i and b_j parameters and the value of the time delay define the model structure $[n (m+1) \delta]$ of a discrete transfer function (Eq. 2.60). It can be shown how certain model structures of discrete transfer functions are analogous to continuous partial differential equations. A first order linear differential equation, e.g., may be represented in discrete time by a $[1 1 \delta]$ time series model. Higher order differential equations can be rewritten as a first order vector differential equation. In terms of system dynamics, this is equivalent to treating higher order dynamics as a series of parallel first order dynamics. Identically, time series models of structure $[n (m+1) \delta]$, restricted to the condition that $n = m+1$, may be written as n parallel $[1 1 \delta]$ models. If $n > m+1$, two dynamical processes in series may occur and if $n < m+1$, a process without dynamics (i.e. only a gain factor, no autocorrelative parameters) results². In this work, the application of time series modelling will be mainly restricted to transfer functions which present one single or two parallel processes (i.e. $n = m+1$). Powerful methods of statistical inference to select the appropriate model struc-

². Instantaneous time series models, i.e. models without an autoregressive decrease of the output after cessation of the input may occur as a consequence of an insufficient resolution of the time series. If for instance, a drainage basin reacts very fast to precipitation (e.g. within minutes), averaged hourly runoff values may not capture these fast dynamics. Consequently, the analysis of the time series will lead to the formulation of instantaneous time series models.

ture are information criteria (see section 2.5.3). In the following of this thesis, we used the YIC information criterion (Eq. 2.67) to distinguish between [1 1 δ] and [2 2 δ] models.

The estimated parameters of the transfer function as well must have a physical significance. An excellent method for examining the physical relevance of a transfer function model is to calculate the impulse response. The impulse response of a dynamical model is the output which is generated after an impulse input has been applied. An impulse input can easily be generated by setting the input u_k to 1 at $k = 1$ and to zero at all $k \neq 1$.

2.7.2 Data Based Modelling of Nonlinear Time Series.

Data series of hydrological processes are typically generated by a nonlinear process. If the nonlinear phenomena underlying the experimental data were mathematically understood, the modelling of the time series would only involve a parameter estimation. However, since a profound physical understanding is rarely available, the modelling process mostly starts with the inspection of a set of nonlinear models. At this stage the analyst can decide to use a priori knowledge as e.g. Jakeman *et al.* (1994), or he can subject a series of nonlinear models to a battery of diagnostic checks (De Gooier and Kumar, 1992). However, in the most general case, a nonlinear process can be approximated by a linear model with time variable parameters. The estimation problem then changes from selecting an appropriate nonlinear model to the estimation of the time variable parameters of a linear model. In the following of this chapter a procedure of statistical linearization is presented. Afterwards, it will be shown how the time series of a time variable parameter can be employed to linearize the rainfall-runoff data.

2.7.3 Derivation of a Time Variable Parameter Model

Consider a time series of output values y_k generated by a nonlinear process

$$y_k = F(\chi_k) + e_k \quad (2.94)$$

$$\chi_k = \left[y_{k-1}, \dots, y_{k-n}, u_{k-\delta}, \dots, u_{k-m-\delta}, \mathbf{U}_{k-\delta^*}, \dots, \mathbf{U}_{k-q-\delta^*}, e_{k-1}, \dots, e_{k-p} \right], \quad (2.95)$$

with $\{e_k\}$ a white noise process, $u_{k-\delta}$ the input variable and F a general (nonlinear) well behaved function. The vectors $\mathbf{U}_{k-\delta^*}$ to $\mathbf{U}_{k-q-\delta^*}$ contain all the other available information on the system. The nonlinear function F can be approximated by a Taylor series about some fixed time instant k_0 and written as

$$\begin{aligned}
 y_k = & F(\mathcal{X}_{k_0}) + \sum_{i=1}^n \left. \frac{\partial F(\mathcal{X}_k)}{\partial y_{k-i}} \right|_{k=k_0} (y_{k-i} - y_{k_0-i}) + \\
 & \sum_{j=0}^m \left. \frac{\partial F(\mathcal{X}_k)}{\partial u_{k-j-\delta}} \right|_{k=k_0} (u_{k-j-\delta} - u_{k_0-j-\delta}) + \sum_{s=0}^q \left. \frac{\partial F(\mathcal{X}_k)}{\partial \mathbf{U}_{k-s-\delta^*}} \right|_{k=k_0} (\mathbf{U}_{k-s-\delta^*} - \mathbf{U}_{k_0-s-\delta^*}) + \\
 & \sum_{l=1}^p \left. \frac{\partial F(\mathcal{X}_k)}{\partial e_{k-l}} \right|_{k=k_0} (e_{k-l} - e_{k_0-l}) + e_k + C(\mathcal{X}_{k_0})
 \end{aligned} \tag{2.96}$$

To proceed, we neglect the higher order terms $C(\mathcal{X}_{k_0})$, as well as the partial derivative of $F(\mathcal{X}_k)$ with respect to $\mathbf{U}_{k-s-\delta^*}$. Certainly, the information contained in the additional system measurements describes the nonlinearity of the system, and should be a substantial part of the linearized process model. However, our attempt is to develop a linear input-output transfer function with time variable parameters. Equation (2.96) can be simplified to

$$y_k = T(\mathcal{X}_{k=k_0}) + \sum_{i=1}^n \left. \frac{\partial F(\mathcal{X}_k)}{\partial y_{k-i}} \right|_{k=k_0} y_{k-i} + \sum_{j=0}^m \left. \frac{\partial F(\mathcal{X}_k)}{\partial u_{k-j-\delta}} \right|_{k=k_0} u_{k-j-\delta} + \sum_{l=1}^p \left. \frac{\partial F(\mathcal{X}_k)}{\partial e_{k-l}} \right|_{k=k_0} e_{k-l}. \tag{2.97}$$

$T(\mathcal{X}_{k=k_0})$ contains all the terms that depend on $k = k_0$ and can be considered as a long term trend of the system. As hydrological time series most often show negligible trend components, it can be neglected. By renaming the partial derivatives as

$$\left. \frac{\partial F(\mathcal{X}_k)}{\partial y_{k-i}} \right|_{k=k_0} = -a_i(k_0) \tag{2.98}$$

$$\left. \frac{\partial F(\mathcal{X}_k)}{\partial u_{k-j-\delta}} \right|_{k=k_0} = b_j(k_0) \tag{2.99}$$

$$\left. \frac{\partial F(\chi_k)}{\partial e_{k-l}} \right|_{k=k_0} = c_l(k_0) \quad (2.100)$$

Equation (2.97) reduces to

$$y_k = \sum_{i=1}^n -a_i(k_0)y_{k-i} + \sum_{j=0}^m b_j(k_0)u_{k-j-\delta} + \sum_{l=1}^p c_l(k_0)e_{k-l}. \quad (2.101)$$

Clearly, the coefficients $a_i(k_0)$, $b_j(k_0)$ and $c_l(k_0)$ depend on the time k_0 and thus Eq. (2.101) may be written in a more general form as

$$y_k = \mathbf{z}_k^T \mathbf{a}(k) + e_k, \quad (2.102)$$

with

$$\mathbf{z}_k^T = \left[y_{k-1}, \dots, y_{k-n}, u_{k-\delta}, \dots, u_{k-m-\delta} \right] \quad (2.103)$$

and

$$\mathbf{a}^T(k) = \left[a_1(k), \dots, a_n(k), b_0(k), \dots, b_m(k), c_1(k), \dots, c_p(k) \right]. \quad (2.104)$$

The deterministic part of the above transfer function model strongly resembles the deterministic part of the linear transfer function equation (2.26), however, now with time variable parameters. Thus, by this procedure of statistical linearization, the estimation problem has been reduced to estimating the temporal change of the parameters of the transfer function model.

2.7.4 Estimation of Linear Transfer Functions with Time Variable Parameters

Equations (2.102) to (2.104) define a transfer function with time variable parameters. To estimate the *parameters* of this transfer function, Young (1993) proposed to apply a fixed interval smoothing technique (Eq. 2.90). But, this fixed interval smoother estimates the temporal variation of a *state vector*. In addition, unlike in the state equation (2.70) of a state space model,

no information on the temporal variation of the parameter vector of transfer function (2.102) is available.

Since there is no a priori information on the temporal variation of the parameters, it is useful to adopt an explicit but simple model which describes the parameter variation. Examples are random walk and integrated random walk models from the family of Gauss-Markov sequences (see Appendix A.3). For the temporal variation of the parameter vector of Eq. (2.102), we can assume a random walk model of the following form

$$\mathbf{a}_k = \mathbf{a}_{k-1} + \mathbf{w}_{k-1}. \quad (2.105)$$

Equation (2.102) can now be recognized as the observation equation and Eq. (2.105) as the state or model equation of a state space model. Actually, there is little difference between the state or model equation of a state space model (Eq. 2.70) and a stochastic difference equation (Eq. 2.105) which describes the temporal parameter variability. If we compare the observation equation of the state space setting Eq. (2.71) with the observation equation of the time variable parameter model Eq. (2.102), the 'observation matrices' \mathbf{H}_k and \mathbf{z}_k^T differ. In state estimation, the matrix \mathbf{H}_k is considered to be deterministic and known, whereas in parameter estimation the \mathbf{z}_k^T vector is usually stochastic. As has been shown before, it contains noisy previous output observations and present and previous input samples. Fortunately this has no further influence on the estimation procedure. Explicit modelling of parameter variation is therefore a sort of state estimation problem, with a parameter vector as the state. The equivalence of parameters and states opens the door to a wide range of state estimation techniques, such as Kalman filtering and Fixed Interval Smoothing. To estimate the time variable parameters $\mathbf{a}(k)$ of Eq. (2.102) the Kalman filter equations can be rewritten as:

- Extrapolation

$$\hat{\mathbf{a}}_{k|k-1} = \hat{\mathbf{a}}_{k-1|k-1} + \mathbf{L}_{k-1} u_{k-1} \quad (2.106)$$

$$\mathbf{P}_{k|k-1}^* = \mathbf{P}_{k-1|k-1}^* + \mathbf{NVR}_{k-1} \quad (2.107)$$

- Updating

$$\hat{\mathbf{a}}_{k|k} = \hat{\mathbf{a}}_{k|k-1} + \mathbf{K}_k [\mathbf{y}_k - \mathbf{z}_k^T \hat{\mathbf{a}}_{k|k-1}] \quad (2.108)$$

$$\mathbf{P}_{k|k}^* = (\mathbf{I} - \mathbf{K}_k \mathbf{z}_k^T) \mathbf{P}_{k|k-1}^* \quad (2.109)$$

$$\mathbf{K}_k = \mathbf{P}_{k|k}^* \mathbf{z}_k^T \mathbf{R}_k^{-1}. \quad (2.110)$$

Thus, in Eqs. (2.106) and (2.108), $\hat{\mathbf{a}}_{k|k}$ and $\hat{\mathbf{a}}_{k|k-1}$ are no longer system states, but the time variable parameter vector of Eq. (2.104). The Kalman filter can be applied to obtain estimations of the time variable parameters for both the backward and the forward filtering problem. The application of the smoother equations (2.89) and (2.90) is then straightforward. Evidently, also alternative forms of the smoother can be applied to estimate the smoothed parameter estimates $\hat{\mathbf{a}}_{k|N}$. An example of the filtering and smoothing procedures applied on rainfall and runoff data of the Erlenbach headwater catchment is given in Fig. (2.3). Rainfall and runoff measured at 10 min intervals are used to estimate the time variable gain parameter of a [1 1 1] linear transfer function. For this purpose, Eqs. (2.106) to (2.110) are employed to obtain the filtered parameter estimation, followed by Eqs. (2.91) to (2.93) for the smoothed estimate. The **NVR** was set to 0.001, and the initial values of the parameter vector and the matrix of the covariances of the estimation error were obtained from estimating a linear time invariant [1 1 1] transfer function. The time series of the time variable parameter, together with additional measurements on the hydrological state of the system, can be applied to linearize the hydrological data series. The complete linearisation procedure, which will be applied later in this work, can be summarized in the following steps (see also Fig.2.4):

1. Estimate a linear transfer function of form (2.23) from the rainfall and the runoff data. To avoid a biased parameter estimation caused by noisy data, an instrumental variable method (section 2.4.2) may be applied. In this work, the applied instrumental variable techniques most often converged after 3 to 5 iterations and the model identification is performed according to the Young Information Criterion (Eq. 2.67).

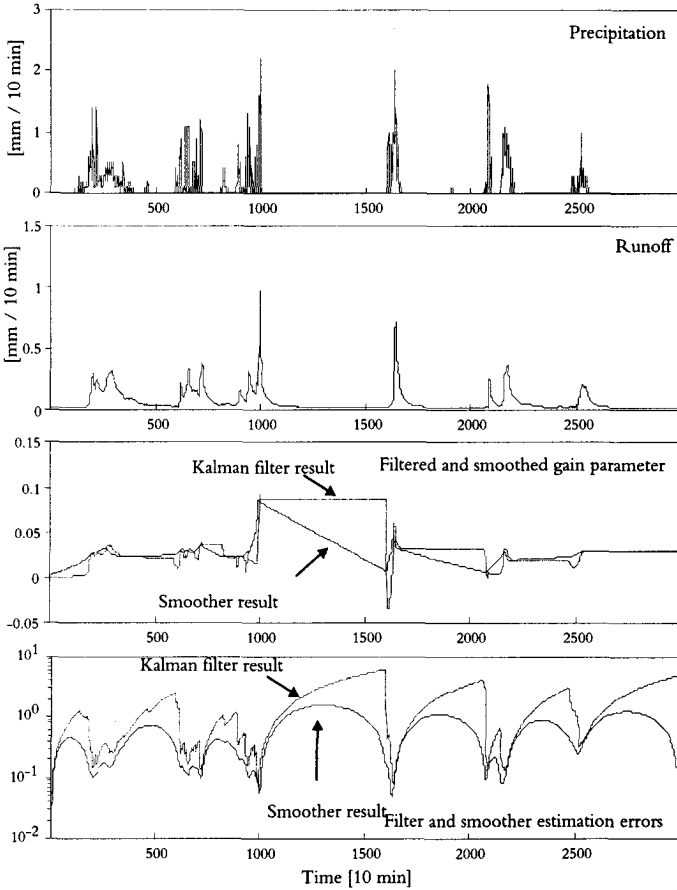


Figure 2.3: Example of the filtering and smoothing procedures. Precipitation and runoff data for the Erlenbach catchment are used to estimate the time variable gain parameter $b_0(k)$ of a $[1 \ 1 \ \delta]$ transfer function. Kalman filter ($\hat{\mathbf{a}}_k$) and smoother ($\hat{\mathbf{a}}_{k|N}$) results are shown together with the covariance of their estimation errors $\mathbf{P}_{k|k}^*$ and $\mathbf{P}_{k|N}^*$.

2. The model residuals, $\hat{\eta}_k$, may be tested for nonlinear and/or nonstationarity effects. The autocorrelation function of the model error, cross-correlogram of $\hat{\eta}_k$ with u_k and y_k , and tests for stationarity can be used. If the residuals do not show any signs of non-stationarity and/or nonlinearity, this estimate is the final solution to the estimation problem.
3. Most of the time, however, a linear transfer function will not describe the rainfall-runoff data well. Studying the estimation errors of the parameter estimates reveals which model parameters are likely to vary in time. For hydrological time series, most often the parameters associated with the exogeneous inputs (i.e. the gain parameter) will show a time variable behaviour. By means of the fixed interval algorithms Eq. (2.90) based on the Kalman forward and backward filter solution or by means of Eqs. (2.91) to (2.93) the temporal variation of the parameters of transfer function (2.102) may be investigated. In most cases, a fairly simple $[1 \ 1 \ \delta]$ time series model with a time variable gain parameter $b_0(k)$ and a fixed autoregressive parameter a_1 is estimated. The time delay δ , the value of a_1 and the initial value for $b_0(k)$ are obtained from the linear transfer function estimation. Identically, the estimation error of the parameter estimates of the linear case may be applied to construct the initial value of the covariance matrix of the estimation error $\mathbf{P}_{k|k-1}^*$ at $k = 1$. This matrix, as the initial value of the parameter vector, is necessary to perform the extrapolation step at the initialization of the Kalman filter algorithm (Eq. 2.72).
4. Examine the nature of the time variable parameter(s) in relation to the additional system information contained in \mathbf{U}_{k,δ^*} (Eq. 2.95). The main interest goes to those additional variables which show a good temporal correlation with the time variable parameters of the transfer function. If a good correlation can be found, a nonlinear equation can be fitted, based on a weighted least squares approach. The weighting follows according to the scaled elements of the covariances of the estimation error in the $\mathbf{P}_{k|N}$ matrix. Thus, in general terms, the nonlinear behaviour of the runoff generation can be modelled as

$$\hat{\mathbf{a}}_{k|N} = \mathfrak{F}(\mathbf{U}_{k,\delta^*}) \quad (2.111)$$

5. This nonlinear equation can be used to transform the measured inputs into effective inputs. By employing the nonlinear equation (2.111), which is most often as power law (see also Section 4.4.3), the effective precipitation is computed as

$$u_{eff,k-\delta} = u_{k-\delta} \mathfrak{S}(U_{k-\delta}^a) \quad (2.112)$$

6. Again, a linear transfer function can be estimated by means of an instrumental variable algorithm. Instead of the measured precipitation, now the effective precipitation is used as the input variable. This procedure can be continued until the residuals of the modelling no longer suggest that the time series are nonlinear.

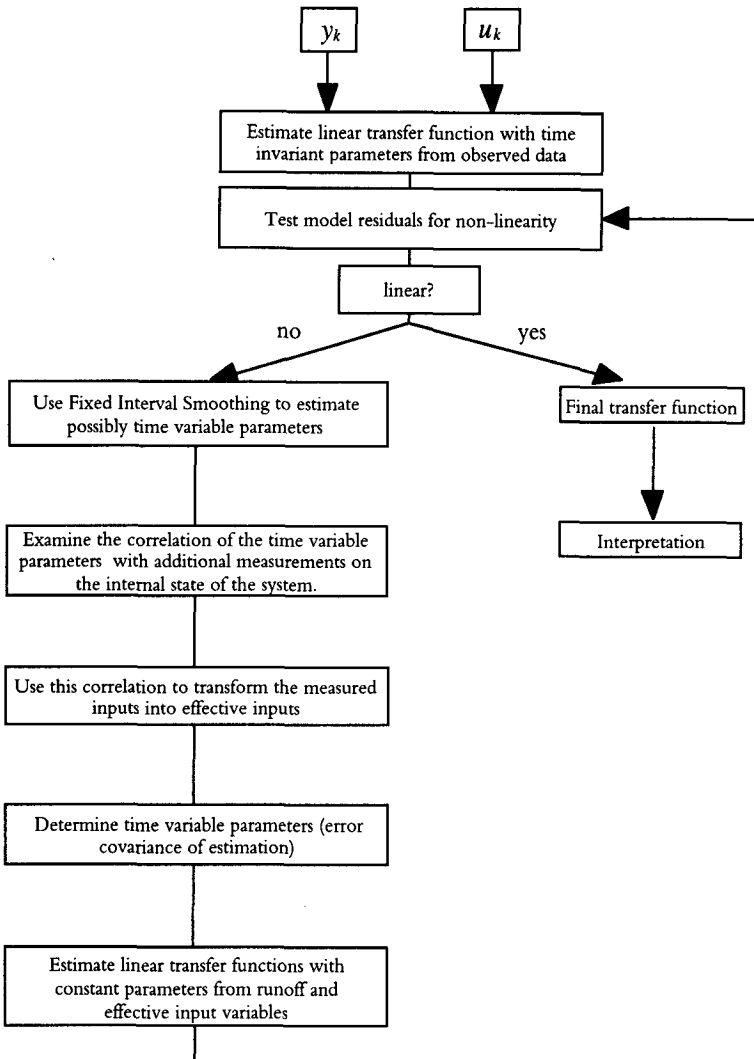


Figure 2.4: Overview of the nonlinear transfer function estimation procedure.

Chapter 3

A Tracer Experiment to Detect Flow Paths in Forest Soils¹



¹Paper by H. Feyen, H. Wunderli, H. Wydler, J. Leuenberger and A. Papritz. 1998. A tracer experiment to detect flow paths in forest soils. *J. Hydrol.*, in prep.

3.1 Introduction

Man-made increase of atmospheric nitrogen deposition may destabilize extensively managed forest ecosystems. As a result of long-term addition of nitrogen beyond the natural demand of plants and micro-organisms, nitrogen species leak from the ecosystem and threaten the quality of surface and groundwater resources (Dise and Wright, 1995).

Wright and Tietema (1995) and Gundersen and Rasmussen (1995) recognized that the leaching of nitrate from nitrogen limited forest ecosystems may be induced by hydrological processes. For catchments treated with ammonium nitrate, an elevated concentration of nitrate in the runoff was observed, whereas ammonium was completely retained. The authors attributed the nitrate losses in the runoff to periods with high flow of water which occurred mainly during the winter and especially during snowmelt periods.

Numerical simulations by Creed *et al.* (1996) supported the so-called flushing hypothesis. Their simulations showed that after a period of low demand by the forest ecosystem, nitrogen was enriched in the upper soil layer and it was flushed from there by intensive rainfall or spring snowmelt. Fast flow processes may cause losses of nitrogen from natural ecosystems, too. The residence times of rainwater and nitrogen species in the vadose zone may be too short to allow plants or micro-organisms to take up readily available nitrogen. Short transit times of water and solutes in a natural catchment have been reported by Nyberg (1995). Thus, the export of nitrogen from natural ecosystems may at least partly depend on the routing of water through the soil.

There is plenty of evidence for the existence of fast flow paths such as macropores or fractures in agricultural soils. It has been shown that a small amount of macropores may dominate the transport of water and dissolved chemicals (e.g. Flury *et al.* 1994, Flury *et al.* 1995).

Runoff generation in forested catchments has been the subject of many hydrological studies (e.g. see Bonell, 1993 for an overview). Since most undisturbed forest soils have a highly permeable surface layer, most of the research concentrates on the subsurface flow. The extent and the residence time of the subsurface flow has a significant influence on the interaction between the soil matrix and the soil water. The nature of subsurface flow depends on the soil

structure and on hydrological controls such as soil moisture content and precipitation intensity. A possible flow mechanism is the displacement of pre-event soil matrix water by the rainfall (Wilson *et al.* 1990). The presence of preferential flow paths in the subsoil which bypass the soil matrix, may drastically change the flow regime (Beven and German, 1982; Turton *et al.*, 1995; Leaney *et al.*, 1993; Mullholland *et al.*, 1990, Hornberger *et al.*, 1991). However, although present in most forest soils, macropores not always play an active role in the subsurface flow. It has been shown by Luxmoore *et al.* (1990) and Tsuboyama *et al.* (1994) that the flow through macropores depends on the rainfall intensity. Additionally, Sidle *et al.* (1995) and Wilson *et al.* (1990) showed that macropore flow was more likely to occur under wet antecedent moisture conditions. If flow through macropores occurs, it can be argued that because of the short residence times of the rainwater in the soil, the runoff will have the chemical signature of the precipitation. This hypothesis has been confirmed by Turton *et al.* (1995), who detected a negative correlation between the rainfall intensity and the amount of pre-event water in the runoff of a small forested soil plot. It has been shown McDonnell (1990) and Bazemore *et al.* (1994), however, that rainfall water can also mobilize pre-event water in the soil and force it into the macropores. As a result, fast runoff will occur with the chemical composition of the soil matrix water. Fast flow processes in the subsoil can also occur if the content of mobile soil water is limited (Lange *et al.*, 1996) or as lateral transport along impermeable bedrock surfaces (Renzetti *et al.*, 1992). Finally, return flow and the resulting saturation overland flow can also be the cause of fast runoff generation (Pearce *et al.*, 1990; Peters *et al.*, 1995).

Thus, to understand the cycling of nitrogen in a biogeochemical environment a sound knowledge of the hydrologic control of the transport of solutes in the subsurface is essential. The experiment we describe in this article was set up to provide this information for an experimental study in which nitrogen was added to a prealpine coniferous forest ecosystem (Schleppi *et al.*, 1998). The study was part of the so-called nitrogen saturation experiments (NITREX, Wright and van Breemen, 1995) in which, at a series of sites, the nitrogen deposition was altered. At the Swiss NITREX site in the Alptal valley, ^{15}N labelled nitrogen was applied to a small, artificially isolated forested catchment (1500m²), and its transfer into the

soil and vegetation and the losses by runoff and denitrification were investigated. Although the concentration of nitrogen species in the runoff seemed to favour the nitrogen flushing hypothesis, the actual hydrologic dynamics were unknown. Visual observations of the artificial catchments during rainstorms led to the hypothesis that the generation of runoff was strongly influenced by the micro-topography. Within the artificial catchment, there is a close correspondence between the micro-topography and the distribution of the two main soil types. A wet umbric Gleysol with a muck humus¹ topsoil is common in the depressions of the catchment, whereas a drier umbric Gleysol with a mor humus topsoil dominates the hills. Our observations suggested that the mor humus soil had a limited hydrological activity and that runoff was mainly generated in the wet dips covered by the muck humus. To put this hypothesis to a test, it was necessary to study the runoff generation in well defined artificial catchments. Because of the high spatial and temporal variability of soil chemical and hydrological processes, even in a small catchment of 1500 m², we decided to concentrate on two soil plots of about 13 m², one representative for the muck and the other for the mor humus soil. To gain insight into the dominating flow processes and to estimate the residence times of water and solutes in the different compartments of the soil, we performed a series of experiments where we applied various tracers on the soil surface and in the subsoil. This article describes the experiments, reports the results we obtained thereof and draws some tentative conclusions about their relevance for the Swiss NITREX project.

¹“muck” and “mor” humus are descriptions of the humic topsoil. Muck humus represents an A_a horizon (mineral soil mixed with organic matter decomposed under hydromorphical conditions) which mainly developed in anaerobic, waterlogged soils. The topsoil we denoted with “mor humus” is much better drained and characterized by a moderately thick litter layer on top of a humic A_h horizon. This humus form has a low pH.

3.2 Material and Methods

3.2.1 Experimental Site

The study site 'Alptal' (47°03N, 8°43E) is situated at the upper end of a blind valley in the Swiss prealps at an altitude of 1200 m a.s.l. The prevailing Northwest winds force wet air into the valley, causing high intensity rainshowers. The annual precipitation is 2300 mm y⁻¹, 30 percent is falling as snow between December and May. The average annual air temperature is 6 °C, with a minimum of -2 °C in February and a maximum of 17 °C in August. The experimental plots are situated on a moderately steep slope (20 percent), with a Southwest orientation. The bedrock at the site is Flysch: a calciferous sediment of the tertiary, composed of schisted, marled clays interchanged with stony or sandy layers.

The soils in the area are generally classified as umbric Gleysols. Closely linked to the microtopography, two different organic topsoils developed (Schleppi *et al.*, 1998). On small ridges, an acid mor humus topsoil formed. The organic surface layers L, F, and H, followed by the humic A_h horizon cover moderately wet and well structured gleyic G_o (fully oxidized) and G_{or} (partly reduced) horizons (Table 3.1). On the mor humus there are two vegetation layers: the understory vegetation consists mainly of *Vaccinium myrtillis* and *Equisetum sp.*, while the tree layer is dominated by Norway spruce (*Picea Abies*, (L.) Karst). On the mor humus plot we selected for this experiment, the understory vegetation was scarce.

Table 3.1: Properties of the mor humus soil.

	Depth ‡ [cm]	Porosity † [m ³ m ⁻³]	particle size distribution †			pH [#]	O.C. [#] [g kg ⁻¹]
			Sand	Silt	Clay		
LFH	0-6	-	-	-	-	3.3	345
A _h	6-30	0.75	47	47	6	3.6	71
G _o						4.2	22
G _{or}	>30	0.74	49	46	5	5.1	25

‡ Depths at the experimental soil plot (profile from 0-60 cm)

† From from Diserens (1992), clay < 2µm, silt 2-63 µm, sand 63 µm-2 mm; percentage by weight

Arithmetic average of 72 samples taken on a regular 4 x 4 m grid on a neighbouring site, measured in CaCl₂ 0.01M

The much wetter muck humus topsoil developed in depressions between the ridges and mounds. The soil profile consists of a thin litter layer L, a well-developed muck humus horizon A₂, a partly reduced gleyic horizon G_{to} and a fully reduced gleyic horizon G_r (Table 3.2). Because of the high soil moisture content throughout the year, there is no tree layer. The understory vegetation is dominated by *Caltha palustris* and grasses. The matrix of the subsoil horizons G_{to} and G_r is barely permeable. However, because of the high number of stones, old tree roots and dead wood in the anoxic soil layers, a partly continuous network of water conducting pores exists.

Table 3.2: Properties of the muck humus soil.

	Depth ‡ [cm]	Poros. † [m ³ m ⁻³]	K _{sat} § [msec ⁻¹]	particle size distribution †		pH#	O.C.# [gkg ⁻¹]	
				sand	silt clay			
L	0-3	-	-	-	-	-	-	
A ₂	3-25	0.90	7.2 10 ⁻⁷	51	45	4	5.4	134
G _r	25-40	0.76	0.2 10 ⁻⁷	44	44	12	5.1	80
o								
G _r	>40	0.71	0.7 10 ⁻⁷	43	42	15	5.9	32

‡ Depth at the experimental soil plot (profile from 0-60 cm)

† From Diserens (1992), clay < 2 μm, silt 2-63 μm, sand 63 μm-2 mm; percentage by weight

§ From Stadler *et al.* (1996)

Arithmetic average of 72 samples taken on a regular 4 x 4 m grid on a neighbouring site, measured in CaCl₂ 0.01M

3.2.2 Soil Plots

For each of the humus types, we selected an experimental plot of about 13 m² with similar plot slope and aspect. To prevent lateral inflow of water, the soil plots were separated from the surrounding soil by means of 80 cm deep trenches. In the downslope open face of the soil plots runoff gutters were inserted 25 cm into the soil profile in three different depths (Fig. 3.1). The gutters were filled with two sand layers of distinct particle sizes to maintain a minimal suction at the profile face and to prevent an early deflection of the flow lines of water. Near the profile face we used a particle size of 0.3-0.9 mm, near the gutter the particle

size was 3–5 mm. The gutters were 2 m wide, their height was 5 cm (surface runoff), 25 cm (interflow) and 30 cm (depth flow). The depths in which the runoff gutters were installed matched approximately the adjacent soil horizons (Fig. 3.1). The gutter outflow was recorded with tipping buckets having a resolution of about 100 ml. The tipping pulses were recorded on a data-logger (Campbell Scientific, CR10). To convert the counts per unit of time to flow rate, each tipping bucket was calibrated separately in the laboratory.

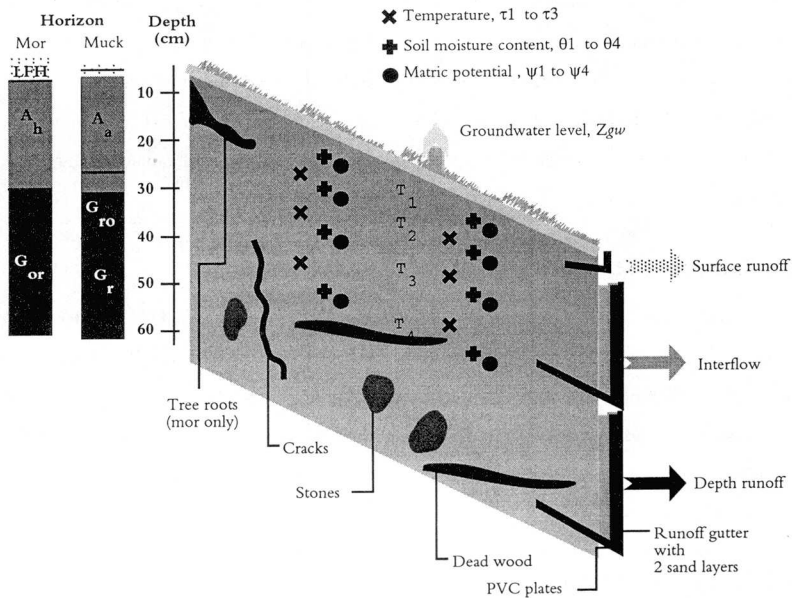


Figure 3.1: Side-view of the soil plots. Temperature (τ_1 to τ_3), soil moisture content (θ_1 to θ_4) and matric potential (ψ_1 to ψ_4) are measured in four depths, T_1 (5 cm), T_2 (10 cm), T_3 (20 cm) and T_4 (40 cm). Three runoff gutters collect the outflow from surface and near surface layers (0–5 cm), from the interflow (5–30 cm) and from the depth runoff (30–60 cm). There is one piezometer per plot (Z_{gw}).

Two small pits (50 x 10 x 60 cm) were excavated within the soil plots to install additional equipment (Fig. 3.1). In total, 16 TDR probes with 2 rods (25 cm long) to record the soil moisture content (θ_1 to θ_4), 16 tensiometers equipped with pressure-transducers to measure the matric potential (ψ_1 to ψ_4) and 9 thermistors (τ_1 to τ_3) were installed in four depths in each plot. The symbols θ_1 to θ_4 and ψ_1 to ψ_4 thus denote the averages of the measurements in T_1 (5 cm), T_2 (10 cm), T_3 (20 cm) and T_4 (40 cm). After the installation, the pits were carefully refilled with the original soil to reduce disturbances of the flow field. In addition, a piezometer was installed in each plot. All instruments were automatically operated, and the observations were recorded using the same data-logging system. The outflow volumes and groundwater levels (Z_{gw}) were recorded every minute, all the other measurements were made at 15 min intervals. Both plots were covered with a plastic roof. The roof served to shield the plots from rain and to minimize the evapotranspiration by reducing net radiation and wind speed near the plot surface.

3.2.3 Sampling and Analyses of the Runoff

Electrical conductivity was continuously monitored (EC Probes LF 323-B, WTW) to adjust the temporal resolution of the runoff sampling. During the early breakthrough of the tracers, water samples (~10 ml) were taken every 2.5 litres of outflow; after passing of the tracer peak the resolution was reduced to one sample per 10 litres of runoff. All water samples were filtered through membrane filters (0.45 μm) and analysed using a Dionex ion chromatograph with a conductivity detector.

3.2.4 Sprinkling Device

Both soil plots were irrigated with a mobile sprinkling device. The device consisted of a spray bar moving on two parallel rails, driven by an electric motor. Irrigation was intermittent: the spray bar was in motion for 20 s an stopping 80 s till the next cycle began. The spray bar had 16 nozzles (TEEJET 110015LP) mounted at an average height of 30 to 40 cm above the ground. The distance between adjacent nozzles was 30 cm and their irrigation angle was set to achieve a 50 percent overlap between two neighbouring nozzles. Tap water was pumped from a 5 m³ freshwater tank installed uphill. The resulting pressure at the nozzle inlet was

200 kPa. The coefficient of temporal variation of the water flow, measured by a water meter at the inlet of the spray bar, was determined as 0.04 for a measuring period of 10 minutes. The uniformity of the irrigation on the soil surface was measured by square cups (10.6 cm x 10.6 cm). The coefficient of spatial variation was 0.04 for the mor humus plot and 0.11 for the muck humus plot. When all nozzles spray at the same rate and if the ground surface is parallel to the spray bar, a uniform irrigation rate can be achieved. The discrepancy between the plots is very likely due to the roughness of the soil surface. On the muck humus plot the micro-topography was much more pronounced due to the ground vegetation and this resulted in the larger spatial variation.

3.2.5 Tracer Application

The tracer experiments lasted from August 20 to 25 for the mor humus soil plot and from September 8 to 13 for the muck humus soil plot (Fig. 3.2). During these periods, we irrigated both plots at a steady rate of $1.75 \cdot 10^{-3} \text{ m}^3 \text{ m}^{-2} \text{ s}^{-1}$ to maintain steady state flow conditions. To faster attain a condition of steady state, we prewetted both plots. The steady irrigation rate remained constant with only minor interruptions during the whole period.

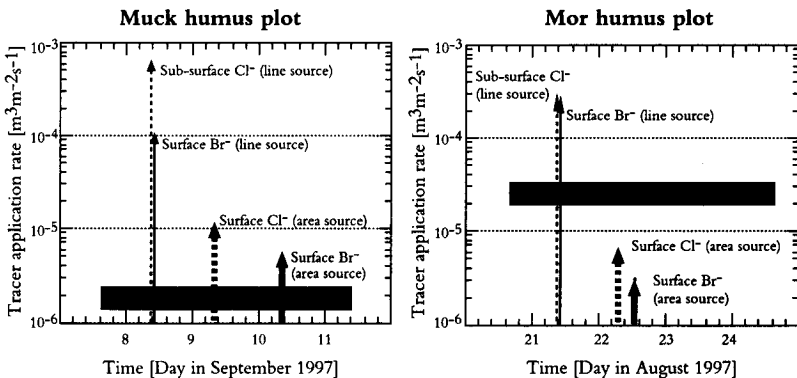


Figure 3.2: Timing of the tracer application on the soil plots. The arrows show the timing and the rate of the tracer application, the grey area indicates the period of steady state irrigation.

To distinguish the processes in the humic topsoil and in the gleyic subsoil, the tracers were applied on the surface (line source of Br^-) and as a subsurface line source (Cl^-). The two line sources were adjacent, at an average distance of about 3.3 m from the runoff gutters. On both lines, the tracers were applied at a high rate (see Table 3.3). To obtain additional information on the flow processes at a lower tracer application rate, the salts were also applied at a smaller rate on two larger areas situated in the lower half of the plot (area sources, see Table 3.3, Fig. 3.3).

For the tracer application onto the soil surface, a small pressure container with a nozzle similar to that of the sprinkler was used. The container was manually moved over the area marked for application and care was taken to keep the distance between the nozzle and the soil surface constant (30 cm). To avoid an interruption of the irrigation regime, the tracers were applied on the surface as fast as possible.

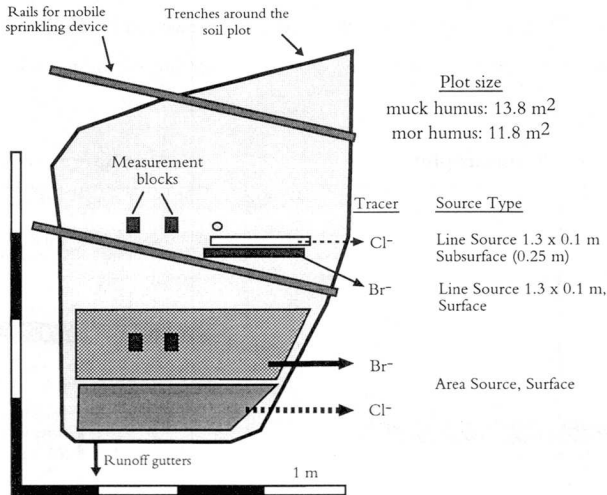


Figure 3.3: Tracer application areas on the soil plots (to scale). The marked areas on the soil plot identify the spatial extent of the tracer application areas. The arrows correspond to those of Fig. 3.2.

For the subsurface application of the tracer at the interface between the humic A horizon and the gleyic G horizon, we developed an alternative technique. To avoid ponding on the horizon interface, it was necessary to apply a small volume of tracer solution. In addition, it was important that the time interval for dispensing the tracer was as short as possible, in order to approximate a pulse input. To meet these requirements, we first excavated a small trench of 10 cm x 130 cm surface area down to the depth of the horizon boundary (~25 cm). Then, 100 hollow steel needles were vertically placed at equal spacing in the trench (Fig. 3.4). The needles, were 25 cm long with a diameter of 3 mm, had a small lateral opening at a distance of 1 cm from the tip.

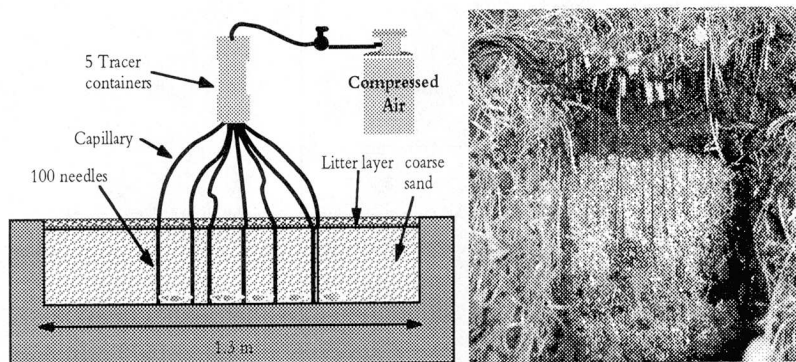


Figure 3.4: *Tracer application in the sub-surface. The drawing on the left is an overview of the equipment used for the tracer application. The picture on the right shows the spreading of the injected tracer in the sub-surface during a pre-experiment. The tracer is visualized by the dye Brilliant Blue FCF.*

The trench was filled with coarse sand (particle size 3–5 mm) and the litter layers on the soil surface were put back in place. Prior to the tracer application, the same irrigation regime was applied to the application area as to the rest of the soil plots. All needles were connected with a flexible capillary tapping one of the five tracer containers. Immediately before the tracer application, the containers were filled with the solution and connected to a bottle with compressed air (50 kPa). The application was stopped at the first sound of air escaping from

the needles. Then, the capillaries were disconnected from the containers and the remaining volume of the tracer solution was measured

Table 3.3: Summary of the tracer application on both plots.

	Tracer	Mass [g]	Solution [10^{-3} m^3]	Application Area [m^2]	Application duration [s]	Application rate [$\text{m}^3 \text{ m}^{-2} \text{ s}^{-1}$]
Mor Humus	Cl^- †	12.7	1.00	0.13	25	$3.1 \cdot 10^{-4}$
	Br^- §	7.8	1.77	0.13	46	$3.0 \cdot 10^{-4}$
	Cl^- ‡	6.3	1.95	2.03	144	$0.067 \cdot 10^{-4}$
	Br^- ‡	18.7	1.95	3.00	215	$0.03 \cdot 10^{-4}$
Muck Humus	Cl^- †	38.7	1.52	0.13	19	$6.2 \cdot 10^{-4}$
	Br^- §	29.5	0.74	0.13	53	$1.1 \cdot 10^{-4}$
	Cl^- ‡	16.9	2.00	2.07	81	$0.12 \cdot 10^{-4}$
	Br^- ‡	19.7	1.99	2.73	129	$0.056 \cdot 10^{-4}$

† Subsurface application, § Surface application as a line source, ‡ Surface Application Area Source

3.3 Results

3.3.1 Results of the Steady State Irrigation

The irrigation of the soil plots at a constant rate of $2.3 \cdot 10^{-6} \text{ m}^3 \text{ m}^{-2} \text{ s}^{-1}$ induced hydrological conditions that were close to steady state. Figure 3.5 shows that the total runoff, but not its components, was almost constant on the muck humus plot. The outflow from the deepest soil layer increased slightly at the expense of interflow and surface runoff. Surface runoff contributed only minimally to the total flow. Since the soil was already very wet at the beginning of the steady irrigation, only minor changes in the volumetric water content of $\Delta\theta = +0.02$ for θ_1 and of $\Delta\theta = +0.01$ for θ_2 were observed. In the subsoil, no changes could be detected. The soil water content was 0.77 (θ_3) and 0.64 (θ_4). Similarly, apart from small fluctuations, the groundwater table Z_{gw} remained at a constant level of 8 cm below the soil surface. The

readings of the tensiometers in 5, 10 and 20 cm indicated that, in the average, soil water was under hydrostatic pressure. Direct solar radiation on the plastic roof in the early afternoon warmed the tensiometers and lead to a sharp drop in the matric potential measurements. The periodic fluctuations are therefore artefacts induced by rapidly changing temperature. During the steady irrigation on the muck humus plot, on average 34.5 percent of the irrigated water could be collected as runoff, 26 percent in the depth runoff and 8.5 percent in the interflow.

On the mor humus soil plot, initially only the depth runoff reacted to the steady irrigation (see Fig. 3.6). After 1.5 days, however, the interflow set on but its contribution to the total runoff remained small. During the further course of the experiment, the depth flow increased slightly although the level of the groundwater table Z_{gw} remained constant at 56 cm. Somehow contradictory is the steady increase of the soil water content in the upper humic layers (θ_1). The soil water content increased from 40 to 46 percent (θ_1) cm and from 45 to 50 percent (θ_2). In the subsoil (θ_3 and θ_4), the soil water content remained constant at 58 percent. The matric potentials ψ_1 , ψ_2 and ψ_3 were generally constant, but decreased briefly after the irrigation stopped. Radiation and changing temperature had a less pronounced influence on the pressure transducer readings because the tree crowns above the plastic roof shaded the plot. Although the mor humus plot was much drier than the muck humus plot, we collected 45 percent of the applied water, mainly in the depth runoff.

The runoff data suggest that on both plots a steady hydrological state could be attained. This is supported by the fairly constant levels of the groundwater table. The small fluctuations in the level of the groundwater, especially for the muck humus plot, are highly likely due to the intermittent character of the irrigation regime. However, there are some striking characteristics that need to be discussed.

Despite the reduced and clayey gley horizon in the subsoil, the water balance was not closed for either of the two plots. Since the evapotranspiration was minimized by the plastic roof and since we did not observe any water flowing into the lateral trenches, seepage into deeper soil layers is the most likely explanation for the water losses. Despite the trenches around the soil plot, the groundwater in the plot was hydraulically linked with the direct surroundings

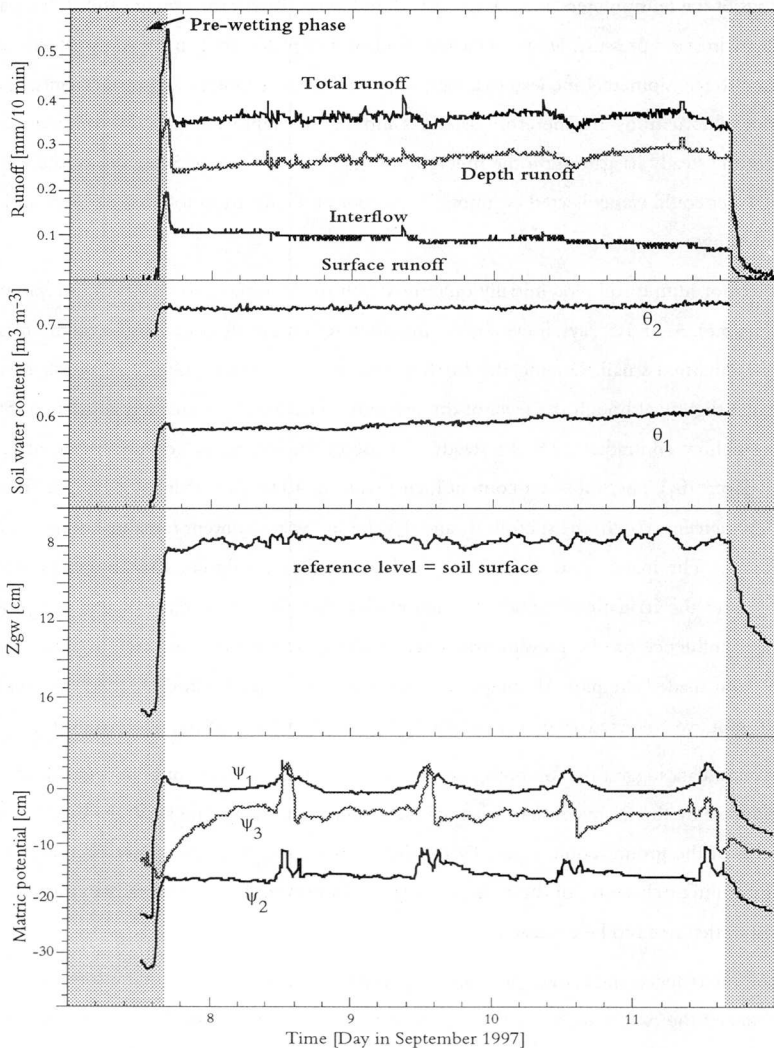


Figure 3.5: Results of the steady state irrigation ($2.3 \cdot 10^{-6} \text{ m}^3 \text{ m}^{-2} \text{ s}^{-1}$) on the muck humus plot. Partitioning of the total runoff, depth of the groundwater layer Z_{gw} , volumetric soil water content in the upper two layers (θ_1 and θ_2) and the matric potential in the upper three layers (ψ_1 , ψ_2 and ψ_3).

of the plot. Because the level of the groundwater was constant during the steady irrigation, we can assume that there was an equilibrium between the groundwater flow within the experimental soil plots and the surroundings of the soil plots. This may have caused the considerable losses of irrigation water into deeper groundwater layers.

Notwithstanding the constant levels of runoff and groundwater table, we observed a noticeable increase in the soil water content of the upper two layers on the mor humus plot. We hypothesize that the large amount of partly decomposed organic matter (see Table 3.1) took up some water. But the obvious increase in the water storage does not question our conclusion that the system reached a steady state. If we assume that the soil water content increased by $\Delta\theta=0.055$ in the upper 15 cm during the phase of steady irrigation, then the storage change corresponds to about 1.6 percent of the total volume of irrigation water.

Finally, the matric potential ψ_1 of both plots indicate that the surface layers were close to saturation. The next deeper layer had slightly smaller matric potential ψ_2 , whereas the two deepest layers were very close to saturation. Here again, we hypothesize that this is an effect of the high organic matter content of the upper layer of both plots. However, the smaller matric potential in the second soil layer (ψ_2) may have been caused by trapped air, too.

3.3.2 Tracer Breakthrough on the Muck Humus Plot

An overview of the results of the tracer experiments on both plots is given in Fig. 3.9. After the beginning of the irrigation, we waited about one day for steady hydrological conditions. Then we applied the subsurface line tracer Cl^- (as CaCl_2) and immediately afterwards the surface line tracer Br^- (as CaBr_2). We monitored the breakthrough of both solutes by measuring the electrical conductivity of the runoff. As soon as the electrical conductivity of the runoff reached its initial value, the tracers Br^- and Cl^- were again applied (as CaBr_2 and CaCl_2), but now on a larger surface in the lower half of the plot (Fig. 3.2 and 3.3).

On the muck humus plot, very little surface runoff was observed. As a consequence, to few samples were collected to estimate reliably the breakthrough of the tracers in the surface runoff. The Cl^- tracer, which was injected along a line of 130 cm at the boundary of the humic topsoil and the mineral subsoil (see Fig. 3.1), appeared shortly after its injection in the depth

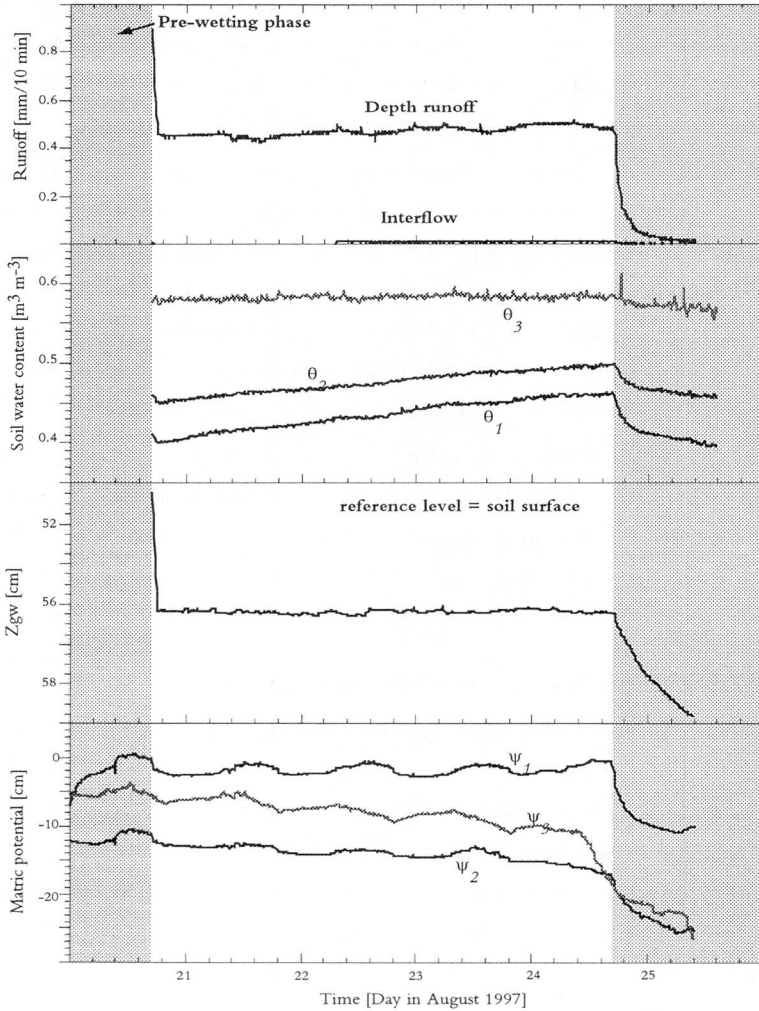


Figure 3.6: Results of the steady state irrigation ($2.3 \cdot 10^{-6} \text{ m}^3 \text{ m}^{-2} \text{ s}^{-1}$) on the mor humus plot. Partitioning of the total runoff, depth of the groundwater layer Z_{gw} , volumetric soil water content (θ_1 , θ_2 and θ_3) and the matric potential (Ψ_1 , Ψ_2 and Ψ_3) in the upper three soil layers.

runoff (Fig. 3.7). After about five minutes only, the Cl^- concentration began to increase, and the maximum concentration was observed 18 minutes after the tracer application. Afterwards, the tracer concentration in the depth runoff decreased exponentially and the background level of Cl^- was attained after about 200 minutes. During this interval of time, 31 percent of the applied mass of chloride was recovered from the depth runoff (Fig 3.9).

The tracer concentrations in the interflow (5–30 cm) showed only a very small increase. Less than one percent of the applied mass was collected in the interflow. This is somewhat surprising because the groundwater table remained in a depth of 10 cm below the surface throughout the experiment (cf. Fig. 3.5). Evidently, the streamlines of water were not parallel to the soil surface but were diverted to greater depths between the application strip and the downslope boundary of the plot. The lateral distance from the tracer source to the runoff gutters was about 3.3 meter. Thus, the velocity of the tracer peak in the water conducting soil pores was 0.0031 m s^{-1} . If we assume that the hydraulic gradient was equal to $\sin\alpha$ with α the slope angle, and if the transport of water follows Darcy's law, the estimated hydraulic conductivity K_{unsat} of the G_{Or} and G_t soil horizons at the prevailing moisture content equals 0.0062 m s^{-1} . This value is about 10^3 times larger than the measured saturated hydraulic K_{sat} conductivity in Table (3.2). This discrepancy suggests there was a network of fast flow paths in the subsoil horizons. During the 18 min between tracer injection and peak detection, less than 0.5 percent of the total amount of water present in the soil volume between tracer injection line and lower plot boundary took part in the transport of the tracer.

Also the Br^- tracer, applied along a line directly onto the surface, rapidly reached the depth runoff gutter (Fig. 3.7). The first sample with increased Br^- concentration was collected 9 min after the end of the tracer application, and the concentration reached its maximum after 22 minutes. Then, the concentration dropped sharply, but reached a second smaller peak after about 400 minutes. Finally, the concentration of Br^- decreased to the background level at a rate much slower than that of the subsurface tracer Cl^- . Thus, it appears that the breakthrough of Br^- was a superposition of a fast pulse and a slower secondary breakthrough. The Br^- recovery in the depth runoff was 28 percent, which compared well with the 31 percent of the total recovery found for Cl^- (Fig. 3.9). No breakthrough of Br^- was observed in the

interflow which suggests that a small portion of the surface tracer Br^- quickly bypassed the upper soil layers along the many roots of plants and through mice holes. The largest fraction of the tracer mass, however, seeped slowly through the soil matrix. The absence of Br^- in the interflow confirms that there was no lateral preferential flow at the interface between the humic and the subsoil horizons.

As described in the section 3.2.5, the tracers Br^- and Cl^- were also sprayed evenly onto the surfaces in the lower half of the plot. The flux densities were about 50 (Cl^-) and 20 times (Br^-) smaller during the application (cf. Table 3.3) as compared to those of the line sources.

Despite the small distance between the area treated with Cl^- and the runoff gutters there was no fast breakthrough, neither in the interflow nor in the depth runoff. The first sample with an elevated Cl^- concentration appeared 40 minutes after the tracer application in the interflow and after 70 minutes in the depth flow. Then the concentration steadily increased in both runoff fractions reaching peak concentrations after 500 min in the interflow and after 400 min in the depth flow. Although the tracer concentration in the interflow was more than twice as large as in the depth runoff, the total recovery of Cl^- in both runoff fractions was almost equal: 22 percent in the interflow and 21 percent in the depth runoff.

Following the application of Br^- over an area of 2.7 m^2 at an average distance of 1.7 m from the lower plot boundary, we obtained Br^- breakthrough curves which closely resembled those of Cl^- . The first breakthrough appeared 60 min after the application in the depth runoff and after 120 min in the interflow; the peak concentrations in both runoff components were similar and simultaneous. The recovery from depth runoff was 26 percent, but only 5 percent of the applied mass was recovered in the interflow (cf. Fig. 3.9). Again, no indications for fast transport of water and solutes were detected.

The differences between the maximum concentrations of Cl^- and Br^- in the interflow can be attributed to the larger soil volume which had to be passed by Br^- . The equal concentrations of Cl^- and Br^- in the depth runoff can be explained by macropore short-cuts in the subsoil. The residence times of both solutes in the upper humic soil is comparable and long compared to those in the subsoil. Therefore, both tracers are equally diluted in the upper

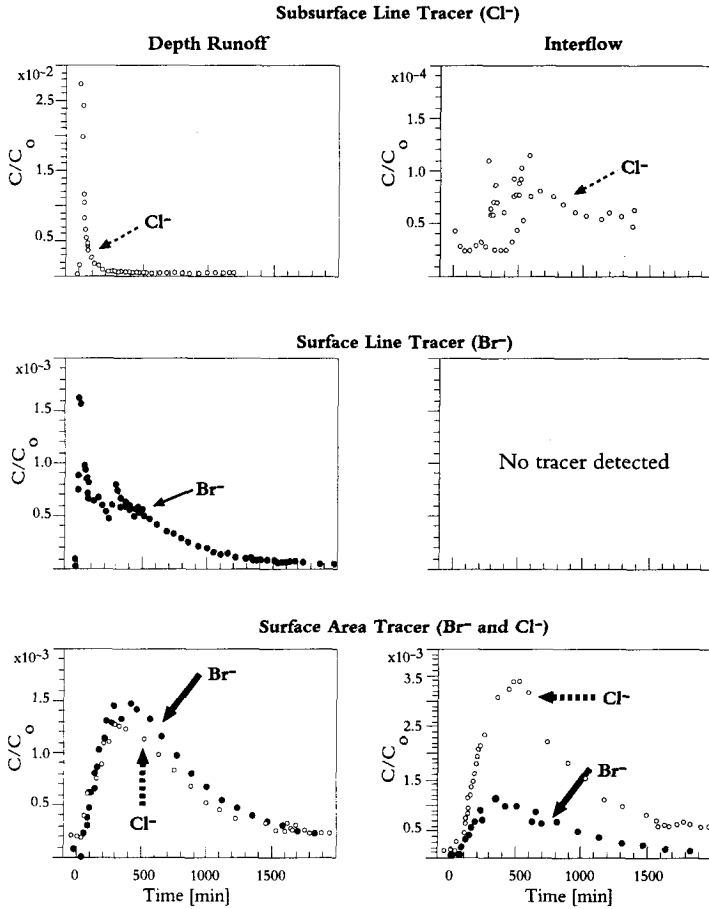


Figure 3.7: Tracer breakthrough on the muck humus plot. The left column is the depth runoff, the interflow is at the right. The arrows define the timing and location of the tracer application and correspond to Figs. 3.2 and 3.3. C/C_0 is the relative tracer concentration.

layer. In the macropores of the subsoil the tracer concentrations undergo only minimal changes. Furthermore, the absence of preferential flowpaths in the humic horizons might explain the slight delay of the breakthrough in the interflow as compared to the leaching by the depth runoff.

We postulate that the relatively slow transport in the matrix of the upper soil layers dominates the breakthrough of the tracers. Given the timing of the surface applied Cl^- breakthrough in the depth runoff, it can be argued that the tracer is transported at an angle intermediate between the surface slope and the vertical downward direction. This is supported by the breakthrough of Br^- , which is equally fast but less concentrated in the interflow. In summary, the solutes moved slowly through the upper muck horizon towards the subsoil. After the solutes reached the clayey subsoil, they were rapidly transported laterally through a network of macropores in the subsoil.

3.3.3 Tracer Breakthrough on the Mor Humus Plot

Since the mor humus soil plot was generally much better drained and had a lower ground-water table than the muck humus plot, hardly any interflow and no surface runoff was observed. Therefore, only the tracer breakthrough curves in the depth runoff will be discussed.

The breakthrough of Cl^- , injected in the subsoil at a distance of 3.3 m from the gutter, was initially as fast as on the muck humus plot: the first increase of the concentration was recorded after 7 min, but thereafter, the concentration rose more slowly reaching a maximum after only 70 min (18 min on the muck humus plot). The subsequent decrease was slow. After 2000 min the concentration was still larger than the background concentration. The maximum concentration of Cl^- in the depth runoff was more than 10 times smaller than the maximum concentration observed on the muck humus soil plot. This indicates that the dilution and thus the volume of water which participated in the transport process was much larger for the drier mor humus soil as compared to the wetter muck humus soil. Based on the travel time of the peak concentration, only 1.3 percent of the total volume of water participated in the flow. Although this is a very small portion of the soil water, it is still about

three times larger than on the muck humus plot. The mass recovery of Cl^- in the depth runoff, however, was 40 percent (Fig. 3.9), clearly exceeding the value observed for the muck humus soil (31 percent).

In the mor humus subsoil the water flows in a more complex network of small cracks and fissures. These flow conditions contrast with those on the muck humus plot where only few but large macropores exist. The runoff and the low level of the groundwater indicate that the water drains vertically down to the groundwater table. Then, in the saturated zone the flow direction is parallel to the slope. This explains the quick passage of the tracer through the subsoil, however, with a delayed arrival of the bulk of the tracer and a slow recession rate. Since in the mor humus plot the tracer plume reaches the saturated zone, it can be diluted much more than within a few (saturated) macropes in the subsoil of the muck humus plot.

The Br^- tracer, which was applied along a line onto the soil surface did not show a quick breakthrough (Fig. 3.8). The leading edge of the Br^- breakthrough reached the depth runoff gutter after about one hour. Thereafter, the Br^- concentration increased slowly reaching a maximum after 400 min. The subsequent decrease of the tracer concentration was comparable to that of the Cl^- tracer applied in the subsurface. The mass recovery of Br^- was 31 percent (Fig. 3.9). Despite the high application rate of the tracer solution, there was no evidence for a fast transport of the Br^- anions in the humic topsoil. This indicates that the transport in the topsoil was predominantly matrix flow. Because of the low soil pH, the transport of Br^- was possibly delayed by adsorption.

The breakthrough of Br^- and Cl^- , applied onto the surface in the lower half of the plot, was even more delayed. The two tracers, which were applied only after the tracers concentrations of the previous experiment (line sources) were close to the background, appeared in the runoff two hours after their application. Then, the concentrations continuously increased reaching a maximum after 620 minutes (Br^-) and 700 minutes (Cl^-). The subsequent decrease was in both cases comparably slow. These results support the hypothesis of vertical transport down to the saturated layer. The stronger retardation of Cl^- is probably due to the more pronounced adsorption of Cl^- in the humus layer. The recovery of the applied tracer mass equaled 49 percent for Cl^- and 37 percent for Br^- (Fig. 3.9). This difference was most prob-

ably caused by the longer travel distance and thus more pronounced mixing with the groundwater.

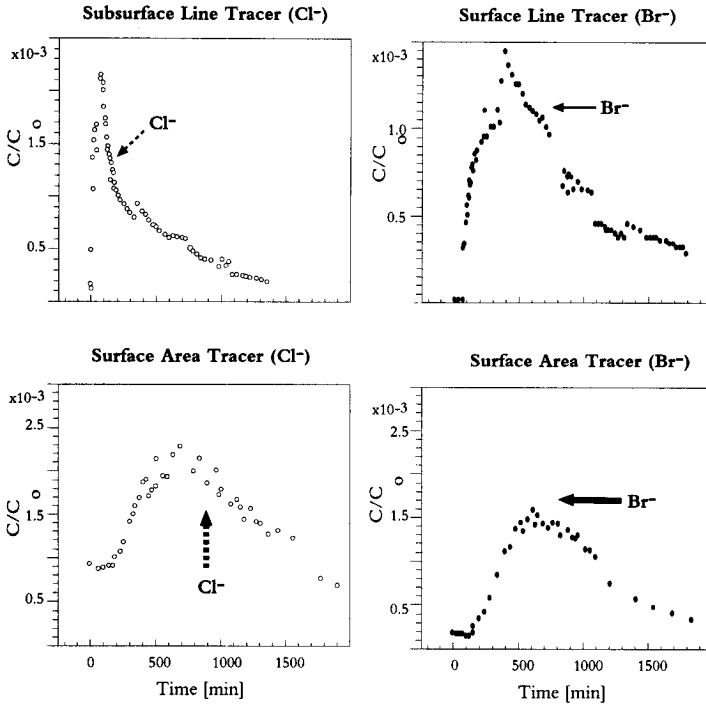


Figure 3.8: Tracer breakthrough in the mor humus plot. Depth runoff only. The arrows define the timing and location of the tracer application and correspond to Figs. 3.2 and 3.3. C/C_0 is the relative tracer concentration.

3.4 Discussion

In the framework of a study on the cycling of nitrogen in forested ecosystems, we performed a tracer experiment on two different soils under controlled hydrological conditions. Based on preliminary results (Schleppi *et al.*, 1998) we hypothesized that the observed leaching of nitrate might be possibly due to fast flow of water. We performed a tracer experiment on two soil plots which were irrigated to attain a steady hydrological state. To identify the influence of the different soil horizons on the transport of water and solutes, we injected tracer solutions into the subsoil and applied tracers as line and are sources at different distances from the runoff gutters (Figs. 3.2 and 3.3, Table. 3.3). Figure 3.9 presents a summary of the results.

A Cl^- tracer, which was injected along a line (130 x 10 cm) at the interface between the humic topsoil and the gleyic subsoil horizon, appeared quickly in the depth runoff of both plots. Still, the maximal tracer concentration was reached four times faster in the muck humus plot (18 minutes) compared to the mor humus plot (70 minutes). In addition, the peak concentration of the breakthrough curve was more than ten times higher on the muck humus. We attribute this discrepancy to the differences in the structure of the subsoil of both plots. The subsoil of the muck humus plot has almost no structure. Because of the almost impermeable soil matrix, the transport in the subsoil of the muck humus plot occurs mainly in the few preferential flow paths such as mice holes or along tree roots and buried wood. Therefore, the transport of water and solutes was very fast and because of the small amount of water taking effectively part in the transport process, the dilution of the tracer was not as pronounced as in the subsoil of the mor humus plot. On the much drier mor humus plot, the subsoil is more structured, which is expressed by a large amount of small cracks and fissures. As a consequence, the bulk of the tracer mass was transported more slowly and the recession of the tracer concentration was slower, though the early breakthrough of the tracer was equally fast as on the muck humus plot. These assumptions were confirmed by an estimation of the fraction of the soil water which took part in the runoff generation between tracer application and peak arrival; for the muck humus plot this was less than 0.5 percent, for the mor humus plot about 1.3 percent. No significant breakthrough of the Cl^- tracer was

detected in the interflow. Because of the slightly higher depth runoff and the clearly slower decrease of the tracer concentration on the mor humus plot, the recovery of Cl^- from the muck humus plot (31 percent) was lower than on the mor humus plot (40 percent). The recovery of the tracer corresponds nicely to the results of the water balance.

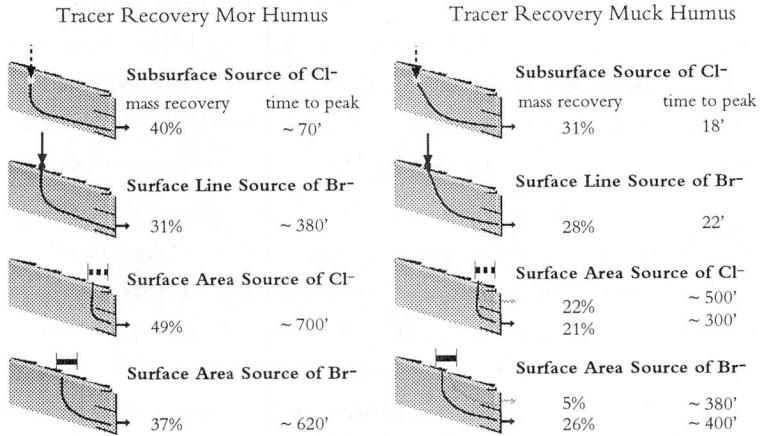


Figure 3.9: Summary of the tracer experiments on the soil plots. The mass recovery of the tracer (percentage of weight), the arrival of the peak concentration in the runoff gutters (in minutes after the application), the approximate location of the application areas and the possible flowpaths of the tracer are displayed.

The breakthrough of the Br^- tracer, which was applied as a line source (130 cm x 10 cm) on the soil surface, confirmed the results of the subsurface application of Cl^- . For the muck humus plot, an equally fast arrival of the Br^- anions in the runoff gutters (peak arrival after 22 minutes) resulted as for the subsurface Cl^- application. Afterwards the concentration of Br^- sank exponentially to 50 percent of the first peak, rose again to form a secondary peak at about 400 minutes after the tracer application and finally decreased much more slowly to the background concentration. Also on the mor humus plot a fast arrival of the tracer was observed, however, followed by a much slower increase of the concentration to reach a max-

imum value after about 400 min. Then, the tracer concentration decreased with a rate which was comparable to that of the subsurface Cl^- tracer, so that after 2000 min the concentration in the depth runoff was still higher than the background concentration. Summarized, apart from the early arrival on the muck humus soil, the breakthrough curves on both soil plots show the same characteristics. We conclude that in the upper humic soil layers of both soil plots flow processes in the soil matrix dominated the transport of the tracer. The early peak in the muck humus soil can be attributed to the higher tracer application rate, which probably caused fast flow processes in mice holes or along roots. From the total mass of Br^- which was applied onto the soil surface, 31 percent was gathered again in the depth runoff of the mor humus plot and 28 percent in the depth runoff of the muck humus plot. No tracer was observed in the interflow of the muck humus plot. The slightly lower tracer recovery of Br^- applied as a surface line source compared to the subsurface source of Cl^- might be due to adsorption of the tracer in the humic soil layers.

After the application of the line sources of Br^- and Cl^- we monitored the electrical conductivity of the different runoff fractions to determine when the tracer concentration in the runoff had reached a constant value or was at the background level. This allowed us to reuse Br^- and Cl^- as tracer anions. In the second part of the experiment, both tracers were applied to the plots again, however, now as areal tracers onto the soil surface. On the muck humus soil plot, a breakthrough of both tracers was observed in the interflow (runoff from the humic horizon) and in the depth runoff (subsoil). The smoothness of the tracer arrival suggested that in the upper humic soil layer flow through the matrix was the prevailing transport mechanism. The ratio of the concentrations of Br^- and Cl^- in the interflow and the depth runoff indicated that the main transport direction was intermediate to the slope direction and vertically downward. The timing of the maximum concentration was comparable to the second peak that was generated after the application of the surface line source of the Br^- tracer. The total mass recovery was 31 percent (5 percent interflow, 26 percent depth runoff) for Br^- and 43 percent for Cl^- (22 percent interflow, 21 percent depth runoff). On the mor humus soil plot both surface area tracers had very similar breakthrough curves, however, with larger arrival times than on the muck humus plot. The maximal concentration of Cl^-

in the subsurface runoff was reached after 700 minutes, the maximal Br^- concentration after 620 minutes. The mass recovery was 37 percent for Br^- and 49 percent for Cl^- . Here again, the results of the line tracer experiments were confirmed. The slower arrival of the peak concentrations is probably due to adsorption effects in the humic layers.

In our initial hypothesis on the hydrology of the artificial NITREX catchments, we assumed that the mor soil on the ridges was not hydrologically active and that the main part of the runoff would flow through the muck humus in the depressions. Given the results of the steady irrigation and the tracer experiments, this hypothesis has to be adjusted. The flow processes in the gleyic subsoil of the muck humus mainly occur in a network of macropores. Therefore, the flow in this soil layer is fast. Also in the humic topsoil fast flow processes can occur, however, most likely during a high intensity rainshower. Given the topographical position of the muck humus soil, in slightly depressed catchment areas, a concentration of the water from upslope areas can also cause fast flow processes in this soil. Since at least part of the macropore network in the subsoil is continuous in the lateral direction, the muck humus soil will most probably be the catchment area where fast runoff is generated. From the water balance during the steady irrigation (34.5 percent of the applied water could be collected), and from the tracer recovery of the subsurface tracer (no adsorption), we can conclude that about one third of the amount of water which passes the soil surface contributes to the fast part of the runoff generation.

Surprisingly high levels of the groundwater and output flow from the mor humus plot could be detected. These results were confirmed by the mass recovery of the subsurface tracer (40 percent). The timing of the runoff generation in the mor humus soil is clearly slower than in the muck humus soil. This is due to the lower soil water content and the more pronounced network of vertical pores. In addition, also the deeper level of the groundwater level in the mor humus might delay the tracer breakthrough. Thus, also on the mor humus plot runoff processes occur which may contribute to the total runoff of the sub-catchments. The upper soil layers of the mor humus areas have a very low pH. For this reason, part of the humic substances may have a positive load, allowing for nitrate to be removed from the soil solution by anion exchange. In addition, the flow processes in the well structured mor humus soil are

clearly slower than in the muck humus soil, which allows the microbial activity to play a more important role. The fast flow processes in the muck humus subsoil cause relatively short residence times of nitrate in at least part of the catchment. As soon as the precipitation water reaches the macropores of the reduced subsoil, it will quickly leave the catchment. As a result, in the muck humus plots an incomplete microbial immobilisation of nitrate may occur. Still, the residence time of the rainwater in the soil is large enough for ammonium to be removed from the soil solution by anion exchange. The findings of this study correspond well with the results and the hypothesis presented by Schleppe *et al.* (1998).

- Acknowledgements

We wish to thank all the members of the Soil Physics group at the Institute of Terrestrial Ecology for their intensive help during the field experiments-especially during the night, B. Studer for the laboratory analysis, and the Swiss National Science Foundation for financing the project '*Modelling runoff from forested hillslopes with a small scale topography*'.

Leer - Vide - Empty

Chapter 4

Runoff Processes at Three Spatial Scales



4.1 Introduction

Scale and spatial variability strongly affect the hydrological response of a drainage basin. Recent research on the effects of spatial variability and catchment scale on runoff generation revealed the importance of two processes: the production of runoff and the transport (routing) of runoff water through the catchment (Blöschl, 1995). The production of runoff on hillslopes, i.e. the partitioning of rainfall into runoff, soil water content changes and evapotranspiration determines the hydrological response of small catchments (Beven and Wood, 1993; Wooding, 1965). For larger catchments, the relative role of the stream network geometry and thus of the transport of the runoff increases. Robinson *et al.* (1995) showed that the transition between hillslope dominated and stream network dominated hydrologic response depends on the Peclet number of the catchment. Both production and transport of runoff are influenced by the spatial variability inherent to natural catchments. Spatially varying soil properties, geology, geomorphology and vegetation, together with the heterogeneity of the rainfall (Seyfried and Wilcox, 1995) affect the runoff generation in a drainage basin. Distributed deterministic catchment models are often used to study the importance of spatial variability for the generation and routing of runoff. Since deterministic models are based on a mathematical description of transport processes, they can be extrapolated beyond the calibration conditions (Beven, 1985). In addition, spatially distributed information on soil and geomorphological properties may be incorporated in the model to optimize its performance. These advantages, however, are outweighed by a serious drawback. Since spatially distributed information is seldom available, the parameterization of a distributed deterministic model is a non-trivial task (Beven, 1995; Grayson *et al.*, 1992). Furthermore, large uncertainties about the estimated parameters arise since numerous parameters have to be estimated from a generally incomplete data base. This limits the suitability of distributed deterministic models for exploring and detecting hydrological processes in heterogeneous study areas.

An alternative to distributed models are lumped-parameter approaches. Wilcox *et al.* (1990) and Grayson *et al.* (1992) showed that the use of lumped-parameter models yield a good correspondence between measured and simulated hydrographs. Although the spatial compo-

nents of the runoff generating processes are often lost in such models, the small number of parameters can be more reliably estimated. Beven (1989) pointed out that three to five parameters should be sufficient to describe the main characteristics of a hydrological data record. Most lumped-parameter models are, as deterministic models, based on certain assumptions on the processes in the real world. A lumped parameter approach which makes no a priori assumptions on the properties of the hydrological processes was presented by Young and Beven (1994). They proposed to use information on the soil moisture state to linearize the rainfall-runoff data and to apply algorithms of times series analysis to estimate discrete linear transfer functions from these data. Wang *et al.* (1981) suggested that the non-linear properties of the runoff generation disappear with increasing scale. This hypothesis was, however, only partly supported by the findings of Robinson *et al.* (1995). Although the hydrological gain (i.e. the amount of runoff caused by one unit of water input) became more linear, the timing of the runoff peaks depended more strongly on the hydrological state with increasing catchment size.

The purpose of this study is to detect the dominating runoff processes at three spatial scales and to discuss the effect of the spatial variability of soil and vegetation and the size of the drainage basin on the generation of runoff. To this end, we used hydrological data of catchments of three different sizes. This chapter addresses the following questions:

- Which processes dominate the runoff generation at a particular scale? Are these processes effective at the next larger scale? How do soil type and vegetation distribution affect these processes on different scales? To identify the various runoff processes, we used non-linear time series modelling
- What is the origin of the runoff water? Is it possible to link the origin of the runoff water with a process detected by the time series modelling? To this aim, we measured the electrical conductivity of the runoff water and determined whether it was 'event' or 'pre-event' water.

4.2 Theory

4.2.1 Linear Reservoir Model

A dynamic storage equation models the rate of change of the catchment storage $S(t)$ as the weighted difference of the volume of rainfall $u^*(t-\delta)$ entering the catchment minus the volume of runoff $x(t)$ leaving the catchment at the outlet

$$\frac{dS(t)}{dt} = -x(t) + \beta u^*(t-\delta). \quad (4.1)$$

The weighting factor is the so-called 'gain constant' β , accounting for losses of water from the basin which are not due to runoff processes. In Eq. (4.1), the input and output variables are true, but unmeasurable quantities, and δ denotes the time delay between the onsets of the input and the output. If we make the physically reasonable assumption that the catchment runoff depends linearly on the storage $S(t)$ or,

$$x(t) = \alpha S(t) \text{ or } S(t) = t_c x(t), \quad (4.2)$$

with $t_c = \frac{1}{\alpha}$ a catchment specific Time Constant. Equation (4.1) can be written as

$$t_c \frac{dx(t)}{dt} = -x(t) + \beta u^*(t-\delta). \quad (4.3)$$

This differential equation describes a first order linear dynamic system, and its response to a step input is equal to

$$x(t) = \beta (1 - e^{-t/t_c}). \quad (4.4)$$

4.2.2 Linear Time Series Model

It can be shown, that the discrete analogue of the differential equation (4.3) takes the following form

$$x_k = -a_1 x_{k-1} + b_0 u_{k-\delta}^*, \quad (4.5)$$

where a_1 describes the autoregressive decline of the hydrograph after cessation of the rainfall and b_0 is the hydrological gain of the basin. As before, x_k is the true catchment outflow and $u_{k,\delta}^*$ the true rainfall. The discrete time equation represents the continuous time description only if the input and the output variables are constant over the sampling interval t_s . The time constant $t_c = \frac{1}{\alpha}$ of this transfer function can be calculated as

$$\alpha = \frac{t_s}{\ln(-a_1)}. \quad (4.6)$$

If the system is stable, i.e. if $|a_1| < 1$, then x_k will reach a steady value if $u_{k-1,\delta}^*$ is chosen as a unit step function. This steady value is the so-called Steady State Gain (G_{stat})

$$G_{stat} = \frac{b_0}{1 + a_1}. \quad (4.7)$$

If two parallel processes occur, both driven by the same input and both contributing to the same output, then Eq. (4.5) can be extended to

$$x_k = -a_1 x_{k-1} - a_2 x_{k-2} + b_0 u_{k,\delta}^* + b_1 u_{k-1,\delta}^* \quad (4.8)$$

or, after application of the backward shift operator z^{-1} ($z^{-1}x_k = x_{k-1}$)

$$x_k = \frac{(b_0 z^0 + b_1 z^{-1})}{(1 + a_1 z^{-1} + a_2 z^{-2})} u_{k,\delta}^*. \quad (4.9)$$

The transfer function form of this type of time series model is obvious from Eq. (4.9). By means of partial fraction expansion, Eq. (4.9) can be decomposed into two parallel first order processes (cf. Eq. 4.5). In short notation, transfer function models are denoted by $[n(m+1)\delta]$, with n the number of autoregressive parameters and $(m+1)$ the number of parameters associated with the exogeneous inputs.

4.2.3 Estimation of the Transfer Function Model

Since the true values of the input and output variable are not measurable, Eq. (4.9) can not be used in this form. It is extended with a stochastic term $\{\xi_k\}$, taking into account stochas-

tic disturbances caused by measurement errors and the inadequateness of the time series model

$$y_k = \frac{(b_0 z^{-0} + b_1 z^{-1})}{(1 - a_1 z^{-1} - a_2 z^{-2})} u_{k,\delta} + \xi_k. \quad (4.10)$$

Here, y_k and $u_{k,\delta}$ are the measured output and input variables. Equation (4.10) is equivalent to an autoregressive time series model with exogeneous inputs (ARX). The unknown stochastic disturbances $\{\xi_k\}$ strongly influence the parameter estimation procedure. Young (1984) showed that a least squares procedure results in biased estimates of the parameters. The method of the so-called 'instrumental variables' (Ljung, 1987; Sönderström and Stoica, 1983) may be applied to bypass these problems. We applied the 'Simplified Refined Instrumental Variable' method by Young (1985) to estimate the parameters of the transfer function model recursively. To identify the optimal number of parameters of the linear transfer function, we used the Young Information Criterion (YIC). This criterion combines the goodness of fit of the transfer function with the uncertainty of the parameter estimations. The transfer function which results in the smallest value for the YIC is preferable.

4.2.4 Nonlinear Time Series

Catchment runoff is typically a nonlinear function of the amount of rainfall. If the nonlinear phenomena underlying the runoff generation were mathematically understood, then the modelling of the time series would only involve parameter estimation. But a profound physical understanding is rarely available and the modelling process starts with the inspection of a set of nonlinear models. In the most general case though, a nonlinear process can be approximated by a linear model with time variable parameters. This operation, which is called statistical linearization, is extensively used by Young (1993). The result is a linear transfer function with time dependent parameters which can be written as

$$y_k = \frac{(b_0(k)z^{-0} + b_1(k)z^{-1} + \dots + b_m(k)z^{-m})}{(1 - a_1(k)z^{-1} - a_2(k)z^{-2} - \dots - a_n(k)z^{-n})} u_{k,\delta} + \xi_k. \quad (4.11)$$

Equation (4.11) is an extension of Eq. (4.10), allowing time varying parameters. The estimation problem changed to inferring the time variable parameters of this linear model. For hydrological series, mostly the parameters associated with the input variables will vary. Therefore, at this stage of the analysis, mainly a transfer function of order $[1 \ 1 \ \delta]$ with a single time variable parameter $b_0(k)$ will be used.

The estimation of the time variable parameters bases on a smoothing algorithm which estimates $b_0(k|N)$ from all N available observations. In this work, we employed the smoothing algorithms developed by Rauch *et al.* (1965), which are based on the outcome of the Kalman filter (Kalman and Bucy, 1961). The correlation of $b_0(k|N)$ with additional information on the hydrological state of the catchment can be explored and used to model $b_0(k|N)$ as a non-linear function \mathfrak{S} of the state variables.

$$b_0(k|N) = \mathfrak{S}(\text{State}) . \quad (4.12)$$

This function can be used to transform the measured precipitation into an effective, runoff generating input variable $u_{\text{eff}, k-\delta}$

$$u_{\text{eff}, k-\delta} = u_{k-\delta}^* \times \mathfrak{S}(\text{State}) . \quad (4.13)$$

To use Eq. (4.13), additional measurements on the hydrological state of a basin (e.g. soil water content, groundwater depth, soil water potential) are necessary. Most often, though, the runoff y_k , too, can be employed as a surrogate variable for the soil moisture state. Based on the output measurements and the effective input variable $u_{\text{eff}, k-\delta}$, linear transfer functions can be estimated, e.g. by means of an instrumental variable technique.

4.3 Material and Methods

All experimental catchments are situated in the upper part of the valley of Alptal, (northern Swiss prealps), between 1100 m and 1530 m a.s.l. The prevailing westwinds force wet air into the valley, causing the high annual sum of precipitation of 2300 mm. The seasonal distribution of the precipitation shows a maximum during summer (270 mm in June) and a

minimum during fall (135 mm in October). About 30 percent of the rainfall is snow, which mainly falls between December and May. The average annual air temperature measured at an elevation of 1150 m is 4.9 °C (Burch, 1994). The geological substrate at the Alptal valley is Flysch, a formation with layers of calcareous sandstone alternating with shists. On this bedrock, an umbric Gleysol (Diserens, 1992) developed with, depending on the soil moisture conditions and the micro-topography, two distinct humus rich topsoils.

4.3.1 Erlenbach

Dating back to 1978, the Erlenbach catchment is one of the best documented alpine catchments in Switzerland. It has a surface area of about 0.7 km² and an average slope of 20 percent with aspect west. The catchment is very narrow with permanent brooks deeply cut into narrow and steep gorges. About 39 percent of the catchment is covered by forests, the other 61 percent are wetland dominated by grasses. To prevent landslides from the wetland areas, a network of drainage trenches covers the catchment. These trenches quickly rout the runoff of the wetland zones to the permanent brooks.

Precipitation is measured at two sites: at one wetland site near the catchment boundary the pluviometer is at a standard elevation of 2 meter, and at the other site, well within the forested part of the catchment, precipitation is measured above the canopy layer (30 m) (see Fig. 4.1). Runoff of the Erlenbach is measured by a large V-notch weir, where also the electrical conductivity (EC) of the runoff is recorded. The average annual water balance (1978-1993) shows that the total amount of precipitation of 2190 mm is partitioned into 1740 mm of runoff (79.5 percent) and 450 mm of evapotranspiration (20.5 percent).

4.3.2 Artificial Sub-catchments

Within the Erlenbach headwater catchment, three artificial sub-catchments of about 1500 m² were set-up to study the cycling of nitrogen (Schleppi *et al.*, 1998). One wetland (N₃) and two forested areas (N₁, N₂) were delineated by 80 cm deep trenches (Fig. 4.1). To avoid erosion of the gleyic subsoil, the trenches were covered with a geo-textile. The runoff of the sub-catchments is caught in the trenches and the trench water collected by a V-notch

weir where the water level is recorded automatically every 10 minutes by an ultrasonic device. In addition, the EC of the runoff is recorded (LF95, Tetracon, WTW) with the same frequency. Since 1996, both sub-catchments have been sprinkled with rainwater, collected on a plastic sheet upslope the sub-catchments. The sprinkling is simultaneous and proportional to the natural rainfall and the applied quantity is continuously logged. The water balance of both forested plots from April 1994 to March 1995 was nearly closed (Schleppi *et al.*, 1998).

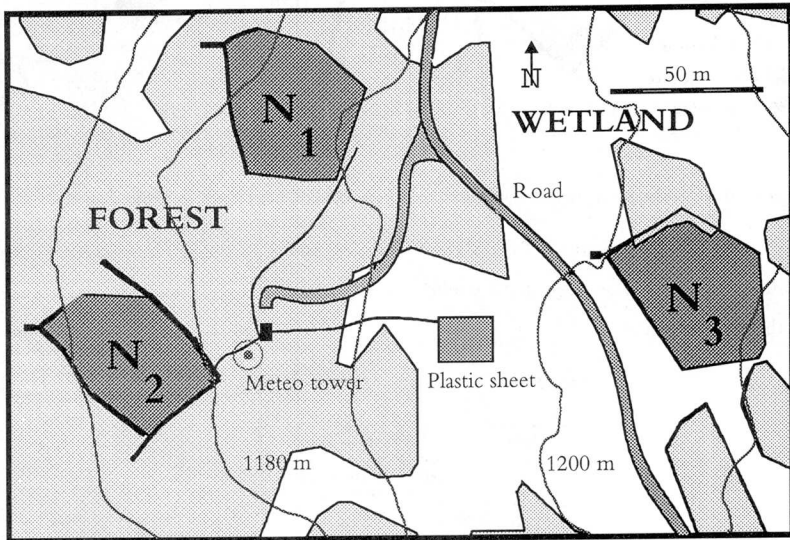


Figure 4.1: Overview of the sub-catchments N_1 , N_2 and N_3 . The shaded areas are forests, the white areas indicate the wetland.

On the forested sub-catchments, closely linked to the micro-topography, two distinct humus topsoils occur on top of the gleyic subsoil. In small depressions, a wet muck humus topsoil prevails and on the drier ridges a mor humus top layer is present. Figure 4.2 shows the distribution of both soils and waterlogged areas in the forested sub-catchment. Table 4.1

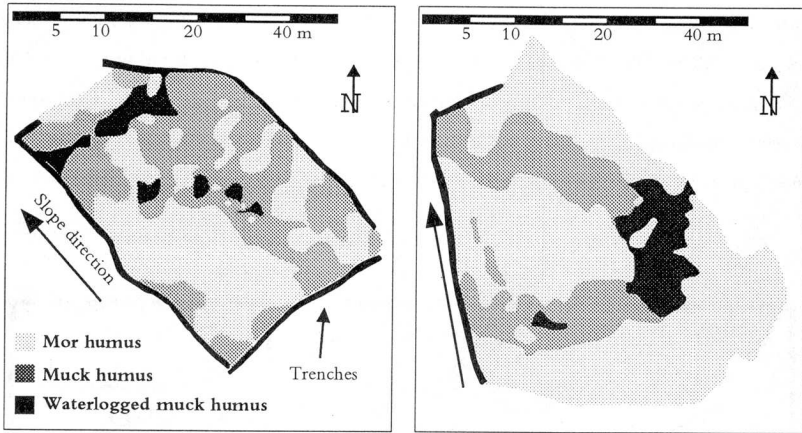


Figure 4.2: Soil type distribution on the sub-catchments N₁ (right) and N₂ (left).

presents the soil type statistics for the three sub-catchments. In the sub-catchment N₂, the groundwater levels were recorded separately for both soil types with readings synchronized to those of the runoff.

Table 4.1: Properties of the sub-catchments.

	Area (m ²)	Fraction covered by	
		Muck humus	Mor Humus
N ₁	1974	32	78
N ₂	1449	53	47
N ₃	1473	100	0

In the wetland sub-catchment N₃, only the muck humus topsoil occurs. On the mounds and ridges, where the upper soil layers are mor humus, Norway spruce (*Picea Abies* (L.) Karst.) and silver fir (*Abies Alba* Mill.) grow. The ground vegetation is dominated by *Vaccinium* species on mor humus and mainly grasses and *Equisetum* species on muck humus. To obtain detailed information on the runoff generation in both humus types, we monitored runoff

from and carried out experiments on two soil plots each representing one of the humus types.

4.3.3 Soil Plots

The soil plots were, as the sub-catchments, separated from their environment by trenches of 80 cm deep. Both plots were situated near the sub-catchments, having the same slope (20 percent) and aspect (SW). The muck soil plot had a surface area of 13.8 m² and was covered with grasses. The mor soil plot (area 11.8 m²), situated between large Pine trees, was covered by needle litter and free of any vegetation. For both plots, we used the same precipitation records as for the sub-catchments, i.e. those which were collected above the tree canopy. Thus, by using these precipitation records, the runoff reflects also the influences of the vegetation. Tables 3.1 and 3.2 show physical and chemical soil properties of the plots. A detailed description of the soil properties of both plots and of the instrumentation is given in section 3.2.

During the installation of the sensors in the soil plots, we recognized differences in the soil structure of the subsoil. The gleyic subsoil horizons in the muck humus plot were almost completely reduced. In addition to stones from calcareous layers of the Flysch bedrock, we detected old tree roots and a few channel-like pores. The subsoil of the mor humus plot was much better structured. Beside stones and tree roots, the partly reduced subsoil had complex micro-structures with varying sizes, but mainly extending in the vertical direction. Generally, the mor humus plot was much better drained and the groundwater was much deeper.

In addition to monitoring natural rainstorms, we performed an irrigation experiment under controlled boundary conditions. This had a dual purpose. First, from earlier experiments (Feyen *et al.*, 1997), we had evidence that the subsoil was definitely more permeable than assumed (see also section 1.1) and that water from upslope areas could possibly enter the plots through the subsoil. By irrigating the plots during periods free of natural rainfall (i.e. no sub-surface inflow of water into the plots), it was possible to study the runoff generated within the soil plots only. Second, since the precipitation was measured above the tree canopy layer, the real input of water into the mor humus plot is overestimated because of the interception

by trees. This again limits the possibilities to compare the influence of the soil type on the runoff generation on both plots. During the irrigation experiments, we covered the plots with a plastic roof to shield them from rainfall and to reduce evapotranspiration. Thus, by controlling the boundary conditions of the soil plots, we studied the influence of soil type on the runoff generation.

An irrigation device, developed to impose a constant and spatially uniform flux of water drove over the inclined plots on two rails. Attached to the device was a bar with 16 nozzles which sprayed the water onto the ground from an average height of 30 cm. In the laboratory, we tested whether the sprinkler sprayed uniformly and we obtained a coefficient of spatial variation of 0.04. Under field conditions, the same coefficient of variation was obtained on the mor humus plot, which had a smooth surface. On the muck humus, however, the rougher soil surface caused a much larger coefficient of variation of 0.11.

4.3.4 Event and Pre-event water

To partition the runoff into 'event' and 'pre-event' water, we used the approach of Pearce *et al.* (1986), who suggested to compute

$$Q_{old} = Q_{tot} \frac{(C_a - C_{new})}{(C_{old} - C_{new})} \quad (4.14)$$

where Q is the runoff, C the electrical conductivity of the runoff and the subscripts denote the event (new), the pre-event (old) water and the actual electrical conductivity of the runoff (a). Conceptually, this corresponds to the existence of two reservoirs, one with the soil water ('old' or 'pre-event') and the other with rainwater ('new' or 'event'). The total runoff Q_{tot} is a mixture of the outflow of both reservoirs. For the electrical conductivity of the rainfall, we used the average of six weekly samples, half of which were taken from the throughfall.

Clearly, considering a natural drainage basin as composed of two reservoirs is an extreme simplification. More sophisticated methods to trace the origin of the runoff water are the use of dissolved organic carbon (DOC) and ^{18}O (Peters *et al.*, 1995), K^+ content and acid neutralizing capacity of the runoff (Elsenbeer and Lorieri, 1995), or the use of end member

mixing analysis (EMMA, Hooper *et al.*, 1990). Still, Caissie *et al.* (1996) showed that the electrical conductivity can be an adequate parameter to separate the contribution of surface flow and groundwater flow to the total runoff.

4.4 Results

4.4.1 Irrigation of the Soil Plots

4.4.1.1 Muck Humus Plot

The irrigation experiment on the muck humus plot was performed on September 1, 1997. We applied nine subsequent pulses of water to the plot with an average application rate of $0.31 \cdot 10^{-4} \text{ m}^3 \text{ m}^{-2} \text{ s}^{-1}$. The duration of the pulses varied between 136 s and 540 s, and the time interval between the start of the pulses was on average equal to 16 minutes.

The total amount of water applied to the soil surface equalled 998 litres. At the begin of the experiment, the groundwater was 12 cm below the soil surface (Fig. 4.3) and the average soil moisture content equal to 0.46 (θ_1) and 0.69 (θ_2 and θ_3). The average matric potential of 23 cm (ψ_1), 14 cm (ψ_2) and 0 cm (ψ_3) indicate that the upper horizons of the plot were fairly wet, though not saturated. Immediately after the first pulse of water was applied, the level of the groundwater started to rise, as did the soil moisture content. Briefly after the application of the last pulse, the groundwater reached a maximal height of about 6 cm below the soil surface, and the soil moisture content increased to 0.64 (θ_1), 0.72 (θ_2) and 0.74 (θ_3). The matric potentials ψ_1 and ψ_2 slowly decreased to a few centimetres, indicating that the upper soil layers, too, were close to saturation. The suction cups in T_3 were mainly below the groundwater table which resulted in positive values for ψ_3 . The discrepancy between the matric head and water table depth are likely due to the micro-relief of the soil plot or to an impeding layer in the subsoil. However, the patterns of the temporal variation of the groundwater table, the soil water content and the matric potential match well (Fig. 4.3).

The irrigation resulted in a fast response of the outflow in the gleyic subsoil (depth runoff). The first 3 pulses produced only a small increase of the interflow. However, as soon as the

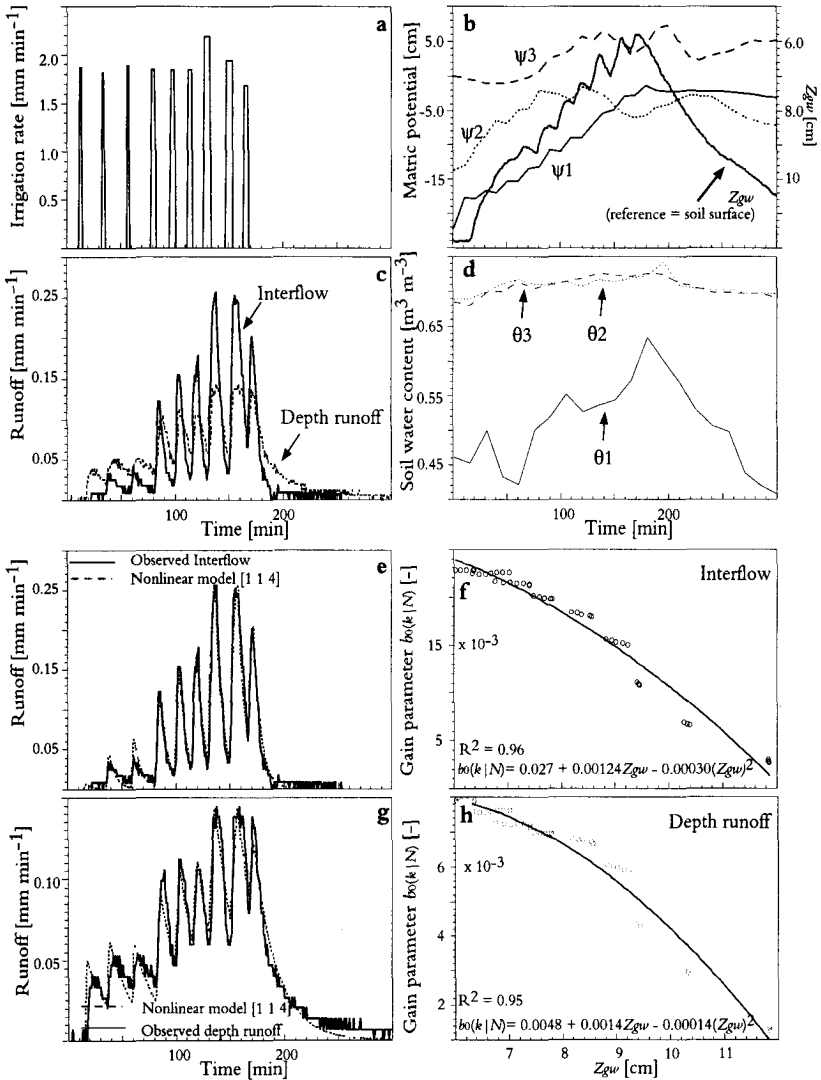


Figure 4.3: Results of the irrigation experiments on the muck humus plot. Water input (a), Z_{gw} and matric head (b), measured outflow (c), soil water content (d), nonlinear models for the interflow (e) and the depth runoff (g), gain parameter vs. Z_{gw} for the interflow (f) and the depth runoff (h).

plot was wetter and the groundwater closer to the soil surface, the interflow reacted more strongly to the irrigation pulses than the depth runoff. After the last pulse, the interflow rapidly decreased and completely vanished after 70 minutes. The depth runoff decreased more slowly and stopped four hours after the end of the irrigation. The total amount of depth runoff equalled 195 litres (20 percent of the irrigation) and 167 litres in the interflow (12 percent). The runoff ratio, i.e. the ratio of the total runoff to the total water input, equalled 32 percent. The other 68 percent of the irrigation water caused changes in the soil water content and flowed from the plot as deep seepage.

Initially, we fitted a discrete linear transfer functions of structure $[1 \ 1 \ \delta]$ to the data (Eq. 4.10, Table 4.2). For both runoff fractions, a time delay $\delta \approx 4$ resulted in the best fit. A linear transfer function described the depth runoff slightly better than the interflow. This is mainly due to the weak initial reaction of the interflow and shows that the production of runoff depends not only on the water input but also on the state of the system, namely, the water storage of the soil.

Table 4.2: *Nonlinear transfer function models for the irrigation experiments. t_c is the Time Constant and G_{stat} the Steady State Gain of the system. YIC is the Young Information Criterion and R_T^2 the Coefficient of determination of the time series model.*

		Runoff fraction	Model	Parameters	t_c [min]	G_{stat}	R_T^2	YIC
Linear transfer function	Muck humus	Depth runoff [mm min ⁻¹]	[114]	$a_1 = -0.9454 \pm 0.0001$ $b_0 = 0.0096 \pm 0.0003$	17.81	0.176	0.93	-10.6
		Interflow [mm min ⁻¹]	[114]	$a_1 = -0.8746 \pm 0.0003$ $b_0 = 0.0231 \pm 0.0007$	7.46	0.184	0.90	-9.9
	Mor humus	Depth runoff [mm min ⁻¹]	[116]	$a_1 = -0.90130 \pm 0.0004$ $b_0 = 0.0519 \pm 0.0002$	9.62	0.528	0.77	-8.4
Nonlinear transfer function	Muck humus	Depth runoff [mm min ⁻¹]	[114]	$a_1 = -0.9538 \pm 0.0001$ $b_0 = 0.0086 \pm 0.0002$	21.14	0.186	0.95	-11.1
		Interflow [mm min ⁻¹]	[114]	$a_1 = -0.8569 \pm 0.0002$ $b_0 = 0.0238 \pm 0.0004$	6.48	0.166	0.97	-12.5
	Mor humus	Depth runoff [mm min ⁻¹]	[114]	$a_1 = -0.8927 \pm 0.0003$ $b_0 = 0.0381 \pm 0.0006$	8.81	0.394	0.84	-9.2

To explore the nonlinear dependence of both runoff fractions on precipitation, linear transfer functions of structure [1 1 4] with a time variable gain parameter $b_0(k|N)$ were fitted to the experimental data. We fixed the autoregressive parameter a_1 of the time-dependent transfer functions to the value obtained for the linear time invariant transfer function (Table 4.2). We modelled the variation of $b_0(k|N)$ by a second order polynomial in the level of the groundwater (see Fig. 4.3). We fitted the polynomial by least squares using only the data records (shifted with the time delay δ) when the plot was irrigated. For these data points the variance of the estimation error of $b_0(k|N)$ was minimal because the estimation depended on information in both the input and output data series. Then, we transformed the irrigation rate to effective (runoff generating) input (cf. Eq. 4.13) and re-estimated the parameters of the time invariant linear transfer function (Table 4.2 and Fig. 4.3). As before, [1 1 4] models described both runoff fractions best. Compared to the linear transfer function estimated from the initial data, the nonlinear models fitted both runoff fractions better. This was most clear for the interflow, indicating that the groundwater level had a more pronounced influence on this runoff fraction.

We calculated time constants (t_c , Eq. 4.3) and a steady state gains (G_{stat} , Eq. 4.5) from the transfer function parameters (Table 4.2). For the depth runoff this resulted in a t_c of 21 minutes and a G_{stat} of 0.19; for the interflow the t_c equalled 6.5 minutes and the G_{stat} was 0.17. The sum of the G_{stat} of both runoff fractions indicates that of the total surface of 13.8 m² about 36 percent is drained by through runoff gutters. This is in close correspondence to the runoff ratio, i.e. the ratio of the total runoff to the total water input.

The first arrival of the outflow in the gutters is much faster than can be expected from the saturated hydraulic conductivity K_{sat} (Table 3.2). The fast reaction of the runoff suggests that flow along preferential flowpaths such as macropores is the most important component of the early runoff. This is in agreement with the tracer experiments described in Chapter 3. Especially in the gleyic sub-soil horizons of the muck humus plot we observed large, channel-like macropores. Preferential flow through macropores also explains the less pronounced nonlinear properties of the depth runoff compared to the interflow. Since in the gleyic sub-

soil most of the water flows through macropores, the impact of the soil water content and the groundwater level on the runoff generation is smaller.

Clearly, the high irrigation rate influences the dynamical properties of the outflow. It may possibly result in an underestimation of the time delay δ and in an overestimation of the hydrological gain of the soil plots. Still, these high rate of water input is not unnatural and can be observed during thunderstorms during fall. Moreover, during the irrigation we never observed ponding, so the water is certainly not forced into macropores. Finally, the dynamical properties of the runoff generation are mainly inferred from the autoregressive components of the estimated time series models. This corresponds temporally with the declining part of the outflow hydrographs, i.e. the draining of the soil plots after the irrigation finished.

4.4.1.2 Mor Humus Plot

On the mor humus plot, the irrigation experiment took place on August 19, 1997. The irrigation regime was comparable to that of the muck humus plot, except a slightly longer time lag between the irrigation pulses. This was a consequence of the slower reaction of the outflow to the irrigation inputs on the mor humus plot. Before the irrigation started, the groundwater was 58 cm below the soil surface (Fig. 4.4) and the soil moisture content was 0.29 (θ_1), 0.47 (θ_2), 0.58 (θ_3) and 0.56 (θ_4). The tensiometer readings at this time indicated that the organic upper soil layer (ψ_1) and the deeper gleyic soil layers (ψ_3 and ψ_4) were close to saturation, whereas the humic A_h horizon was still unsaturated (ψ_2). Except for the first five pulses, the level of the groundwater reacted rapidly to the irrigation. In contrast to the muck humus plot, the groundwater level fluctuated strongly with the irrigation. The tensiometers showed that the topsoil (ψ_1) and the deeper subsoil (ψ_3 and ψ_4) became quickly saturated, whereas the humic A_h layer remained unsaturated (ψ_2) during the experiment. As can be seen from Fig. 4.4, mainly depth runoff drained from the mor humus plot. The first five pulses, however, only induced a small increase of the discharge from the subsoil. This corresponds well to the piezometer measurements, suggesting a direct influence of the groundwater on the runoff generation. After the last irrigation pulse, the depth runoff declined more slowly than on the muck humus plot. Surface runoff flowed only during the irrigation, whereas the interflow reacted on the last two pulses only (Fig. 4.4).

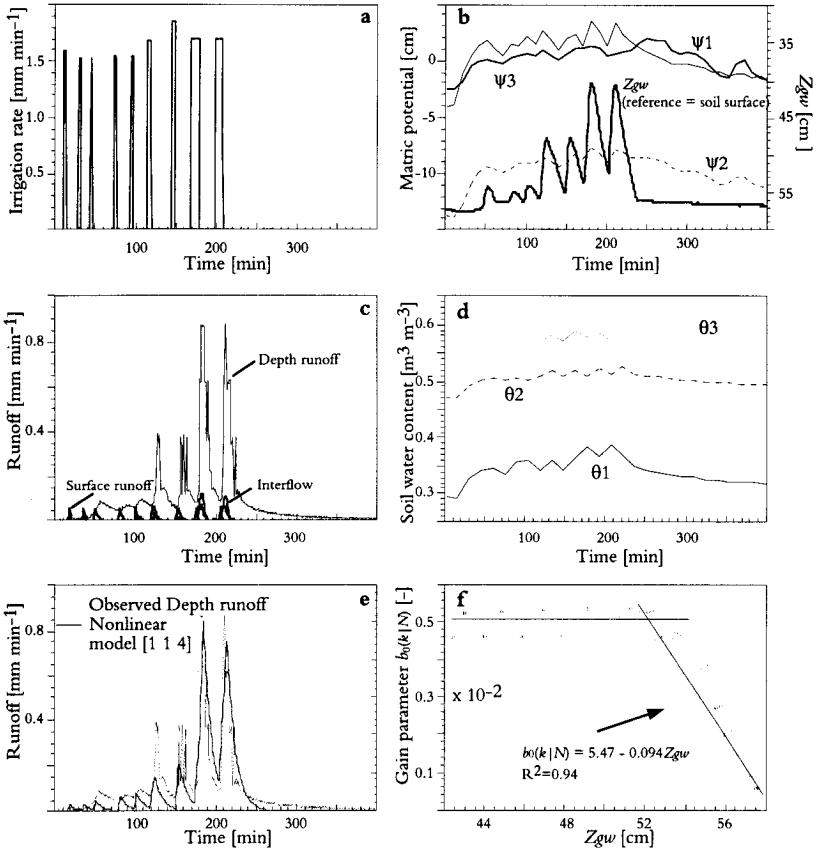


Figure 4.4: Results of the irrigation experiment on the mor humus plot. Water input (a), Z_{gw} and matric head (b), measured outflow (c), soil water content (d), nonlinear model for the depth runoff (e), gain parameter vs. Z_{gw} for the depth runoff (f)

In total, we applied 894 litres of water to the mor humus plot, of which 446 litres (50 percent) were collected as depth runoff, 23.4 litres (3 percent) as surface runoff and 8.5 litres (1

percent) as interflow. Thus, the total runoff ratio of 54 percent was clearly larger than on the muck humus plot.

A linear transfer function of structure [1 1 4] failed to describe depth runoff (Table 4.2). As before, we estimated a linear transfer function with a time variable gain parameter from the data. Almost the complete range of variation of $b_0(k|N)$ could be attributed to changes in the groundwater level between 58 cm and 52 cm. For higher groundwater levels, the influence on $b_0(k|N)$ declined. Two straight lines that intersect at a groundwater level of 52 cm (see Fig. 4.4) served to transform the irrigation into effective input. The resulting re-estimated time invariant transfer function still did not fit the data well (Fig. 4.4). The time delay δ between the begin of the irrigation and the begin of the runoff varied during the experiment. Since this type of nonlinearity is not taken into account by the discrete transfer function models, the estimated parameters are to be interpreted with care.

Similar to the muck humus plot, the fast reaction of the outflow to the irrigation pulses reveals that water flow along preferential flow paths determines the runoff dynamics on the mor humus plot, too. However, the fast outflow occurred only when the soil was very wet. Furthermore, the recession of the hydrographs after stopping the irrigation was smoother than on the muck humus plot, indicating slower runoff dynamics.

4.4.2 Runoff from Natural Rainstorms

The rainfall, runoff, soil water content and matric potential were recorded from fall 1996 to spring 1997. Avoiding snowfall or snowmelt, we selected two data records: S1 from May 28 1997 to July 27 1997 (from May 28 1997 to June 28 1997 for both soil plots) and S2 from October 1, 1996 to November 11, 1996. Since the precipitation and runoff measurements serve to estimate nonlinear transfer functions, snowfall (monitored with heated pluviometers) and snowmelt (runoff without rainfall) would have severely complicated the analysis. The rainfall-runoff data series at all scales are shown in Figs. 4.6, 4.7, 4.10, 4.11 and 4.13.

Table 4.3: *Water balances during S1 and S2. The abbreviations stand for Erlenbach (ERL) and the sub-catchments N₁, N₂ and N₃. The number in brackets is the runoff ratio, i.e. the ratio of the total runoff to the total rainfall, in percentage. The different amount of rainfall on the sub-catchments is due to the additional irrigation: N₁ and N₂ were sprinkled, N₃ received natural rain only.*

spring S1				fall S2			
	Rainfall [mm]	Runoff fraction	Runoff [mm]		Rainfall [mm]	Runoff fraction	Runoff [mm]
May 28 to June 28, 1997				Oct. 01 to Nov. 11, 1996			
Muck	287.0	surface runoff	0.2 (<1%)	Muck	296.4	surface runoff	1.9 (<1%)
		interflow	2.5 (<1%)			interflow	71.9 (24%)
		depth runoff	62.2 (22%)			depth runoff	63.7 (22%)
Mor	287.0	surface runoff	0.1 (<1%)	Mor	296.4	surface runoff	0.7 (<1%)
		interflow	1.1 (<1%)			interflow	0.1 (<1%)
		depth runoff	74.9 (26%)			depth runoff	143.0 (48%)
May 28 to July 27, 1997				Oct. 01 to Nov. 11, 1996			
N₁	704.7		366.9 (52%)	N₁	318.5		182.5 (57%)
N₂	709.4		461.9 (65%)	N₂	320.7		223.3 (70%)
N₃	660.4		471.3 (71%)	N₃	296.4		272.8 (92%)
ERL	662.4		468.8 (71%)	ERL	300.7		234.2 (78%)

We monitored the precipitation at two different locations within the Erlenbach headwater catchment: near the NITREX sub-catchments above the tree canopies (Fig. 4.1), and 2 m above the ground near the boundary of the headwater catchment. From Table 4.3 it can be seen that the total amount of precipitation measured at the two locations was comparable during both measurement periods (ERL vs. N₃). Sprinkling the sub-catchments N₁ and N₂ increased the total precipitation by 7 percent (ERL vs. N₁, N₂). The characteristics of the rainfall slightly differed for the two observation periods. In fall 1996 (S2), the average rate during periods with rainfall equalled 0.3471 ± 0.3475 mm/10 min, whereas the rate was 0.4390 ± 0.7075 mm/10 min for spring 1997 (S1). The higher average precipitation and the larger standard deviation during spring are presumably due to thunderstorms.

The runoff ratio shows that during fall 1996 more runoff flowed from all studied areas than during spring 1997. For both periods, more water flowed from the mor humus plot than

from the muck humus plot. This is in agreement with the irrigation experiments. The runoff from the mor humus plot flowed mainly in the depth runoff. During period S1 the interflow almost completely vanished on the muck humus plot, but during S2 interflow was the most important contribution to the runoff. Surface runoff did in none of all cases contribute significantly to the total runoff, on either plot and for either period. During both observation periods, the runoff ratio was largest for the sub-catchment N_3 . The particularly low runoff ratio on N_1 is most likely due to an overestimation of the catchment area. The upper part of this sub-catchment has a flat topography, which makes an accurate delineation of the watershed boundaries uncertain. Based on the balances for nitrogen and water of N_1 and N_2 , the actual area of N_1 was estimated as 1670 m² (pers. comm. P. Schleppe). This increases the runoff ratio for N_1 to 62 percent (S1) and 67 percent (S2) respectively, which now compares well to N_2 .

4.4.3 Modelling the Nonlinear Properties of the Runoff Generation

To model the variation of the time variable parameter $b_0(k|N)$ we explored the measurements of the soil water content (θ_1 to θ_4), the matric potential (ψ_1 to ψ_4) and the depth of the groundwater table (Z_{gw}). First we studied the correlation of these variables with the time series of $b_0(k|N)$. The variables θ_1 and Z_{gw} correlated best with $b_0(k|N)$. In contrast to the irrigation experiments, however, the selection of a nonlinear equation was less obvious. Therefore, according to the approach by Beven and Young (1995), we modelled the variation of $b_0(k|N)$ by powerlaw relationships of the form

$$u_{eff, k, \delta} = u_{k, \delta}^* \times (\theta_1)^\gamma \quad (4.15)$$

and

$$u_{eff, k, \delta} = u_{k, \delta}^* \times (Z_{gw})^\gamma \quad (4.16)$$

We estimated the parameter γ by minimizing the YIC which results from the estimation of a linear $[1 \ 1 \ \delta]$ transfer function from y_k and $u_{eff, k, \delta}$ (Eq. 4.10). This procedure, although computationally inefficient, resulted in an equation which described the nonlinear depend-

ence of $b_0(k|N)$ on the hydrological status best. As an example of the minimization procedure, Fig. 4.5 shows the relation of γ and the YIC for the depth runoff of the muck humus plot.

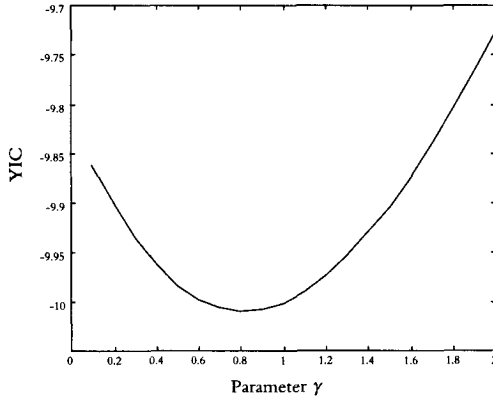


Figure 4.5: Result of the power law optimization for the depth runoff from the muck humus plot.

Table 4.4: Power law exponent γ estimated for both time series and all studied areas. YIC is the Young Information Criterion and R_T^2 the Coefficient of determination of the time series model.

Scale	S1				S2				
	γ	(R_T^2)	YIC	[model]	γ	(R_T^2)	YIC	[model]	
Soil plot	muck interflow	-1.2	(0.80)	-9.4	[1 1 2]	-2.7	(0.65)	-8.4	[1 1 2]
	muck depth runoff	-1.3	(0.77)	-10.0	[1 1 4]	-0.8	(0.88)	-10.5	[1 1 2]
	mor depth runoff	-1.9	(0.76)	-8.5	[1 1 3]	-0.5	(0.74)	-8.7	[1 1 2]
Sub-catchment	N_1	-1.1	(0.89)	-11.9	[1 1 1]	-1.3	(0.94)	-13.2	[1 1 1]
	N_2	-1.1	(0.90)	-12.0	[1 1 1]	-1.8	(0.94)	-13.3	[1 1 1]
	N_3	-0.9	(0.88)	-11.3	[1 1 1]	-1.2	(0.93)	-12.6	[1 1 1]
	ERL	-1.0	(0.88)	-11.6	[1 1 3]	-1.0	(0.87)	-11.2	[1 1 2]

Since it generally resulted in lower values of the YIC, we used the level of the groundwater Z_{gw} to model $b_0(k|N)$. A linear combination of Z_{gw} and θ_1 , obtained by a multiple linear

regression, did not result in a lower value of the YIC, nor did a sequential application of both variables. This is possibly due to the high correlation between the two variables. Table 4.4 lists the results of the powerlaw optimization for all studied areas and time periods.

4.4.4 Runoff from the Soil Plots

Generally, the nonlinear transfer functions performed better for the wetter fall period (S2) than for the spring period (S1). Table 4.5, reports the parameters of $[1 \ 1 \ \delta]$ models, which characterize the runoff of both soil plots best.

In addition to the complete time series S1 and S2, we analysed the hydrological processes using subsets of the data series S1 (S1a, S1b and S1c) and S2 (S2a and S2b) (see Figs. 4.6 and 4.7). Fig. 4.9 presents the resulting nonlinear transfer functions for S1b and S2b. For both soil plots and both time series, the predicted runoff increased sooner after the begin of the rainfall than the observed runoff. In addition, the runoff of small rainfall events was most often overestimated. These findings suggest that the time series were still not completely linear. On the muck humus plot, the peak runoff and the declining limb of the hydrographs were described well, but $[1 \ 1 \ \delta]$ transfer functions failed to describe the depth runoff from the mor humus plot. Thus, the interpretation of the transfer function models for the mor humus plot requires some care.

The optimal time delay δ was slightly smaller in case of the interflow. Still, as for the irrigation experiments on the mor humus plot, the time delay δ was probably not constant during the observation periods. Therefore, the estimated time delays of Table 4.5 are the average δ for all rainfall events within the time series. This nonlinear property, the variable time delay between the beginning of the rainfall and the beginning of the runoff, can not be accounted for by the nonlinear transfer function. The transfer function estimation procedure results in an average δ for the complete time series. It occurs that, according to the YIC criterion, several models with only δ varying are about equally good. The autoregressive parameters vary slightly for those models with δ varying within one unit around the optimal value, but more for δ varying with two or more units. It is therefore important to consequently base the model selection on the YIC.

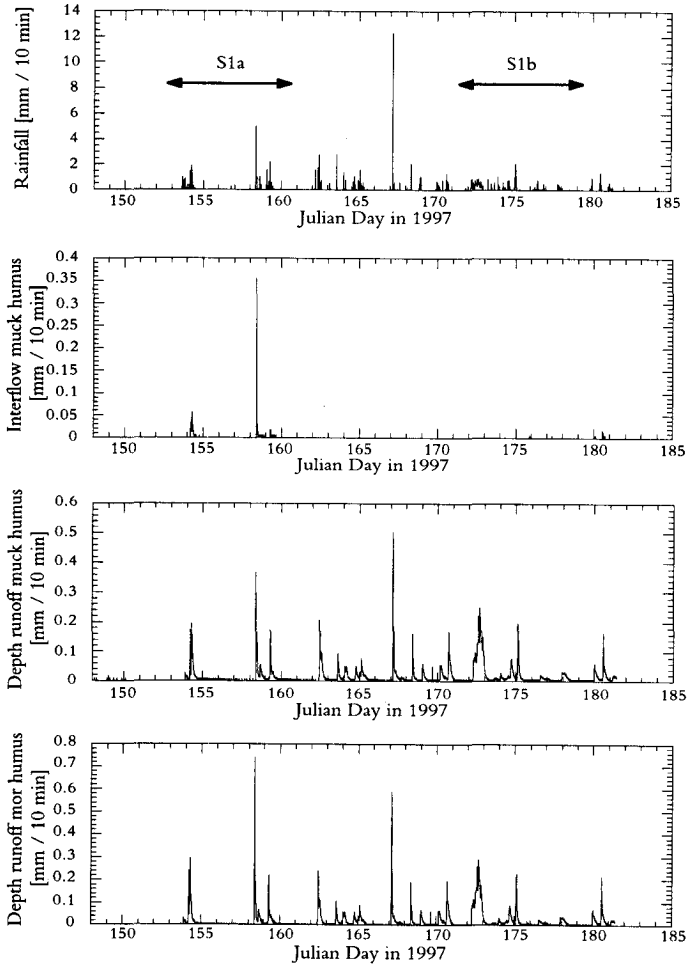


Figure 4.6: Runoff from the soil plots for S1. S1a and S1b denote the subsets studied in detail.

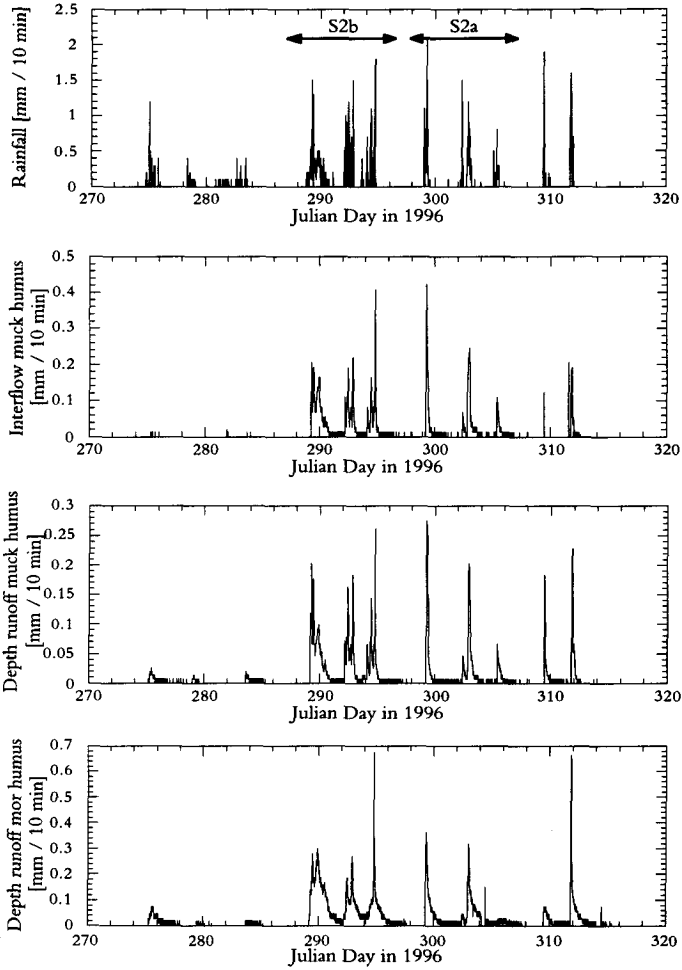


Figure 4.7: *Runoff from the soil plots for S2. S2a and S2b denote the subsets studied in detail.*

The estimated t_c of the depth runoff of the muck humus plot varied between 77.6 and 97.6 minutes (Table 4.5). In the interflow, the t_c were of the same magnitude, but the G_{stat} were larger. The t_c estimated for the depth runoff from the mor humus plot were significantly larger than those of the muck humus plot. In addition, the estimates differed considerably for the series S1 and S2. Since the t_c can be interpreted as an indicator for the average travel time, the water resides longer in the mor humus than in the muck humus plot.

For S2, the combined G_{stat} of the depth runoff and interflow from the muck humus plot was smaller than the G_{stat} of the depth runoff from the mor humus. This agrees with the results of the irrigation experiments and with the runoff ratio under natural rainfall conditions, showing that the total runoff per unit area was larger on the mor humus plot. This result is surprising because the mor humus plot was definitely better drained than the muck humus plot.

During the irrigation experiments on both plots, the soil moisture content was high and the evapotranspiration limited because of the roof. Due to the high irrigation rates, the nonlinear properties of the runoff generation on the muck humus disappeared towards the end of the irrigation experiment. Transforming the time series of water input and subsequently estimating a transfer function provided suitable results. For the natural rainfall events on the muck humus plot, the fit of the transfer functions was slightly worse. Still, it was possible to compare the results of the transfer functions estimated from the irrigation experiments and from the natural rainfall events. On the mor humus plot, the nonlinear dependence of the depth runoff on the rainfall or irrigation was more pronounced and the fit of the transfer function was inferior.

Generally, the time constants t_c of the depth runoff and of the interflow were much shorter for the irrigation experiments than for the natural rainstorms. In addition, the t_c of the interflow and the depth runoff clearly differed for the experiments (faster interflow), but were about the same for the natural events. We think that this is due to sub-surface flow (e.g. rising groundwater or flow through macropores in deep soil layers) entering the plot through the subsoil. Since such fluxes originate from a large volume of water and have longer travel distances than runoff produced on the soil plot, they react more slowly on water inputs and thus

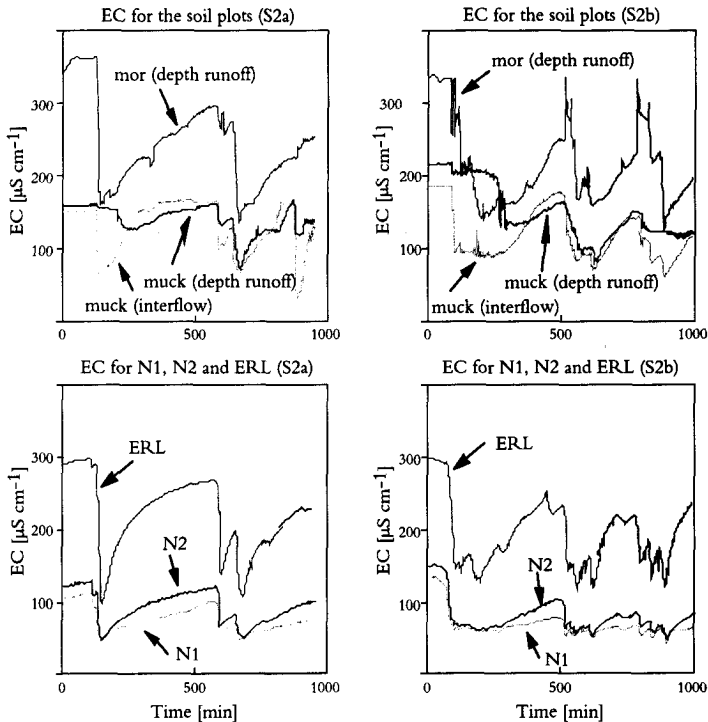


Figure 4.8: Electrical conductivity (EC) for S2a and S2b at all scales. N_1 and N_2 are sub-catchments and ERL denotes the headwater catchment Erlenbach.

result in slower runoff dynamics. If sub-surface fluxes dominate the outflow from a soil plot, it is likely that the t_c of the runoff will be much larger than the t_c detected from the irrigation experiments. Also, the differences in the runoff dynamics from the interflow and the depth runoff from the muck humus will disappear. Both consequences of sub-surface fluxes are obvious from Table 4.5. If the soil moisture state during controlled irrigation experiments and natural rainfall events is comparable, the G_{stat} resulting from the latter will be larger. For the wetter observation period S2, this effect could be observed (cf. Table 4.2 and 4.5).

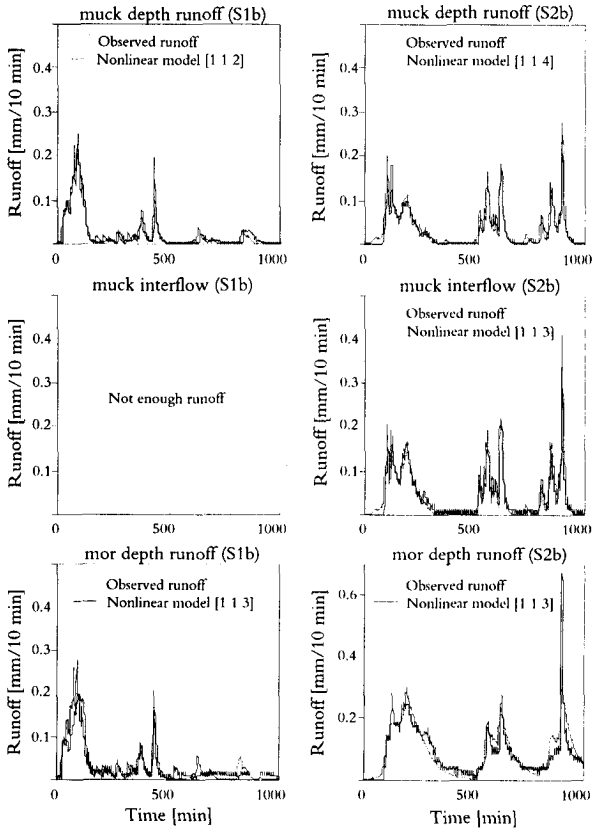


Figure 4.9: Nonlinear transfer functions for the runoff from the soil plots.

Table 4.5: Nonlinear transfer functions for the runoff from the soil plots during natural rainfall events. t_c is the Time Constant and G_{stat} the Steady State Gain of the transfer function. YIC is the Young Information Criterion and R_T^2 the Coefficient of determination of the time series models.

	Time	Model	Parameter	G_{stat}	t_c [min]	R_T^2	YIC	
Depth runoff muck humus	S1a	[1 1 1]	$a_1 -0.8868 \pm 0.0030$ $b_0 0.0133 \pm 0.0003$	0.118	83.3	0.83	-10.10	
	S1b	[1 1 2]	$a_1 -0.9026 \pm 0.0027$ $b_0 0.0159 \pm 0.0004$	0.163	97.6	0.91	-10.39	
	S1	[1 1 2]	$a_1 -0.8705 \pm 0.0037$ $b_0 0.0175 \pm 0.0003$	0.135	72.1	0.77	-9.97	
	S2a	[1 1 3]	$a_1 -0.8791 \pm 0.0028$ $b_0 0.0295 \pm 0.0002$	0.244	77.6	0.91	-10.79	
	S2b	[1 1 4]	$a_1 -0.8932 \pm 0.0028$ $b_0 0.0288 \pm 0.0005$	0.270	88.6	0.90	-10.48	
	S2	[1 1 4]	$a_1 -0.8948 \pm 0.0027$ $b_0 0.0242 \pm 0.0005$	0.230	90.0	0.88	-10.51	
	S1a		Not enough runoff generated					
Interflow muck humus	S1b		Not enough runoff generated					
	S1		Not enough runoff generated					
	S2a	[1 1 1]	$a_1 -0.8821 \pm 0.0028$ $b_0 0.0378 \pm 0.0001$	0.320	79.7	0.91	-10.71	
	S2b	[1 1 3]	$a_1 -0.9025 \pm 0.0001$ $b_0 0.0353 \pm 0.0007$	0.362	97.5	0.91	-11.01	
	S2	[1 1 2]	$a_1 -0.8961 \pm 0.0014$ $b_0 0.0308 \pm 0.0009$	0.362	97.5	0.80	-9.37	
Depth runoff mor humus	S1a		Estimation algorithms did not converge					
	S1b	[1 1 3]	$a_1 -0.9198 \pm 0.0033$ $b_0 0.0136 \pm 0.0005$	0.169	120	0.80	-8.74	
	S1	[1 1 3]	$a_1 -0.9334 \pm 0.0033$ $b_0 0.0114 \pm 0.0040$	0.172	145	0.69	-8.69	
	S2a	[1 1 3]	$a_1 -0.9789 \pm 0.0005$ $b_0 0.0164 \pm 0.0003$	0.775	469	0.87	-10.5	
	S2b	[1 1 3]	$a_1 -0.9832 \pm 0.0004$ $b_0 0.0172 \pm 0.0004$	1.023	590	0.79	-9.82	
	S2	[1 1 3]	$a_1 -0.9588 \pm 0.0024$ $b_0 0.0293 \pm 0.0011$	0.685	241	0.76	-8.54	

The data on the electrical conductivity of the runoff were analysed to determine the origin of the runoff water (cf. Eq. 4.14). In Table 4.6, the analysis of the origin of the runoff is summarized. The percentages of event and pre-event water were obtained by integrating the time series of Q_{old} and Q_{new} (cf. Eq. 4.14). The EC data of the runoff are shown in Fig. 4.8 and Fig. 4.15 presents the corresponding time series of pre-event and event water for S2a and S2b.

Table 4.6: Origin of the runoff water during S2a and S2b at all scales. N_1 and N_2 are sub-catchments and ERL is the Erlenbach headwater catchment. EC is the electrical conductivity ($\mu\text{S cm}^{-1}$) of the rainfall (C_{new}) and of the pre-event soil water (C_{old}). The percentages of old and new water are integrated for the complete subsets S2a and S2b.

Period	EC	Muck humus		Mor humus	N_1	N_2	ERL
		Depth flow	Interflow	Depth flow			
S2a	C_{old}	162	130	300	110	126	294
	C_{new}	10.3	10.3	10.3	10.3	10.3	10.3
	% old	71	43	59	42	40	50
	% new	29	57	41	58	60	50
S2b	C_{old}	205	181	322	135	149	295
	C_{new}	7.2	7.2	7.2	7.2	7.2	7.2
	% old	71	47	54	38	38	48
	% new	29	53	46	62	62	52

At the begin of a rainfall event, water with the chemical signature of pre-event soil water percolated through the clayey subsoil of the muck humus plot. It was only after more than one hour that the EC of this runoff fraction decreased and the dilution with event water was discernible. This resulted in a cumulated amount of 71 percent pre-event water in the depth runoff, for both S2a and S2b. The EC of the interflow of the muck humus plot decreased quickly after the start of the rainfall. Thus, event water contributed more to the runoff in this depth. At the end of the rainfall event, the EC rose more rapidly in the interflow than in the depth runoff. Integrated over the subsets, this results in a larger contribution of event water (57 percent for S2a and 53 percent for S2b) than pre-event water. The different course of the

EC in the interflow and in the depth runoff again suggests that at least part of the runoff flowing from the mor humus is not produced within the soil plot.

The EC of the depth runoff of the mor humus plot is about twice as high as that of the muck humus plot. This can be due to the longer residence time of water in this soil, as indicated by the transfer functions, or by the larger contact area in the better structured mor humus subsoil. This was observed during the tracer experiments of Chapter 3, too: the residence time of the tracers in the mor humus soil was longer than in the muck humus soil. In addition, the relative tracer concentrations in the outflow were lower (= more diluted) in the mor humus soil.

Since no interflow flowed from the mor humus plot, the infiltrated rainfall lowered the EC of the depth runoff mor rapidly than on the muck humus plot. For both subsets S2a and S2b, pre-event water dominated the outflow from the mor humus plot (59 percent pre-event water for S2a, 54 percent for S2b). These values are intermediate between those of the depth runoff and the interflow of the muck humus plot.

4.4.5 Runoff from the Sub-catchments

Table 4.7: Nonlinear transfer functions for the runoff from the sub-catchments N_1 , N_2 and N_3 . t_c is the Time Constant and G_{stat} the Steady State Gain of the transfer function. YIC is the Young Information Criterion and R_T^2 the Coefficient of determination determination of the time series model. The number in brackets is the contribution of the slow and the fast process (derived from [2 2 δ] models) to the total runoff.

Area	Time	Model	t_c [min]		G_{stat}		R_T^2	YIC
			Fast	Slow	Fast	Slow		
N_1	S1a	[1 1 1]	59		0.32		0.83	-9.77
	S1b	[1 1 0]	183		0.40		0.93	-11.09
	S1c	[1 1 0]	109		0.59		0.93	-10.78
	S1	[1 1 1]	83		0.31		0.89	-11.90
	S2a	[2 2 1]	66	550	0.45 (58%)	0.33 (42%)	0.98	-13.58
	S2b	[2 2 2]	52	477	0.42 (49%)	0.44 (51%)	0.96	-11.67
	S2	[1 1 1]	106		0.74		0.94	-13.22
N_2	S1a	[1 1 3]	94		0.56		0.87	-10.60
	S1b	[1 1 1]	258		0.73		0.95	-12.43
	S1c	[1 1 1]	188		0.66		0.93	-10.96
	S1	[1 1 1]	92		0.40		0.90	-11.96
	S2a	[2 2 2]	38	738	0.62 (55%)	0.51 (45%)	0.98	-12.96
	S2b	[2 2 2]	48	475	0.65 (50%)	0.65 (50%)	0.97	-13.07
	S2	[1 1 1]	102		1.03		0.94	-13.28
N_3	S1a	[1 1 1]	62		0.49		0.84	-10.04
	S1b	[1 1 0]	215		0.56		0.95	-11.87
	S1c	[1 1 0]	150		0.48		0.93	-10.82
	S1	[1 1 1]	109		0.45		0.88	-11.29
	S2a	[2 2 1]	46	1020	0.69 (59%)	0.49 (41%)	0.97	-12.62
	S2b	[2 2 2]	34	434	0.60 (51%)	0.58 (49%)	0.95	-10.99
	S2	[1 1 1]	83		0.95		0.92	-12.57

Also at the scale of the sub-catchments, the performance and the results of the transfer function modelling strongly differed between the time series S1 and S2. Whereas for S2a and S2b [2 2 δ] transfer functions described the runoff data best (based on YIC and R_T^2), [1 1 δ]

models agree best with the S1a, S1b and S1c series and the complete S1 and S2 periods. However, Fig. 4.12 and Table 4.7 show that $[1\ 1\ \delta]$ models describe the dynamics of the runoff generation slightly worse than $[2\ 2\ \delta]$ models¹. The question arises why it is not possible to fit $[2\ 2\ \delta]$ models to all time series. In our opinion, this is also caused by the possibly varying time delay between the begin of the rainfall and the begin of the runoff, as discussed before. Transfer functions of structure $[2\ 2\ \delta]$ typically identify two different regions of the declining limb of the hydrograph: the steep upper part (fast process) and the lower, less steep part (slow process). However, to estimate $[2\ 2\ \delta]$ models from data, it is a prerequisite that the timing of the modelled peak runoff (and thus of δ) is appropriate. This is most likely the case for short data series obtained during wet periods.

On average, the time delay δ of the transfer functions is smaller for the sub-catchments than for the soil plots, indicating that the sub-catchments responded faster to rainfall. The runoff dynamics were the same for all sub-catchments. The t_c of the $[1\ 1\ \delta]$ models for the complete data series S1 and S2 are consistent for all the sub-catchments. If we adjust the G_{stat} of N_1 by the corrected surface area of N_1 (cf. section 4.4.2), then the G_{stat} of N_1 and N_2 match well. For the time series S1a, S1b and S1c, the t_c and the G_{stat} varied in a similar fashion for all the sub-catchments and were in the average larger than those obtained for the soil plots. Thus, despite the initially faster reaction of the sub-catchments to rainfall (smaller δ), the runoff dynamics are slower than on the soil plots.

The $[2\ 2\ \delta]$ transfer functions provided more detailed information, showing the existence of fast and slow processes contributing to runoff. As before, all sub-catchments react similarly to rainfall inputs and the t_c of the fast and the slow processes are very similar. For S2a, the t_c of the fast process varied between 38 and 66 minutes, and their contribution to the total runoff was between 55 and 58 percent. The slow process has t_c values between 434 and 1020 minutes. For S2b, the t_c of the parallel flow processes are generally smaller, except for the fast

¹In some cases, the sum of G_{stat} of the fast and slow flow processes is larger than 1. This is an effect of the transformation of the rainfall inputs. For the original data series, G_{stat} approximates the runoff ratio (for the number of data points $N \rightarrow \infty$). By transforming the rainfall inputs, their total amount decreases. As a consequence, the estimated b_0 parameter(s) will be larger and the total G_{stat} might be larger than 1.

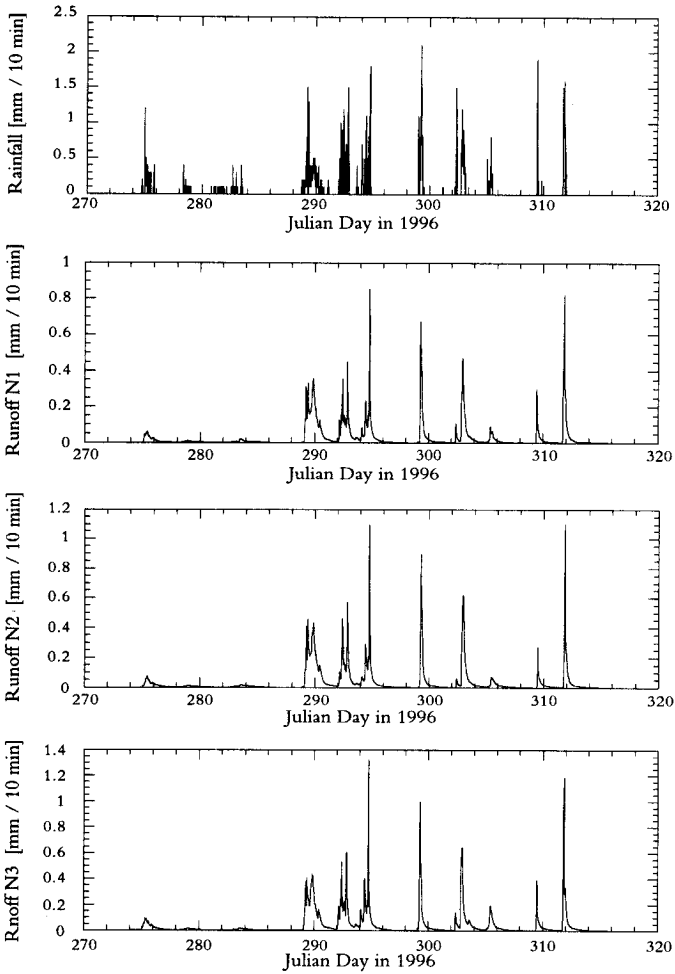


Figure 4.10: *Runoff from the sub-catchments for S2.*

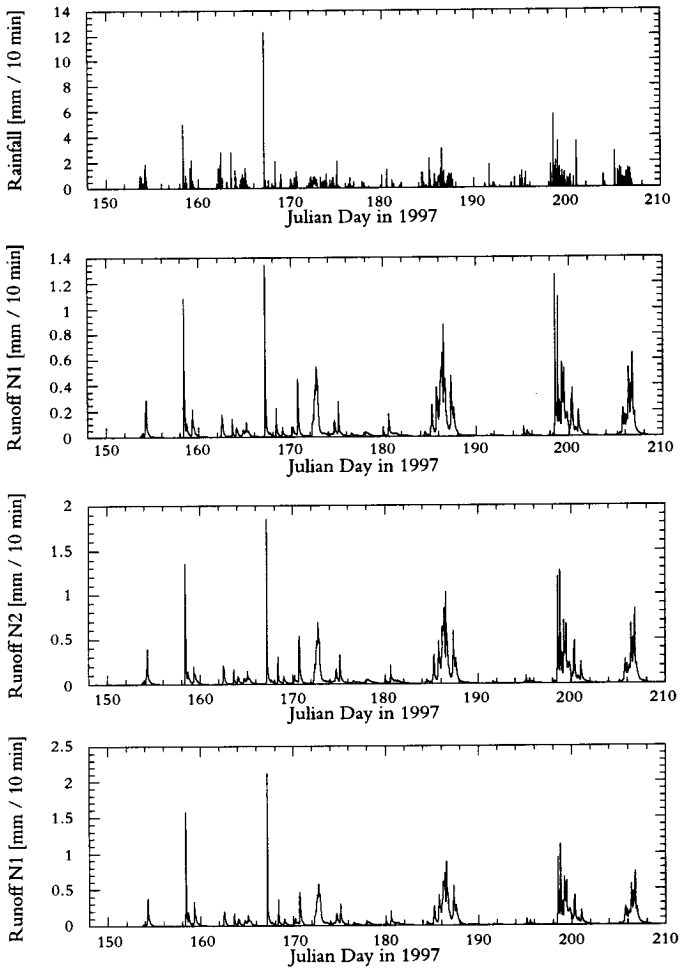


Figure 4.11: Runoff from the sub-catchments during S1.

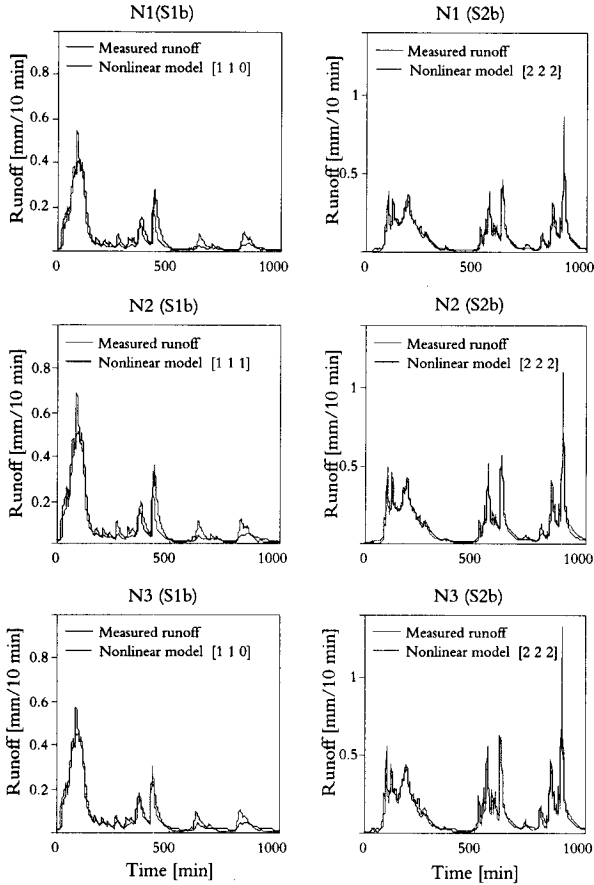


Figure 4.12: Nonlinear transfer functions for the runoff from the sub-catchments N_1 , N_2 and N_3 for the data subsets S1b and S2b.

flow process in N_2 . The contribution of the fast and the slow flow processes to the total runoff for S2b approached 50 percent for all sub-catchments. The time constants of the fast and the slow processes on the sub-catchments approximately matched those of the muck humus plot (fast process) and the mor humus plot (slow process). This corresponds with the

tracer experiments of Chapter 3, where we also concluded that flow processes in the muck humus areas are faster than in the mor humus. In addition, both tracer recovery and runoff ratio were higher on the mor humus plot. Thus, based on these findings we can expect that both soil types contribute to the runoff of the sub-catchments. However, measurements of the EC of the sub-catchment runoff clearly altered this interpretation.

At the begin of a rainfall event, the EC of the runoff from N_1 and N_2 decreased quickly (Fig. 4.8). The minimal value of the EC, about $40 \mu\text{S cm}^{-1}$, was much smaller than that of the soil plots. This is likely due to the effect of saturated areas in the sub-catchments (see Fig. 4.2). With increased rainfall duration, these saturated areas expand. Because of the saturated soil, rainwater can not infiltrate and quickly reaches the outlet of the sub-catchment. As a consequence, the runoff peaks are mainly composed of rainwater.

The absolute values of the EC provide further evidence about the origin of the runoff. For N_1 and N_2 , the EC varied between 40 and $139 \mu\text{S cm}^{-1}$, at the muck humus plots between 50 (interflow) and $205 \mu\text{S cm}^{-1}$ (depth runoff) and on the mor humus plot between 110 and $320 \mu\text{S cm}^{-1}$. Thus, the quality of the runoff from N_1 and N_2 corresponds best to the interflow of the muck humus plot and least to the depth flow of the mor humus. Thus, attributing the fast process of the $[2 \ 2 \ \delta]$ transfer function to the runoff from the muck humus and the slow process to the mor humus plot, as suggested by the results of the transfer function estimation, is incorrect. Rather, both the slow and the fast flow processes on the sub-catchments originate from muck humus zones.

4.4.6 Runoff from the Erlenbach headwater catchment

The nonlinear transfer functions describe the runoff from the Erlenbach catchment slightly worse than that of the sub-catchments. For the complete data series S1 and S2, $[1 \ 1 \ \delta]$ transfer functions performed best, for subsets S2a and S2b $[2 \ 2 \ \delta]$ models. Since δ varied between 2 and 4, the Erlenbach catchments reacted in the average more slowly to precipitation than the sub-catchments. The t_c of the fast and slow flow processes match nicely those obtained for the sub-catchments. The t_c of the slow process varied between 521 and 1163 min, that of the fast processes between 44 and 66 min. Thus, although the headwater catchment is about

Table 4.8: Nonlinear transfer function models for the runoff from the Erlenbach headwater catchment. t_c is the Time Constant and G_{stat} the Steady State Gain of the system. YIC is the Young Information Criterion and R_T^2 the Coefficient of determination of the time series model. The number in brackets is the contribution of the slow and the fast processes (derived from [2 2 δ] models) to the total runoff.

Time	Model	t_c		G_{stat}		R_T^2	YIC
		Fast	Slow	Fast	Slow		
S1a	[1 1 3]	94		0.56		0.87	-10.6
S1b	[1 1 3]	233		0.72		0.95	-12.32
S1c	[2 2 3]	66	719	0.36 (53%)	0.32 (47%)	0.95	-10.94
S1	[1 1 3]	143		0.57		0.88	-11.64
S2a	[2 2 2]	39	1163	0.52 (49%)	0.54 (51%)	0.93	-10.18
S2b	[2 2 4]	44	521	0.48 (44%)	0.60 (56%)	0.92	-10.03
S2	[1 1 2]	145		0.85		0.87	-11.20

460 times larger than the sub-catchments, we identified similar flow processes. Furthermore, the t_c of the fast processes compare well to those of the muck humus plot; the slow processes on the other hand, are in agreement with the depth runoff of the mor humus plot.

Both the results of the transfer function modelling and the EC measurements indicate that there are two components with a distinct origin. At the begin of a rainfall event, the EC of the Erlenbach runoff decreased more slowly than the EC of the sub-catchment runoff. Hence, during the initial phase, both pre-event and event water contributed to the runoff. The EC during low flow was comparable to that of the mor humus, as were the t_c of the slow flow process from Erlenbach and the depth runoff of the mor humus. During peak runoff, when mainly the fast flow process contributed to the runoff, event water dominated catchment runoff. The t_c of the fast flow process correspond well with those of the sub-catchments and of the muck humus plot. Summarized, at the begin of a rainstorm, flow in the subsoil horizons and in the upper layers of muck humus areas contributes to the runoff. Because of the mixing of both runoff sources, the EC in the Erlenbach runoff decreases less fast than on the sub-catchments. If the rainfall continues, the wetland areas start to contribute more to the runoff, enhanced by flow from waterlogged zones. Because of the network of draining trenches, this runoff is routed quickly towards the main brooks. Still, also flow in

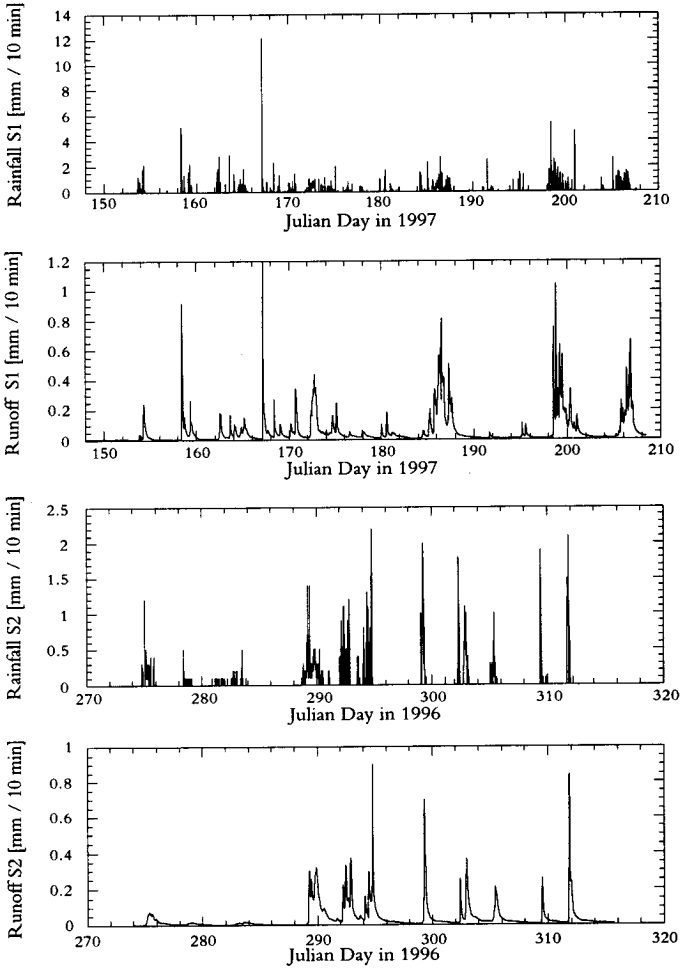


Figure 4.13: Runoff from the Erlenbach headwater catchments for S1 and S2.

the sub-surface leaves the flanks of the gorges and adds to runoff. Therefore, the EC during peak runoff is still much higher than on the sub-catchments. When the rainfall stops, the flow in the sub-surface gains importance and the EC increases again. The amount of pre-event and event water in runoff of the Erlenbach catchment was approximately equal (Table 4.6 and Fig. 4.15).

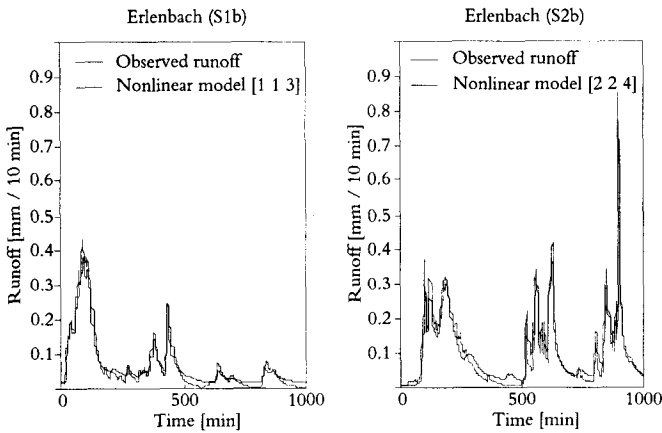


Figure 4.14: Nonlinear transfer functions for the runoff from the Erlenbach headwater catchment.

4.5 Comparison of Scales

An overview of the runoff processes at all scales is presented in Fig. 4.16. Generally, the runoff ratio increased with the size of the drainage area (Table 4.3). We attribute this to flow processes in the subsoil. In the soil plots, the observed level of the groundwater was mostly well above the level of the trenches and both inflow of water from the surroundings and seepage losses from the plot were therefore possible. Although the runoff ratio suggested losses of water, the different t_e during the irrigation experiments and the natural rainstorms

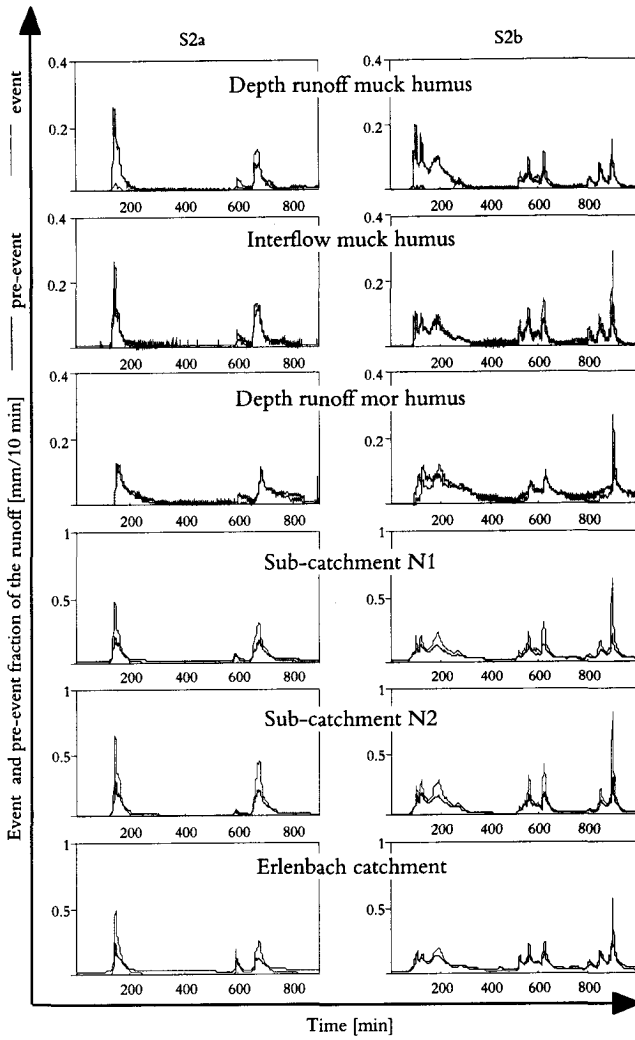


Figure 4.15: Contribution of event and pre-event water to the runoff at all scales for the subsets S2a and S2b.

indicated that water entered the plots through the subsoil. In the sub-catchments too, the groundwater table may occasionally drop below the depth of the trenches. Because of the larger size of the drainage area, the influence of this boundary effect is, however, expected to be smaller than in the soil plots. On the Erlenbach catchment, a water balance deficit of 29 percent during spring and 21 percent during fall did not indicate a significant loss of water as deep seepage. These deficits are in agreement with the long term water balance of the Erlenbach catchment (Burch, 1994) and caused by evapotranspiration.

Irrigation experiments under controlled boundary conditions in both soil plots revealed that runoff generation in all depths was fast (see Table 4.2, Fig. 4.3 and 4.4). The t_c of the depth runoff and the interflow of the muck humus plot were 21 and 6 minutes, respectively, and the G_{stat} equalled 0.19 and 0.17. On the mor humus plot the time delay δ depended on the soil water storage and no good model fit could be obtained. Typically, the runoff ratio was higher on the mor humus than on the muck humus, but the runoff dynamics slower and the residence time of water longer.

The runoff ratio and the G_{stat} of the muck humus plot during the irrigation were intermediate to the results obtained for natural rainstorms during S1 and S2. Independent of the runoff ratio, the t_c of the interflow and the depth runoff of the muck humus plot for natural rainstorms were much larger than those obtained from the irrigation experiment. In addition, for S2 the differences in the t_c of the interflow and the depth runoff disappeared. This was caused by sub-surface fluxes which determined the dynamics of the outflow of the muck humus even if the soil was dry and deep seepage occurred.

The t_c of the depth runoff of the mor humus were larger than those of the muck humus plot (Table 4.6). The slower runoff from the mor humus plot is likely due to the better structured subsoil.

The EC measurements on the soil plots suggest that (cf. Table 4.6, Fig. 4.8 and 4.15): i) the slow decrease of the EC in the depth runoff of both plots shows that at the begin of the rainfall mainly pre-event water, possibly entering the plot through the subsoil, contributed to the depth runoff, ii) in the interflow of the muck humus, rainwater contributed directly to

the runoff, and iii) the residence time of water in the mor humus is larger than in the muck humus. Since no interflow flowed from the mor humus, the depth runoff was diluted more by the rainfall causing a faster decline of the EC.

The t_c of [1 1 δ] transfer functions fitted for the sub-catchments N_1 , N_2 and N_3 were larger than those of the soil plots, but the average time delay δ was smaller (Table 4.7). The faster initial response of the sub-catchments to rainfall is mainly caused by waterlogged areas and wet muck humus depressions near the trenches and the catchment outlet (Fig. 4.2). Rain falling onto these areas may contribute directly to the runoff, causing a rapid decrease of the EC (Fig. 4.8, 4.15). During the course of a rainstorm, the spatial extent of waterlogged areas increases. If the rainfall stops, the contribution of the saturated areas to the runoff quickly vanishes. Such dynamics, which mainly occur in the muck humus depressions, correspond well to the fast process of the [2 2 δ] transfer functions (Table 4.7). Since the EC at peak runoff is very low, the contact of the rainwater with the muck humus layers is minimal. After the rainfall has stopped, the relative contribution of the slow flow processes to the total runoff increases. Comparing the EC of the outflow of the soil plots and of the sub-catchments shows that also this runoff process originates mainly from the upper layers of the muck humus areas. Consequently, flow processes in the subsoil of the mor humus contributed little to the runoff of the sub-catchments.

The characteristics of the runoff from the sub-catchments N_1 , N_2 and N_3 were very similar. Thus, it appears that not the relative abundance of mor and muck humus soil determines the characteristics of the runoff, but rather the spatial arrangement of the muck humus zones and the extent of the areas waterlogged during rainstorms.

The time constants of the Erlenbach headwater catchment runoff were very similar to those of the sub-catchments. This is surprising since the size of the Erlenbach catchment is about 460 times larger. Apart from the spatial distribution of forests and wetland areas, two features of the Erlenbach catchment are of major importance. First, most of the wetland areas are drained by trenches which are directly connected small brooks in deeply incised gorges. Secondly, sub-surface flow leaving the flanks of the gorges may contribute directly to the catchment runoff. At the begin of a rainfall event, the EC of the brook water decreases more

gradually than in the runoff from the sub-catchments. At this stage, both the flow from the wetland areas and that from the sub-surface adds to the runoff and the mixing of both fractions causes the slower decrease of the EC. When the rainstorm continues, flow from the wetland areas contributes more to the runoff. This is favoured by the draining trenches throughout the catchment, routing the runoff to the brooks. Because sub-surface fluxes too add to the total runoff, the EC is not as low as in the sub-catchment runoff. After cessation of the rainfall, the runoff from wetland areas diminishes and sub-surface flow dominates the runoff. The EC rises and reaches the same value as in the depth runoff of the mor humus plot. Because of the drainage trenches runoff produced in the wetland areas reaches the brook quickly so that the t_e of the transfer functions for the headwater catchment and the sub-catchments match well. The slow component of the [2 2 δ] models, as well as the EC, correspond well to the depth runoff of the mor humus plot, too.

4.6 Summary and Conclusions

To support a study on nitrogen cycling of nitrogen in an alpine forest ecosystem, we studied the runoff generation tree spatial scales. Our objective was an improved understanding of the influence of the catchment scale and of the soil and vegetation on the runoff characteristics.

To this end, we used data on the precipitation, runoff and the EC of the runoff water during two observation periods, together with the findings from an irrigation experiment. The largest study area, the Erlenbach headwater catchment (0.7 km²), is a steep and narrow catchment which is drained by permanent brooks. Within this headwater catchment, we delineated three sub-catchments (1500 m²) with distinct vegetation and soil types. Two sub-catchments are covered by forest - they were the subject of a study on nitrogen cycling - and a third is wetland. Within these sub-catchments only two soil types occur, namely a mor and a muck humus soil. We studied runoff generation on two small plots (~13 m²), one representing the muck, the other the mor humus soil. We collected the runoff in three depths and measured the soil moisture regime with a high temporal resolution.

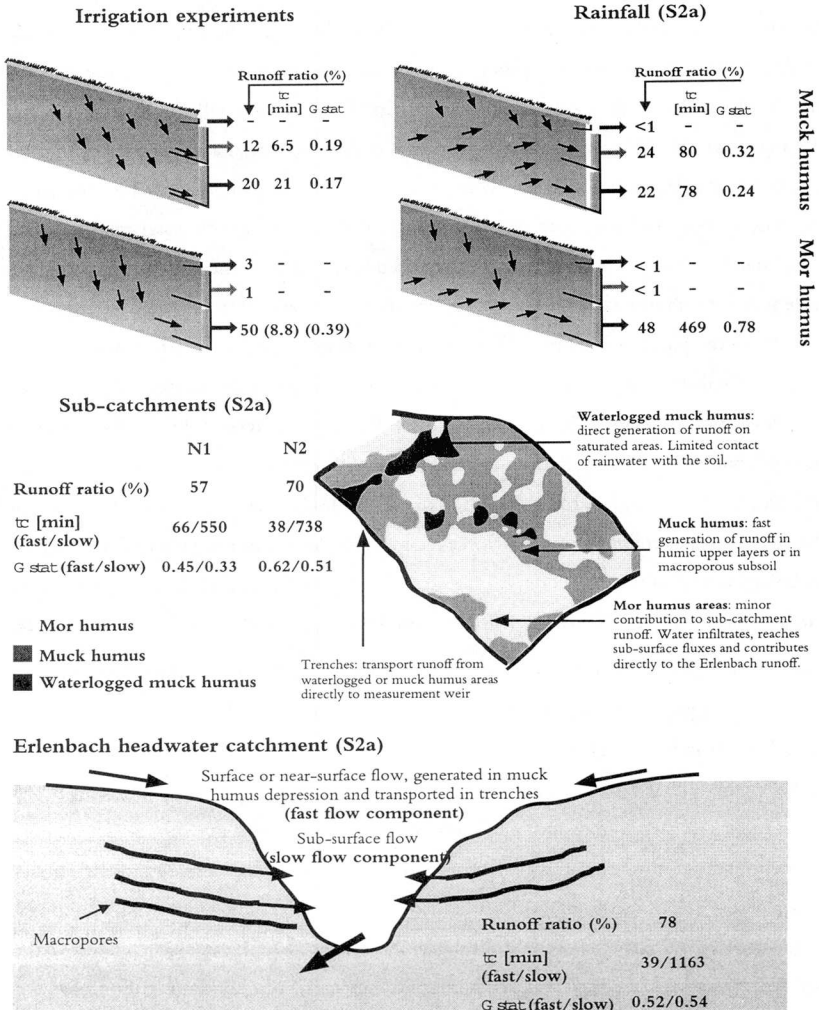


Figure 4.16: Overview of the runoff processes at all scales. The runoff ratio is the ratio of the total amount of runoff to the total amount of water input, t_c is the time constant (min) and G_{stat} the Steady State Gain.

From the rainfall-runoff data, we calculated input-output water balances during two periods, one in fall 1996 and the other in spring 1997 and estimated the contribution of event and pre-event water to the total runoff from the EC data. To model the runoff generation, and to detect the contributing flow processes, we used time series analysis. Since the rainfall-runoff relationship is very often nonlinear, we first linearized the relation by transforming the input variable, using the following procedure. First, we fitted a linear $[1 \ 1 \ \delta]$ transfer function to the data with parameters which were allowed to vary in time. Then we compared the time varying gain parameter to additional data characterizing the water regime of the soils. Since the level of the groundwater correlated best with the time variable gain, this variable served to transform the measured precipitation nonlinearly into an effective precipitation. From the linearized the rainfall-runoff data we estimated linear transfer functions with constant parameters. To avoid overparameterization we selected the model order based on an information criterion. The structure and the fitted parameters of the transfer function models could be physically interpreted and provided information on the dynamics of the hydrological processes. The main drawback of this method is the incomplete description of the nonlinear dependence of runoff on precipitation. Although the effect of the varying hydrological gain of the system can be corrected by a transformation of the precipitation, the varying start of the runoff after rainfall could not be adequately described.

The runoff ratio of both soil plots showed that a considerable fraction of the water input was lost as deep seepage. Surprisingly, the runoff ratio of the wetter muck humus plot was smaller than that of the better drained mor humus plot. Irrigating resulted in fast runoff processes on both plots. The dynamics of the interflow of the muck humus plot were much faster and the influence of the groundwater more pronounced than in the depth runoff. Especially in the depth runoff, flow through large macropores contributed substantially to the runoff. Compared to the muck humus, the sub-surface horizons of the mor humus plot are much better structured, causing slower runoff dynamics. Under natural boundary conditions the difference between interflow and depth runoff disappeared on the muck humus plot. On both plots, the runoff dynamics were clearly slower than during the irrigation experiments. This was due to sub-surface water flow entering the soil plots through the subsoil.

Flow through the sub-surface horizons of the mor humus areas was of a minor importance for the runoff generation from the sub-catchments. Despite the different vegetation and soil, the runoff dynamics were very similar in all sub-catchments. They reacted faster to rainfall than the smaller soil plots, which is likely due to the presence of waterlogged areas. Rain falling onto waterlogged areas cannot infiltrate and contributes directly to the runoff. The electrical conductivity indicated that, particularly during peak runoff, the mixing of event with pre-event water was limited. The results of the transfer function estimation suggest that the sub-catchment runoff is composed of flow in the muck humus areas (fast flow process) and in the subsoil of the mor humus (slow flow process). Still, the EC measurements disprove this interpretation and strongly suggest that most of the runoff from the sub-catchments originates from near surface muck humus soil layers or waterlogged zones. Thus, extrapolating the results obtained in the soil plots did not clarify the runoff processes observed in the sub-catchments. Identically, expanding the findings for the sub-catchments to the Erlenbach headwater catchment may not result in adequate runoff predictions. Although the runoff dynamics are very similar on both scales, partly caused by the network of drainage trenches, the EC of the Erlenbach runoff is much higher. During low flow conditions, the EC compares well to that of the depth runoff of the mor humus plot. During peak runoff the EC is still much higher than on the sub-catchments. We conclude that the runoff of the Erlenbach is composed of water flowing from the wetland areas (fast flow process) and water which flows in the sub-surface soil layers (slow flow process).

Leer - Vide - Empty

Chapter 5

Concluding Remarks

Determined by the structure of the dissertation, Chapters 3 and 4 presented their own conclusions drawn from the specific material given therein. This chapter synthesizes the results and comments the relevance for the NITREX project. I used two independent approaches to identify the dynamics of the governing runoff processes and the origin of the runoff water on the soil plots (13 m^2), the sub-catchments (1500 m^2) and the Erlenbach headwater catchment (0.7 km^2). First, discrete transfer functions estimated from observed data served to detect the nonlinear characteristics of runoff generation. Second, I used tracer experiments on the soil plots and electrical conductivity measurements at all the scales to locate the runoff processes in the studied catchments. The synthesis of the results obtained by the two methods led to the conclusions presented in this chapter.

Irrigation and tracer experiments under controlled boundary conditions showed that on the isolated soil plots different runoff processes prevailed. On the better drained mor humus plot, the recovery of tracer and irrigation water was higher than on the wetter muck humus plot. From the muck humus plot, 32 percent of the irrigation water flowed into the runoff gutters (54 percent from the mor humus plot) and, on average, the tracer recovery on the muck humus plot was 33 percent (39 percent on the mor humus plot). This matches with the runoff ratio of the soil plots during natural rainstorms, which was higher for the mor humus

plot. Tracers injected into the subsoil appeared in the outflow after 18 min (muck humus) and 70 min (mor humus). This implies that the macropores in the subsoil of both plots were important flowpaths. Also from the results of the irrigation experiments of chapter 4 could be concluded that preferential flowpaths contributed to the runoff generation.

The transfer functions, estimated from the irrigation experiment and from the natural rainstorms, indicate slower runoff generation on the mor humus than on the muck humus plot. The slower dynamics in the mor humus soil are very likely due to the effect of the more pronounced network of small pores in the subsoil. The tracer experiments support these findings. Tracers, either applied onto the surface or at the interface of the humus layers and the clayey subsoil broke through much faster on the muck than on the mor humus plot.

The monitoring of natural rainfall events revealed that both soil plots were affected by sub-surface fluxes, too. In general, the dynamics of the runoff processes on both soil plots were slower during natural rainfall events than during the irrigation experiments. On the muck humus plot, sub-surface flow reduced the difference between the dynamics of the depth runoff and the interflow, detected during the irrigation experiments. Data on the electrical conductivity of the outflow from both plots indicated that at the beginning of a rainstorm mainly pre-event water flowed from the soil plots. This agrees well with the notion of sub-surface (pre-event) water from upslope areas which enters the soil plots. The first initial hypothesis can be adapted as follows

Conclusion 1

Due to differences in the structure of the subsoil, the flow processes on the muck humus plot much are faster than on the mor humus plot. Despite the lower soil water content and the rainfall interception by vegetation, the runoff ratio of the mor humus plot is higher. There is a clear influence of sub-surface water flow, originating from upslope areas, which enters the soil plots.

By transferring the information from the runoff processes on the soil plots to the sub-catchments, we would hypothesize that the major (but slowest) component of the sub-catchment runoff would flow through the porous clayey subsoil of the mor humus. The fast component of the runoff would be generated in the muck humus depressions. Data on the EC of the runoff of the sub-catchments, however, completely altered this view.

Despite the differences in the distribution of the soil type, the generation of runoff was comparably fast on all sub-catchments. The spatial distribution of the wetter muck humus zones and the waterlogged areas appeared to be more important than the area covered by each of the soil types. Waterlogged areas in the muck humus depressions caused surface or near-surface runoff, responsible for the very low electrical conductivity during runoff peaks. The dynamics of fast and slow flow processes in the sub-catchments, identified by estimating transfer functions, agreed well with those of the soil plots, suggesting faster generation of runoff process in the muck humus and slower generation in the mor humus soil. Electrical conductivity measurements, however, revealed that the largest part of the runoff of the sub-catchments had a limited contact time with the soil and originated from surface or near-surface muck humus zones. Consequently, our initial hypothesis can be refined as

Conclusion 2

Surface and near-surface flow in muck humus soil layers contributes most to the runoff of the sub-catchments. Waterlogged areas in the muck humus depressions cause a direct contribution of rainfall to the runoff. Consequently, the residence time of rainwater and solutes in the sub-catchments is short.

Conclusion 3

Despite the different abundance of mor and muck humus areas within the three sub-catchments, the runoff processes are almost similar. Thus, the spatial distribution of muck humus and waterlogged areas influences the runoff generation more strongly than the area covered by each of the soil types.

By extrapolating this information to the Erlenbach catchment we would predict slower flow processes at this scale because of the longer travel distance, but the runoff would still be dominated by water flowing from the muck humus layers.

Despite the larger area of the headwater catchment, the dynamics of the runoff processes resembled those of the sub-catchments closely. The nonlinear transfer functions and the measurements of the electrical conductivity indicate that two components add to the runoff. The slow component flows mainly in the subsoil, having an electrical conductivity comparable to that of the depth runoff of the mor humus soil plot. This component contributes directly to the brook through the flanks of the gorges. The fast component originates from wet or even waterlogged muck humus areas and is led to the brooks by the drainage trenches.

Conclusion 4

Erlenbach runoff is composed of two parallel flow processes: the first is surface or near-surface runoff from wetland areas (fast process) and the second is sub-surface flow in the macroporous subsoil (slow process).

At the beginning of the NITREX project, the subsoil was assumed to be impermeable. The water balances of the sub-catchments support this hypothesis. But, the different origin of the runoff of the sub-catchments and of the Erlenbach catchment casts some doubt on this assumption. In the sub-catchments, flow in the subsoil contributes little to the runoff. Rain-water falling on the trees and infiltrating into the mor humus has therefore little importance for the runoff of the sub-catchments, but flows through the subsoil to the brooks of the Erlenbach catchment. Thus, the trenches surrounding the sub-catchments might not capture all of the runoff leaving (or entering) through the subsoil. This limits the possibility to extrapolate the findings of NITREX to 'natural' catchments. On the other hand, the findings about runoff generation in the sub-catchments suggest, that the runoff of the sub-catchments is mainly generated in surface or in near-surface muck humus zones. This, off course, agrees to the regions where most of the nitrogen transformation occurs. Thus, it is possible that the NITREX programme, which focuses on the nitrogen turnover at the spatial scale of the sub-catchments, captures the most important nitrogen fluxes.

Preliminary results on the export of ammonium and nitrate showed that the leaching of inorganic nitrogen occurred mainly as nitrate (Schleppi *et al.*, 1998). This agrees well with the runoff characteristics of the sub-catchments. The fast flow processes in the muck humus depressions and on waterlogged areas cause relatively short residence times of water and solutes in the sub-catchments. As a result, in the muck humus areas an incomplete microbial immobilisation of nitrate may occur. Still, the residence time of the rainwater in the soil may be large enough for ammonium to be removed from the soil solution by anion exchange.

However, detailed information on the runoff generation in the sub-catchments and on the permeability of the subsoil can only be obtained from a multiple tracer experiment with inert tracers applied at various locations within and near the sub-catchments.

Appendix A

A.1 Recursive Formulation of the Least Squares Solution

To derive a recursive algorithm, we first set the initial recursive solution after k observations equal to the 'en bloc' solution, or $\hat{\mathbf{a}}^{(k)} = \hat{\mathbf{a}}_k$. We can write the equation (2.29) as

$$\hat{\mathbf{a}}_k = \left[\sum_{i=1}^k \mathbf{z}_i \mathbf{z}_i^T \right]^{-1} \sum_{i=1}^k \mathbf{z}_i y_i = \mathbf{P}_k \mathbf{F}_k, \quad (\text{A.1.1})$$

with

$$\mathbf{P}_k = \left[\sum_{i=1}^k \mathbf{z}_i \mathbf{z}_i^T \right]^{-1} \quad (\text{A.1.2})$$

and

$$\mathbf{F}_k = \sum_{i=1}^k \mathbf{z}_i y_i. \quad (\text{A.1.3})$$

A recursive expression for \mathbf{P}_k^{-1} is obtained by splitting the total sum of Eq. (A.1.2) as follows

$$\mathbf{P}_k^{-1} = \left[\sum_{i=1}^{k-1} \mathbf{z}_i \mathbf{z}_i^T \right] + \mathbf{z}_k \mathbf{z}_k^T \quad (\text{A.1.4})$$

$$\mathbf{P}_k^{-1} = \mathbf{P}_{k-1}^{-1} + \mathbf{z}_k \mathbf{z}_k^T. \quad (\text{A.1.5})$$

Equation (A.1.5) can be pre-multiplied by \mathbf{P}_k and then post-multiplied by \mathbf{P}_{k-1} to give

$$\mathbf{P}_{k-1} = \mathbf{P}_k + \mathbf{P}_k \mathbf{z}_k \mathbf{z}_k^T \mathbf{P}_{k-1}. \quad (\text{A.1.6})$$

By post-multiplying Eq. (A.1.6) by \mathbf{z}_k and rearranging the terms one finds

$$\mathbf{P}_{k-1} \mathbf{z}_k = \mathbf{P}_k \mathbf{z}_k \left[\mathbf{1} + \mathbf{z}_k^T \mathbf{P}_{k-1} \mathbf{z}_k \right]. \quad (\text{A.1.7})$$

Again, Eq. (A.1.7) can be postmultiplied by $\left[\mathbf{1} + \mathbf{z}_k^T \mathbf{P}_{k-1} \mathbf{z}_k \right]^{-1} \mathbf{z}_k^T \mathbf{P}_{k-1}$ to obtain

$$\mathbf{P}_{k-1} \mathbf{z}_k \left[\mathbf{1} + \mathbf{z}_k^T \mathbf{P}_{k-1} \mathbf{z}_k \right]^{-1} \mathbf{z}_k^T \mathbf{P}_{k-1} = \mathbf{P}_k \mathbf{z}_k \mathbf{z}_k^T \mathbf{P}_{k-1}. \quad (\text{A.1.8})$$

By substituting Eq. (A.1.6) into Eq. (A.1.8) the final recursive expression for \mathbf{P}_{k-1} can be obtained

$$\mathbf{P}_k = \mathbf{P}_{k-1} - \mathbf{P}_{k-1} \mathbf{z}_k \left[\mathbf{1} + \mathbf{z}_k^T \mathbf{P}_{k-1} \mathbf{z}_k \right]^{-1} \mathbf{z}_k^T \mathbf{P}_{k-1}. \quad (\text{A.1.9})$$

In a similar way, \mathbf{F}_k can be expressed as

$$\mathbf{F}_k = \mathbf{F}_{k-1} + \mathbf{z}_k \mathbf{y}_k. \quad (\text{A.1.10})$$

Combining Eqs. (A.1.1, A.1.9 and A.1.10) gives the recursive form of the least squares estimation:

$$\hat{\mathbf{a}}_k = \left[\mathbf{P}_{k-1} - \mathbf{P}_{k-1} \mathbf{z}_k \left[\mathbf{1} + \mathbf{z}_k^T \mathbf{P}_{k-1} \mathbf{z}_k \right]^{-1} \mathbf{z}_k^T \mathbf{P}_{k-1} \right] \left[\mathbf{F}_{k-1} + \mathbf{z}_k \mathbf{y}_k \right], \quad (\text{A.1.11})$$

or, after expanding the product

$$\begin{aligned} \hat{\mathbf{a}}_k &= \mathbf{P}_{k-1} \mathbf{F}_{k-1} - \mathbf{P}_{k-1} \mathbf{z}_k \left[\mathbf{1} + \mathbf{z}_k^T \mathbf{P}_{k-1} \mathbf{z}_k \right]^{-1} \mathbf{z}_k^T \mathbf{P}_{k-1} \mathbf{F}_{k-1} + \mathbf{P}_{k-1} \mathbf{z}_k \mathbf{y}_k \\ &\quad - \mathbf{P}_{k-1} \mathbf{z}_k \left[\mathbf{1} + \mathbf{z}_k^T \mathbf{P}_{k-1} \mathbf{z}_k \right]^{-1} \mathbf{z}_k^T \mathbf{P}_{k-1} \mathbf{z}_k \mathbf{y}_k \end{aligned} \quad (\text{A.1.12})$$

$$\begin{aligned} \hat{\mathbf{a}}_k &= \hat{\mathbf{a}}_{k-1} + \mathbf{P}_{k-1} \mathbf{z}_k \left[\mathbf{1} + \mathbf{z}_k^T \mathbf{P}_{k-1} \mathbf{z}_k \right]^{-1} \mathbf{z}_k^T \hat{\mathbf{a}}_{k-1} + \mathbf{P}_{k-1} \mathbf{z}_k \mathbf{y}_k \\ &\quad - \mathbf{P}_{k-1} \mathbf{z}_k \left[\mathbf{1} + \mathbf{z}_k^T \mathbf{P}_{k-1} \mathbf{z}_k \right]^{-1} \mathbf{z}_k^T \mathbf{P}_{k-1} \mathbf{z}_k \mathbf{y}_k \end{aligned}, \quad (\text{A.1.13})$$

or

$$\hat{\mathbf{a}}_k = \hat{\mathbf{a}}_{k-1} - \mathbf{K}_k \left[\mathbf{z}_k^T \hat{\mathbf{a}}_{k-1} + \mathbf{z}_k^T \mathbf{P}_{k-1} \mathbf{z}_k y_k - \left[\mathbf{1} + \mathbf{z}_k^T \mathbf{P}_{k-1} \mathbf{z}_k \right] y_k \right], \quad (\text{A.1.14})$$

with

$$\mathbf{K}_k = \mathbf{P}_{k-1} \mathbf{z}_k \left[\mathbf{1} + \mathbf{z}_k^T \mathbf{P}_{k-1} \mathbf{z}_k \right]^{-1}. \quad (\text{A.1.15})$$

Thus, we can write

$$\hat{\mathbf{a}}_k = \hat{\mathbf{a}}_{k-1} - \mathbf{K}_k \left[y_k - \mathbf{z}_k^T \hat{\mathbf{a}}_{k-1} \right]. \quad (\text{A.1.16})$$

Equation (A.1.16) is the recursive solution to the estimation problem.

A.2 Derivation of the Kalman Filter

The best estimate of the state vector is denoted as $\hat{\mathbf{x}}_{k|k}$ and is defined as the conditional expectation of \mathbf{x}_k given \mathbf{U}_{k-1} and \mathbf{Y}_k , i.e.

$$\hat{\mathbf{x}}_{k|k} = \mathbb{E} \left[\mathbf{x}_k | (\mathbf{Y}_k, \mathbf{U}_{k-1}) \right]. \quad (\text{A.2.1})$$

Before proceeding with the derivation of the Kalman filter, a few assumptions on the noise and state vector are necessary.

a) \mathbf{x}_0 is a Gaussian random vector with known mean and covariance matrix:

$$\mathbb{E}[\mathbf{x}_0] = \bar{\mathbf{x}}_0$$

$$\text{Cov}[\mathbf{x}_0, \mathbf{x}_0] = \Sigma_0$$

This assumption defines the initial condition of the state vector at time $k = 0$, before the Kalman filter estimation procedure starts.

b) \mathbf{w}_k is a Gaussian random vector with zero mean and serially uncorrelated

$$\mathbb{E}[\mathbf{w}_k] = 0$$

$$\text{Cov}[\mathbf{w}_k, \mathbf{w}_\tau] = \mathbf{Q}_k \delta_{k\tau}$$

with \mathbf{Q}_k a known positive semi-definite covariance of \mathbf{w}_k and $\delta_{k\tau}$ the Kronecker delta.

c) \mathbf{v}_k is a Gaussian random vector with zero mean and serially uncorrelated

$$\mathbb{E}[\mathbf{v}_k] = 0$$

$$\text{Cov}[\mathbf{v}_k, \mathbf{v}_\tau] = \mathbf{R}_k \delta_{k\tau}$$

with \mathbf{R}_k a known positive semi-definite covariance matrix of \mathbf{v}_k and $\delta_{k\tau}$ the Kronecker delta.

d) $\mathbf{v}_k, \mathbf{w}_k, \mathbf{x}_0$ and \mathbf{v}_τ are mutually independent for all values of k and τ .

e) \mathbf{U}_{k-1} and all matrices $\mathbf{L}_k, \Phi_{k-1}, \mathbf{G}_{k-1}$ and \mathbf{H}_k are deterministic, noise free vectors or matrices.

Since the system and the observation equation are linear, and because of the first four assumptions, the conditional probability density of $\hat{\mathbf{x}}_{k|k}, p[\mathbf{x}_k | (\mathbf{Y}_k, \mathbf{U}_{k-1})]$, is also Gaussian. Thus, it is completely characterized by its conditional expectation $\hat{\mathbf{x}}_{k|k}$ and by the conditional covariance matrix $\Sigma_{k|k} = \text{Cov}[\mathbf{x}_k^T, \mathbf{x}_k | (\mathbf{Y}_k, \mathbf{U}_{k-1})]$.

The Kalman filter is a powerful discrete algorithm applied to compute an estimate of $\hat{\mathbf{x}}_{k|k}$. The Kalman filter approaches the estimation problem with a two-stage solution. To derive the Kalman filter equations, we assume, without loss of generality, that the system noise transition matrix \mathbf{G}_{k-1} equals the identity matrix of the right dimensions. To avoid a growing memory filter, the estimate is sought for in a linear recursive form

$$\hat{\mathbf{x}}_{k|k} = \mathbf{K}_k^* \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k \mathbf{Y}_k, \quad (\text{A.2.2})$$

where \mathbf{K}_k^* and \mathbf{K}_k are time varying weighting matrices and $\hat{\mathbf{x}}_{k|k-1}$ a prior estimate of the system state vector at time instant k , based on all observations up to $k-1$. Because of the second assumption, the prior estimate of the system state can easily be obtained and by extrapolating the system state at $k-1$ as

$$\hat{\mathbf{x}}_{k|k-1} = \Phi_{k-1} \hat{\mathbf{x}}_{k-1|k-1} + \mathbf{L}_{k-1} \mathbf{U}_{k-1}. \quad (\text{A.2.3})$$

Equations (A.2.2) and (A.2.3) show the multiple stage properties of the Kalman filter: first, the system state is extrapolated forward in time (Eq. A.2.3), afterwards this extrapolation is updated (Eq. A.2.2) with the newly available data at time instant k . The Kalman filter equations can be derived by optimizing the linear estimator (Eq. A.2.2) for \mathbf{K}_k^* and \mathbf{K}_k . The optimality criteria are the unbiasedness of the estimator and the minimum variance of the estimation error. If the assumptions on the measurement noise and the system noise cohere

with reality, the Kalman filter will be *the* optimal filter. The estimation errors $\tilde{\mathbf{x}}_{k|k}$ and $\tilde{\mathbf{x}}_{k|k-1}$ are defined as

$$\tilde{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k} - \mathbf{x}_k \quad (\text{A.2.4})$$

$$\tilde{\mathbf{x}}_{k|k-1} = \hat{\mathbf{x}}_{k|k-1} - \mathbf{x}_k. \quad (\text{A.2.5})$$

Combining Eqs. (A.2.4), (A.2.5) and (A.2.2) results in

$$\tilde{\mathbf{x}}_{k|k} = \left[\mathbf{K}_k^* + \mathbf{K}_k \mathbf{H}_k - \mathbf{I} \right] \mathbf{x}_k + \mathbf{K}_k^* \tilde{\mathbf{x}}_{k|k-1} + \mathbf{K}_k \mathbf{v}_k. \quad (\text{A.2.6})$$

Since by definition $E[\mathbf{v}_k] = 0$ and if $E[\tilde{\mathbf{x}}_{k|k-1}] = 0$, this estimator will be unbiased for any given state vector only if the term in brackets is zero, or

$$\mathbf{K}_k^* = \mathbf{I} - \mathbf{K}_k \mathbf{H}_k \quad (\text{A.2.7})$$

Equation (A.2.2) can now be written as a function of \mathbf{K}_k , the Kalman gain matrix, as

$$\hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k \left[\mathbf{Y}_k - \mathbf{H}_k \hat{\mathbf{x}}_{k|k-1} \right], \quad (\text{A.2.8})$$

with the corresponding estimation error

$$\tilde{\mathbf{x}}_{k|k} = \tilde{\mathbf{x}}_{k|k-1} + \mathbf{K}_k \left[\mathbf{v}_k - \mathbf{H}_k \tilde{\mathbf{x}}_{k|k-1} \right]. \quad (\text{A.2.9})$$

Equation (A.2.9) can be used to calculate the covariance matrix of the estimation error of the system state, which is defined as

$$\mathbf{P}_{k|k} = \text{Cov} \left[\tilde{\mathbf{x}}_{k|k}, \tilde{\mathbf{x}}_{k|k} \right] = E \left[\tilde{\mathbf{x}}_{k|k} \tilde{\mathbf{x}}_{k|k}^T \right]. \quad (\text{A.2.10})$$

Also the covariance matrix of the estimation errors can be extrapolated between two measurements as (prove not shown)

$$\mathbf{P}_{k|k-1} = \phi_{k-1} \mathbf{P}_{k-1|k-1} \phi_{k-1}^T + \mathbf{Q}_{k-1}, \quad (\text{A.2.11})$$

and updated when a new measurement is employed as

$$\mathbf{P}_{k|k} = (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k) \mathbf{P}_{k|k-1} (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k)^T + \mathbf{K}_k \mathbf{R}_k \mathbf{K}_k^T. \quad (\text{A.2.12})$$

An optimal expression for \mathbf{K}_k can be determined by choosing and minimizing an appropriate cost function. This can be achieved by choosing the cost function to minimize as a weighted scalar sum of the diagonal elements of the covariance matrix of the estimation errors $\mathbf{P}_{k|k}$

$$J_k = \|\tilde{\mathbf{x}}_{k|k}\|_{\mathbf{S}}^2 = \mathbb{E}[\tilde{\mathbf{x}}_{k|k}^T \mathbf{S} \tilde{\mathbf{x}}_{k|k}], \quad (\text{A.2.13})$$

with \mathbf{S} being a symmetric and positive semi-definite matrix of order n . Taking the first derivative of the cost function Eq. (A.2.13) with respect to \mathbf{K}_k and comparing it to zero results in the following expression for \mathbf{K}_k

$$\mathbf{K}_k = \mathbf{P}_{k|k-1} \mathbf{H}_k^T \left[\mathbf{H}_k \mathbf{P}_{k|k-1} \mathbf{H}_k^T + \mathbf{R}_k \right]^{-1} \quad (\text{A.2.14})$$

A disadvantage of the calculation of the Kalman gain matrix \mathbf{K}_k is the matrix inversion. The computational efficiency of algorithm (A.2.14) would increase considerably if it could be rewritten avoiding the matrix inversion. By using the matrix inversion relationship

$$\mathbf{P}_{k|k} = \mathbf{P}_{k|k-1}^{-1} + \mathbf{H}_k^T \mathbf{R}_k^{-1} \mathbf{H}_k, \quad (\text{A.2.15})$$

combined with Eq. (A.2.14) and after expanding and rearranging terms, an alternative expression results

$$\mathbf{K}_k = \mathbf{P}_{k|k} \mathbf{H}_k^T \mathbf{R}_k^{-1}. \quad (\text{A.2.16})$$

The complete two stage Kalman solution to the estimation problem now takes the form:

- Extrapolation

$$\hat{\mathbf{x}}_{k|k-1} = \phi_{k-1} \hat{\mathbf{x}}_{k-1|k-1} + \mathbf{L}_{k-1} \mathbf{U}_{k-1} \quad (\text{A.2.17})$$

$$\mathbf{P}_{k|k-1} = \phi_{k-1} \mathbf{P}_{k-1|k-1} \phi_{k-1}^T + \mathbf{Q}_{k-1} \quad (\text{A.2.18})$$

- Updating

$$\hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k \left[\mathbf{Y}_k - \mathbf{H}_k \hat{\mathbf{x}}_{k|k-1} \right] \quad (\text{A.2.19})$$

$$\mathbf{P}_{k|k} = [\mathbf{I} - \mathbf{K}_k \mathbf{H}_k] \mathbf{P}_{k|k-1}. \quad (\text{A.2.20})$$

Equation (A.2.20) can be obtained by substituting Eq. (A.2.14) into Eq. (A.2.12).

A.3 Gauss-Markov sequences

If no physical information on the dynamics of the system state is available, it still can be explicitly modelled by assuming that the state vector evolves in a fairly general stochastic manner. It is possible to assume that the state vector at time k is a realization of a random process. To completely describe the time series of the state vector \mathbf{x}_k for $k = 1$ to N , the joint probability density function of all state vectors needs to be specified

$$p(\mathbf{x}_1 \dots \mathbf{x}_N). \quad (\text{A.3.1})$$

However, since in almost all applications the state vector at all time instants is estimated from only one realization of the input and output data series, this huge amount of information is not available. We can simplify the joint probability density function by assuming that the temporal variation of the state vector is a Markov random sequence:

$$p(\mathbf{x}_k | \mathbf{x}_{k-1} \dots \mathbf{x}_1) = p(\mathbf{x}_k | \mathbf{x}_{k-1}). \quad (\text{A.3.2})$$

This means that the probability density function of \mathbf{x}_k depends on deterministic or stochastic knowledge of \mathbf{x}_{k-1} only. A Gauss-Markov random sequence can be obtained with the additional requirement that $p(\mathbf{x}_k)$ and $p(\mathbf{x}_{k-1})$ have Gaussian probability distributions. The variation of the state vector can now be described by a general Gauss-Markov stochastic difference equation of the form

$$\mathbf{x}_k = \phi_{k-1} \mathbf{x}_{k-1} + \mathbf{G}_{k-1} \mathbf{w}_{k-1}. \quad (\text{A.3.3})$$

The simplest and probably best known form of this stochastic difference equation is a random walk process:

$$\mathbf{x}_k = \mathbf{x}_{k-1} + \mathbf{w}_{k-1}. \quad (\text{A.3.4})$$

The random walk model can be slightly changed to obtain an integrated random walk process, a smoother model for the parameter variation.

$$\mathbf{x}_k = \mathbf{x}_{k-1} + \mathbf{s}_{k-1} \quad (\text{A.3.5})$$

$$\mathbf{s}_k = \mathbf{s}_{k-1} + \mathbf{w}_{k-1}. \quad (\text{A.3.6})$$

The smoothness of the model for the parameter variation can be manipulated in a smoothed integrated random walk model

$$\mathbf{x}_k = \rho \mathbf{x}_{k-1} + \mathbf{s}_{k-1} \quad (\text{A.3.7})$$

$$\mathbf{s}_k = \mathbf{s}_{k-1} + \mathbf{w}_{k-1}. \quad (\text{A.3.8})$$

References

- Anderson, M.G. and T.P. Burt. 1990. Process studies in hillslope hydrology: an overview. In: *Process studies in hillslope hydrology*. J. Wiley & Sons, Chichester, 1-8.
- Akaike, H. 1974. Markovian representation of stochastic processes and its application to the analysis of autoregressive moving average processes. *Ann. Inst. Statist. Math.*, **26**: 363-387.
- Astatkie, T., D.G. Watts and W.E. Watt. 1997. Nested threshold autoregressive (NeTAR) models. *Int. J. of Forecasting*, **13**: 105-116.
- Bazemore, D.E., K.N. Eshleman and K.J. Hollenbeck. 1994. The role of soil water in storm-flow generation in a forested headwater catchment: synthesis of natural tracer and hydro-metric evidence. *J. Hydrol.*, **162**: 47-75.
- Beven, K. 1982. On subsurface stormflow: An analysis of response times. *Hydrological Sciences J.*, **4**: 505-521.
- Beven, K. and P. Germann. 1982. Macropores and water flow in soils. *Water Resour. Res.* **18**(5): 1311-1325.
- Beven, K.J. 1985. Distributed models. In: M.G. Anderson and T.P. Burt (eds.), *Hydrological forecasting*. Wiley, Chichester, pp. 405-45.
- Beven, K. J. and P.C. Young. 1988. An aggregated mixing zone model of solute transport through porous media. *J. Contam. Hydrol.*, **3**: 129-143.
- Beven, K.J. 1989. Changing ideas in hydrology. The case of physically-based models. *J. Hydrol.*, **105**: 157-172.
- Beven, K.J. 1995. Linking parameters across scales: subgrid parameterization and scale dependent hydrological models. *Hydrol. Process.*, **9**: 507-525.
- Blöschl, G., and M. Sivapalan. 1995. Scale issues in hydrological modelling: a review. *Hydrol. Process.*, **9**: 251-290.
- Bonell, M. 1993. Progress in understanding of runoff generation dynamics in forests. *J. Hydrol.*, **150**: 217-275.
- Box, G.E.P., and G.M. Jenkins. 1970. *Time Series Analysis, Forecasting and Control*. Revised edn published 1976. San Francisco, Holden-Day.
- Bras, R.L. 1990. *Hydrology*. Reading, MA: Addison-Wesley Publishing Co.

- Burch, G.J., I.D. Moore and J. Burns. 1989. Soil hydrophobic effects on infiltration and catchment runoff. *Hydrol. Processes*, **3**: 211-222.
- Burch, H. 1994. Ein Rückblick auf die hydrologische Forschung der WSL im Alptal. In: *Beiträge zur Hydrologie der Schweiz Nr. 35*. SGHL, Bern.
- Caissie, D., T.L. Pollock, and R.A. Cunjak. 1996. Variation in stream water chemistry and hydrograph separation in a small drainage basin. *J. Hydrol.*, **178**:137-157.
- Chao-Lin Chiu. 1978. *Applications of Kalman filter to hydrology, hydraulics, and water resources*. Proc. AGU Chapman conference, Pittsburgh.
- Conell, P.E., and E. Todini. 1996. Modelling of rainfall, flow and mass transport in hydrological systems: an overview. *J. Hydrol.* **175**: 3-17.
- Creed, I.F., L. E. Band, N.W. Forster, I.K. Morrison, J.A. Nicolson, R.S. Semkin and D.S. Jeffries. 1996. regulation of nitrate-N release from temperate forests: A test for the N flushing hypothesis. *Water Resour. Resour.*, **32**: 3337-3354.
- Cushman, 1987. More on stochastic models. *Water Resour. Res.* **23**, 1668-1676.
- de Gooijer, J.G., B. Abraham, A. Gould and L. Robinson. 1985. Methods for Determining the Order of an Autoregressive-Moving Average Process: A Survey. *Int. Stat. Rev.*, **53**(3): 301-329.
- De Gooijer, J.G. 1992. Some recent developments in non-linear time series modelling, testing and forecasting. *Int. J. of Forecasting*, **8**: 135-156.
- Dise, N.B. and R.F. Wright. 1995. Nitrogen leaching from European forests in relation to nitrogen deposition. *For. Ecol. Managem.*, **71**: 153-161.
- Diserens, E. 1992. Etude de quelques aspects pédologiques liés aux dépositions acides dans une pessière humide de suisse centrale, Diss. 9857 ETH Zürich.
- Dunne, T. and R.D. Black. 1970. Partial area contributions to storm runoff in a small New England watershed. *Water Resour. Res.* **6**(4): 1296-1311.
- Elsenbeer, H. and D. Lorieri. 1995. Mixing model approaches to estimate storm flow sources in an overland flow-dominated tropical rain forest catchment. *Water Res. Res.* **31**:2267-2278.
- Ferraresi, M., E. Todini and R. Vignoli. 1996. A solution to the inverse problem in ground-water hydrology based on Kalman filtering. *J. Hydrol.* **175**:567-581.
- Feyen, H., J. Leuenberger, A. Papritz, M. Gysi, H. Flühler and P. Schleppei. 1996. Runoff processes in catchments with a small scale topography. *Phys. Chem. Earth.* **21**: 177-181.
- Flury, M., H. Flühler, W.A. Jury and J. Leuenberger. 1994. Susceptibility of soils to preferential flow of water: A field study. *Water Res. Res.*, **30**: 1945-1954.
- Flury, M., J. Leuenberger, B. Studer and H. Flühler. 1995. Transport of anions and herbicides in a loamy and sandy field soil. *Water Resour. Res.*, **31**: 823-835.
- Freeze, 1972. Role of subsurface flow in generating surface runoff. II. Upstream source areas. *Water Resour. Res.* **8**(5): 1272-1283.

- Grayson, R.B., I.D. Moore and T.A. McMahon. 1992. Physically based hydrological modelling, 1. A terrain-based model for investigative purposes. *Water Resour. Res.* **28**: 2639-2658.
- Gundersen, P. and L. Rasmussen. 1995. Nitrogen mobility in a nitrogen limited forest at Klosterhede, Denmark, examined by NH_4NO_3 addition. *For. Ecol. Managem.* **71**: 75-88.
- Gysi, M. 1995. *Abflussprozesse in einer kleinen Untersuchungsfläche am Hang*. Diploma Thesis, ETH Zürich.
- Hooper, R.P., N. Christophersen, and N.E. Peters. 1990. Modelling streamwater chemistry as a mixture of soilwater end-member--an application to the Panola Mountain catchment, Georgia, U.S.A. *J. Hydrol.* **116**:321-343.
- Hornberger, G.M., P.F. Germann, and K.J. Beven. 1991. Throughflow and solute transport in an isolated sloping soil block in a forested catchment. *J. Hydrol.*, **124**: 81-99.
- Horton, R.E. 1933. The role of infiltration in the hydrologic circle. *Trans. Amer. Geophys. Union*, **14**: 446-460.
- Ihringer, J. 1985. Vorhersage von Hochwasserflüssen aus Niederschlägen mit dem Kalman Filter. Mitteilungen IHW No. 25, Karlsruhe.
- Jakeman, A. J., I. G. Littlewood and P. G. Whitehead. 1990. Computation of the instantaneous unit hydrograph and identifiable components of flows with application to two small upland catchments. *J. Hydrol.*, **117**: 275-300.
- Jury, W.A. 1982. Simulation of solute transport using a transfer function model. *Water Resour. Res.* **18**(2): 363-368.
- Kachroo R.K. and L. Natale. 1992. Non-linear modelling of the rainfall-runoff transformation. *J. Hydrol.* **135**: 341-369.
- Kalman, R.E., and R. Bucy. 1961. New Results in Linear Filtering and Prediction, *Journal of Basic Engineering (ASME)*, **83D**: 95-108.
- Kirkby, M.J. 1988. Hillslope runoff processes and models. *J. Hydrol.* **100**: 315-339.
- Lange, H., G. Lischeid, R. Hoch and M. Hauhs. 1996. Water flow paths and residence times in a small headwater catchment at Gardsjön, Sweden, during steady state storm flow conditions. *Water Resour. Res.*, **32**: 1689-1698.
- Leaney, F.W., K.R.J. Smettem and D.J. Chittleborough. 1993. Estimating the contribution of preferential flow to subsurface runoff from a hillslope using deuterium and chloride. *J. Hydrol.*, **147**: 83-103.
- McDonnell, J.J. 1990. A rationale for old water discharge through macropores in a steep, humid catchment. *Water Resour. Res.*, **26**: 2821-2832.
- Moldan, F., H. Hultberg, U. Nyström and R.F. Wright. 1995. Nitrogen saturation at Gardsjön, southwest Sweden, induced by experimental addition of ammonium nitrate. *For. Ecol. Manage.*, **71**: 89-97.
- Mosley, P.M. 1982. Subsurface flow velocities through selected forest soils, South Island, New Zealand. *J. Hydrol.* **52**: 321-335.

- Mulholland, P.J., G.V. Wilson and P.M. Jardine. 1990. Hydrogeochemical response of a forested watershed to storms: Effects of preferential flow along shallow and deep pathways. *Water Resour. Res.*, **26**:3021-3036.
- Müller, N. 1997. *Short-term response of the ground vegetation in a montane forest ecosystem under increased nitrogen deposition--Influence of light and competition*. PhD Dissertation No. 12388, ETH Zürich.
- Natale, L. and E. Todini. A constrained parameter estimation technique for linear models in hydrology. In: *Mathematical Models for Surface Water Hydrology*. Wiley, Chichester.
- Nyberg, L. 1995. *Soil- and Groundwater Distribution, Flowpaths, and Transit Times in a Small Till Catchment*. PhD Dissertation, Uppsala University, Sweden.
- Parlange, M.B., G. G. Katul, R. H. Cuenca, M. L. Kavvas, D.R. Nielsen, and M. Mata. Physical basis for a time series model of soil water content. *Water Resour. Res.* **28**(9): 2437-2446.
- Pearce, A.J., M.G. Stewart and M.G. Sklash. 1986. Storm runoff generation in a humid headwater catchment 1. Where does the water come from, *Water Resour. Res.*, **22**(8): 1263-1272.
- Pearce, A.J. 1990. Streamflow generation processes. An Austral view. *Water Resour. Res.*, **26**: 3037-3047.
- Peters, D.L., J.M. Buttle, C.H. Taylor and B.D. Lazerte. 1995. Runoff production in a forested, shallow soil, Canadian shield basin. *Water Resour. Res.*, **31**: 1291-1304.
- Priestley, M.B. 1981. *Spectral Analysis and Time Series*. Academic Press, London.
- Priestley, M.B. 1988. *Non-linear and Non-stationary Time Series Analysis*. Academic Press, London.
- Rauch, H.E., F. Tung and C.T. Striebel, 1965. Maximum Likelihood Estimates of Linear Dynamic Systems, *AIAA Journal*, **3**(8): 1445-1450.
- Rawlins, B.G., A.J. Baird, S.T. Trudgill and M. Hornung. 1997. Absence of preferential flow in the percolating waters of a coniferous forest soil. *Hydrological Processes*, **11**: 575-585.
- Renzetti, A.V.E., C.H. Taylor and J.M. Buttle. 1992. Subsurface flow in a shallow Canadian shield watershed. *Nordic Hydrology*, **23**: 209-226.
- Robinson, J. S., M. Sivapalan and J. D. Snell. 1995. On the relative role of hillslope processes, channel routing, and network geomorphology in the hydrologic response of natural catchments. *Water Resour. Res.*, **31**: 3089-3101.
- Rodriguez-Iturbe, I., J.B. Valdes and J.M. Velasquez. 1978. Applications of Kalman filter in rainfall runoff studies. In Chao-Lin Chiu (ed.) *Applications of Kalman filter to hydrology, hydraulics, and water resources*. Proc. AGU Chapman conference, Pittsburgh.
- Seyfried, M.S. and B.P. Wilcox. 1995. Scale and the nature of spatial variability: Field examples having implications for hydrologic modeling. *Water Resour. Res.*, **31**: 173-184.
- Schleppi P., N. Muller, H. Feyen, A. Papritz, J.B. Bucher and H. Flühler. 1998. Nitrogen budgets of two small experimental forested catchments at Alptal, Switzerland. *For. Ecol. Manage.* **101**: 175-187.

- Sidle, R.C., Y. Tsuboyama, S. Noguchi, I. Hosoda, M. Fujieda and T. Shimizu. 1995. Seasonal hydrologic response at various spatial scales in a small forested catchment, Hitachi Ohta, Japan. *J. Hydrol.*, **168**: 227-250.
- Sönderström, T. and P.G. Stoica. 1983. *Instrumental Variable Methods for System Identification*. Lecture Notes in Control and Information Sciences. Springer-Verlag, Berlin.
- Stadler, D. 1996. *Water and Solute Dynamics in Frozen Forest Soils---Measurement and Modeling*. PhD Dissertation No. 11574, ETH Zürich.
- Tong, H. 1983. *Threshold Models in Non-linear Time Series Analysis*. Springer Verlag, New York.
- Tong, H. 1990. *Nonlinear Time Series: A Dynamical System Approach*. Oxford University Press, New York.
- Tsuboyama, Y., R.C. Sidle, S. Noguchi and I. Hosoda. 1994. Flow and solute transport through the soil matrix and macropores of a hillslope segment. *Water Resour. Res.* **30**: 879-890.
- Turton, D.J., D.R. Barnes Jr. and J. de Jesus-Navar. 1995. Old and new water in subsurface flow from a forest soil block. *J. Environ. Qual.*, **24**: 139-146.
- van Genuchten, M. T. and P.J. Wierenga 1976. Mass transfer studies in sorbing porous media. I. Analytical solutions. *Soil Sci. Soc. Am. J.*, **40**: 473-480.
- Wang, C.T., V.K. Gupta and E. Waymire. 1981. A geomorphological synthesis of non-linearity in surface runoff. *Water Resour. Res.*, **17**: 545-554.
- Ward, R.C. and M. Robinson. 1990. *Principles of Hydrology*. McGraw-Hill, London.
- Wilson, G.V. and R.J. Luxmoore. 1988. Infiltration, macroporosity, and mesoporosity distributions in two forested watersheds. *Soil Sci. Soc. Am. J.*, **52**: 329-335.
- Wilson, G.V., P.M. Jardine, R.J. Luxmoore and J.R. Jones. 1990. Hydrology of a forested hillslope during storm events. *Geoderma*, **46**: 119-138.
- Wood, E.F., M. Sivapalan and K.J. Beven. 1990. Similarity and scale in catchment storm response. *Rev. Geophys.*, **28**: 1-18.
- Wooding, R.A. 1965. A hydraulic model for the catchment-stream problem. 1. Kinematic wave theory. *J. Hydrol.*, **3**: 254-267.
- Wright, R.F., J. G. M. Roelofs, M. Bredemaier, K. Blanck, A.W. Boxman, B.A. Emmett, P. Gundersen, H. Hultberg, O.J. Kjonaas, F. Moldan, A. Tietema, N. van Breemen and H.F.G. van Dijk. 1995. NITREX: responses of coniferous forest ecosystems to experimentally changed deposition of nitrogen. *For. Ecol. Manage.*, **71**: 163-169.
- Wright, R.F. and N. van Breemen. 1995. The NITREX project: an introduction. *For. Ecol. Manage.*, **71**: 1-5.
- Wright, R.F. and A. Tietema. 1995. Ecosystem response to 9 years of nitrogen addition at Sogndal, Norway. *For. Ecol. Manage.*, **71**: 133-142.
- Wright, R.F. and L. Rasmussen. 1998. Large-scale ecosystem experiments: ecological research and European environmental policy. *For. Ecol. Manage.* **101**(1-3): 353-363.
- Young, P.C. 1984. *Recursive Estimation and Time-Series Analysis*. Springer, Berlin.

- Young, P. 1985. The instrumental variable method: a practical approach to identification and system parameter estimation. IFAC Identification and System Parameter Estimation, York, UK.
- Young, P.C. and S.G. Wallis. 1986. The Aggregated Dead Zone (ADZ) model for dispersion in rivers. In: Proceedings of Int. Conf. on Water Quality Modelling in the Inland Natural Environment. BHRA, Bournemouth.
- Young, P.C. 1989. Time variable and state dependent modelling of nonstationary and non-linear time series. In Subba Rao, T. (ed.), *Developements in Time Series Analysis*. Chapman and Hall, London. pp.374-413.
- Young, P.C. 1993. Time variable and state dependent modelling of nonstationary and non-linear time series. In *Developments in Time Series Analysis* (ed. T. Subba Rao), Chapman and Hall, London. p. 374-413.
- Young, P.C. and K. J. Beven. 1994. Data-based mechanistic modelling and the rainfall-flow non-linearity. *Environmetrics*, 5: 335-363.

Curriculum Vitae

1970	Born on November 14 in Hasselt (Belgium)
1976 - 1982	Primary school in Zonhoven
1982 - 1988	Gymnasium 'Sint Jan Berchmansinstituut', Zonhoven
1988 - 1993	Catholic University of Leuven Graduated as Agricultural Engineer (ir. Agr.) Specialization Environmental Sciences and Soil Science
1993 - 1994	Research assistant at the Institute of Land and Water management, Catholic University of Leuven
1994 - 1998	PhD student at the Institute of Terrestrial Ecology, Soil Physics, Swiss Federal Institute of Technology (ETH) in Zürich