

Final Report of the DAUFIN project

E.E. van Loon and P.A. Troch (eds.)

Rapport 115

Sectie Waterhuishouding
Nieuwe Kanaal 11, 6709 PA Wageningen
Internet: www.dow.wau.nl/whh

ISSN 0926-230X

1679312

Contributors

P. Troch

P. Torfs

E. van Loon

Wageningen University
Hydrology and Quantitative Water Management Group
Nieuwe Kanaal 11
6709 PA Wageningen
The Netherlands

M. de Neef

M. Hassanizadeh

P. Reggiani

T. Rientjes

Technical University Delft
Faculty of Civil Engineering and Geosciences
Stevinweg 1
PO Box 5048
2600 GA Delft
The Netherlands

M. Marrocu

C. Paniconi

Centro di Ricerca, Sviluppo e Studi Superiori di Sardegna
VI strada ovest z.i. Macchiareddu
PO Box 94 09010 Uta (CA)
Italy

A. de Roo

G. Somma

Space Applications Institute
Agricultural and Regional Information Systems Unit
Via E. Fermi
PO Box TP 262 21020
Ispra (VA) Italy

E. Roulin

Institut Royal Meteorologique de Belgique
Recherche Meteorologique et Developpement
Avenue Circulaire 3
1180 Bruxelles
Belgium

S. Brouyère

A. Dassargues

University of Liege (LGIH)
Lab. de Geologie de l'Ingenieur, d'Hydrogeologie et de prospection geophysique
B19 - Sart Tilman, 4000 Liege
Belgium

S. Ignar

D. Miroslaw-Swiatek

Warsaw Agricultural University (SGGW)
Department of hydraulic structures
Nowoursynowska 166
02-787 Warsaw
Poland

P. Botti

Ente Autonomo del Flumendosa
Via Mameli N. 88
09123 Cagliari
Italy

A. Nalberczynski

Regional Water Authority Wroclaw
Wybrzeze Wyspianskiego
50-370 Wroclaw
Poland

Abstract

E.E. van Loon and P. A. Troch (eds.) 2002. *Final Report of the DAUFIN project*, Wageningen University, The Netherlands. 157 pp.

This report summarizes the main results of a 20-month EU Research and Development project (the DAUFIN project), aiming at the development of improved tools for hydrological modeling. The tools concentrate on a new catchment-based modeling approach - the Representative Elementary Watershed model - and new data assimilation techniques. The main strength of the REW approach is the combination of being general and physically based and at the same time operating at the sub-catchment level. Data assimilation techniques have the potential to integrate observations and models in an automated way, so that automatic model calibration as well as online prediction with state corrections is possible. Data assimilation can be useful to parameterize and correct the REW-model, while the REW-model has a size and the numerical properties that are ideal for data assimilation purposes. In this way both approaches capitalize each others strengths.

In this report the coding of the REW model is explained, and it is shown in detail how it is applied to the Geer catchment, a sub-catchment of the Meuse. Next, a number of new data assimilation techniques, specifically suited for hydrological problems, are described. In the first place a nudging technique, aimed at application in 2D or 3D Richards-based models, is explained and tested. Secondly, a technique based on cross-validation for application in a wide class of lumped or distributed hydrological models is described. Both techniques give satisfactory results. The use of data assimilation techniques for the automatic calibration of hydrological models is studied in the Geer and the Ourthe (another tributary to the Meuse). Finally, a statistical correction technique is applied to the the assimilation of Meteosat-based evapotranspiration data in a distributed hydrological model. The efficiency of the last method is tested on data from the Mehaigne, Ourthe Orientale, Eau Blanche and Loison (Meuse sub-catchments).

The conclusions of the project are that both the REW modeling technique as well as the newly developed data assimilation techniques appear to be very useful already at an early stage of development. The techniques need to be further developed and tested for operational use. In the future there should be particular attention to develop data assimilation algorithms that can handle problems typically encountered in hydrological applications: more robust handling of unobservable system parts and multi-scale heterogeneity.

Preface

The DAUFIN project started as an ambitious enterprise, focusing on fundamental hydrological modeling issues while closely linking-up with existing hydrological modeling techniques and practical applications. All project partners were dedicated for 100% to reach those project goals. The project had to cope with problems of finding trained staff for the short project duration of 20 months, and collecting a variety of hydrological data from much more institutes than initially foreseen. The fact that those problems were swiftly solved and the original project goals were still reached, illustrate the dedication of each of the partners. In this respect the project team wishes to express also thanks for the good support from the administrations of the Universities, and partner institutes as well as the responsible EU offices. It is due to the Polish and Sardinian end-users that the project did receive good feed-back on the applicability of the models developed by the DAUFIN project team. At the end of the project a larger group of end-users took the effort to provide feed-back to the project by participating in an end-user workshop. The DAUFIN team wishes to thank all end-users for their interest and constructive comments. Finally, the DAUFIN team found great inspiration in the response to the project results by scientific colleagues from all over the world at the CAHMDA workshop in Wageningen, September 2001.

Contents

1. Introduction	1
1.1. Background	1
1.2. Objectives	2
1.3. Project structure and overview of the report	2
I. Theory	5
2. Development of the Representative Elementary Watershed modeling code	7
2.1. Introduction	7
2.2. Fundamental concepts	9
2.3. Review of balance equations	11
2.4. Balance equations and constitutive relationships	11
2.5. Parameter estimation	13
2.6. Numerical applications	14
2.6.1. Study site	14
2.6.2. Software architecture	16
2.6.3. Numerical experiments and results	17
2.6.4. Sensitivity analysis	17
2.6.5. Model Simulations	19
2.6.6. Future Research	20
2.7. Conclusions	21
3. A review of existing Data Assimilation techniques	25
3.1. Introduction	25
3.2. A Review of Soil Moisture Data Assimilation	26
3.2.1. Overview	26
3.2.2. Optimal approaches	27
3.2.3. Sub-optimal approaches	28
3.2.4. Discussion	29
3.3. A prototype for 4D hydrological data assimilation algorithms	31
3.3.1. Algorithmic structure	31
3.3.2. The practical consequences for model forms and data handling	32
4. Developing and testing a nudging technique in a 3D-Richards model	33
4.1. Introduction	33
4.2. Selection of a data assimilation technique	34
4.3. Newtonian relaxation	35
4.3.1. Weighting functions	35
4.4. Implementation of nudging for the hydrological model	37
4.4.1. Model description	37

4.4.2. Numerical discretization with the nudging term	38
4.5. Numerical experiments	39
4.5.1. Description of the test case	39
4.5.2. Perturbation of atmospheric forcing boundary conditions	41
4.5.3. Influence of nudging parameters	42
4.5.4. Perturbation of initial conditions	44
4.6. Conclusions	45
5. Application of a nudging technique to the Brisy catchment with a 2D-Richards model	55
5.1. Introduction	55
5.2. Mathematical model of hillside subsurface flow	55
5.3. Numerical approximation of the Richards equation	58
5.4. Newtonian nudging as data assimilation method	59
5.5. Application to the Brisy catchment	61
5.6. Discussion and Conclusions	64
6. Developing and testing a regularization technique	67
6.1. Introduction	67
6.2. Material and Methods	68
6.2.1. Description of data	68
6.2.2. Formulating the distributed hydrological model	68
6.2.3. Formulating the geo-statistical model	70
6.2.4. Formulating the data assimilation procedure	71
6.2.5. Comparing data assimilation with alternative estimation techniques	73
6.3. Results	74
6.3.1. Distribution of soil moisture in space	74
6.3.2. Characteristics of the various models	75
6.3.3. Predicting discharge and ground water level	77
6.4. Discussion and Conclusions	77
II. Application of control models to sub-catchments of the Meuse	91
7. Automatic calibration of a groundwater model for the Geer catchment	93
7.1. Introduction	93
7.2. General structure of MOHISE	93
7.3. Improvement of the Geer groundwater model and presentation of results	93
7.4. Transient state calibration and validation	95
7.5. Conclusions	96
8. Calibration of a flood prediction model for the Ourthe	99
8.1. Description of the LISFLOOD model	99
8.2. Finding optimum parameter sets for the Ourthe catchment	102
8.3. Evaluation of optimal parameter values	103
8.4. Summary of the results	104
8.5. Conclusions	104
9. Statistical correction applied to a water-balance model for the Meuse	113
9.1. Introduction	113

9.2. Soil moisture and evapotranspiration in SCHEME	113
9.3. Evapotranspiration from Meteosat data	114
9.4. Assimilation method	116
9.5. Soil moisture data and validation	117
9.6. Actual evapotranspiration assimilation and soil moisture	118
9.7. Comment on weather radar data	119
9.8. Actual evapotranspiration and the water balance	123
9.9. Conclusions	126
III. Conclusions and Recommendations	131
10. Conclusions	133
10.1. Achievements by the DAUFIN project	133
10.2. Project output	133
10.3. Dissemination of results	134
11. Recommendations	135
Appendix A	
DAUFIN project output: publications, presentations and data.	137
Bibliography	147

1. Introduction

1.1. Background

The study of the response of a river basin to atmospheric forcing is of critical importance to applied hydrologists and water resource managers and remains a major research challenge. Continued progress in our scientific understanding of hydrological processes at the catchment scale relies on making the best possible use of advanced simulation models and the large amounts of environmental data that are increasingly being made available. Processes at the interface between the land surface and the atmosphere, for instance, determine the partitioning of rainfall into infiltration and runoff and the redistribution of water between the surface, soil, underlying aquifers, and streams. Understanding and predicting these exchanges is important to agriculture (irrigation planning and vegetation growth), climate studies (weather prediction and global change), natural hazards prevention and mitigation (floods, droughts, erosion, landslides), and water quality management (point and non-point source pollution in ground and surface waters).

A wide variety of distributed hydrological models has been developed over the past decades. The common feature of such models is that they can incorporate the spatial distribution of various inputs and boundary conditions, such as topography, vegetation, land use, soil characteristics, rainfall, and evaporation, and produce spatially detailed outputs such as soil moisture fields, water table positions, groundwater fluxes, and surface saturation patterns. A major factor contributing to the popularity of the distributed modeling approach is the availability of digital terrain data and simple algorithms for extraction of hydrologically relevant information from this data. One of the major problems plaguing distributed modeling is parameter identifiability, owing to a mismatch between model complexity and the level of data which is available to parameterize, initialize, and calibrate models, and to uncertainty and error in both models and observation data [Beven, 1996; Gupta et al., 1998a; Perrin et al., 2001].

Continued progress in our scientific understanding of hydrological processes at the catchment scale relies on making the best possible use of advanced simulation models and the large amounts of environmental data that are increasingly being made available [Hornberger and Boyer, 1995; Grayson and Blöschl, 2001]. New data sources for observation of hydrological processes can alleviate some of the problems facing the validation and operational use of hydrological models. In situ or ground-based measurement has become more feasible with the advent of simpler and cheaper sensors, gauges, and loggers, while satellite and airborne remote sensing has begun to fulfill some of its potential for hydrological applications, allowing monitoring and measurement of rainfall, snow, soil moisture, vegetation, surface temperature, energy fluxes, and land cover over large areas [Dozier, 1992; U. S. Congress Office of Technology Assessment, 1994]. Data assimilation provides a means of integrating these data in a consistent manner with model predictions. In general terms, geophysical data assimilation is a quantitative, objective method to infer the state of the

earth-atmosphere-ocean system from heterogeneous, irregularly distributed, and temporally inconsistent observational data with differing accuracies, providing at the same time more reliable information about prediction uncertainty in model forecasts. In general terms, geophysical data assimilation is a quantitative, objective method to infer the state of the earth-atmosphere-ocean system from heterogeneous, irregularly distributed, and temporally inconsistent observational datasets with differing accuracies, providing at the same time more reliable information about prediction uncertainty in model forecasts [Daley, 1991; U. S. National Research Council, 1991; McLaughlin, 1995]. Data assimilation is by now routinely used in research and operational meteorology, although many scientific challenges remain for improving and extending existing methodologies. Very recently, data assimilation was introduced in the oceanographic and hydrologic disciplines, owing to the trend towards better and more regular observation of a wide range of parameters of interest to the Earth sciences, beyond those traditionally used.

Notwithstanding the fact that there are many spatially distributed models in surface and sub-surface hydrology which could provide a basis for data assimilation, the most advanced domain of application has been atmospheric hydrology, due to the important role this "fast component" of the global hydrological cycle plays in weather forecasting. Other applications of data assimilation in hydrology have concentrated on the evaluation of new satellite sensors, and were often based on synthetic experiments. There are even no case studies available yet, where data assimilation has been applied at a large-scale. The reason for this is that in surface as well as groundwater hydrology the technology is not yet ready to be applied operationally.

1.2. Objectives

The aim of the DAUFIN project was to design and apply methodologies that are able to improve the predictive skill of a hydrological model on the basis of scattered measures of various kinds (satellite, in situ, indirect measures, etc.). The project aimed at studying the available data assimilation routines in depth, and applying these to existing distributed hydrological models. In addition to this, the DAUFIN project also intended to unify of those different models through the Representative Elementary Watershed (REW) concept. This unification will simplify future hydrologic models and in this way open the way to apply more advanced data assimilation routines. Eventually it will lead to a better understanding of the hydrologic systems being studied and an enhanced predictability of those systems. The modest size and the short project period of 20 months is of course not sufficient to fully realize such an ambitious goal, therefore only initial steps are set towards a unification. These are: the development of a fully operational REW modeling code, the inter-comparison of this code with an existing distributed model, and finally the design and application of data assimilation routines that can be combined with the REW model.

1.3. Project structure and overview of the report

The project was organised along four main activities that took place more or less in parallel. The first was the development of the REW computer code. The second project activity was the development of data assimilation routines. A third project activity was the application of existing distributed hydrological models (so-called control models) to

sub-catchments of the meuse. And a fourth activity was the inter-comparison of the REW computer model with one of the control models. In support of these activities there was a data collection effort to make a good set of meteorologic and hydrologic data available to all the partners, and several efforts to disseminate the project results in the scientific community and among end-users.

Originally, the DAUFIN project intended to also apply data assimilation to the REW-model. Towards the completion of the project, it became clear that this task could not be finished within the project. The work that remains to be done to achieve this aim will be discussed in the last chapter of this report.

This report is subdivided into three main parts. The first part (Chapters 2 - 6) provides an overview over the theoretic studies that have been performed. First the development of the REW-modeling code is presented in Chapter 2. Then a review of existing data assimilation techniques is presented in Chapter 3, the application of nudging to full 3D and 2D Richards models is studied in Chapters 4 and 5 respectively. The possibility to apply a more data-based assimilation technique is explored in Chapter 6.

The second part of this report comprises a number of applied modeling studies with control models in sub-catchments of the meuse. Chapters 7 presents a study on automatic model calibration in the Geer catchment. In Chapter 8 reports a calibration study in the Ourthe catchment. And Chapter 9 details a study where Meteosat-based evapotranspiration data is assimilated in a distributed hydrological model, for four sub-catchments of the Meuse.

The last part of this report summarises the main project results, provides a comprehensive discussion, and gives an overview over the research that needs to be done to achieve further progress in hydrological data assimilation.

Part I.

Theory

2. Development of the Representative Elementary Watershed modeling code

Tom Rientjes; Michiel de Neef; Majid Hassanizadeh and Paolo Reggiani

2.1. Introduction

Traditionally three runoff modeling approaches are available to the scientific community: i) physically based models, ii) conceptual models and iii) empirical models. Physically based runoff models apply a mathematical model based on point scale conservation equations of mass, momentum, and energy for the discretised model domain representing the hydrological system. Freeze [1969] proposed a "Blue print" of such distributed physically based watershed modeling approach in the seventies, based on rigorous numerical solution of differential equations, governing flow through the soil (Richards' equation, Darcy's law), overland flow (kinematic wave equation) and open river flow (Saint-Venant equations). These equations also form the basis of other distributed physically based models developed such as the *Système Hydrologique Européen (SHE)* model [Abbott *et al.*, 1986a,b]. Others to follow similar approaches were *Beven* [1987]. There has been considerable discussion regarding the advantages and disadvantages of such model approaches by *Beven* [1989] and *Klemeš* [1986]. The authors of these papers come to similar conclusions: firstly, the numerical solution of the mathematical model generally is computationally demanding and models require a vast amount of physiographic, topographic and catchment geometric data, that for most natural watersheds, are not available at appropriate spatial and temporal scales. Other conclusions relate to the performance of the model that suffer from overparameterisation and non-uniqueness [Beven and Wood, 1993; Bathurst and Wicks, 1991].

The second type, and by far the largest group, of runoff models that have been proposed are conceptual watershed models. Among these we cite the *Stanford Watershed Model (SWM)* described by *Crawford and Linsley* [1966], *TOPMODEL* [Beven and Kirkby, 1979], *Nash cascade models* or the *LASCAM* model developed by *Viney and Sivapalan* [1999]. Conceptual models either apply a lumped (SWM; NASH) or distributed model domain (TOPMODEL). Models apply a mathematical model that is based on mass conservation and simple water transport and mass exchange equations and, when combined, make up the mathematical model. The latter equations often lack physical meaning and often (only) have empirical validity. Model parameter sets generally lack physical meaning although the number of parameters required by the model generally is small compared to distributed physically based models.

The third group of runoff models are empirical models that apply a lumped model domain and apply a mathematical model based on regression type equations. In literature, this group of models often is referred to as *black-box* or *transfer models* since the model

structure and the required model parameters (i.e. regression coefficients) have no meaningful physical interpretation. Well known models are the ARIMA models while the most popular group nowadays are the Artificial Neural Network models.

In view of the first two modeling approaches, *Reggiani et al. [1998, 1999]* proposed a novel approach, with the idea that it could form the basis for an alternative watershed model for runoff modeling. Although the mathematical model is based on real world physics and as such can be classified as physically based, the manner the model domain is discretised can be put into perspective as an intermediate approach between lumped and distributed models. Based on topographic analysis, a number of sub-catchments, called Representative Elementary Watersheds (REWs), are defined yielding a semi-distributed model domain. Within each REW, conservation equations for mass, momentum, energy and entropy are applied and spatially lumped into calculation units such as applied in distributed physically based models. In these models it is common to represent such units by nodal points in the spatially continuous model domain and this approach also has been adopted in our approach. On the other side the real world physics is conserved by averaging rigorously and defining exchange terms for mass and energy as well as forces for the various REWs. By the semi-distributed model approach the number of required model parameters reduces significantly compared to distributed physically based models and yet the real world physics ruling the production and transport of runoff water is maintained. Also required model parameters have physical meaning that can be defined by use of observation, data-bases or maps. Optimisation of such parameter values can be achieved by "trial and error" [*Refsgaard and Knudsen, 1996*] or automated model calibration procedures such as presented by *Q. Duan and Gupta [1992]*; *Gupta et al. [1998b]*; *Binley and Beven [1992]*.

In previous research on the REW approach, the authors have elaborated the main modeling concepts and applied them to two particular flow systems. They calculated the annual water balance for a single REW using the balance equations for the unsaturated and the saturated zone *Reggiani et al. [2000]*. The purpose of this simple exercise was to show that the set of balance equations could capture the annual water balance of the hydrological system for different geometrical configurations, climatic forcing conditions and soil types. Secondly *Reggiani et al. [2001]* applied the balance equations for the river network to a natural watershed reproducing instantaneous unit response functions for rainfall events of different intensity.

The objective of the work within the DAUFIN-project is to develop a computer code based on the REW modeling approach solving the balance equations for mass and momentum for an ensemble of mutually connected REWs. In this work the code is applied for the first time to an entire watershed that is drained by a river network and that is underlain by an aquifer that extends beyond the watershed boundary. The presentation of the work is carried out in a sequence of steps: Firstly the main concepts of the REW approach are briefly introduced. Secondly the balance equations are simplified, adapted to the particular flow situation, and projected along the axes of an appropriate reference system. Thirdly the closure scheme for the balance equations are presented. Finally it is shown how the necessary parameters are estimated for a natural watershed. The actual implementation, application and test of the model are shown in the last part of this chapter.

2.2. Fundamental concepts

In the approach derived by *Reggiani et al. [1998, 1999]*, a watershed is separated into a series of discrete spatial units called Representative Elementary Watersheds (REWs). REWs are defined by analysis of the topography and constitute a set of the interconnected REWs that are organized around the tree-like structure of the stream river network. The boundaries coincide with the topographic divides and they delineate a well defined area of the land surface that captures the precipitation. A REW is defined in three spatial dimensions and is delimited externally by a prismatic mantle, which coincides with the shape of the topographic divides circumscribing the sub-watershed. A schematic representation of an ensemble of three REWs is given in Figure 2.1. On top, the REW is delimited by the atmosphere while the bottom is bounded by either an impermeable substratum or an assumed limiting depth. The volume making up a REW contains all flow domains encountered within a watershed: 1) the unsaturated zone, 2) the saturated zone, 3) the saturated overland flow area, 4) the concentrated overland flow area, and 5) the river reach. The flows within the various domains are characterised by vastly different temporal scales and include different flow phenomena, such as multiphase unsaturated and saturated porous media flow (subsurface zones), as well as single phase flow (overland and river flow).

In distributed physically based watershed models, partial differential equations governing water flow in various zones are discretized and solved for the spatial scales of the calculation units. In our approach, balance laws for mass and momentum for each flow zone are averaged over an REW that, generally, is much larger than the grid element scale. Up-scaled balance laws are derived for each REW and resulting conservation equations for each flow domain of every REW constitute Ordinary Differential Equations (ODEs) which have the following general form:

$$\frac{d\Psi}{dt} = \sum_i e_i^\Psi + U + G \quad (2.1)$$

where Ψ represents a generic thermodynamic property such as mass or momentum, e_i^Ψ is a generic watershed-scale exchange term for Ψ , U is an external supply term for Ψ and G is its internal production. The exchange terms account for the transfer of Ψ among phases, zones and REWs. We note in this occasion that all equations of type 2.1 constitute global balance laws, in opposition to partial differential equations of the earlier mentioned distributed physically based models. Equations such as 2.1 no longer contain any spatial information and are referred to as mega-scale equations. In this context the exchange terms e_i^Ψ are referred to as mega-scale exchange terms and these terms constitute unknowns of the hydrological state estimation problem. As such, the mega-scale exchange terms require appropriate closure schemes in the form of constitutive relationships. These were obtained by *Reggiani et al. [1999]* in a single and physically consistent procedure where the second law of thermodynamics serves as a constraint.

In *Reggiani et al. [2000]* the main practical concepts concerning the modeling of a single REW were introduced. In that particular case, the REW has been considered in an isolated fashion, allowing for a considerable simplification of the governing equations. The aim was to simulate annual water balances and thus the flow processes of overland flow and river flow were neglected. In the current work we aim at modeling an entire watershed system that is underlain by an aquifer that extends beyond the watershed boundary. The purpose of this section is to introduce the reader to the basic geometrical

aspects which are necessary for solving the set of governing equations underlying the mathematical model and to highlight the model variables which will be used to describe the hydrological system. For the modeling of flow phenomena, it is necessary to adapt the governing equations and constitutive relationships derived in *Reggiani et al. [1998, 1999]* to the specific flow conditions, and to project the vectorial momentum conservation equations along the axes of a suitable reference system. Therefore we introduce a Cartesian reference system positioned at the outlet of the watershed. In Figure 2.2 such a projection is given for one REW. The x and y -axes form a horizontal plane delimiting the watershed at the bottom where the zero elevation mark is assumed to coincide with an arbitrary reference elevation. We also introduce three unit geo-reference vectors pointing along the axes of the reference system, e_x , e_y and e_z , respectively. In addition we introduce two unit model vectors, denoted with n^r and n^o respectively that are specific for each REW. n^r is a resultant unit vector tangent to the river axis and pointing downstream and n^o is a resultant unit vector pointing along the direction of steepest descent of the saturated overland flow area. These two vectors are required to project the balance equations of momentum for the river reach and the saturated overland flow area within each REW.

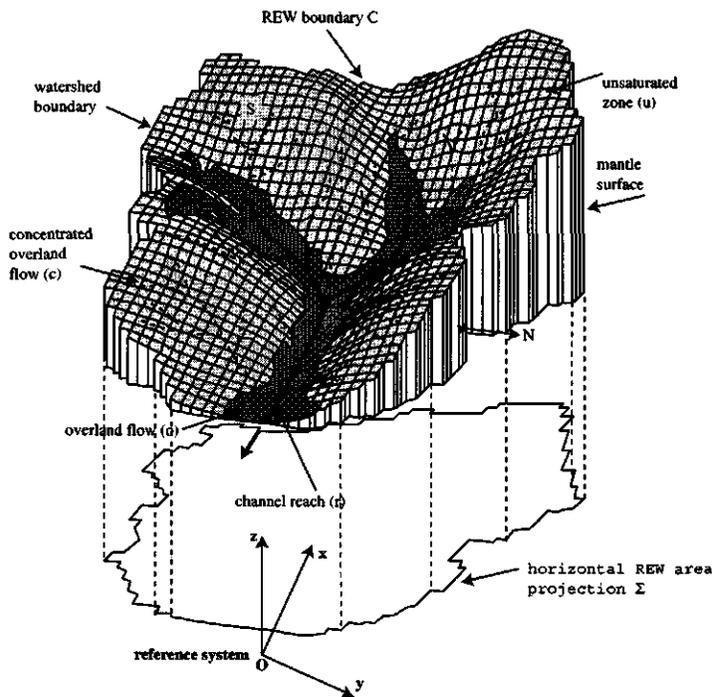


Figure 2.1.: A birds-eye view of an ensemble of REWs and all possible lateral mass exchange fluxes between a REW and its neighbours.

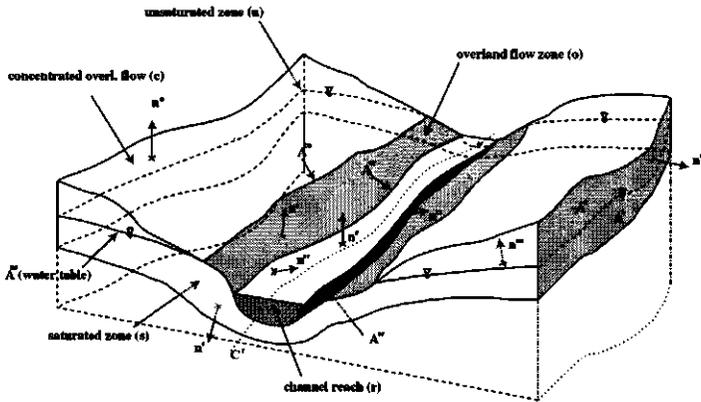


Figure 2.2.: An overview over the zones and fluxes in an REW.

2.3. Review of balance equations

REW-scale conservation laws for mass, momentum, energy and entropy for the five flow domains of an REW have been presented in *Reggiani et al. [1998]*. In the present work we aim at solving the equations of mass and momentum to an ensemble of REWs constituting an entire watershed. In the approach, all REWs balance equations are solved for the five flow domain and thus for each domain a specific mathematical model is developed. In the subsequent paragraphs we introduce the balance equations for mass and momentum for the unsaturated and saturated zones, the river reach and both overland flow source areas and project the vectorial momentum balance equations along the reference system, in order to provide scalar equations, which can be solved numerically. For the subsurface zones, the momentum balance equations are projected along the horizontal axes of the global reference system O while for the saturated overland flow areas and the river reach appropriate balance equations are projected along average resultant flow directions n^o and n^r for each REW (see Figure 2.2).

2.4. Balance equations and constitutive relationships

In this section the mass and momentum balance equations, forming at the core of the REW model, are briefly listed and The constitutive relations by which each balance term is quantified are briefly listed thereafter. For a detailed overview of all the equation the reader is referred to *Reggiani et al. [1998, 1999, 2000]*.

There REW model considers 5 mass and 5 momentum balance equations equations, namely for: the unsaturated zone, the saturated zone, a river reach, saturated overland flow, and concentrated overland flow

The balance equations need to be supplemented with constitutive relationships in order to quantify the unknown mass fluxes and force exchanges which appear in the mass conservation equations and the momentum balance equations. A pre-requisite of the use mass

and momentum balance equations for each flow domain in a REW is that the combined equations are conservative implementing that the set of balance equations requires closure. The development of a closure scheme has been carried out by *Reggiani et al. [1999]* and has partially been tested in *Reggiani et al. [2000, 2001]*. We note that we have not yet pursued a rigorous study of the bare soil evaporation and plant transpiration components of the water balance but instead, we have decided to include a rather simplistic parameterisation of evapotranspiration (combination of evaporation and transpiration) as a preliminary step to close the hydrologic cycle as a whole at the REW scale. A more rigorous representation of evapotranspiration is an urgent issue vis à vis the theoretical development which is being pursued here, and is left for further research.

All mass fluxes must satisfy the continuity conditions, i.e., jump conditions, across both internal and external boundaries of the REW. Such boundaries are e.g. the water table, the saturated overland flow source area, the unsaturated land surface, the free river area or the wetted river bottom. This is equivalent to stating that the exchange from one region towards another (e.g. unsaturated zone towards the saturated zone) has to equal the exchange in the opposite direction: . These mass exchange terms generally are functions of the mean values of the velocities on both sides of the boundary and the difference in hydraulic potentials. Ideally, such functions have to be determined from field observations, however, in physics and for practical applications, a linear dependence is commonly assumed as a first approximation. This approximation is also applied in the REW presently.

In the mass balance equations constitutive relations are being used to describe the following processes:

- Infiltration at the land surface
- recharge to and capillary rise from the saturated zone
- groundwater flow across REW mantle surface
- ex-filtration at the saturation overland flow source area
- evaporation from the unsaturated zone
- saturation overland flow into river reach
- river inflow and outflow
- atmospheric stress at the river free surface
- groundwater - river interactions
- concentrated overland flow into Saturated overland flow source area
- atmospheric stress at concentrated overland flow area

In analogy to the parameterisations for the mass exchange terms, suitable expressions for the momentum exchange terms need to be introduced. Momentum exchange equations are developed by *Reggiani et al. [1999, 2000]* and exchange terms are expressed as the sum of an equilibrium and a non-equilibrium term. The equilibrium component is attributable to pressure forces while the non-equilibrium part accounts for friction effects. A series of assumptions are made to allow a significant simplification of the momentum balance

equations. The assumptions are justified from a physical as well as hydrological viewpoint, within certain space and time scales. Such scales can be defined a-priori by physical reasoning but upper limits to these scales must be defined by model sensitivity analysis with regard to the applied REW catchment as subject to the selected Strahler order discretisation.

The simplifying assumptions with regard to the momentum equations in the REW model are:

- the average vertical flow in the saturated zone is ignored and the flow is mainly horizontal
- the slope of the land surface within the REW is uniform and small such that the horizontal force acting on the water in the unsaturated and the saturated zones across the land surface can be considered negligible
- within each REW the water table is near-horizontal and therefore horizontal components of the forces are negligible
- the impermeable surface delimiting the aquifer at the bottom of each REW is near-horizontal
- the river edge is orthogonal to the unit normal vector tangent to the overland flow zone

Now, using these assumptions, in the momentum balance equations constitutive relations are being used to describe the following processes:

- flow in the unsaturated zone
- flow in the saturated zone
- saturated overland flow
- river flow

2.5. Parameter estimation

After specification of all the balance equations and constitutive relationship, the model parameterisation remains.

The relationships describing flow through a river reach are parameterised by the equations given by geometrical laws provided by *Leopold and Maddock Jr. [1953]*. These relationships allow to express the average cross-sectional velocity, the top width and the average flow depth for a given stream reach as power laws of the discharge. Strictly speaking, they are valid under steady-state conditions but are in fact applied to a non-steady state flow situation. These relationships have been successfully employed for real world watershed situations, as shown in *Naden et al. [1999]*. Details of the derivations of the relationships used in the REW model are shown by *Snell and Sivapalan [1995]*.

The parameterisation of flow in the unsaturated zone is done by using the van Genuchten relationships *van Genuchten [1980]*. This relationship requires that soil types are characterised by two pressure shape factors, a pressure scaling parameter, a conductivity shape

factor and a conductivity scaling parameter. In addition, the porosity, is necessary for complete characterisation of the soil types. The Van Genuchten equation expresses the pressure head as a function of moisture content. The effective hydraulic conductivity in the unsaturated zone under unsaturated conditions can then via the Brooks and Corey equation *Brooks and Corey [1966]* be related to the saturated hydraulic conductivity. These point-scale equations will apply to our situation only if the soils in the subsurface zone can be assumed to be homogeneous and isotropic. For heterogeneous and anisotropic soils the above relationships will change from point to point and the modeling requires detailed information on the heterogeneity and anisotropy. In general such variability of soil properties cannot be modelled explicitly and as such also are ignored in our REW modeling approach. Instead, parameter values of the soil properties will be defined by model calibration and are assumed to be REW-effective. This implements that all variables and parameters used within each REW are REW-scale quantities which cannot vary within the soil profile or the REW.

The parameterisation of flow in the saturated zone comprises the inverse estimation of the coefficients that determine the partition of mass flux from a REW towards neighbouring REWs. The inverse modeling approach here is based on the Kirchhoff laws. It is assumed that the aquifer is at steady state over a short period of time. As shown in Figure 2.3, the aquifer underlying the watershed and separated into a finite number of REWs, can be envisaged as a network of pipes interconnecting nodes, which correspond to the centroids of the REWs. These pipes, representing REW interconnections, identify a series of non-redundant loops, whose total number can be calculated. With reference to Figure 2.3, we can formulate the first Kirchhoff law for a network node, which states that the sum of fluxes entering or exiting the node must be zero under steady state conditions. The second Kirchhoff law states that the sum of head losses along a closed loop must equal zero. The application of these laws to a system of non-redundant loops leads to a determined system of as many equations as pipes (or REW interconnections) in the network. By applying the proportionality of the discharge exchanged along a pipe to the head difference along the pipe, the required unknowns can be calculated.

To obtain parameterized balance equations, which can be solved numerically for a real-world application, it is necessary to introduce the expressions for the mass exchanges into the respective balance equations for mass. In the same manner we introduce the parameterised non-zero momentum exchange terms into the projected momentum balance equations. In this fashion we obtain a mass and a momentum balance law for each zone within the REW, parameterised through several coefficients, which need to be either measured or identified through calibration or inverse modeling techniques. The final set of balance equations constitutes a set of non-linear coupled ordinary differential equations which are solved through a fourth-order Runge-Kutta solution algorithm [*Press et al., 1992*]. For reasons of brevity we omit the reporting of the parameterized equations and move on to the results produced on a synthetic watershed.

2.6. Numerical applications

2.6.1. Study site

The numerical model implemented for the resolution of the balance equations has been tested on a real world watershed, the Geer basin covering 494 km^2 . The basin is drained by the river Geer that is a tributary of the river Meuse. The Geer basin can be characterised

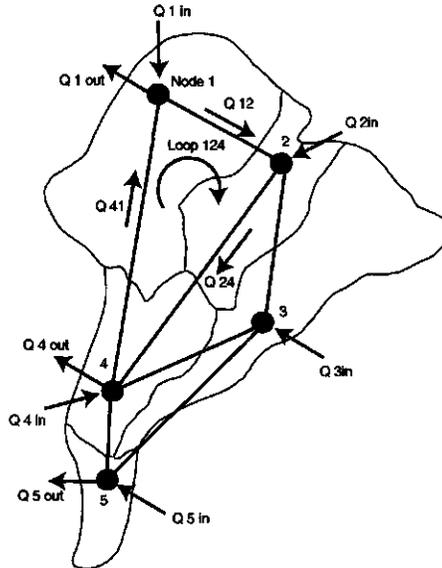


Figure 2.3.: The connection between REWs envisaged as nodes (in the REW-centers) linked by pipes.

as a "deep" groundwater system that us is under-lain by an aquifer that extends beyond the catchment boundaries and consists of Cretaceous chalks with a thickness varying from few meters in the south to about 100 meters in the North East. This aquifer is under-lain by a layer of Smectite, which inclines north wards with a gradient of 1% to 1.5% and can be considered as the hydrological base of the basin. The vadose zone in the basin can reach a thickness of up to 40 meters where the sub-strata are made up by several layers of chalk stone that can be characterised by a relatively low permeability. The aquifer underlying the Geer basin has a water divide at the topographic ridges in the southern part of the basin. Water flow across the northern topographic divide most likely takes place. Geological surveys have estimated the following water balance between 1975 and 1999 for the catchment:

$$P = \text{EvapoTransp.} + Q_{\text{Geer}} + Q_{\text{out}} + \delta Q_{\text{Reserves}} + Q_{\text{Losses}} \quad (2.2)$$

for which the various quantities have been expressed in millimeters as follows

$$810 \text{ mm} = 508 \text{ mm} + 145 \text{ mm} + 88 \text{ mm} + 7.5 \text{ mm} + 61.5 \text{ mm} \quad (2.3)$$

The last three terms constitute the portion of rainfall which is transferred to the aquifer and from there is lost through lateral outflow and pumping wells and drainage galleries, which are installed in the aquifer. In fact the Geer watershed serves as a water supply catchment for the city of Liege and has been exploited in this respect since the nineteenth century. For the purpose of the numerical modeling exercise, we ignore the presence of pumping wells and drainage galleries and assume that the entire mean annual recharge of the aquifer, i.e. $Q_{\text{out}} + Q_{\text{Losses}} \approx 150 \text{ mm}$ is extracted from the aquifer through losses

across the northern flux boundary, hereby implying that the aquifer has reached a long-term equilibrium and therefore is in a steady-state. In this manner we have an exact estimation for the supply and demand discharges to be extracted at the various nodes of the pipe network representing the lateral inter-connections of the REW (see Section 6.2.5).

2.6.2. Software architecture

For the implementation of the REW approach four processing stages are identified. Stages deal about DEM analysis, REW analysis, pre-processing and a numerical solver that serves as a processor. The software for the Preprocessor and the Solver is implemented in an object-oriented architecture using the language C++. The REW analysis software, the pre-processing software and the numerical solver are developed and coded within the DAUFIN-project.

1. For the analysis of the DEM the public domain software TARDEM has been used. TARDEM is a freely available software which performs DEM analysis and extraction of the drainage network through series of modules. The underlying algorithms are implemented according to research in DEM analysis as explained by *Tarboton [1997]*. TARDEM calculates for a given Horton-Strahler threshold order a corresponding number of sub-basins through a flow accumulation matrix procedure. In addition the river pixels for each sub-basin are marked and their coordinates calculated and written to file.
2. For the extraction of the REW geometry and the calculation of the geometrical attribute parameters such as REW mantle segment surfaces, REW elevations, slopes and aspects, the REWANALYSIS software module has been developed. REWANALYSIS reads the marked pixel map for each sub-basin and identifies the neighbouring REWs for each REW. The contour curves separating two neighbouring sub-basins are calculated and their projections on the x and y axis of the global coordinate system are extracted. In addition mean values such as the average surface elevation, the average river elevation and the average surface slope are calculated. The latter one is given by equation (35), whereby the surface area and the surface area projection of each REW need to be calculated. Finally the results are written to file following the topology of the network tree.
3. For the solution of the balance equations a series of parameter values as well as boundary and initial conditions need to be implemented. This is done through an appropriate routine called PREPROCESSOR. First the geometrical information generated by REWANALYSIS are read. Subsequently the parameter values for the river geometry are calculated and stored for each REW. Next the soil parameters (hydraulic conductivity, Brooks-Corey and Van-Genuchten parameters) are read and stored. Next boundary conditions are imposed by indicating which external boundaries of the watershed allow a head dependant flow boundary conditions and which ones are no-flow boundaries. The results are once more written to file.
4. The actual solution of the balance equations is performed in a software module SOLVER. SOLVER reads in the output of PREPROCESSOR and specific calculated information required for the calculation of rainfall or evapotranspiration. The interconnected system of REWs and the corresponding zones are set up and subsequently the coupled system of non-linear ODEs are solved with a Runge-Kutta

ODE-solver. The results of the calculations (storages, velocities) are written to a respective output file for each REW at a given output time step.

2.6.3. Numerical experiments and results

The REW code has been tested on a synthetic watershed model of the Geer basin. For the basin a raster based DEM was available that, by use of the TARDEM software, served to generate a polygon map with REWs of the basin. In our work, a partitioning based on a second order Strahler network was applied resulting in 73 REWs of various sizes and shapes (see Figure 2.4). Specific topographic properties as average elevation and slope and REW attributes like size and length of the REW mantle were calculated by the REWANALYSIS software. With the Geer basin partitioned in 73 REWs, the model tests were performed. Model test focuses not only on interpreting the calculated hydrograph at the basin outlet but also on analysing the internal model behaviour that causes the production of runoff. Such behaviour is e.g. expressed by the change of the (averaged) saturation of the unsaturated zone by infiltration and the change of the size of the saturated overland flow source area in space over the simulation time. Model test were set-up with uniform distributions of model parameter values and atmospheric stresses over all 73 REWs. Soil properties like saturated hydraulic conductivity, porosity and depth of the soil profile were assumed homogeneously distributed over the basin and for each REW equal values are applied. Also the parameters required to model the soil characteristics curves of Brooks-Corey are modelled as uniformly distributed over all REWs. As such, each of the model parameter values of the subsurface model domains are defined a-priori and required no optimisation through model calibration. A similar approach has been applied to modeling atmospheric stresses. For modeling overland flow and river flow, the values of model parameters (geometric and physical) change with each REW and so over the entire basin. As stated in Section 6.2.5, the river geometry is calculated by the Leopold and Maddock hydraulic geometric relationships while the model parameters for the overland flow areas are calculated through analysis of the topography by TARDEM and REWANALYSIS.

2.6.4. Sensitivity analysis

When the code became operative after debugging, a large number of model tests were executed and the objective was to "understand" the model behavior by "physical reasoning". The model behavior was analysed by simple model sensitivity analysis and especially focused on understanding the internal model behavior that causes the generation of runoff production. Model test generally were executed over a 5-years period with constant rainfall input and evaporation loss and analysis were performed for various rainfall depth and model parameter values. Parameters selected for the sensitivity analysis are commonly applied calibration parameters in known physically based models such as SHE. In such models, generally the most optimum values are defined through model calibration, either by Trial and Error or by automated procedures.

Parameters selected are the horizontal hydraulic conductivity and the porosity. Changes in model parameter values are done prior to model simulation and model results are analysed for each model run. After the first tests applying constant rainfall inputs throughout the simulation, it appeared that the model became in steady-state after a period exceeding 10 year. Due to the calculation burden, however, we decided to limit the simulation runs

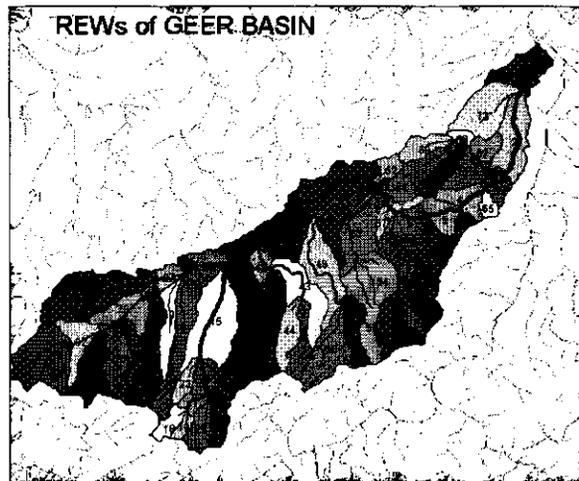


Figure 2.4.: Subdivision of the Geer basin into 73 REWs.

to a 5 year period this after water balance calculations proved that the model was mass conservative. The water balance loss was less than 1% which, in numerical modeling, is a common guide line to accept modeling results. In our work we defined a reference model case where the rainfall input was set to 1 mm. per day and where the saturated hydraulic conductivity was set to 1 meter per day. In Figure 2.5, comprising 6 windows, the modeling results are presented. In each window 6 graphs are presented for REWs with numbers 1, 9, 10, 53, 54 and 55. In Figure 2.4 it can be observed that these REWs make up a cross-section from the south boundary of the basin up to the river Geer. In the first window the channel discharge at the outlet of each REW is presented. It is shown that at the lower REWs near the outlet of the river Geer a steady increase of channel discharges takes place due to the accumulation of discharges from upstream REWs. The graphs show no discontinuity or sudden change in channel discharge and as such it was concluded that the overall model behavior, that resulted in the calculated channel discharge at the basin outlet, was correctly described by the model. In the third graph the area fraction of saturated overland flow is described. It is shown that for the lower REW this area fraction is significantly higher than for the upper REW which implements that runoff production by the saturation overland flow runoff mechanism [Dunne, 1978] is well described. It is concluded that a "build up" of groundwater storage in the lower REWs takes place by saturated flow from interconnected REWs. Groundwater is transported from the upper REWs to the lower REWs and results in large saturation overland flow source area fractions in the lower REWs. By the graphs it can be concluded that, within each REW, the saturation overland flow mechanism is well described sine the area fraction shows a continuous increase in time up till a steady-state flow situation is simulated. Therefore it is concluded that the runoff production mechanism by saturation overland flow for each REW as well as the entire model is well described by the model. In the windows 4 and 5 the internal model behaviour with respect to the unsaturated zone saturation and the unsaturated zone depth are presented. Both variables

are interpreted as model variables that have great effect on runoff production. At first, in both windows discontinuities in a number of graphs can be observed that are regarded as model initialisation effects. Similar effects are observed in many (other) REW trial runs and, moreover, such effects also are often observed in (complex) numerical models in which highly dynamic model behavior is simulated. In order to by-pass this phenomenon in model simulations it is common in numerical modeling to use intermediate results as the initial model conditions. Conclusions with regard to both variables are as follows. After initialisation of both variables it can be observed that for all REW both variables only show small changes as the simulation progresses over time. This implements that the nett volume of infiltration, evaporation and recharge only has a minor effect and that the volume of water stored in the unsaturated zone remains constant. As a consequence of this observed behavior it is concluded that infiltration fluxes, with some undefined but probably very small time delay, are added to the saturated zone of the model. In window 6 it is observed that water table elevations, especially for the lower REWs increase slowly but constantly over time. In the real world, however, the unsaturated zone acts as a large storage volume where rainfall is stored and, after some time delay, water is recharged to the saturated zone. By this reasoning the authors conclude that the modeling of the unsaturated zone requires modification. Extensive sensitivity tests with modified parameter and rainfall input values also showed similar model behavior.

2.6.5. Model Simulations

In the second stage of model testing the focus was on modeling the runoff behavior with model input of observed rainfall and evaporation data. At first, for these simulations also a five year period (1992-1996) was selected and model behaviour was analysed. In these model simulations all model parameters and atmospheric stresses by rainfall and evaporation were taken homogeneously distributed over the basin. In Figure 2.6 the calculated channel flow as caused by observed rainfall input is presented. By a quick visual inspection it is clear that the model responded well to changes in rainfall intensity and that the time delay between rainfall input and model output is relatively small considering the size of the basin. In Figure 2.7 a comparison is made between the observed and calculated discharges at the basin outlet at Kanne. By visual inspection it is obvious that significant deviations exist between both hydrographs and that base flow discharges as well as peak flow discharge do not match. Further testing focused on analysing the effects model input (i.e. rain) had on the calculated output and on analysing the internal model behavior that caused the generation of the high channel discharges. After extensive simulations and tests, it appeared that the model response due to changes in rainfall input is over-stressed by the size of the fully saturated model surface (i.e. river and saturated overland flow source area). In Figure 2.5 it is shown that area fractions of saturated overland flow for all REW rapidly increase from 0.1, after the effect of model initialisation were dampened out, to 0.18 for REW_1 and 0.65 for REW_55 after a 500 days simulation period. This rapid increase is due to the internal model behaviour that results in the size of saturated overland flow area and as such has great effect on the simulated model results (see Figure 2.6). Although the model was mass conservative it can be concluded and observed that to little storage takes place in both subsurface model zones that, in real world catchments, generally cause a dampening of the output discharges. Specific test on modeling rainfall input on the river surface only, also supported our observations that the size of the model area that receives rainfall as "direct" rainfall dominated the model output behaviour. During these test, however, the base flow discharges for these simulations remained much to

high and indicated that the interaction between river and the groundwater zone was incorrectly modelled. Both observations indicate that the average water table elevations are incorrectly calculated and cause on the one hand that the saturated overland flow area fractions are too large while, on the other hand, the base flow discharges are too high. Theoretically, groundwater drainage by river interaction causes a reduction of the water stored in the saturated zone and as such causes a decrease of average water table elevation and a decrease of the saturated area fraction. It is clear that both observations support the conclusion that the calculated water table elevation is over-estimated resulting in high base flow calculations and peak runoff rates by saturation overland flow. In further model tests the interaction between the river reach and the saturated zone was modified and resulted in significant model improvements as shown in Figure 2.8. Details on the modified approach are not presented yet since they are still under revision by the authors. In the figure it is clear that the peak discharges still are largely over-estimated and remains subject to future research. With regard to this aspect the authors like to stress that the Geer basin is a very deep basin with, especially at higher elevation, a very deep unsaturated zone. Such system characteristic has major effect on the catchment runoff behavior and is, to the authors, not well modeled by the REW code. (see also previous comments on unsaturated zone modeling). To the authors, this aspect makes it difficult to judge whether the mathematical model of the REW-approach is insufficient or that the current model structure in terms of the applied vertical discretisation is insufficient to model the specific runoff behavior of deep river basin systems such as the Geer basin.

2.6.6. Future Research

In order to be conclusive on the performance of the REW modeling approach, much more test runs are required. Such runs must extend the applied sensitivity analysis and effects of the applied catchment partitioning must be researched. Also the model must be tested on, at first, smaller scale basins where detailed information is available on physiographic and groundwater table data. Such data is required to allow the modeler to analyse the performance of the model on other hydrological state variables than the channel flow discharge that, basically must be interpreted as an integrated model output variable of all upstream flow processes. When groundwater table data is available also the internal model behavior can be analysed in more detail.

By the results it has become clear that the modeling of the unsaturated zone is insufficient with regard to the transport and water storage in the unsaturated zone. Infiltration fluxes and recharge fluxes, by approximation, are equally large and implements that changes in the unsaturated zone saturation are very small. It can be observed that, during periods of rainfall, the depth of the unsaturated zone decreases by the rise of the water table. Future research must focus on improving the modeling of the unsaturated zone allowing the storage of larger water volumes and allowing a time delay of infiltration water to become recharge. The authors therefore suggest that the unsaturated zone must be subdivided in multiple model layers.

With regard to the calculation of the saturated overland flow area fractions, it is foreseen that the approach needs some minor modification. By the modeling results it has become clear that, at least for the Geer basin, the saturated overland flow areas are too largely calculated. Future research must focus on developing modified expressions for either the calculation of the groundwater table depth or the manner the average land surface elevation is calculated. The assumption that the land surface elevation and the slope of each REW

may be averaged and that these parameter values may be used as "model effective" values probably is a too strict pre-requisite.

By the modeling results an urgent need is identified to developed and incorporate evaporation in a physically consistent and sound manner. At this stage of code development, evaporation is modelled in a very simple manner that requires improvement in order to allow the more detailed modeling of evaporation during inter-rainfall periods. Evaporation greatly effects the soil moisture saturation prior to a rainfall event and as such effects the model behaviour.

Future research also must focus on combining the REW approach to 4DDA and automated multi-objective model calibration to improve model simulation and forecasting. So far, due to the enormous amount of time required for code development, debugging and testing, the REW model could not be combined with 4DDA techniques. In the near future, however, after the code has been further developed, the use of multi-objective model calibration and 4DDA is foreseen and regarded an absolute necessity to future model applications.

2.7. Conclusions

The present chapter introduces the governing equations required to model a whole watershed underlain by an aquifer through the representative elementary watershed approach. The chapter states the governing equations for mass and momentum. The momentum balance equations are vectorial and are therefore projected along the axes of an appropriate reference system. Subsequently the constitutive relationships are presented, which are used to close the unknown mass and momentum exchange terms found in the governing equations. Finally there are still a series of parameters appearing in the balance equations, which need to be expressed in terms of either the independent variables for which the governing equations are solved, or in terms of quantities which are directly measurable in the field. In addition a theoretical procedure has been pointed out, which allows to estimate the exchange coefficient for the partition of mass between a REWs and its neighbours.

In the numerical modeling, the high potential of the REW-approach could not fully be explored. At this stage of code development it most likely may be concluded that the model structure of the REW-model requires some minor modifications. First however, much more model tests on smaller, and preferably, shallow catchments must be executed in order to understand the model behavior in more detail. The overall model behavior, however, has proven to be successful in modeling the runoff production mechanism of saturation overland flow and the routing of water in the channel network system. As such, the semi-distributed modeling approach with interconnected REWs and the applied 5 model domains is a very promising approach.

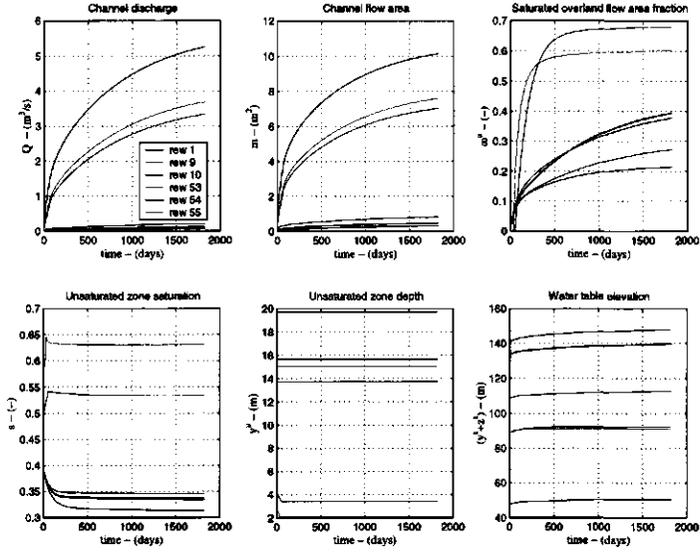


Figure 2.5.: Sensitivity of 6 REW state variables, for some selected REWs (numbers 1, 9, 10, 53, 54 and 55).

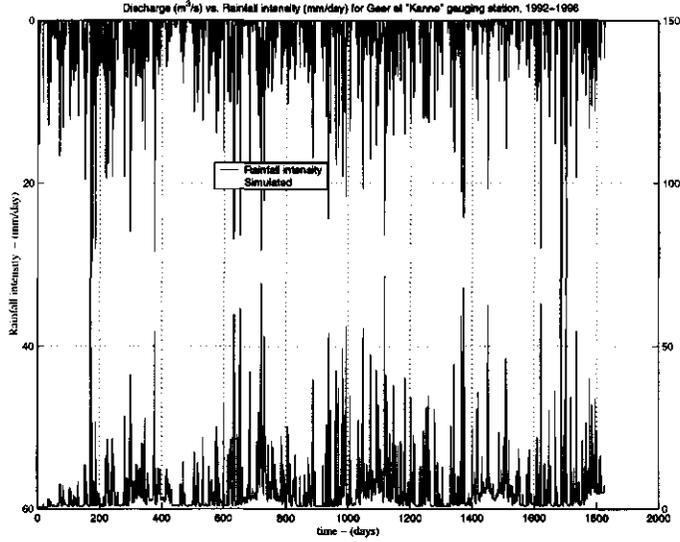


Figure 2.6.: Rainfall and simulated channel flow for the Geer at Kanne over the simulation period.

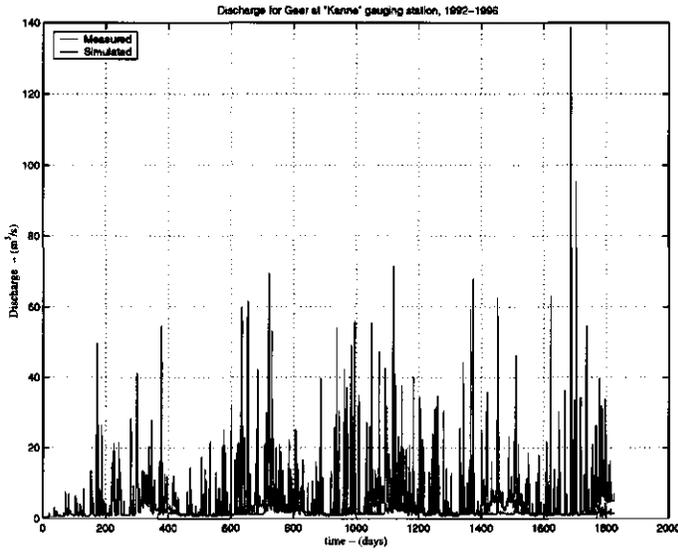


Figure 2.7.: Measured and simulated channel flow for the Geer at Kanne over the simulation period.

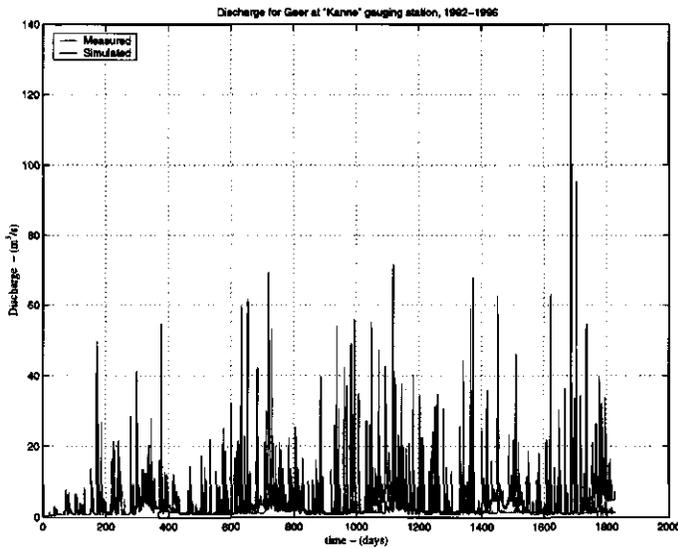


Figure 2.8.: Measured and simulated channel flow for the Geer at Kanne over the simulation period, after model-modification.

3. A review of existing Data Assimilation techniques

Emiel van Loon and Peter Troch

3.1. Introduction

The significance of soil moisture is due to its large influence on the partitioning of radiant energy at the ground surface into sensible and latent heat (evapotranspiration), as well as its important role in partitioning of rainfall into infiltration and runoff [Eagleson, 1970; Schmugge, 1985]. As a result soil moisture in the root zone is a key parameter in numerous environmental studies. Part of those aim at the prediction of soil water storage and movement in a heterogeneous landscape to improve the management of water resources for agricultural or domestic purposes. Another group of studies aim at the improved estimation of surface soil moisture as a boundary condition for weather systems to enhance weather and climate prediction. The latter type of studies typically use more data from airborne and satellite remote sensing sources, the models operate at coarser resolutions and cover a larger study area than the former [Entekhabi and Njoku, 1994; Jackson et al., 1996; van Loon and Troch, 2002]. An enormous variety of models is being used for these purposes, but these generally yield unsatisfactory results in open-loop form (using only atmospheric inputs and initial conditions) [van Dam, 2000]. This implies that these models will have to be conditioned with additional soil moisture observations for adequate performance. Fortunately such observations are available at increasing quality and quantity, owing to innovations in field monitoring as well as airborne and satellite remote sensing techniques [Georgakakos and Baumer, 1996a]. The combination of these factors create a need for a formal methodology to integrate simulation models and data in a more automated way than what has been the practice to date, leading to a much more intensive use of the various resources and hence more accurate predictions and scenario analyses. In this way data and models become more useful and reliable for decision makers and others involved in operational water resources management [U. S. National Research Council, 1991].

Data assimilation is a technique to provide physically consistent estimates of spatially distributed environmental variables. It does so in an automated way via a methodology that, mathematically, can be viewed as a state estimation problem [McLaughlin, 1995]. By incorporating error terms into the state equations and using model ensembles, data assimilation also addresses some of the problems related to uncertainty in models and data. This helps to assess the worth of various model structures and data sources and to use this data in an optimal way to validate or drive hydrological models [U. S. National Research Council, 1991; van Loon and Troch, 2002].

Following its conception in the systems and control disciplines and preceding the application to hydrology, data assimilation has been successfully applied to improve the predictive power of short-term weather forecasting models and to provide the formal framework of dynamic analysis of oceanographic models [Daley, 1991; Bennett, 1992; Ghil and Malanotte-Rizzoli, 1991]. McLaughlin [1995] provides an excellent review of the state of the art in hydrological data assimilation and explains the origins of the various techniques in the oceanographic and meteorologic disciplines. Despite the fact that this study already touches upon many of the issues that will be dealt with here, it is not very specific on soil moisture. Moreover, a lot of new work has been done in the field of soil moisture data assimilation since the appearance of this review.

The studies on soil moisture data assimilation are still limited to one-dimensional (vertical) soil water movement, utilizing only rain, evapotranspiration, and mostly remotely sensed soil moisture estimations as observations. Examples are found in Calles et al. [1998], Calvet et al. [1998], Galantowicz et al. [1999], Hoeben and Troch [2000a], Houser et al. [1998b], Katul et al. [1993] and Mahfouf [1991]. As such, the present data assimilation techniques do still not fulfill their promise [McLaughlin, 1995]. However, as shown in this study, there are various options to extend the current techniques to two or three spatial dimensions and to incorporate other hydrological observations such as discharge, ground water level as well as static soil and terrain properties. We aim at outlining the problems with the different extensions and suggest possible solutions.

To this end we first review the state of the art in soil moisture data assimilation in Section 6.3.2 with an overview of recent studies. In Section 6.3 an idealistic data assimilation approach is described. The section discusses the consequences of extending the present techniques towards the suggested approach as well as the main problems concerning data handling.

3.2. A Review of Soil Moisture Data Assimilation

3.2.1. Overview

Various classification schemes can be applied to characterize soil moisture estimation studies, e.g. those according to: problem size (small scale vs large scale), statistical rigour (optimal vs various sub-optimal methods), computational implementation (recursive vs non-recursive), or type of data to be assimilated (many choices).

In general, it appears that the larger the problem, the less statistical rigour can be applied, due to the computational requirements of statistically optimal techniques. By statistically optimal (the antithesis of sub-optimal) we mean that the estimations result from the optimization of an objective function, and that the model equations are derived through a consistent simplification such that the approximation error can be quantified in some way. The computational implementation, although relevant when techniques are implemented for operational purposes, are less relevant from a hydrological viewpoint. And albeit there are many possibilities with regard to the type of data to be assimilated, nearly all existing studies use either direct measurements of soil moisture or remotely sensed brightness data which are very closely related to surface soil moisture.

As a result of this we will use a subdivision according to statistical rigour and consider only selected works, focusing on studies that apply statistically optimal or non-trivial sub-optimal estimation techniques. A first group of studies use synthetic data [Entekhabi and

Njoku, 1994; Milly and Kabala, 1986a; Hoeben and Troch, 2000a], or small-scale field or laboratory data [*Katul et al., 1993; Parlange et al., 1993; Galantowicz et al., 1999; Calvet et al., 1998; F. Castelli and Caporali, 1999; Hoeben and Troch, 2000b*]. A second group considers large-scale field data sets [*Houser et al., 1998b; Reichle et al., 2000, 1999*]. All studies integrate the data in spatially one-dimensional physical models and do not include any prior information other than model parameters (i.e. no smoothness constraints or other regularizing information). In the first group of small-scale studies, the dimensions of the state vector and the observation vector are small (up to a few dozens of state variables), and hence optimal estimation is applied. The models in the second group are horizontally distributed and of high dimensionality (starting from hundreds of state variables). Consequently, only sub-optimal estimation techniques have been implemented in these studies. We will discuss the literature for these two classes in sections 3.2.2 and 6.2.3.

3.2.2. Optimal approaches

Optimal estimation techniques are applicable if the dimension of the soil moisture system is relatively small, e.g. if it contains only a single vertical column. Algorithmically the problem of optimal estimation can be solved by both recursive and non-recursive techniques. The recursive algorithm for optimal estimation is the (extended) Kalman filter, which has been used by many investigators [*Entekhabi and Njoku, 1994; Galantowicz et al., 1999; Hoeben and Troch, 2000a*]. The non-recursive technique, often called the variational approach [*Bouyssel et al., 1999; Callies et al., 1998; Mahfouf, 1991; Reichle et al., 2000, 1999*], has been used much less in soil moisture data assimilation.

In cases where optimal estimation has been applied, the soil moisture system has been described as a simple linear multi-layer model [*Milly, 1986*], a non-linear bucket model, obtained from the depth-integration of a one-dimensional version of Richards' equation [*Katul et al., 1993; Parlange et al., 1993*], or a multi-layer version - often in combination with the heat equation, and a model for the radiative transfer [*Milly, 1986; Entekhabi and Njoku, 1994; Galantowicz et al., 1999; Hoeben and Troch, 2000a*].

The pioneering work in applying optimal estimation to soil moisture data assimilation has been done by Milly [*Milly and Kabala, 1986a; Milly, 1986*]. Milly and Kabala [*1986a*] applied a Kalman filter to determine the relative merits of the accuracy and the sampling frequency of the observations with a remote soil moisture sensor, using a linear soil moisture model. Milly [*1986*] demonstrate the use of an Extended Kalman filter in combination with a two-layer Richards' equation model and the heat equation .

A follow-up from this initial work is found in Katul et al. [*1993*], who use an extended Kalman filter for the estimation of the soil moisture state in a simple bucket model, derived from the depth-integration of a one-dimensional Richards' model. In addition to the state estimation, Katul et al. [*1993*] also estimate the two soil hydraulic parameters in their model, the initial estimation error variance, and the model error of the state-space formulation. Parlange et al. [*1993*] and Cahill et al. [*1999*] apply the same techniques to solve the diffusion equation linked with a water balance model. In this way they estimate diffusivity at field scale as well as the initial estimation error variance and the model error of the state-space formulation.

A next refinement has been introduced by Entekhabi and Njoku [*1994*] and Galantowicz et al. [*1999*]. They apply a multi-layer model of soil moisture and temperature dynamics,

and use a Kalman filter to update the temperature and moisture profile from brightness temperature observations. The spatially one-dimensional model is entirely physically-based, making use of Richards' equation, the heat equation, and radiative transfer equations. *Entekhabi and Njoku [1994]* are the first to show that it is possible to infer information about the temperature and the moisture at depths below the penetration depth of the microwaves. The study of *Entekhabi and Njoku [1994]* uses synthetic data, representing a single soil column without vegetation. In their study the updates of the brightness and infrared temperature data are made hourly, which is not a very realistic situation. At this point *Hoeben and Troch [2000a]* and *Hoeben and Troch [2000b]* provide a follow-up, using synthetic data as well as multi-frequency active microwave observations to determine the observation interval whereby the algorithm breaks down. *Galantowicz et al. [1999]* apply their study to data from the Beltsville Agricultural Research Center, over a seven-day period of drying in July 1994 [*Jackson, 1997*]. They arrive at similar conclusions as *Entekhabi and Njoku [1994]*. In addition *Galantowicz et al. [1999]* conclude that updating the brightness temperature only once every three days suffices for the estimation of the soil moisture profile.

Calvet et al. [1998] present a comprehensive study on the feasibility of retrieving root zone soil moisture from surface soil moisture or surface soil temperature observations. They use the Interaction between Soil, Biosphere, and Atmosphere surface scheme of the French weather forecast system, which models soil moisture in a shallow surface layer and a deep reservoir. The assimilation technique is a strong-constraint variational method. The uncertain parameter is the initial soil moisture of the deep reservoir. The data originate from two months of field observations in Spring and Fall 1995 in southern France. The assimilation period is either a moving fifteen-day window or a moving five-day window during the thirty-day observation periods. In a series of assimilation experiments, observations are available to the estimation algorithm from twice daily to once every four days. *Calvet et al. [1998]* state that deep soil moisture can indeed be retrieved with reasonable accuracy from surface soil moisture observations once every three days, but concede that soil moisture estimation from soil temperature measurements can at best work under dry conditions. In addition they conclude that the length of the assimilation window should not be less than ten days.

F. Castelli and Caporali [1999] and *Lakshmi [2000]* use ground temperature observations for soil moisture estimation. In these studies the input is treated as an uncertain time-dependent parameter which is called soil moisture index. In both studies a variational technique is applied whereby the surface energy balance is included as a physical constraint in the objective function. Mathematically, the estimation of the soil moisture index amounts to the estimation of a state-dependent model error term. Estimates of the surface heat flux and the soil moisture index are derived from the FIFE. The experiments cover the summer months of 1987 and 1988, but the individual assimilation windows are limited to one day. Daily averages of the estimated surface heat flux compare well to independent latent heat flux observations. However, *F. Castelli and Caporali [1999]* conclude that there is a need to discriminate between soil moisture and aerodynamic contributions to the surface control over evaporation.

3.2.3. Sub-optimal approaches

So far, the most comprehensive studies on soil moisture data assimilation have been carried out by *Houser et al. [1998b]* and *Reichle et al. [2000]*. *Houser et al. [1998b]* modified

and extended the TOPLATS land-atmosphere transfer scheme [Famiglietti and Wood, 1994b,a].

TOPLATS is a spatially distributed hydrologic model to predict the diurnal dynamics of the water and energy fluxes at the land surface as well as the local vertical recharge into the underlying aquifer. The basic components of TOPLATS are water balance equations for the canopy and the soil as well as an energy balance equation at the surface. The original model describes the unsaturated zone with two layers, a root zone and a transmission zone. Houser *et al.* [1998b] added a shallow third soil layer at the top, the moisture content of which can potentially be inferred from remotely sensed microwave observations. The soil moisture dynamics are based on an approximate analytical solution of Richards' equation using infiltration and ex-filtration capacities. Horizontal flow exists only in the underlying saturated layer. In the unsaturated zone lateral flow is completely neglected. The model is applied to the Walnut Gulch watershed in southeastern Arizona. Houser *et al.* [1998b] apply five different data assimilation methods (from simple to more complex): direct insertion, statistical corrections, two forms of nudging and optimal interpolation in combination with two ways of data reduction.

A more advanced data assimilation study to a large study domain has been performed by Reichle and co-workers [Reichle *et al.*, 2000, 1999]. They apply a model of coupled moisture and heat transport from Ács *et al.* [1991], using six layers in the vertical direction. For computational efficiency the force-restore method [Hu and Islam, 1995] rather than the full heat equation are used to describe temperature dynamics. In addition a vegetation sub-model is used analogous to the Simplified Biosphere Model by Xue *et al.* [1991]. The downward flux out of the bottom soil layer is described by gravitational drainage. The brightness temperature is related to the system states with the Radiative Transfer model described by Galantowicz *et al.* [1999]. As in the study of Houser *et al.* [1998b], these studies also neglect lateral unsaturated horizontal moisture and heat transport, assuming that in the relatively flat study area the vertical fluxes dominate in the unsaturated zone. For the soil parameters and the meteorological forcings the error covariances are assumed to be correlated exponentially in space as well as time.

3.2.4. Discussion

Rather than summarizing the achievements, we will concentrate on points where progress is still needed.

First of all there is the very limited use of different data sources: nearly all studies only utilize field or remotely sensed brightness data. There have been many interesting investigations on soil moisture data assimilation from low-level atmospheric parameters such as air temperature and relative humidity at 2 m above the ground [Bouyssel *et al.*, 1999; Bouttier *et al.*, 1993a,b; Calies *et al.*, 1998; Hu *et al.*, 1999; Mahfouf, 1991; Rhodin *et al.*, 1999]. However, these parameters are only weakly and indirectly related to surface soil moisture. In addition, the studies using these data are geared towards improving numerical weather prediction and treat soil moisture rather as a tuning parameter. For this reason these were not considered in this review. Still these studies contain interesting contributions that may be applied in other soil moisture data assimilation studies.

The reason for not including data such as catchment discharge, evapotranspiration from different vegetation patches or ground water levels is twofold. In the first place the 1D-models that have been applied so far do not lend themselves for the inclusion of any lateral fluxes. Secondly, no suitable data set has been compiled so far that includes such a variety

of field and RS-data - a problem that has been noticed previously [*Kostov and Jackson, 1993*].

None of the above data assimilation studies has incorporated any regularizing information [*Tarantola, 1987*]. Regularization is a generic term for mathematical techniques that enable to solve an ill-posed problem by varying problem resolution and include prior information in a structured and objective way. Such techniques were not required until now simply because the 1D-problems were not ill-posed. Nonetheless, a clear-cut regularization technique will be required for solving higher dimensional to decide which prior information to utilize, how to weigh it against (posterior) observations and, most importantly, how to vary the resolution of the modeling problem.

The present data-assimilation studies concentrate on the mathematical and computational implementation of the proposed techniques, and on the relationship between remote sensing brightness data and surface soil moisture (i.e. the measurement equation). Much less effort has gone into the re-design of hydrological models or the extension to more spatial dimensions. In addition, no advantage has been taken of the existing work which has quite successfully related soil moisture to static soil and terrain properties with statistical techniques [*Greminger et al., 1985; Schiffler and Bárdossy, 1991; Unlu et al., 1990; Yeh et al., 1986*]. In some of these studies systematic components have been identified and linked to topographic characteristics, [*Moore et al., 1991, 1993; Hanna et al., 1982; Hairston and Grigal, 1991*], soil morphological features [*Kreznor et al., 1989*], or chemical and physical attributes [*Brubaker et al., 1993*]. And the integration of both systematic and stochastic components has partially been achieved by conditioning geo-statistical techniques with secondary data such as topographic indices via (indicator) co-kriging [*Lehmann, 1995; Western et al., 1998, 1999*]. Especially these latter (geo-statistical) techniques closely fit into the data assimilation framework. The lack of this integration can (again) be explained from a lack of suitable field studies.

Very much related to the previous problem, the focus in assimilation studies has been on using only a single (assumably unbiased) model, and never on model ensembles. A consequence has been that the role of model-error and bias has never been assessed, and techniques for building and using model ensembles have not been developed. In the meteorologic and oceanographic disciplines the use of model ensembles to reduce model bias and estimate the forecast-error statistics is gaining acceptance since the work by Evensen [*Evensen, 1992; Houtekamer and Mitchell, 1997*]. Recently it has also been shown that within a Kalman-filter framework, ensemble techniques can help to reduce the computational burden for realistic higher dimensional problems to a reasonable level [*Mitchell and Houtekamer, 2000*].

However the knowledge and technical means are available, the first operational application of soil moisture data assimilation has still to be made. On one hand one can argue that it is still too early to ask for applications when the data assimilation techniques are still so immature and data not even operationally available. But on the other hand it is quite viable to state that data assimilation is pre-eminently a branch of science which should be application-driven, and that the availability of operational tools may stimulate the development of the appropriate observing systems. In addition, one could argue that it is more appropriate to start from simple operational models and data assimilation techniques, and enhance these stepwise, than first developing advanced models and data assimilation techniques while implementing these in an operational setting at a later stage. The (successful) development of data assimilation in meteorology is exemplary for the first strategy [*Daley, 1991; Courtier et al., 1993*].

The main achievements over the past decade since the work of Milly [*Milly and Kabala, 1986a; Milly, 1986*] have been made in establishing that optimal data assimilation algorithms can indeed be applied for inferring soil moisture from remote sensing observations of passive or active microwave data. Physically based models, combining the energy and water balance equations as well as radiative transfer equations, have been successfully applied in data assimilation procedures. The following generalizations can be made with regard to the work that still remains to be done in soil moisture data assimilation:

1. making use of different observation types, i.e. not only from different RS-sources (visible light, thermal infrared and microwave measurements), but also from discharge and piezometric observations;
2. incorporating a regularization component into the data assimilation algorithms;
3. applying data assimilation to distributed problems (i.e. beyond one dimension) of a realistic size;
4. utilizing the known relations between static soil and terrain properties and soil moisture distribution;
5. applying ensemble techniques; and
6. implementing simple data assimilation techniques in an operational setting.

3.3. A prototype for 4D hydrological data assimilation algorithms

3.3.1. Algorithmic structure

A possible algorithm to include many of the aspects discussed in section 6.3.2 is shown in Figure 3.1. The figure shows that multiple data sources form input to different data assimilation systems, and that the output of these may constrain each other. As an example, in Figure 3.1 the data assimilation systems give: 1) catchment runoff by a lumped conceptual model, 2) vertical fluxes in the unsaturated zone through a 1D physically based model, and 3) lateral movement in the unsaturated zone through a 2D statistical model. The output of sub-system 1 places a hard constraint on sub-systems 2 and 3, to impose e.g. mass conservation at the catchment scale. The outputs of sub-systems 2 and 3 provide boundary conditions to each other. The choice to give subsystem 1 its important role is given in by the uncertainty of this system component relative to the others, which is determined by the quality of observations and system resolution. In case there are two models which estimate the same system component, their relative importance can be weighted by using regularization techniques such as generalized cross-validation, singular value decomposition or a so-called l-curve criterion [*Tarantola, 1987; Hansen, 1992, 1998*]. Imposing the hard constraints can be achieved by a.o. a singular value decomposition [*Wunsch and Minster, 1982; Menke, 1989*]. A demonstration that the scheme of Figure 3.1 may work effectively for realistic problems, is given in *van Loon and Troch [2002]*.

3.3.2. The practical consequences for model forms and data handling

It is important to notice that the structure proposed in Section 6.3.1 is not very specific: each sub-system may comprise a single model but equally well an ensemble of models, and each sub-system may in addition comprise an open-loop model but also a data assimilation scheme. However, with regard to one point the proposed generic structure is explicit: it relies on recursive model structures - otherwise the different sub-systems can not effectively interact. As a consequence, the algorithm will function best when using models that are easy to linearize, or can provide the Jacobian of the system equation analytically. It has been shown that good way to achieve this, may be via linear models which are variable in the parameters [Young and Beven, 1994; van Loon et al., 2000] or via ensemble filtering techniques [Evensen, 1992; Houdekamer and Mitchell, 1997; Mitchell and Houdekamer, 2000]. The latter technique kills two birds with one stone: it makes the redefinition of non-linear systems redundant, and is computationally also very effective.

Assuming that the required data are available in sufficient quantities, properly stored and documented, the primary requirement with regard to data handling is a set of robust pre-processing routines. With pre-processing we mean a set of empirical tools to detect outliers, statistical tools to aggregate or disaggregate observations to the model resolution, and static (possibly physically based) equations to relate observations to model-state variables (i.e. measurement equations).

For post-processing similar tools could be required, but then applying to the output from the data assimilation algorithm. Presently no generic pre/post-processing routines exist for hydrologic variables, implying that a significant investment is required before soil moisture data assimilation can be implemented operationally. It is notable that, in contrast with data assimilation schemes, the majority of pre/post-processing routines will be applicable to a wide range of different systems (with the exception of some measurement equations).

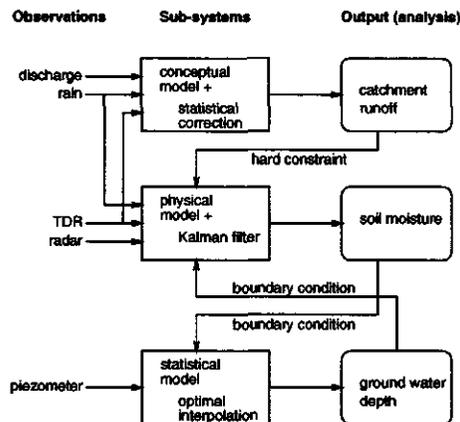


Figure 3.1.: Schematic representation of a data assimilation algorithm, comprising three sub-systems for soil moisture estimation in a catchment, using multiple models.

4. Developing and testing a nudging technique in a 3D-Richards model

Claudio Paniconi and Marino Marrocu

4.1. Introduction

In this chapter a fairly simple data assimilation technique, Newtonian nudging, will be implemented for a detailed numerical model of coupled surface and subsurface flow. The Darcy, Richards, and convection-diffusion equations that form the basis of the coupled model are widely accepted mathematical representations of verifiable conservation principles, so this physically-based distributed model, capable of reproducing the spatial and temporal patterns of pressure head, soil moisture content, saturation, and other variables, is a good candidate for data assimilation. Coupling in the model is via interaction between surface ponding heads generated by the subsurface model, computed as the residual between potential and actual atmospheric inputs (rainfall and evaporation), and the overland accumulation (in lakes and reservoirs) or routing (on hillslopes and in channels) of this ponded water [Putti *et al.*, 2002]. The nudging technique has been implemented for hydrological applications by Houser *et al.* [1998a] and Pauwels *et al.* [2001], who used TOPMODEL-based soil-vegetation-atmosphere transfer models. This paper focuses instead on a more numerically complex Richards equation-based model, and we also wish to examine some of the computational aspects of the implementation and performance of the nudging method that have not yet been addressed in detail.

The state variable selected for assimilation in this work is surface soil moisture. The importance of soil moisture in hydrology is such that the study of its spatial and temporal variability continues to receive a lot of attention, supported by efforts to improve the observation database for soil moisture across a range of scales [Georgakakos and Baumer, 1996b; Islam and Engman, 1996; Western and Grayson, 1998; Yu *et al.*, 2001; Robock *et al.*, 2000]. The structure of the coupled numerical model used in this study is such that, for surface soil moisture as the assimilation variable, only the subsurface module of the code needs to be adopted in order to introduce the nudging algorithm. This and other advantageous features of the implementation will be pointed out in the development.

The overall aims of this study are to describe the implementation of Newtonian nudging for a coupled Richards equation-based model, to demonstrate in a simple synthetic application how data assimilation can add value with respect to stand-alone data and model predictions, and to assess the numerical performance of the assimilation model, including the sensitivity to nudging parameters and the impact of nudging on convergence, mass balance, time stepping, and computational efficiency.

4.2. Selection of a data assimilation technique

Data assimilation is routinely used in research and operational meteorology, and has more recently been introduced in the oceanographical and hydrological sciences, owing to the trend towards better and more regular observation of a wide range of parameters of interest to the Earth sciences and to a strong need to address global change and other environmental problems. Although data assimilation is still relatively new to hydrology, a number of papers on closely related topics go back as far as 20 years, for instance studies on the incorporation of remotely sensed data into hydrological models using algorithms such as the extended Kalman filter, in particular with regards to the problem of soil moisture profile estimation from near-surface observations [Newton et al., 1983; Milly and Kabala, 1986b]. As argued in Chapter Indeed soil moisture continues to be a common focus for data assimilation in hydrology, with applications of importance to agriculture, water management, and meteorology such as root zone or depth-averaged soil moisture estimation, water table extrapolation, and land surface initialization for climate models [Entekhabi et al., 1994; Hoeben and Troch, 2000c; Walker and Houser, 2001]. Less has been done, however, on methodological development and inter-comparison of data assimilation techniques covering a broad range of hydrological models, scales, and state variables, so it is felt that experiments with a detailed numerical surface–subsurface flow model using even simple data assimilation techniques will provide some useful insights. Moreover the implementation lends itself readily to utilization of information from different observation sources, both satellite (grid-based) and ground (scattered) data.

One of the factors still hindering progress in hydrological data assimilation and in the operational use of hydrological models is incomplete knowledge of the spatial and temporal variability of important hydrological processes and state variables such as soil moisture, rainfall, evapotranspiration, and hydraulic conductivity. Adequate characterization of this variability is needed in the measurement equations and interpolation/extrapolation methods used in many data assimilation schemes. For surface soil moisture, for example, widely varying correlation lengths have been reported in the literature, and indeed it is entirely reasonable even at the hillslope scale to expect a range of characteristic scales owing to diurnal and seasonal cycles in rainfall and evapotranspiration and to topographic and geomorphologic factors that influence soil wetness patterns [Capehart and Carlson, 1997; Grayson and Blöschl, 2001; Yu et al., 2001]. As will be described, the weighting functions used in the nudging method provide a simple and intuitive way of incorporating any available knowledge regarding the characteristic length scales of the state variables being assimilated, and of attributing this information in a spatially and temporally variable manner so that the data assimilation technique is not restricted to invariant correlation structures.

Data assimilation methodologies range from simple, suboptimal techniques such as direct insertion, statistical correction, and statistical interpolation to more sophisticated, optimal algorithms such as variational techniques and a family of methods based on the Kalman filter [Daley, 1991; U. S. National Research Council, 1991; Reichle et al., 2001]. We selected nudging for a first implementation of data assimilation in our model since it possesses, in part, some of the advantages of four-dimensional data assimilation, such as dynamic consistency and propagation of observation information, while remaining relatively straightforward in formulation and much less computationally intensive than other schemes, both important considerations for three-dimensional nonlinear numerical models.

4.3. Newtonian relaxation

Newtonian relaxation is a four-dimensional data assimilation procedure whereby model variables are driven (nudged) toward observations by adding to the forcing F of a model equation $\partial v/\partial t = F(v)$, where v is the state variable and t is time, an additional term with the aim to relax the actual model state to the observed one [Davies and Turner, 1977; Stauffer and Seaman, 1990; Houser et al., 1998a]. This relaxation term can be taken to be active for a certain period of time t_a , called assimilation time, either preceding the observation time t_o or both before and after t_o . After time t_o (in the case of a backward time influence window) or $t_o + t_a/2$ (centered window) is reached, the relaxation term is deactivated and the model equation is integrated in its original form. The main objective is to obtain, at least partially, a dynamical consistency between measured data and numerical solution of the model.

When a gridded analysis of the observations is available, i.e., an interpolation of observation data to the model grid, our dynamical model with nudging can be expressed in the form

$$\frac{\partial v}{\partial t} = F(v, \mathbf{x}, t) + GW(\mathbf{x}, t)\epsilon(\mathbf{x})(v_o - v) \quad (4.1)$$

where the nudging term active during the assimilation time is proportional to the difference between the model solution v and the observed state of the system v_o . In this equation $\mathbf{x} = (x, y, z)^T$ is the Cartesian spatial coordinate vector, G determines the relative strength of the nudging term with respect to the physical forcing term, $W(\mathbf{x}, t)$ are weights to be specified as functions of space and time (see below), and $\epsilon \leq 1$ is a factor that reflects the accuracy of the observation (equal to 1 for perfect measurements).

In the case where observations are not available at all grid points, a formulation of the nudging term that allows the system equation to be relaxed to individual scattered observations is [Stauffer and Seaman, 1990; Houser et al., 1998a]

$$\frac{\partial v}{\partial t} = F(v, \mathbf{x}, t) + G \frac{\sum_{k=1}^{N_T} \sum_{i=1}^{N_X} W_{ki}^2(\mathbf{x}, t)\epsilon_i(v_{o_{ki}} - v_i(t))}{\sum_{k=1}^{N_T} \sum_{i=1}^{N_X} W_{ki}(\mathbf{x}, t)} \quad (4.2)$$

where N_T is the number of observation times, N_X is the number of observation points, and $v_i(t)$ is the model variable interpolated to the observation point. A variant of equation (4.2) is proposed by Houser et al. [1998a] that is more appropriate for cases where there is a high density of observation data. We have discretized and implemented equation (4.2) for our model, but special cases to handle gridded or high-density data can be readily included as well. In this way the nudging technique can assimilate both remote sensing grid-based data or in situ measurements that are typically not uniformly distributed in space. The two types of data are of course complementary, and ground data is particularly important in small watershed applications and in providing temporal coverage that may not be obtainable from remotely sensed data alone.

4.3.1. Weighting functions

The nudging module has been implemented using Cressman-type (distance-weighted) functions [Stauffer and Seaman, 1990; Houser et al., 1998a]. We write $W(\mathbf{x}, t) = W_1(x, y)W_2(z)W_3(t)$ and define the three components as (Figure 4.1)

$$\begin{aligned} W_1(d) &= (R_{xy}^2 - d^2)/(R_{xy}^2 + d^2), & d^2 \leq R_{xy}^2 \\ W_1(d) &= 0, & d^2 > R_{xy}^2 \end{aligned} \quad (4.3)$$

$$\begin{aligned} W_2(z) &= 1 - (|z - z_o|)/R_z, & |z - z_o| \leq R_z \\ W_2(z) &= 0, & |z - z_o| > R_z \end{aligned} \quad (4.4)$$

$$\begin{aligned} W_3(t) &= 1, & |t - t_o| < \tau/2 \\ W_3(t) &= (\tau - |t - t_o|)/(\tau/2), & \tau/2 \leq |t - t_o| \leq \tau \\ W_3(t) &= 0, & |t - t_o| > \tau \end{aligned} \quad (4.5)$$

where R_{xy} and R_z are, respectively, the horizontal and vertical radii of influence, $d^2 = (x - x_o)^2 + (y - y_o)^2$, x_o , y_o , and z_o are the spatial coordinates of the observation point, and τ is the half period of the time influence window (i.e., $\tau = t_a/2$ when the assimilation time t_a is centered with respect to the time of observation t_o).

In equation (4.2) we write W_{ki} and ϵ_i because the implementation allows spatially (and for τ also temporally) variable nudging parameters ϵ , R_{xy} , R_z , and τ . Thus for instance within drainage or contributing areas along a channel network where soil moisture is characterized by high connectivity or spatial organization *Western et al. [2001]*, it may be of interest to use a higher R_{xy} value than in drier, less connected upslope recharge areas. Similarly when the upper and lower soil moisture zones can be considered to be decoupled owing to geomorphological or seasonal factors [*Zecharias and Brutsaert, 1988; Capehart and Carlson, 1997; Grayson et al., 1997*] a low value of R_z can be used. Temporally, during extended inter-storm periods or for dry soils (where large changes in pressure head cause only small changes in soil moisture) t_a can take on larger values. Finally ϵ can be tailored to the reliability and accuracy of different measuring stations or data sources distributed over the catchment.

It should be remarked that the assimilation time parameter t_a interacts in an important manner with the relaxation term ($v_{o_{ki}} - v_i(t)$) and should thus be selected with care. If the time influence window is so large that the simulated values at time t are uncorrelated with the observations at t_o , ($v_{o_{ki}} - v_i(t)$) will represent a large error that gets introduced into the model dynamics because the time weighting function at time t is still active (nonzero). At the other extreme this interaction also justifies the choice of a plateau-shaped time weighting function as opposed to the hat-shaped functions for spatial weighting. This shape assigns a high weight (equal to 1) in the immediate neighborhood of the observation time t_o to compensate for the comparatively small values of ($v_{o_{ki}} - v_i(t)$) expected for t close to t_o .

Note in equation (4.2) that since the relaxation term $v_{o_{ki}} - v_i(t)$ uses the computed solution at the current rather than observation time, the numerical implementation developed in the next section retains its time marching nature (there is no need to simultaneously solve Richards equation at all time steps which occur within the time influence window), and moreover, by assuming that the observation times are known a priori (i.e., can be read in at the start of the simulation) we avoid having to march backward to recompute the solution, with the nudging term activated, for all time steps from $t_o - \tau$ to t_o . Finally, prior knowledge of the observation times makes it relatively straightforward to handle the case of overlapping (concurrent) observation times, useful in applications with frequent observations and/or large time influence windows relative to the numerical time step of the model. The contributions from overlapping observation datasets at any give time are simply summed in the computation of the nudging forcing term.

4.4. Implementation of nudging for the hydrological model

4.4.1. Model description

The CATHY (CATchment HYdrological) model simulates overland and subsurface flow by coupling the three-dimensional Richards equation for variably saturated porous media with a one-dimensional convection-diffusion equation for surface water dynamics [Bixio *et al.*, 2000; Putti *et al.*, 2002]

$$\sigma(S_w) \frac{\partial \psi}{\partial t} = \nabla \cdot [K_s K_r(S_w) (\nabla \psi + \eta_z)] + q_s(h) \quad (4.6)$$

$$\frac{\partial Q}{\partial t} + c_k \frac{\partial Q}{\partial s} = D_h \frac{\partial^2 Q}{\partial s^2} + c_k q_L(h, \psi) \quad (4.7)$$

where $\sigma(S_w) = S_w S_s + \phi(\partial S_w / \partial \psi)$ is the general storage term, $S_w(\psi) = \theta / \theta_s$ is water saturation, θ is the volumetric moisture content, θ_s is the saturated moisture content (generally equal to the porosity ϕ), S_s is the aquifer specific storage coefficient, ψ is pressure head, t is time, ∇ is the gradient operator, K_s is the saturated hydraulic conductivity tensor, $K_r(S_w)$ is the relative hydraulic conductivity function, $\eta_z = (0, 0, 1)^T$, z is the vertical coordinate directed upward, and q_s represents distributed source (positive) or sink (negative) terms (volumetric flow rate per unit volume). The surface water is routed using (4.7) along each single hillslope or channel link using a one-dimensional coordinate system s defined on the drainage network. In this equation, Q is the discharge along the channel link, c_k is the kinematic wave celerity, D_h is the hydraulic diffusivity, and q_L is the inflow (positive) or outflow (negative) rate from the subsurface to the surface. We note that q_s [L^3/L^3T] and q_L [L^3/LT] are both functions of the ponding head h , and that h can be easily derived from the discharge Q via mass balance calculations.

The soil hydraulic properties are specified by K_s and by families of characteristic (constitutive) relationships $S_w(\psi)$ and $K_r(\psi)$ such as those of *Huyakorn et al.* [1984] where the water saturation is expressed in terms of effective saturation S_e , in the form $S_w(\psi) = (1 - S_{wr}) S_e(\psi) + S_{wr}$, with S_{wr} the residual water saturation. The characteristic relationships are then written as

$$\begin{aligned} S_e(\psi) &= [1 + \alpha^\beta (\psi_a - \psi)^\beta]^{-\gamma}, & \psi < \psi_a \\ S_e(\psi) &= 1, & \psi \geq \psi_a \end{aligned} \quad (4.8)$$

$$k_{rw}(\psi) = k_{rw}(\psi) (S_e(\psi)) = S_e^\mu \quad (4.9)$$

In the above expressions ψ_a is the capillary or air entry pressure head and α , β , γ , and μ are constants.

Once discretized (see Appendix), the system of equations (4.6) and (4.7) must be solved simultaneously for (Q, ψ) or (h, ψ) . Nonlinearities arise from the soil hydraulic functions, the dependence of q_s on the ponding head, and the dependence of q_L on ψ . Spatial discretization proceeds from a DEM (digital elevation model) representing the catchment surface. These DEM cells are triangulated and replicated vertically to form a three-dimensional tetrahedral grid of the underlying aquifer,

Precipitation fluxes during storm events and potential evapotranspiration during inter-storm periods are the main driving forces of the model. The catchment partitions this atmospheric forcing into surface runoff, groundwater flow, actual evapotranspiration, and

changes in storage. Surface runoff involves different phenomena such as hillslope and channel flow and retardation and storage effects due to pools and lakes, while groundwater flow includes infiltration to and ex-filtration from the vadose zone. In the CATHY model hillslope flow is assumed to concentrate in rills or rivulets confined to cells for which the upstream drainage area A does not exceed some prescribed threshold A^* , while channel flow is assumed to occur on all those cells for which A equals or exceeds A^* [Montgomery and Foufoula-Georgiou, 1993]. Retardation and storage effects due to lakes or topographic depressions are implemented via the lake boundary-following procedure of Mackay and Band [1998].

For the treatment of atmospheric boundary conditions, the input flux values are considered potential rainfall or evaporation rates, and the actual rates, which depend on the prevailing flux and pressure head values at the surface, are dynamically calculated by the code during the simulation. This automatic “switching” of surface boundary conditions from a specified flux (Neumann) to a constant head (Dirichlet) condition, and vice versa, is implemented to correctly reproduce the physical phenomena occurring at the surface, and allows us to track the four possible states for any surface node: water-stressed, unsaturated, saturated, and ponded [Putti et al., 2002]. Surface saturation or ponding can occur via infiltration excess or saturation excess mechanisms, and both of these are accounted for by the same switching algorithm.

An explicit time discretization based on the Muskingum-Cunge scheme is used for the overland flow equation [Orlandini and Rosso, 1996], allowing us to adopt a non-iterative algorithm for the solution of equations (4.6) and (4.7). The “bookkeeping” for the coupled model is done by the subsurface flow module, which determines the current state of each surface node, partitions the actual inflows/outflows, and calculates the overland fluxes to be passed to the surface flow module. The isolation of this water accounting task makes it possible to avoid complications in the code due to, for instance, our use of a node-based subsurface module and a cell-based surface module. More importantly, for implementation of data assimilation in CATHY, having all this logistics in one module means that introducing the forcing term for nudging will not affect the structure of the surface routing module. Thus in the next section we need only examine the numerical discretization of the subsurface, Richards equation component to describe how the data assimilation module is introduced into the coupled numerical model.

4.4.2. Numerical discretization with the nudging term

We express equation (4.6) with a generic source/sink term q and with the nudging term from equation (4.2) as

$$\sigma \frac{\partial \psi}{\partial t} = \nabla \cdot [K_s K_r (\nabla \psi + \eta_z)] + q + G \frac{\sum_{k=1}^{N_T} \sum_{i=1}^{N_X} W_{ki}^2(\mathbf{x}, t) \epsilon_i (\theta_{o_{ki}} - \theta_i(t))}{\sum_{k=1}^{N_T} \sum_{i=1}^{N_X} W_{ki}(\mathbf{x}, t)} \quad (4.10)$$

where θ_o are the soil moisture observations.

Initial conditions and Dirichlet, Neumann, or Cauchy boundary conditions are added to complete the mathematical formulation of the flow problem.

$$\psi(\mathbf{x}, 0) = \psi_o(\mathbf{x}) \quad (4.11)$$

$$\psi(\mathbf{x}, t) = \psi_D(\mathbf{x}, t) \quad \text{on } \Gamma_1 \quad (4.12)$$

$$\mathbf{v} \cdot \mathbf{n} = -q_n(\mathbf{x}, t) \quad \text{on } \Gamma_2 \quad (4.13)$$

where superscript T is the transpose operator, ψ_o is the pressure head at time 0, ψ_p is the prescribed pressure head (Dirichlet condition) on boundary Γ_1 , \mathbf{n} is the outward normal unit vector, and q_n is the prescribed flux (Neumann condition) across boundary Γ_2 . We use the sign convention of q_n positive for an inward flux and negative for an outward flux.

The details of the spatial and temporal discretization of equation (4.10) are reported in [7]. The resulting algebraic system of equations, for the case of the backward Euler time discretization used in this study, is

$$\left(H^{k+1} + \frac{P^{k+1}}{\Delta t_k} \right) \hat{\Psi}^{k+1} = \frac{P^{k+1}}{\Delta t_k} \hat{\Psi}^k - \mathbf{q}^{*k+1} - \mathbf{r}^{*k+1} \quad (4.14)$$

where index k denotes time level, Δt_k is the time step size, H and P are the stiffness and mass matrices, respectively, \mathbf{q}^* contains boundary condition and gravitational gradient components, and \mathbf{r}^* is the discretized nudging term. Equation (4.14) is linearized using Picard iteration and solved using a standard conjugate gradient method.

In the derivation given in *Panicconi et al. [2003]*, it is straightforward to substitute or add other state variables (pressure head, groundwater velocity, etc) in the nudging term without altering the basic formulation. The implementation could perhaps also accommodate indirect observations of hydrological state variables, for example backscattering measurements obtained from active microwave remote sensing of the land surface. One possible incentive to doing so would be to avoid having to solve a complex or intractable inverse model to back observed soil moisture values out of backscattering data, and instead simply use a forward model to convert computed soil moisture into backscattering values.

4.5. Numerical experiments

4.5.1. Description of the test case

The CATHY model with and without the nudging algorithm has been applied to a small hypothetical test catchment made up of 6×11 grid cells of 50×50 m of which 9 cells represent an internal depression or lake. Figure 4.2 shows the catchment DEM along with the schematized representation of the flow paths as determined from topographic analysis augmented by the lake boundary-following procedure. The outlet cell for the catchment is labeled “O” while the reservoir cell designated for the lake (accumulation point for the lake cells plus buffer cells) is labeled “R”.

Furnished with spatio-temporal observation data representing the “true” distribution of a state variable (we assume that $\epsilon = 1$ in these synthetic datasets), a model simulation or integration enabled with data assimilation should generate results that are closer to the true field than a prior integration where no assimilation is used. In addition to assessing the hydrological performance of the nudging algorithm on this principle, we will also compare the numerical performance of assimilation and prior simulations in order to assess the computational “cost” of the nudging-based procedure.

The numerical experiments consisted in first running the CATHY model to generate physically reasonable initial conditions from an arbitrary starting point (such as a fully saturated catchment in hydrostatic equilibrium). A numerical simulation is then run using these initial conditions to generate the true integration (“base run”). The prior integration (CATHY model without nudging, or “no-nudging run”) is generated by varying or perturbing the initial or boundary conditions of the base run. The assimilation integration

("nudging run") is the prior integration augmented by surface soil moisture observation data at selected points in space and time extracted from the true integration.

For the test catchment of Figure 4.2 we used five observation nodes and two observation times, at 1 and 3 hours, for a four-hour simulation period. The location of the five surface observation points is indicated by the black dots in Figures 4.3, 4.7, and 4.10. Note that the test catchment in these figures is flipped (outlet cell at the bottom) with respect to Figure 4.2 (outlet at the top). Three of the observation points are clustered immediately downstream of the lake, while the remaining two are situated within the buffer cells of the lake, and thus remain wetter in our simulations than the other three observation nodes. The soil moisture time plot results will be shown for only two of the five observation points, one downstream of the lake and one in the buffer zone. The results for the other three points are very similar to these two.

Table 4.1 gives the model discretization and parameter values used for the numerical experiments. The nudging parameters were set to the following values: $G = 0.1$, $\epsilon = 1$, $\tau = 1800$ s, $R_{xy} = 100$ m, and $R_z = 2$ m. Variations of G , τ , and R_{xy} to examine the influence of these three parameters will be explicitly referred to in the following sections.

Two test case scenarios were run, the first in which the prior integration differs from the true integration in the atmospheric forcing boundary conditions applied during the simulation, and the second in which it is the initial conditions rather than the atmospheric boundary terms that are perturbed.

The initial conditions used for the true, prior, and assimilation integrations of the atmospheric forcing perturbation scenario and for the true integration of the initial conditions perturbation scenario were generated from a CATHY run in which the fully saturated catchment with vertical pressure head profiles in hydrostatic equilibrium is allowed to drain and dry under an atmospheric forcing regime of 20000 s (5.56 hrs) of rainfall at a spatially uniform rate of 6×10^{-5} m/s followed by a 5000 s period during which the rainfall rate drops linearly from 6×10^{-5} m/s to zero and a 1000 s period during which the soil evaporates at a linearly increasing rate from zero to 1.2×10^{-5} m/s. The initial conditions used for the prior and assimilation integrations of the initial conditions perturbation scenario were generated from a CATHY run in which these previously generated initial conditions are imposed at time zero, and the catchment is allowed to drain and dry for an additional period of 1800 s (0.5 hrs) during which the soil evaporates at a linearly increasing rate from 1.2×10^{-5} m/s to 3.06×10^{-4} m/s.

The atmospheric boundary conditions used for the true simulation of the atmospheric forcing perturbation scenario and for the true, prior, and assimilation simulations of the initial conditions perturbation scenario consist of a linearly increasing evaporation rate from 1.2×10^{-5} m/s at time zero to 6×10^{-5} m/s at time 4000 s, after which the evaporation rate stays fixed at 6×10^{-5} m/s until the end of the simulation at 14400 s (4 hrs). The atmospheric boundary conditions used for the prior and assimilation integrations of the atmospheric forcing perturbation scenario represent significantly drier conditions and consist of a linearly increasing evaporation rate from 1.2×10^{-5} m/s at time zero to 6×10^{-4} m/s at time 3600 s, followed by a linearly decreasing rate from 6×10^{-4} m/s to 5.73×10^{-6} m/s during the next hour, followed by a one-hour period during which the evaporation rate linearly increases to 6×10^{-5} m/s, after which the evaporation rate stays fixed at this rate for the final (fourth) hour of the simulation.

4.5.2. Perturbation of atmospheric forcing boundary conditions

In Figures 4.3 and 4.4 we compare the spatial and temporal surface soil moisture (or water saturation) results between the prior (no-nudging) and assimilation (nudging) integrations. We can see from Figure 4.3, where the differences in surface saturation between observed data (true or base run) and computed values are plotted for both the prior and assimilation runs, that nudging produces a significant reduction in model prediction errors. This is true at the two observation times of 1 and 3 hours, i.e. in the center of the time influence windows, but also at the end of the simulation which is outside of the second time window. Spatially, the improvement in the computed solution is, as could be expected, highest at and in the immediate vicinity of the five observation points, but there is also a notable reduction in model prediction error farther away from the observation points, showing that the model dynamics are able to successfully propagate or advect information contained in the observation dataset beyond the horizontal radius of influence of 100 m. The mean model prediction errors, calculated from the water saturation differences over all nodes of the catchment surface, are 32% lower for the assimilation run compared to the prior run at time 3600 s, 38% lower at time 10800 s, and 29% lower at the final time.

The temporal behavior of the model with and without nudging, shown in Figure 4.4 at two of the observation points, confirms the success of the nudging algorithm in significantly reducing model prediction errors, but also highlights the important role played by the time weighting function and its time window parameter τ . The activation of nudging after the simulation time t becomes greater than $t_o - \tau$ can be clearly seen for the second observation period in the buffer node plot (for the first observation period the true and prior solutions remain identical well into the time window for this very wet node) and for the first observation period in the downstream node case (for the second observation period this effect is hidden by a lake discharge effect described below). The correction that nudging is able to give the model is quite dramatic for this test scenario given that the perturbed evaporation boundary conditions applied to the prior run result in a much drier soil than for the true integration, as already described. Indeed as can be seen in Figure 4.4 the hillslope portion of the catchment (downstream node) dries out to a soil moisture value of about 0.175 within the first hour of simulation, while the wetter lake buffer zone takes a little bit longer to reach this value. This soil moisture level at the surface is maintained except for a brief period between about the second and third hour of simulation wherein the soil moisture increases for the prior integration (and consequently also for the assimilation integration) due to the lake reservoir draining its water through subsurface discharge to the buffer and downstream nodes after the surrounding and underlying soil has apparently reached a dry enough state. That this lake discharge effect does not occur during the true integration, which for the buffer zone node has a soil moisture value roughly equal to that of the assimilation run during this brief period, is due to the fact that the surrounding nodes also maintain relatively high soil moisture values, compared to the surrounding nodes in the assimilation run, especially those farther away from the influence of the nudging weighting functions, which are more likely to have soil moisture values that correspond to those of the much drier prior simulation.

The critical influence of the time parameter and weighting function in the nudging algorithm can be evidenced in the downstream node plot of Figure 4.4, where the plateau shape of the time weighting function is probably the main contributor to undershoot in the assimilation integration during the period $t_o - \tau$ and overshoot for $t_o + \tau$ for this case where instead the true soil moisture shows a monotonically decreasing trend, and in the buffer node plot, where we get a rapid deterioration in the assimilation solution once we are

outside of the time influence window, suggesting that temporal advection of information in the nudging algorithm may not be as effective as the spatial (horizontal) propagation shown earlier. We will return to this point in the next section and for the initial conditions perturbation scenario, where more encouraging results were obtained.

In terms of computational performance, the assimilation run for this test scenario was surprisingly more efficient than the prior simulation without the added nudging term, requiring 25% fewer time steps and 26% less CPU. This and other numerical aspects will be explored further in the following section.

4.5.3. Influence of nudging parameters

For the atmospheric forcing perturbation scenario we examined the influence of some nudging parameters, namely G , R_{xy} , and τ , on the computational performance of the model.

The nudging factor G determines the relative strength of the nudging term with respect to the physical forcing term, and an appropriate value for this coefficient will depend on the particular model equation and model application being considered. Too low a value should result in nudging having a negligible and therefore ineffectual influence on the solution, whereas too high a value can introduce spurious, non-physical behavior in the model, and perhaps numerical convergence problems associated with this noise. Optimization of G should be based on dimensional and sensitivity analyses of the system equation, to determine the relative contribution of the various model coefficients and terms to the solution under varying conditions. Thus some theoretical analysis of the hydrological model equations is needed to determine an appropriate value of G . In meteorology *Stauffer and Seaman [1990]* suggest that the nudging factor should be set such that the time scale of the slowest model process and the nudging term are similar, and that a numerical stability criterion of $G \leq 1/\Delta t$ must also be satisfied. As they point out, these two requirements can be in conflict. Moreover for the Richards equation model, where time step determination is an adaptive process that can produce widely differing time step sizes over the course of a simulation, the stability criterion suggests that G should be time variant. For the hydrologic model used in *Houser et al. [1998a]*, a value of $G = 1/(10\Delta t)$ was used with satisfactory results. For this paper we conducted a heuristic, numerical sensitivity analysis based on our test case.

Figures 4.5 and 4.6 and Table 4.2 confirm these general aspects regarding the influence of G . In Figure 4.5 the temporal variation of soil moisture at the buffer and downstream observation points shows that for values of G below about 0.1, nudging is unable to bring the solution close to the true values during the time periods which correspond to the assimilation windows. At values of G of 0.1 and 0.5 and, in Figure 4.6 of 1.0 and 10.0 as well, nudging performs well in a hydrological sense, bringing the computed solution in line with the true or observed soil moisture distribution. The small improvements as G is increased from 0.1 to 10.0 come at a high computational cost, however, as reported in Table 4.2 where we can see that the number of time steps, back-steps (failed time steps), nonlinear iterations, and CPU all increase rapidly for values of G larger than about 0.3 or 0.5. The serious numerical difficulties encountered for $G = 10.0$ made it unrealistic to simulate still higher values, and in any event Figure 4.6 suggests that there would be little gain, in terms of reducing model prediction errors, in doing so. Based on a combination of numerical and hydrological performance criteria, it therefore appears that the best value of G for this test case is on the order of 0.1. For the smallest and average time step sizes

reported in Table 4.2 (and also in Tables 4.3 and 4.4) the stability criterion $G \leq 1/\Delta t$ is satisfied for $G = 0.1$, although for some time steps where Δt reaches values up to the maximum prescribed step size of 32 s, the criterion is not met.

As an aside, and related to the temporal advection comment made earlier, Figure 4.5 shows clearly the important re-initialization function of data assimilation. Qualitatively the nudging solutions for the different values of G computed outside of the time influence windows are as different from each other as they are inside of the time windows, when nudging acts to effectively re-initialize the model to soil moisture conditions close to the observed values.

The effect of the horizontal radius of influence parameter R_{xy} is shown in Table 4.3 and in Figures 4.7 and 4.8. The computational impact of R_{xy} is not nearly so great as that of parameter G , although it does appear that the model is able to adapt to larger (and thus fewer) time steps as R_{xy} is increased, suggesting a smoothing influence of the weighting functions as the forcing term from nudging is extrapolated or spread over a larger area. Interestingly the gain in computational efficiency that can be seen from Table 4.3 is not only between smaller and larger values of R_{xy} but also, except for the smallest R_{xy} of 50 m, between the assimilation and prior runs, resulting in 26% less CPU for $R_{xy} = 100$ as mentioned earlier and as much as 56% less CPU ($R_{xy} = 200$). Thus the "relaxation" capacity of the Newtonian nudging algorithm is also apparent, improving the temporal evolution of the numerical solution in perhaps the same way that relaxation is often applied to improve nonlinear convergence of a linearization scheme. This is an issue that should be investigated further.

Nudging's ability to spatially propagate information contained in observation data via model dynamics, shown in Figure 4.3, is further explored in Figures 4.7 and 4.8. It is clear from Figure 4.7 that model prediction errors, measured in terms of surface water saturation differences between assimilation and true runs, are further reduced when R_{xy} is increased from 50 to 150 to 250 m, at observation times but also outside of the time influence windows (14400 s). As expected the spatial area influenced by the nudging term, i.e., the extent of the area which is affected by an observation at a given point, increases as the radius of the area which is affected by an observation increases. This is confirmed in Figure 4.8 where the average soil moisture differences over all nodes of the catchment surface are plotted as a function of R_{xy} for the three simulation output times of 3600, 10800, and 14400 s. The behavior for all three times is similar, with the biggest drops in model prediction error occurring between $R_{xy} = 0$ (prior run) and $R_{xy} = 100$ m (between 30% and 38% for the three output times) and with more gradual improvements for $R_{xy} > 100$ m. It is to be expected that the improvement in spatial agreement between the computed and observed results should tail off, and even deteriorate, as R_{xy} values larger than the spatial correlation lengths of surface soil moisture are used. Larger scale test cases will be used in future studies to test this hypothesis.

The effect of the time influence window half period parameter τ is shown in Table 4.4 and Figure 4.9. Computationally, the best results are obtained for the run with $\tau = 1350$ s; at lower (including the prior run) and higher values, CPU is higher due to smaller time steps or increased occurrences of back-stepping. At high τ values this behavior may be due to numerical inconsistencies between a high correlation implied by large τ and the low temporal correlation corresponding to the rapidly changing soil moisture that results from the strong forcing in the perturbed atmospheric boundary conditions and the abrupt lake discharge event in the middle of the simulation; in other words contrasting signals between the model dynamics and its natural forcing and the forcing introduced by the

nudging term.

Hydrologically, Figure 4.9 shows that model prediction errors are reduced as τ is increased, but this behavior cannot be generalized and it is important that τ and also the frequency of nudging observations be compatible with the correlation scales of the true or expected soil moisture patterns. In this figure the shape and width of the time weighting function can clearly be seen, and we also see that the gains in performance are obtained mainly outside of the time influence windows, where a wider window (larger τ) gives the model less time over which to reduce soil moisture values to the levels implied by the prior solution, so that the assimilation runs retain soil moisture values closer to the higher levels of the true solution. It is easy to envisage scenarios, however, where both the true and prior solutions change abruptly, and in the same direction, after an observation time. In such cases model prediction errors would undoubtedly be higher for an assimilation run with a large time influence window than one using a smaller τ . An effective assimilation in such cases would of course also benefit from a new observation during the period of rapid soil moisture change. Figure 4.9 also suggests that care must be taken to avoid excessive under and overshoot in the assimilation solutions, since, for example, we can observe that for the downstream node overshoot causes the run with $\tau = 2700$ s to produce a higher model prediction error than the run with $\tau = 2250$ s.

4.5.4. Perturbation of initial conditions

The initial condition perturbation scenario represents a less stringent, but in some sense also more realistic, test of the nudging algorithm than the atmospheric forcing perturbation scenario. In this scenario it is presumed that the prior simulation begins with a poor estimate of the true initial conditions, and at the two observation times the model is essentially re-initialized via data assimilation. Model parameters and atmospheric boundary conditions are assumed to be known exactly, and the re-initialization is a partial one only, based on only five observation points on the catchment surface. The initial state for the prior and assimilation runs, as described earlier and as can be seen in Figure 4.11 for the downstream node, represents a drier soil than the true conditions.

Qualitatively, the influence of nudging on the spatial and temporal surface soil moisture distribution (Figures 4.10 and 4.11, respectively) is similar to what was seen for the atmospheric forcing perturbation scenario. This time the mean model prediction errors, again calculated from the water saturation differences over all nodes of the catchment surface, are 32% lower for the assimilation run compared to the prior run at time 3600 s, 34% lower at time 10800 s, and 32% lower at the final time. In some respects however the performance of nudging for this scenario is better than for the previous one. In these figures we observe that reductions in model prediction error obtained when the nudging weighting functions are active are better conserved outside of the space and time influence windows, leading to more continuous and accumulated improvements in the assimilation run as soil moisture observations at different times are assimilated. In this situation it is even more important to select an appropriate time weighting function, since as can be clearly seen for the downstream node in Figure 4.11, the symmetric plateau-shaped weighting function has a strong influence on the solution error outside of the assimilation windows.

4.6. Conclusions

We have selected Newtonian relaxation, or nudging, as a suitable scheme for conducting some initial investigations into the applicability and effectiveness of data assimilation for a three-dimensional finite element Richards equation-based distributed hydrological model. In the nudging procedure an additional forcing term is added to the model, so the implementation requires significant, though not complicated, code modifications. Nonetheless the overheads associated with nudging, including computation of the new forcing term, add negligible cost to the model simulation. The simplicity and efficiency of the Newtonian relaxation scheme are the two main advantages of this approach over more sophisticated, optimal data assimilation methods. Another appealing feature of the algorithm, and our implementation of it for the CATHY hydrological model, is its flexibility. It is always forward-marching in time (we assume prior knowledge of the observation times), spatial and temporal variability in the nudging parameters can be handled (to account for seasonal and geomorphological effects on the state variables, for instance), and concurrent observation datasets distributed over the entire grid or at scattered points can be assimilated (allowing utilization of information from different measurement sources).

The relative merits and tradeoffs between factors such as computational efficiency, flexibility, and hydrological performance (the ability to reduce model prediction errors and incorporate knowledge of model and data uncertainties) between simple schemes such as Newtonian relaxation and more advanced data assimilation methods, in the context of a detailed, three-dimensional, nonlinear numerical catchment model, will require inter-comparison studies. The test cases presented here nonetheless show that the nudging technique is successful in decreasing the error between the computed and observed values, these latter derived from a numerically generated surrogate true integration with ϵ , the observation data quality factor, set to one. Surprisingly, nudging was also found to generally improve the numerical performance of the model in comparison to prior integrations, in terms of convergence, time stepping, and computational efficiency. This can perhaps be attributed to a smoothing influence of the weighting functions in the nudging relaxation term. A contrasting effect of deterioration in the numerical solution was shown for relatively high values of G , suggesting that care is required in selecting G to ensure that the nudging term does not become dominant in Richards equation. The limited sensitivity trials we conducted indicated, for our test case, an appropriate value of around 0.1 for this parameter. The effectiveness of adapting G to the model time step size throughout the simulation in order to satisfy the stability criterion $G \leq 1/\Delta t$ can easily be tested. More effort will be needed to investigate model time scales in relation to the nudging term.

More rigorous sensitivity analysis will be needed to assess not only the impact of G , but also of the weighting function parameters τ , R_{xy} , and R_z . Of these three parameters, τ has perhaps the subtlest effect because of its close interaction with the relaxation difference term, its relationship with time step size in an adaptive time evolution scheme, the possibility of overlapping time windows of different lengths, and the evidence of numerical under and overshoot seen with the centered time influence windows used in our tests. Increasing the horizontal radius of influence leads to a considerable decrease in the mean spatial difference between observed and computed surface soil moisture values. Though not investigated, similar results are expected for the vertical radius of influence parameter R_z , with an important bearing on soil moisture profile and water table estimation. These parameters, however, must in some way be consistent with or reflect the correlation scales of the state variable being assimilated. Indeed the four-dimensional weighting functions used in the nudging approach provide a simple and intuitive way of incorporating any avail-

able knowledge regarding characteristic length scales and the spatio-temporal variability of state variables. As our understanding of this variability and how to represent it improves, the weighting functions will acquire more physical relevance, taking on shapes that conceivably will be quite different from the Cressman functions used in this and previous studies, and the nudging technique will become a little less heuristic in the process.

In addition to inter-comparison studies, more detailed sensitivity analyses, and improved weighting function formulations, other topics for more immediate future research include: implementation of the nudging algorithm for the Newton linearization option already available in the CATHY code; implementation of a nudging term better suited to gridded and/or high-density data such as remotely sensing observations of surface state variables; application of the model to an actual catchment with assimilation of real observation data; and evolution to more advanced data assimilation methods for the Richards equation-based coupled surface–subsurface model, using the nudging implementation as a basic framework.

Table 4.1.: Model Discretization and Parameter Values for the Test Case

Saturated hydraulic conductivity K_s	10^{-4} m/s
Aquifer specific storage S_s	10^{-7} m $^{-1}$
Porosity ϕ ($= \theta_s$, saturated moisture content)	0.35
Soil hydraulic properties (equations (4.8) and (4.9))	$S_{wr} = 0.333$, $\psi_a = 0$, $\alpha = 0.02$, $\beta = 2$, $\gamma = 2$, $\mu = 1$
Number of soil layers	10
Soil depth	10 m
Soil layer thicknesses (top to bottom)	0.1, 0.1, 0.1, 0.5, 0.5, 1.0, 1.0, 2.2, 2.2, 2.3 m
Number of nodes in the three-dimensional grid	924
Number of tetrahedral elements in the grid	3960

Table 4.2.: Influence of the Nudging Parameter G on the Numerical Performance of the Model

G	no nudging	0.001	0.005	0.01	0.05	0.1	0.3	0.5	0.8	1.0	10.0
Total mbe ¹ (%)	-5.7	-6.2	-7.1	-7.4	-9.4	-9.1	-9.6	-9.4	-9.6	-9.5	-9.8
# time steps ²	9655	11173	6827	10593	6959	7196	12535	9656	14624	16104	54157
# back-steps	1	1	12	2	4	10	5	452	1631	2121	31299
Smallest time step (s)	0.125	0.25	0.00012	0.031	0.0078	0.00012	0.00049	0.00049	0.00049	0.00049	0.00024
Average time step (s)	1.5	1.3	2.1	1.4	2.1	2.0	1.1	1.5	0.98	0.89	0.26
Avg NL iter/step ³	4.2	4.0	4.3	4.0	4.3	4.3	4.3	6.6	9.8	10.7	32.1
Total CPU (s)	1408	1315	837	1393	1165	1040	1612	1901	5564	5137	81385 (22.6 hrs)

¹Total (global) mass balance error

²Excludes back-steps

³Average nonlinear (Picard) iterations per time step

Table 4.3.: Influence of the Horizontal Radius of Influence R_{xy} on the Numerical Performance of the Model

R_{xy} (m)	no nudging	50	100	150	200	250
Total mbe ¹ (%)	-5.7	-7.1	-9.1	-10.9	-11.0	-10.9
# time steps ²	9655	12079	7196	7076	4565	6771
# back-steps	1	2	10	11	4	11
Smallest time step (s)	0.125	0.0312	0.00012	0.00012	0.0039	0.00012
Average time step (s)	1.5	1.2	2.0	2.0	3.2	2.1
Avg NL iter/step ³	4.2	4.1	4.3	4.3	4.5	4.3
Total CPU (s)	1408	1486	1040	910	622	954

¹Total (global) mass balance error²Excludes back-steps³Average nonlinear (Picard) iterations per time stepTable 4.4.: Influence of the Time Influence Window Half Period τ on the Numerical Performance of the Model

τ (s)	no nudging	900	1350	1800	2250	2700
Total mbe ¹ (%)	-5.7	-7.8	-8.8	-9.1	-9.0	-9.0
# time steps ²	9655	11294	7099	7196	11248	11284
# back-steps	1	3	0	10	11	0
Smallest time step (s)	0.125	0.0039	0.0625	0.00012	0.00012	0.0625
Average time step (s)	1.5	1.3	2.0	2.0	1.3	1.3
Avg NL iter/step ³	4.2	4.1	4.3	4.3	4.2	4.2
Total CPU (s)	1408	1173	816	1040	1529	1564

¹Total (global) mass balance error²Excludes back-steps³Average nonlinear (Picard) iterations per time step

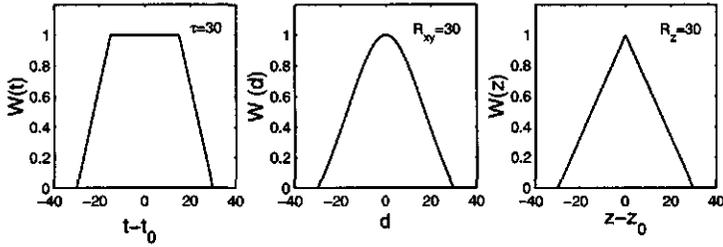


Figure 4.1.: An example of the temporal, horizontal ($d^2 = (x - x_o)^2 + (y - y_o)^2$), and vertical components of the four-dimensional weighting functions for the nudging algorithm.

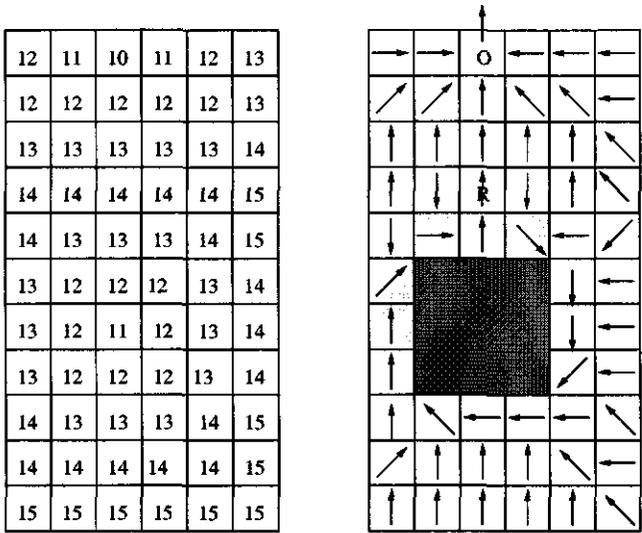


Figure 4.2.: 50 × 50 m DEM surface cell elevations (left) and surface flow paths (right) for the test case representing a small catchment with an internal lake.

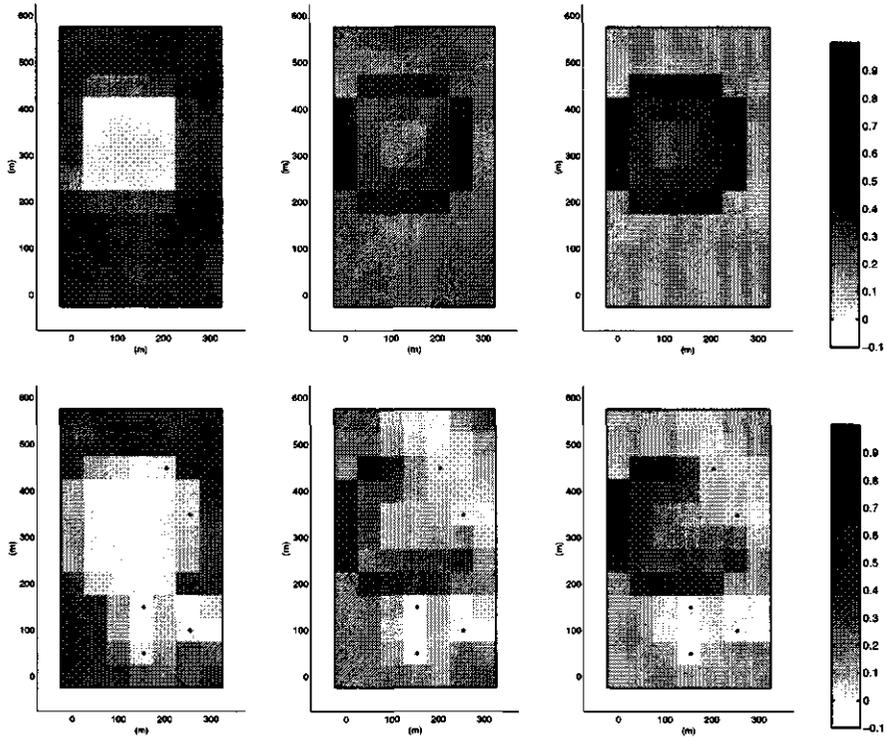


Figure 4.3.: Difference in water saturation at the surface nodes between the base run and the no-nudging run (top row) and between the base run and the nudging run (bottom row) at three different times (3600, 10800, and 14400 s from left to right) for the atmospheric forcing perturbation scenario.

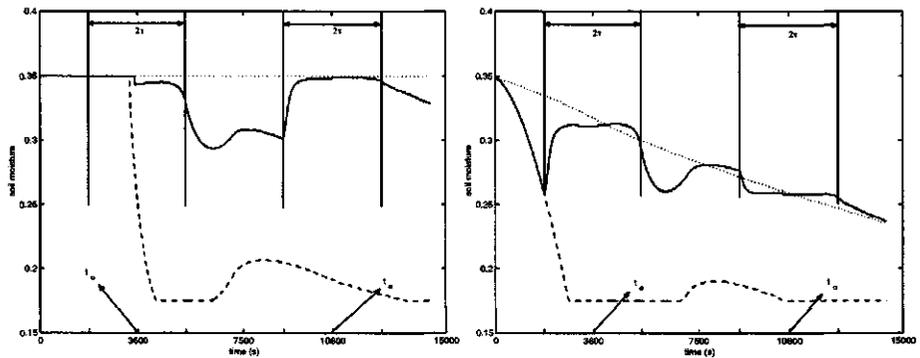


Figure 4.4.: Temporal soil moisture behavior at two observation points (left: lake buffer zone; right: downstream of the lake) for the base run (dotted line), no-nudging run (dashed line), and nudging run (solid line) for the atmospheric forcing perturbation scenario.

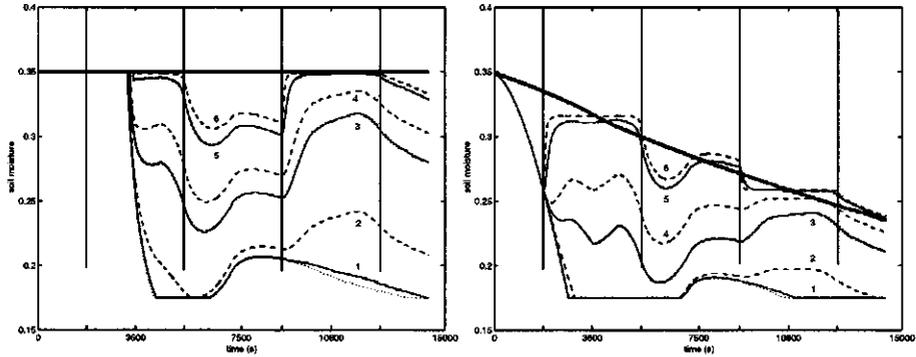


Figure 4.5.: Temporal soil moisture behavior at two observation points (left: lake buffer zone; right: downstream of the lake) for the base run (thick line), the no-nudging run (dotted line), and the nudging runs for six values of the nudging parameter G (1: $G = 10^{-4}$; 2: 10^{-3} ; 3: 0.005; 4: 0.01; 5: 0.1; 6: 0.5).

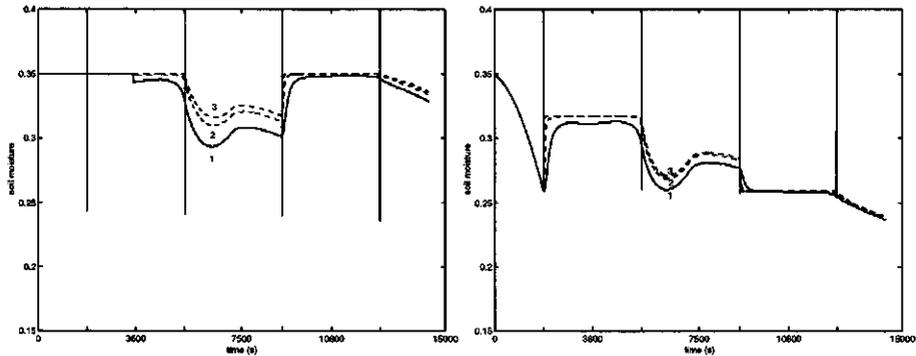


Figure 4.6.: Temporal soil moisture behavior at two observation points (left: lake buffer zone; right: downstream of the lake) for relatively high values of the nudging parameter G (1: $G = 0.1$; 2: 1.0; 3: 10.0). The base run solution, not plotted here, is the same as in the previous figure.

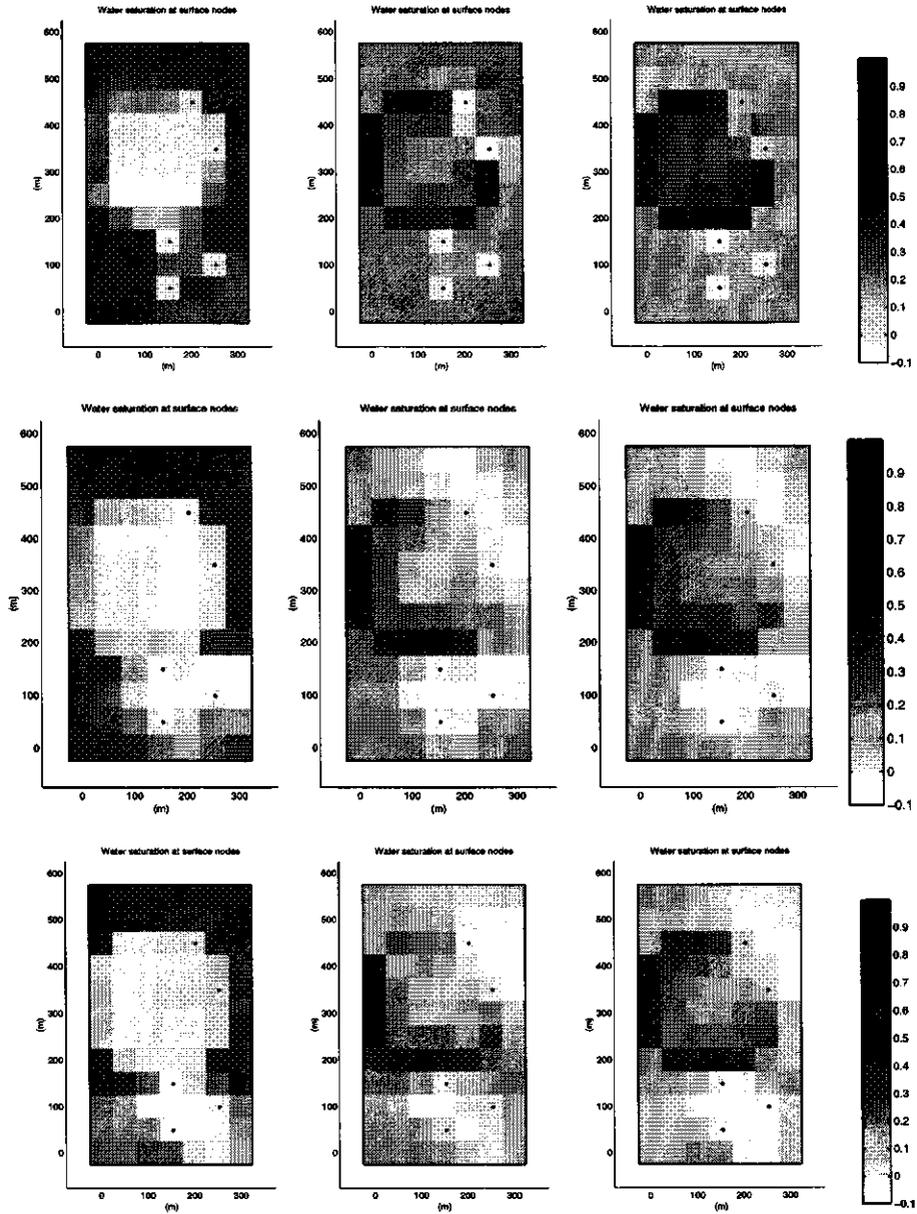


Figure 4.7.: Difference in water saturation at the surface nodes between the base run and the nudging run at three different times (3600, 10800, and 14400 s from left to right) for different values of the horizontal radius of influence parameter R_{xy} (top row: $R_{xy} = 50$ m; middle row: 150 m; bottom row: 250 m).

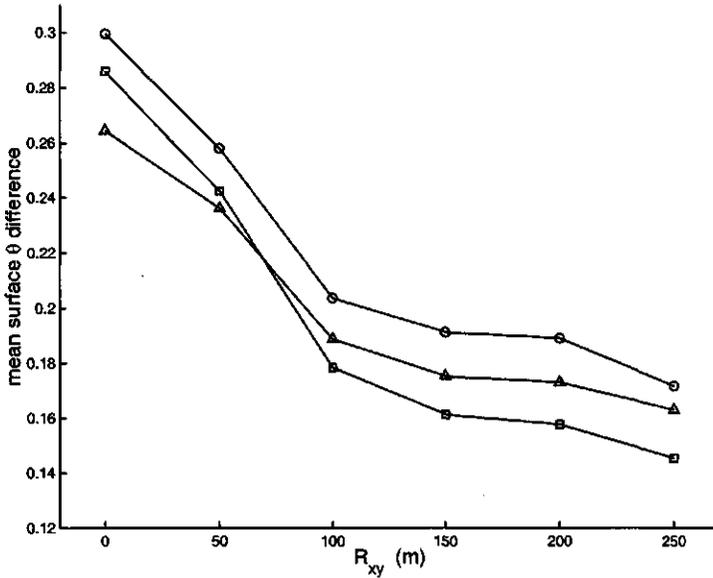


Figure 4.8.: Mean spatial difference in surface soil moisture between the base run and the nudging run for different values of the horizontal radius of influence parameter R_{xy} ($R_{xy} = 0$ corresponds to the no-nudging run) at three different times (circle: time 3600 s; square: 10800 s; triangle: 14400 s).

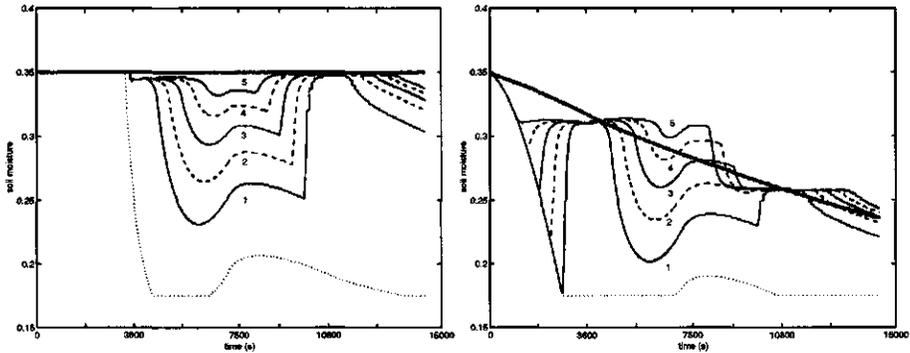


Figure 4.9.: Temporal soil moisture behavior at two observation points (left: lake buffer zone; right: downstream of the lake) for the base run (thick line), the no-nudging run (dotted line), and the nudging runs for five values of the time influence window parameter τ (1: $\tau = 900$ s; 2: 1350 s; 3: 1800 s; 4: 2250 s; 5: 2700 s).

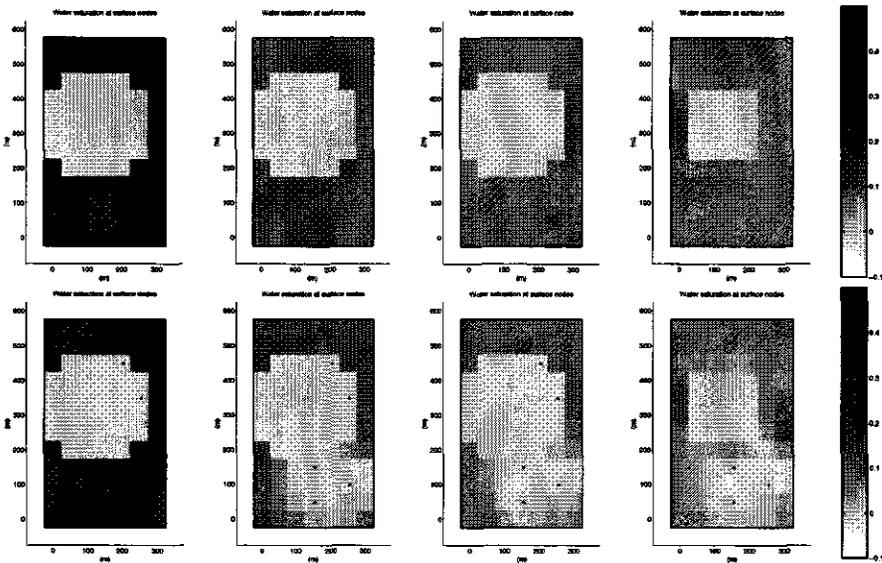


Figure 4.10.: Difference in water saturation at the surface nodes between the base run and the no-nudging run (top row) and between the base run and the nudging run (bottom row) at four different times (0, 3600, 10800, and 14400 s from left to right) for the initial conditions perturbation scenario.

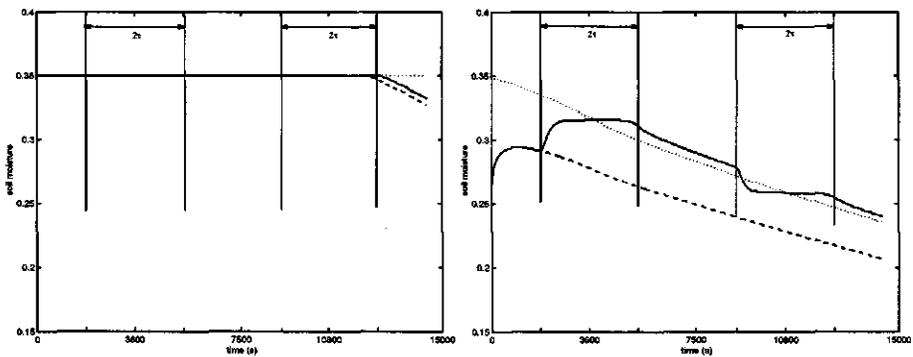


Figure 4.11.: Temporal soil moisture behavior at two observation points (left: lake buffer zone; right: downstream of the lake) for the base run (dotted line), no-nudging run (dashed line), and nudging run (solid line) for the initial conditions perturbation scenario.

5. Application of a nudging technique to the Brisby catchment with a 2D-Richards model

Dorothea Mirosław-Swiątek

5.1. Introduction

In this chapter a 2D-Richards finite element model is used to describe the subsurface-flow processes on a hillside. The applicability of a nudging technique in this model is investigated [Houser, 1996]. A 2D hydrological model can be applied to a vertical cross-section taken in a direction parallel to the downslope flow, or roughly perpendicular to the stream. It allows to define the heterogeneity of soil and terrain in a flexible way. It can handle a variable-head lower-boundary condition (the seepage face on the stream bank), as well as time and space-dependent atmospheric forcing (rain and evapotranspiration).

The Richards equation is the standard mathematical description of saturated-unsaturated flow phenomena at the hillslope scale [Freeze, 1978]. In addition to the actual equation of subsurface flow, also important are: the size and shape of the region of flow, the boundary conditions around the boundaries of the region, and their spatial and temporal distribution, the initial conditions and their spatial distribution, the spatial and temporal distribution of the soil hydraulic parameters and the mathematical solution method. In this study the Brisby data, as generated by the Cathy model (see Chapter 4), were used [Oosterhoff and Paniconi, 2001]. From this data set a hillslope slope is selected for which the 2D Richards model is applied. The initial- and boundary conditions, as well as all the states are available for this synthetic data set. It is investigated how the assimilation of observations (by the nudging technique), improves the model performance.

5.2. Mathematical model of hillside subsurface flow

Two-dimensional unsteady subsurface flow in a heterogeneous (and possibly anisotropic) medium can be described by [Zarudny, 1993].

$$\nabla \cdot [k_r(h) K \nabla H] + S = [C(h) + S_w S_s] \frac{\partial H}{\partial t} \quad (5.1)$$

where: H is the total head ($h+z$) [cm], h is the pressure head [cm], z is the vertical coordinate [cm], $k_r(h)$ is the relative hydraulic conductivity [cm/s], $C(h)$ is the specific moisture

capacity $(d\theta/dh)$ [cm^{-1}], K is the symmetric positive-definite tensor, representing the conductivity at saturation, S_w is the degree of volumetric water saturation, $(\theta(h)/\theta_s)$, S_s is the specific storage [cm^{-1}], S is the internal source [cm^3/s], $\theta(h)$ is the volumetric moisture content [cm^3/cm^3], θ_s is the volumetric moisture content at saturation [cm^3/cm^3], $k(h)$ is the hydraulic conductivity [cm/s], and t is time [s].

The above equation is valid for soils that do not swell or shrink, for the assumptions that processes are isothermal and that the osmotic and the so called pneumatic components of soil water potential are negligible. Equation 6.3 is called the general Richards equation and describes the water flow in unsaturated-saturated porous medium, with S as a function of water intake. The unknown in Equation 6.3 is the pressure head h , which is a continuous function for unsaturated and saturated zone. There is a variation in value h with the transition from one zone to another. Value $h = 0$ (the contours of zero pressure head) corresponds to the water table position. Above the water table h values are negative, and below positive. The dependence of volumetric moisture content θ on the pressure head h (neglecting hysteresis effects), can be described by the following empirical formula [van Genuchten, 1980].

$$\frac{\theta - \theta_r}{\theta_s - \theta_r} = \left[\frac{1}{1 + (\alpha \cdot |h|)^n} \right]^\beta \quad (5.2)$$

where θ_r and θ_s are as defined previously, and α , n and β are soil-specific coefficients.

Using formula (5.2) the dependence on specific moisture capacity $C(h)$ - measuring the capacity of soil for water sorption - is given by

$$C(h) = \frac{d\theta}{dh} = \frac{(n-1) \cdot \alpha^n \cdot (\theta_s - \theta_r)}{[1 + (\alpha \cdot |h|)^n]^{\beta+1}} \cdot |h|^{n-1} \quad (5.3)$$

Relative hydraulic conductivity $k_r(h)$ is defined by the following formula:

$$0 \leq k_r(h) = \frac{k(h)}{k_s} \leq 1 \quad (5.4)$$

In the unsaturated zone, hydraulic conductivity $k(h)$ is a rising function of the pressure head and is here defined by the empirical relationship formulated by van Genuchten [1980]:

$$k(h) = k_s \left[\frac{1}{1 + (\alpha \cdot |h|)^n} \right]^{\beta \cdot \alpha} \cdot \left\{ 1 - \left[\frac{(\alpha \cdot |h|)^n}{1 + (\alpha \cdot |h|)^n} \right]^\beta \right\}^p \quad (5.5)$$

where α , n are soil-specific coefficients (viz. equations 5.2 and 5.3).

The parameters in the Richards equation take different values depending on whether water flow takes place in unsaturated or saturated zone. Water flow in saturated zone is a special case where hydraulic conductivity and water moisture content take maximum values ($k_r(h) = 1$, $\theta(h) = \theta_s$, $C(h) = 0$). Specific water storage S_s enables to describe unsteady water flow in saturated and unsaturated medium using the same type of differential equation. Equation 6.3 is, in both zones, a non-linear second order parabolic differential equation.

From a mathematical point of view, solving the problem of the unsteady subsurface flow means that a function $h(x, z, t)$ must be found that satisfies equation 6.3 in a two-dimensional flow area Ω for $t > 0$, where the spatial distribution of the pressure head at $t = 0$, as given by:

$$h(x, z, t = 0) = h_0(x, z) \quad (5.6)$$

is known, and on the boundary Γ of boundary area flow Ω the boundary conditions are met.

The boundary conditions may be specified with a prescribed pressure head (Dirichlet type) on boundary Γ_1 :

$$h(x, z, t) = f(x, z, t) \quad (5.7)$$

where f is defined in time function at the appropriate part of boundary Γ .

The boundary condition can also be specified by a prescribed flux (Neumann type) normal to the boundary Γ_2 :

$$-k_r(h) \left[\left(k_{xx} \frac{\partial h}{\partial x} + k_{xz} \frac{\partial h}{\partial z} + k_{xz} \right) \cos \alpha + \left(k_{zx} \frac{\partial h}{\partial x} + k_{zz} \frac{\partial h}{\partial z} + k_{zz} \right) \cos \gamma \right] = q_1(x, z, t) \quad (5.8)$$

A third boundary condition may be given by a prescribed flux (Cauchy type) normal to the boundary Γ_3 :

$$q_n(x, z, t) = q_2(x, z, t) + a(x, z, t)h(x, z, t) \quad (5.9)$$

In the case of the unsteady subsurface flow, special attention should be given to the so-called potential boundary conditions, which are of the Dirichlet or Neumann types [Neuman, 1973]. Here, the boundary conditions are not prescribed. The occurrence of a given type of boundary condition is influenced by the actual instantaneous surface distribution of pressure head and the ability of the flow system to transport water. Infiltration and evapotranspiration phenomena, defined in the form of the highest possible flow rate Q_{pot} , provide such a boundary condition. For this kind of boundary, while solving equation (6.3), the following conditions must be met:

$$|k_r(h) \left[\left(k_{xx} \frac{\partial h}{\partial x} + k_{xz} \frac{\partial h}{\partial z} + k_{xz} \right) \cos \alpha + \left(k_{zx} \frac{\partial h}{\partial x} + k_{zz} \frac{\partial h}{\partial z} + k_{zz} \right) \cos \gamma \right]| \leq |Q_{pot}| \quad (5.10)$$

where $h_{Lim} \leq h \leq 0$, and h_{Lim} is the minimum permissible pressure head on the soil.

In the case of infiltration on the surface of the boundary Γ_2 , with prescribed flow rate Q_{pot} , the values of the calculated pressure head cannot be higher than $h = 0$ (assuming free surface outflow), and for the evapotranspiration lower than h_{Lim} . From condition (5.10) it follows that, at the contact between soil and atmosphere, the boundary is described

during the calculating process by boundary conditions of the Neumann ($q_1 = Q_{pot}$) or Dirichlet type ($h = 0$ for infiltration or $h = h_{Lim}$ for evapotranspiration), depending on the distribution of the pressure head on the surface.

The length of the boundary with the seepage face is not known 'a priori' during the simulation of the subsurface flow process. An outflow of water from the area must occur, and the pressure head on the boundary should equal zero:

$$\varphi = h + z = z; \quad q < 0 \quad (5.11)$$

In unsaturated zone, however, the water cannot flow out of the system, and the pressure head on the boundary must be negative:

$$\varphi = z + h < z; \quad q = 0 \quad (5.12)$$

Each of the states described by equations 5.11 and 5.12 can turn one into another depending on the boundary restrictions and the distribution of water potentials in the subsurface area. Meeting the conditions of free outflow means changing the type of boundary condition, from a Neumann into a Dirichlet type, and the other way round.

5.3. Numerical approximation of the Richards equation

The non-linear differential equation 6.3 with initial condition 5.6 and boundary conditions 5.7-5.9 cannot be solved analytically. In the numerical solution the finite element method (FEM) was applied. The orthogonalization Galerkin finite element method was used in order to derive the FEM equations for the Richards equation [Zienkiewicz, 1971].

Let the area Ω , where equation (6.3) is defined, be divided into linear triangle elements L_e , where nodes are placed in the corners of the elements in such a way that they are common for the adjoining elements. For each element (e) the pressure head h can be approximated in the following way:

$$h^{(e)}(x, z, t) = \sum_{i=1}^3 N_i(x, z) h_i(t) \quad (5.13)$$

where N_i is the base function for linear triangle elements (e), and $h_i(t)$ is a nodal value of pressure h in time t in element (e).

Having defined the residual and forced the weighted residual to zero, the Richards equation is approximated in each triangle element as:

$$\int \int_{\Omega^e} N_i \left\{ \frac{\partial}{\partial x} \left[k_r(h) \left(k_{xx} \frac{\partial h^e}{\partial x} + k_{xz} \frac{\partial h^e}{\partial z} + k_{xz} \right) \right] + \frac{\partial}{\partial z} \left[k_r(h) \left(k_{zx} \frac{\partial h^e}{\partial x} + k_{zz} \frac{\partial h^e}{\partial z} + k_{zz} \right) \right] + S - (C(h) + S_w S_s) \frac{\partial h^e}{\partial t} \right\} dx dz = 0 \quad (5.14)$$

where $i = 1, 2, 3$. It is assumed that in element (e) coefficients of hydraulic conductivity at saturation and specific storage S_s are constant, and $k_r(h)$, $C(h)$ as well as $S_w(\theta)$ vary linearly and are interpolated by the base function N_i . Having calculated each integral in

equation (5.14) and having conducted the aggregation for the whole area the following non-linear first-order differential equations are derived:

$$A_{ij}h_j + F_{ij}\frac{\partial h_j}{\partial t} = Q_i - B_i + D_i \text{ with } i, j = 1, 2, \dots, m \quad (5.15)$$

where

$$A_{ij} = \sum_e \frac{Le}{4\Delta} \bar{k}_r [k_{xx}b_i b_j + k_{zz}(b_i c_j + b_j c_i) + k_{zz}c_i c_j]$$

$$F_{ij} = \sum_e \frac{Le}{12} \frac{\Delta}{\Delta} [(2C_i + C_{j1} + C_{j2}) + S_s (2S_{wi} + S_{wj1} + S_{wj2})] \text{ for } i = j$$

$$F_{ij} = 0 \text{ for } i \neq j$$

$$Q_i = - \sum_e \frac{Le}{2} \frac{(Lq)_i}{2}$$

$$B_i = \sum_e \frac{Le}{e} 0.5 \bar{k}_r (k_{xx}b_i + k_{zz}c_i)$$

$$D_i = \sum_e \frac{Le}{e} \frac{\Delta}{3} \cdot S_i$$

$$\bar{k}_r = \frac{1}{3} (k_{ri} + k_{rj1} + k_{rj2})$$

where $(Lq)_i$ is the flow rate (5.8) across any side of the triangle, of length L , which includes nodal point n , $D_i \neq 0$ for nodes where internal source $S \neq 0$, j_1, j_2 are numbers of the other nodes in an element, m is the total number of nodes, Le is the total number of elements, Δ the area of a triangle element.

A fully implicit backward difference scheme is used for the time integration in terms of h . The matrix equation 5.15 is non-linear because the hydraulic conductivity, the water capacity and the degree of volumetric water saturation are functions of the pressure head h . In order to solve the non-linear matrix equation the Picard method is applied. The Successive Over Relaxation iteration method is used to solve the linearized matrix equation [Yeh, 1981].

5.4. Newtonian nudging as data assimilation method

Newtonian nudging relaxes the model state toward the observed state. This relaxation is performed by adding a term to the prognostic equation that is proportional to the difference between the two states. The model can be nudged toward regularly spaced observation or toward randomly spaced observation during a period of time and space. Therefore

two assimilation techniques are distinguished for the Newtonian nudging method. One of them, being connected with the observation net, is called 'nudging to the analysis', and the other one—'nudging to individual observations' [Houser et al., 1998a]. The 'nudging to the analysis' method means interpolating observation data to the model grid. The following expression is added to the model's predictive equation:

$$G \cdot W(x, y, z, t) \cdot \varepsilon(x, y, z) \cdot (h'_o - h) \tag{5.16}$$

where G is a factor that determines the magnitude of nudging term, W is a weighting function, ε is an analysis quality factor, and h_o is the observation at the model grid.

The nudging factor G determines the magnitude of the nudging term relative to all the other model processes. The four-dimensional weighting function W specifies its spatial and temporal variation. The analysis quality factor $0 < \varepsilon < 1$ is based on the quality and distribution of the observations.

In the 'nudging to individual observations' method the difference between the simulated and observed state is determined at the location of the observation, then interpolated back onto the model grid. The nudging term takes the following form:

$$G \cdot \frac{\left[\sum_{i=1}^N W_i^2(x, y, z, t) \gamma_i(x, y, z) (h_o - h'_i) \right]}{\sum_{i=1}^N W_i(x, y, z, t)} \tag{5.17}$$

where γ is an observational quality factor $0 < \gamma < 1$, h_o is a locally observed value of h , h' is the model's prognostic variable interpolated in three dimensions to the location of the observation.

The quality factor γ accounts for characteristic errors in measurement systems and representativeness. The subscript i indicates all of N observations, on whose area of influence model's prognostic variable is located.

The Newtonian nudging weighting function W at time t and location (x, y, z) for each observation i is a combination of the horizontal weighting function w_{xy} , the vertical weighting function w_z , and the temporal weighting function w_t :

$$W(x, y, z, t) = w_{xy} \cdot w_z \cdot w_t \tag{5.18}$$

Functions w_{xy} , w_z , w_t are defined as in Stauffer and Seaman [1994] (and as given in equations 4.3, 4.4 and 4.5, and illustrated in Figure 4.1).

The data assimilation method in the form of 'nudging to individual observations' was applied in the analysed model of subsurface flow. So the model's predictive equation, assuming that the influence areas of given observations are disjoint, takes the following form:

$$[C(h) + S_w S_s] \frac{\partial H}{\partial t} = \nabla \cdot [k_r(h) K \nabla H] + S + G_h W(x, z, t) \gamma_i(x, z) (h_o - h') \tag{5.19}$$

The values of the pressure head h' , being obtained from the model, can be directly interpolated in a triangular element at the observation point (x_o, y_o) by base functions N_i that are defined in the finite element method, which leads to the following expression:

$$h'(x_o, z_o) = N_1(x_o, z_o) \cdot h_2 + N_2(x_o, z_o) \cdot h_2 + N_3(x_o, z_o) \cdot h_3 \tag{5.20}$$

where h_i is a nodal value of pressure h in time t in element e .

When using the FEM, the referenced Newtonian nudging weighting function $w(x, z)$ is approximated in each linear triangular element by:

$$w_{xz}^{(e)} = N_1 \cdot w_2 + N_2 \cdot w_2 + N_3 \cdot w_3 \quad (5.21)$$

Having accepted the estimation (5.21) and applied the FEM to the orthogonalization Galerkin formulation, the introduced 'nudging to individual observations' term will make the expression that defines vector B_i in non-linear first-order differential equations (16) take the following form:

$$B_i = \sum_e^{Le} \left(\frac{1}{2} \bar{k}_r (k_{xz} b_i + k_{zz} c_i) + G_h \gamma (h' - h_0) w_t \frac{\Delta}{12} (2w_i + w_{j_1} + w_{j_2}) \right) \quad (5.22)$$

where: w_t is a temporal weighting function, $w = w_x w_z$ is a combination of the horizontal and vertical weighting functions, i, i_1, j_2 are subscripts which refer to the three corners of each triangle.

The pressure heads obtained from the tensiometer measurement, the water table positions measured by piezometers or the observed surface soil moisture by remote-sensing can constitute the model's assimilated value. In the case of the observed water table position the assimilated values of the pressure head at the observation point are $h_0 = 0$. An interesting case is presented by the assimilation of the observed surface soil moisture, whose values in the model are calculated on the basis of boundary conditions of the potential type, defined in the form of infiltration or evapotranspiration rate.

5.5. Application to the Brisy catchment

The discussed model of subsurface flow employing the data assimilation was applied to calculate flow processes in a cross-section of the Brisy catchment. The length of the cross-section was 220m, the thickness 3m, and the average hillslope was about 20° (See Figure 5.1). In this area four soil layers were distinguished. The thickness of particular layers

Table 5.1.: Soil parameters, per layer.

Layer nr	Ks [cm/day]	θ_s [cm ³ /cm ³]	S_s [1/cm]	θ_r [cm ³ /cm ³]	α	n
1	25	0.45	0.00016	0.030	0.0345	1.25732
2	25	0.45	0.00016	0.030	0.0345	1.25732
3	0.37	0.49	0.00016	0.005	0.01736	1.41398
4	0.37	0.49	0.00016	0.005	0.01736	1.41398

and the adequate soil hydraulic properties in the equations (5.2) and (5.5) are shown in Table 5.1. The cross-section fragment in the neighbourhood of the river Brisy are shown in Figure 5.1. Three piezometers P_1, P_2, P_3 , with the coordinates $x_1 = 20\text{ m}$, $x_2 = 90\text{ m}$, $x_3 = 180\text{ m}$, are located in the analysed cross-section. On the terrain surface we assumed the potential boundary conditions (equation (5.10)) which were defined by the infiltration or evapotranspiration rate. Along the river bottom we assumed the boundary condition Dirichlet type ($H_w(t) = h + z$, where H_w water level in the river). There is a possibility of

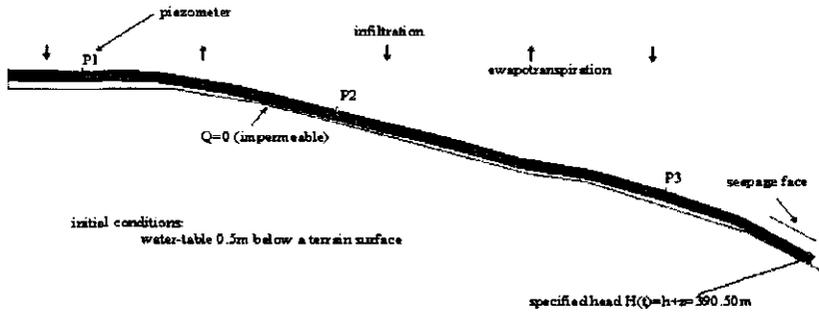


Figure 5.1.: The cross-section through Brisy catchment with boundary condition scheme and numerical mesh.

a seepage face on the slope above the river, along the 10m-section. The other banks of the river were assumed to be impermeable. The initial position of the water table was 50 cm below the terrain surface. Then, for 10 days the evapotranspiration at 0.6 cm/day was simulated. The overall simulation took place for the period 01.02.1993-27.09.1993, and the variations in the values of infiltration and evapotranspiration at this time are shown in Figure 5.2.

The observations (measurement) in piezometers P_1 , P_2 and P_3 were generated by numerical simulation in which K_s for layers 1 and 2 is 2.5 cm/day. Table 5.2 (column: $Sym - P_i$, $i=1,2,3$) presents the variations in the position of the water table in piezometers between simulation ($K_s = 25$ cm/day) and measurement (results from previous simulation). The simulated water table is below the observed one by 143 cm. In the following simulation the procedure of piezometer readings assimilation was started. Assumed that observations of the piezometers are conducted every other week. In the 'nudging to individual observations' method, for given piezometers we assumed the radius of influence for the horizontal weighting function $R_1 = 20$ m, $R_2 = 50$ m and $R_3 = 40$ m, as well as $R_z = 1.5$ m for the vertical weighting function. For the temporal weighting function the half period of a predetermined observation-influencing time window was 168 hours. Table 5.2 (columns $Nud - P_i$, $i=1,2,3$) presents the obtained findings, which resulted from the application of the model of the data assimilation procedure, in the form of variations of the water table position between simulation and measurement in given piezometers. In this case the maximum variation in all piezometers was not higher than 0.08 cm. Figure 5.3 shows how the water table is shaped in the neighbourhood of piezometer P_2 after 5952 hours of simulation. We can observe that at a distance longer than 30 to 40 meters the influence of measurement on the values calculated during simulation is positively decreasing.

Figure 5.4 shows the influence of the data assimilation process on the vertical distribution of pressure head in the neighbourhood of piezometers P_1 , P_2 and P_3 . It can be noticed

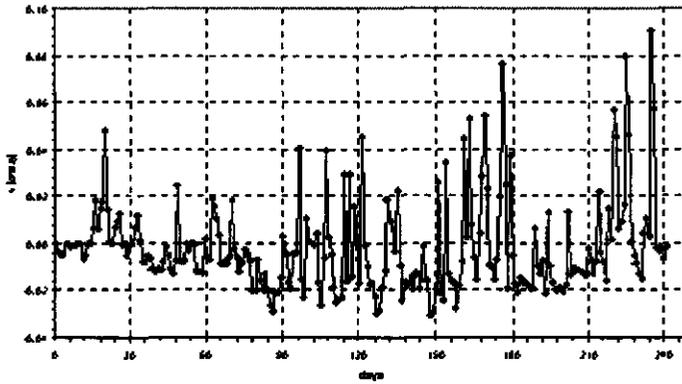


Figure 5.2.: Variation of infiltration (evapotranspiration) rate in period (01.02.93-27.09.93).

Table 5.2.: Comparison results from simulations with observed in piezometers

T [days]	Date	<i>sym</i> - P_1 [cm]	<i>nud</i> - P_1 [cm]	<i>sym</i> - P_2 [cm]	<i>nud</i> - P_2 [cm]	<i>sym</i> - P_3 [cm]	<i>nud</i> - P_3 [cm]
10	93-02-01	-18.26	-1.47E-01	-18.55	-1.09E-01	-15.09	-1.14E-01
24	93-02-15	-25.05	-1.29E-02	-23.64	-1.13E-02	-20.46	-1.23E-02
38	93-03-01	-6.14	2.57E-02	-1.73	1.10E-02	-1.62	1.31E-02
52	93-03-15	-6.19	-4.75E-02	-5.72	-4.51E-02	-4.38	-4.11E-02
66	93-03-29	-13.55	-1.54E-02	-14.03	-1.19E-02	-15.16	-1.35E-02
80	93-04-12	-15.86	-7.13E-03	-13.67	-6.27E-03	-14.75	-7.91E-03
94	93-04-26	-19.35	-6.67E-02	-21.57	-5.78E-02	-17.06	-5.91E-02
108	93-05-10	-22.25	-2.81E-02	-24.20	-1.68E-02	-19.27	-3.94E-02
122	93-05-24	-23.71	-3.59E-02	-22.65	-4.44E-02	-21.54	-1.62E-02
136	93-06-07	-22.78	-1.55E-02	-24.13	2.62E-03	-25.45	-1.48E-02
150	93-06-21	-25.75	-2.96E-02	-26.42	-2.68E-02	-28.03	-2.50E-02
164	93-07-05	-31.46	-5.19E-02	-31.53	-4.15E-02	-31.37	-3.77E-02
178	93-07-19	-36.58	-9.21E-03	-38.14	-3.91E-03	-37.22	-9.20E-03
192	93-08-02	-35.67	4.80E-02	-36.29	3.85E-02	-37.26	3.58E-02
206	93-08-16	-32.39	-1.93E-01	-36.18	-2.04E-02	-37.95	-1.75E-02
220	93-08-30	-38.83	-1.92E-02	-38.29	-2.72E-02	-40.35	-2.48E-02
234	93-09-13	-41.98	3.30E-02	-42.85	-4.01E-03	-44.15	-7.44E-03
248	93-09-27	-42.20	1.23E-01	-39.48	7.81E-02	-43.77	7.15E-02

that the water table position in the piezometer is higher than its position obtained from simulation without assimilating their readings. The introduction of the data assimilation procedure to the model results in raising the water table, which leads to the rise in pressure both in the unsaturated and saturated zones.

5.6. Discussion and Conclusions

The example discussed in this Chapter indicates the high potential of the nudging technique for complex models like the one employed here. Naturally, it is important to answer the question to what degree the observations conducted are correct. In the 'nudging to individual observations' method the measurement quality can be controlled by parameter γ . The other parameters (G , R_z , R_x as well as τ , also do not possess optimal values and should be selected by means of numerical experiments - or perhaps a Kalman filter could be used for that purpose. It is worth stressing that the data assimilation procedure does not replace the stage of model verification. Improving the model, which never reflects the complexity of real processes, remains important. In practice the terrain geometry as well as the system of soil layers (along with the hydraulic parameters) are set on the basis of several scattered observations, or, with the lack of proper examination, these parameters are taken from literature. In such situations the data assimilation procedure also allows to correct the influence of these factors on the results obtained from the model. When observations are taken in the area of the influence of real boundary conditions (e.g. soil moisture in the surface soil layer), data assimilation may be particularly efficient. In that respect it is fortunate that Remotely Sensed observations particularly measure that part of the hydrologic system. The model together with the data assimilation procedure presented in this Chapter, can be applied to virtually any 2D hydrologic flow problem.

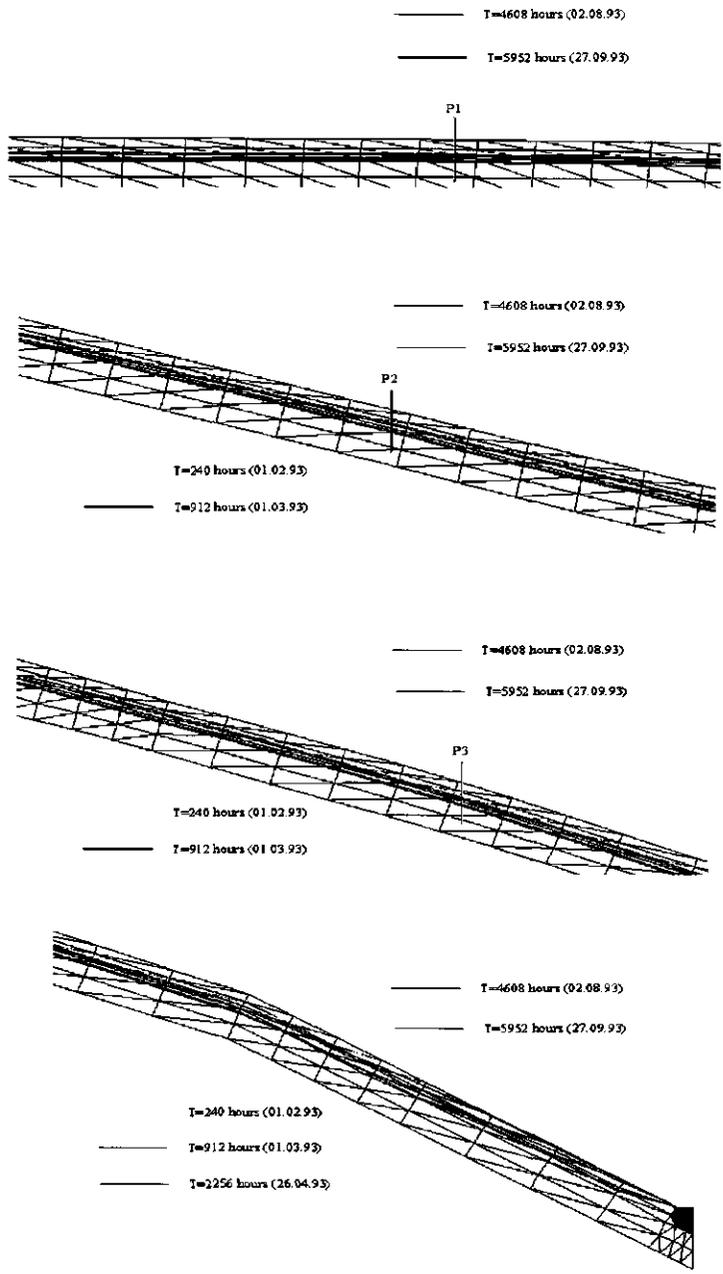


Figure 5.3.: Water-table position at selected time instants, in the neighbourhood of P1, P2, P3 and the river.

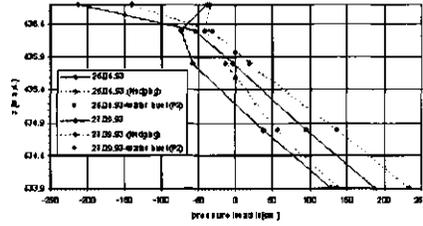
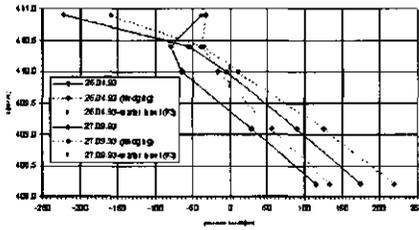
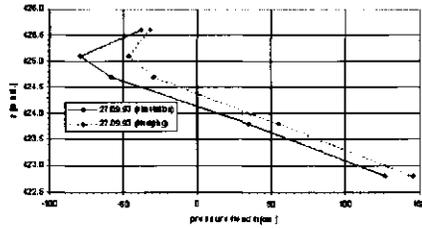
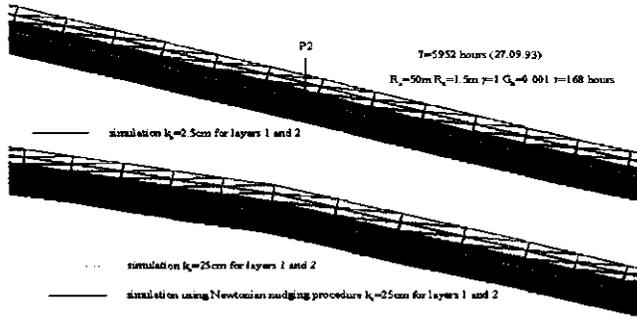


Figure 5.4.: Top: Differences between water-table position for simulation with and without Newtonian nudging. Middle: Variation of pressure head with depth (distance 37.6m from piezometer P2, $x=127.6m$). Lower left: Variation of pressure head with depth near piezometer P3, $x=179.6m$. Lower right: Variation of pressure head with depth near piezometer P1, $x=21m$.

6. Developing and testing a regularization technique

Emiel van Loon and Peter Troch

6.1. Introduction

The ability to predict soil water storage and movement in a heterogeneous landscape is important to manage water resources. Many studies have addressed spatial variability of soils and soil water by recognizing the stochastic nature of local variability [Greminger *et al.*, 1985; Unlu *et al.*, 1990; Yeh *et al.*, 1986]. Also systematic components have been identified and linked to topographic characteristics, [Moore *et al.*, 1991; Hanna *et al.*, 1982; Hairston and Grigal, 1991], soil morphological features [Kreznor *et al.*, 1989], or chemical and physical attributes [Brubaker *et al.*, 1993]. The integration of both systematic and stochastic components has partially been achieved by conditioning geo-statistical techniques with secondary data such as topographic indices via (indicator) co-kriging [Lehmann, 1995; Western *et al.*, 1998, 1999]. However, geo-statistical techniques do not explicitly incorporate the knowledge about system dynamics and cannot easily take advantage of additional conditioning information at various scales such as catchment discharge or evapotranspiration from different vegetation patches. Such an integration of observations from various sources with a dynamic model is known as data assimilation. Examples of data assimilation techniques applied to soil moisture estimation are found in Callies *et al.* [1998], Calvet *et al.* [1998], Galantowicz *et al.* [1999], Hoeben and Troch [2000a], Houser *et al.* [1998b], Katul *et al.* [1993] and Mahfouf [1991]. All these data assimilation studies have in common that they consider one-dimensional soil water movement (i.e. in the vertical direction), while utilizing only rain and remotely sensed soil moisture estimations as observations.

In summary we can say that there are two distinct types of soil moisture studies: those that focus on the lateral soil moisture distribution and use static models, versus those that consider the vertical distribution of soil moisture while using dynamic models. The first type of studies use mainly field observations of soil moisture in combination with soil and terrain properties, while the latter almost exclusively use remote-sensing observations. It is the aim of this study to combine elements from both areas, as a first step towards an integration of the two approaches. More precisely, in this study a data assimilation technique is applied to a system where both lateral and vertical soil water movement take place and for which only ground-based observations are available. First a distributed hydrological model (which will be called the d-model) as well as a geo-statistical soil moisture model (s-model) is developed. Then a data assimilation algorithm (the da-model) is developed to combine the results from both models. The relative efficiency of the da-model is compared with with the alternative methods of state estimation through

the straightforward application of the d-model, the re-calibrated distributed hydrological model (the dc-model) or the s-model.

Each model is parameterized for a number of resolutions. The reason for considering different resolutions is that *a priori* it is unclear at which resolutions the different methods will perform best. Different model parameterizations are required at different resolutions because the models for the hydrological system under consideration can generally not be defined at multiple resolutions [van Loon et al., 2000].

6.2. Material and Methods

6.2.1. Description of data

The data for this study have been collected in a 44 ha catchment. A series of low-resolution measurements is available from 20 July 1997 till 21 December 1997 and within this period a set of high-resolution measurements is available for 4 October till 21 December. At four locations rain has been measured using tipping buckets. At two locations discharge has been measured at 1-min. time instants, using v-crest weirs. Ground water depth has been observed manually in 20 piezometers at hourly instants during and just after rain, and daily between rainfall events. During the period 20 July - 4 October, volumetric soil moisture has been measured at 40 locations once every 4 days, and during the period 4 October - 21 December at 60 locations once every 2 days. For the soil moisture measurements a Trime TDR system (in plastic tubes) has been used, enabling the measurement of soil moisture over 20 cm layers down to 80 cm. The location of the various instruments is shown in Figure 6.1. Close to the Trime tubes about 150 soil moisture measurements have been taken with in the 0-10 cm topsoil every 10 days, using a TDR system which was directly inserted into the ground. These measurements have been used to correlate soil moisture observations with soil and terrain properties and for checking the other measurements.

Terrain and soil have been mapped in detail. The terrain has been measured using a kinematic GPS technique in combination with a conventional ground-based survey. Soil colour, soil depth, the dimensions of cracks, stability of soil aggregates, organic matter content, and texture have been determined at 90 locations, and the hydraulic permeability has been determined at 30 of these locations, using a Guelph permeameter (both at 10 and 20 cm depths). In addition soil colour, the dimension of cracks, areal density of cracks and texture (field-determined) have been observed at a regular spacing of 20×20 m. The purpose of the soil data is to either relate these directly to soil moisture behaviour or group them into a limited number of model units.

A contour map and the distribution of upstream area (A) and the $\ln\left(\frac{A}{\tan\beta}\right)$ index are shown in Figures 6.2 and 6.3. The data collected at this site has been extensively described and analyzed with regard to overland flow in van Loon and Keesman [2000] and van Loon et al. [2001]. In this study the emphasis will be on soil moisture.

6.2.2. Formulating the distributed hydrological model

The soil moisture distribution in the research catchment is described by a model that considers three horizontally distributed state variables: surface water (q), moisture in the top 0.4 m of the soil (w) and water in the deeper soil layer (0.4 m down to parent

material, which varies in depth from 0.4 to 3 m, d). In the model w represents the unsaturated zone and d represents the saturated zone, which implies that soil moisture observations are related to the value of w and piezometric observations to the ground water level in d . Horizontally the terrain is subdivided by a regular grid. The deeper soil layer is assumed to be impermeable at the bottom, and there is assumed to be no flow at the catchment boundaries. Both surface flow as well as flow in the upper soil is assumed to be one-dimensional, following the direction of steepest descent as defined by a D8 algorithm [Marks et al., 1984]. Flow in the deeper soil layer is assumed to be driven by water table differences and is two-dimensional. Atmospheric forcings are rain (p) and potential evapotranspiration (pe). The surface elevation (z), the maximum infiltration capacity (p_{max}) and the maximum storage capacity of the upper soil and deeper soil layers (w_{max} and d_{max} respectively) are time invariant variables in the model. The couplings between the different model-components are schematically drawn in Figure 6.4.

The fluxes indicated in Figure 6.4 are computed based on the method described fully in van Loon et al. [2000]. In essence, the fluxes are determined by empirical relationships made up of linear segments (so-called kernel functions). Figure 6.5 illustrates the form of the kernel functions. In this figure the variable at the horizontal axis (v_i) is a scaled model variable ($w_i/w_{max,i}$, $d_i/d_{max,i}$), a scaled forcing ($p_i/p_{max,i}$) or an average slope between locations i and j ($\Delta z_{i,j}/\Delta x_{i,j}$), depending on the model state variable to which the kernel function applies. The variable on the vertical axis, $\theta_{st}(i, j)$ is a transport fraction ($0 \leq \theta_{st}(i, j) \leq 1$). For transport in the vertical direction $\theta_{st}(i, j)$ gives the fraction of transport from the system compartment s_i to t_i (s_i and t_i refer to any two state variables with a connection in Figure 6.4, f.i. q_i and w_i). For transport in the horizontal direction, $\theta_{st}(i, j)$ gives the fraction of transport from the system compartment s_i to s_j (e.g. from q_i to q_j in Figure 6.4). Note that v_i is always defined by the state s_i (not by t_i). The value of $\theta_{st}(i, j)$ is not treated as deterministic but is a stochastic variable with an identified unknown-but-bounded probability distribution (illustrated by the shaded area around the black lines in Figure 6.5), rendering the model kernel function stochastic in the parameters [van Loon et al., 2000].

The value of $\theta_{st}(i, j)$ is computed through the following steps:

- applying the appropriate kernel function in Figure 6.5 (I, II or III);
- finding the appropriate parameter range as the ordinate value associated with the corresponding v_i value;
- randomly drawing a value for $\theta_{st}(i, j)$ out of the parameter range.

A flux is then given by

$$f_{st}(i, j) = \theta_{st}(i, j)s_i \quad (6.1)$$

where s_i [L] is the value of the state variable s at grid cell i and $f_{st}(i, j)$ [L] is the flux computed at a given time step from a state variable s_i to a state variable t_i (vertical flux), or from a state variable s_i to an adjoining state variable s_j (lateral flow in the deeper soil) or a state variable downslope from s_i (lateral flow at the surface flow or flow in the upper soil). The model is chosen to be discrete in time, so that the flux $f_{st}(i, j)$ in Equation 6.1 is defined at time instant k if s_i is taken for $k - 1$. The precise formulation of the model is given in [van Loon and Troch, 2001].

The reason for using this model form is twofold. In the first place this particular definition allows to capture a large variety of system representations within a single mathematical framework, as explained in *van Loon et al. [2000]*. In the second place, at each time step the model is linear but can easily represent non-linearities since the kernel functions form in fact the linearization of the non-linear model. Both properties are useful in this study.

The above model is defined at a range of time and space resolutions. The temporal resolution is set to 5, 10 and 15 minutes, and the spatial resolution is set to 20×20 m, 40×40 m and 60×60 m. In total, there are thus 9 different models. These temporal and spatial resolutions are selected from a wider spectrum of model structures.

The models have been derived with the identification method described in *van Loon et al. [2000]*, using 3/4th of the observations (randomly selected) over the period 20 July till 20 October. On the basis of the remaining data over this period, the models are calibrated with the technique described in *Klepper and Rouse [1991]* and *Stol et al. [1992]*, using the average RMSE of the predicted soil moistures as a goodness-of-fit criterium. Note that the 9 models differ in three respects: the time step used in the calculations, the spatial resolution, and the shape of the calibrated kernel functions. All models contain the same state variables (s , w and d), forcings (p and pe), additional variables (p_{max} , w_{max} , d_{max} and z) and fluxes (f_{st}).

6.2.3. Formulating the geo-statistical model

The geo-statistical model used in this study is based on independent spatial and temporal prediction. A soil moisture estimate is the weighted average of the spatial and the temporal models. Both statistical models are linear. The temporal model predicts soil moisture at a certain location for the same (or closest) locations on the basis of observations at a previous time instant, and the spatial model predicts the soil moisture at a certain location on the basis of observations within the catchment at the same time instant. To simplify the notation of the geo-statistical model we first concatenate the soil moisture of the individual grid-cells i at a time instant k ($w_{i,k}$) in a vector \mathbf{w}_k . An additional vector \mathbf{w}_k^o contains the values of the individual measurements at a time instant k . The soil moisture estimation, which is based on \mathbf{w}_{k-1} and \mathbf{w}_k^o , is named \mathbf{w}_k^* . The model can then be summarized as follows

$$\mathbf{w}_k^* = (1 - o_k)\mathbf{K}\mathbf{w}_{k-1} + (o_k)\mathbf{L}\mathbf{w}_k^o \quad (6.2)$$

where \mathbf{K} and \mathbf{L} are cross-correlation matrices for \mathbf{w} at a previous time instant and at distant locations respectively. o_k is a parameter which is set zero for time instants where no observations are available, and has a fixed value when observations are available. Through \mathbf{L} Equation 6.8 implicitly converts point observations to averaged grid-cell values. The matrices \mathbf{K} and \mathbf{L} are fixed for the entire simulation period. The structure of these matrices is determined by trial and error and will be explained in Section 6.3.1. The parameter o_k is determined (after fixing \mathbf{K} and \mathbf{L}) on the basis of a calibration procedure, using the observations over the period 20 July till 20 October and the average RMSE of the predicted soil moistures as a goodness-of-fit criterium.

The geo-statistical model is defined at the same resolutions as the distributed hydrological model, i.e. temporal resolutions of 5, 10 and 15 minutes, and spatial resolutions of 20×20 m, 40×40 m and 60×60 m. The temporal resolutions are so small, relative to the sampling intervals of two or four days, that that these do not influence the model parameterization.

For that reason only three different geo-statistical models are defined (one for each spatial resolution).

6.2.4. Formulating the data assimilation procedure

The distributed hydrological model outlined in Section 3.2.2 can be formulated as a time-varying linear discrete state-space model of the following form.

$$\mathbf{x}_k = \mathbf{A}_k \mathbf{x}_{k-1} + \mathbf{B}_k \mathbf{u}_k + \mathbf{v}_k \quad (6.3)$$

$$\mathbf{y}_k = \mathbf{H}_k \mathbf{x}_k + \mathbf{v}_k^* \quad (6.4)$$

In these equations the vector \mathbf{x}_k contains the state variables of surface water (q), moisture in the top 40-cm of the soil (w) and water in the deeper soil layer (0.4 m to parent material, which varies in depth from 0.4 to 3 m, d), for each grid cell. The vector \mathbf{u}_k contains the atmospheric forcings at time instant k , rain (p) and potential evapotranspiration (pe). The matrices \mathbf{A}_k and \mathbf{B}_k contain the parameters $\theta_{st}(i, j)$ (as explained earlier these are known functions of \mathbf{x}_{k-1} and \mathbf{u}_k). \mathbf{A}_k can be seen as an operator that defines spatial redistribution of soil moisture in both the lateral and the vertical direction. \mathbf{B}_k is an operator that defines the spatial redistribution of rain and evapotranspiration, i.e. movement in the vertical direction only. The vector \mathbf{y}_k contains the observations, which are in this study discharge and water table depth. \mathbf{H}_k is the observation matrix that relates the model state variables to observations. The vector \mathbf{v}_k contains model errors and \mathbf{v}_k^* observation errors. In what follows we will assume that the distributions of \mathbf{v}_k and \mathbf{v}_k^* are unknown but bounded, so that \mathbf{x}_k as well as \mathbf{y}_k are characterized by a minimum and a maximum ($\underline{\mathbf{x}}_k > \mathbf{x}_k > \bar{\mathbf{x}}_k$, $\underline{\mathbf{y}}_k > \mathbf{y}_k > \bar{\mathbf{y}}_k$). The initial condition for this system, \mathbf{x}_0 , is known. *van Loon et al. [2000]* describe a system to identify models with the structure of Equation 6.3. A more elaborate description of Equations 6.3 and 6.4 is given in Appendix 1.

The estimation of \mathbf{x}_k can be achieved by combining Equations 6.3 and 6.4 with a Kalman filter. The drawback of this approach is that it requires accurate knowledge of the error structure of both the model and the observations to give a relatively good performance. Moreover it is not very robust in cases where an incorrect model is used [*Anderson and Moore, 1979*]. In this study the structure of the model error is non-Gaussian and largely unknown (as will be illustrated later). Moreover, due to the indeterminacy of the problem in this study, it is not the *optimal* weighting of the observations and model predictions (one of the assets of the Kalman filter) but rather the *robust* weighting, which is crucial for good performance. *Optimality* is namely not defined for indeterminate problems. For these reasons a different approach was chosen to combine model predictions with observations, making use of a technique called Tikhonov regularization [*Tikhonov and Arsenin, 1977; Johansen, 1997*]. Recall that Equations 6.3 and 6.4 can be used straightforwardly for prediction by supplying the necessary inputs and draw realizations of the stochastic parameter vectors in Equation 6.3. If there are, in addition, observations (in \mathbf{y}_k) available, Equation 6.4 can be inverted in order to derive $\hat{\mathbf{x}}_k$. The weighted average of this result and the state vector estimated by 6.3 gives the desired state estimate. This problem of state estimation is ill-posed since the observation matrix \mathbf{H}_k (and consequently $\mathbf{H}_k \mathbf{A}_k$ as well as $\mathbf{H}_k \mathbf{B}_k$) is not of full rank. Assumptions have to be made about the variables to be estimated in order to find a solution to an ill-posed problem. In this context the making of assumptions is called regularization.

Before we will apply regularization, Equation 6.3 is substituted into 6.4. Hence,

$$\mathbf{y}_k - [\mathbf{H}_k \mathbf{B}_k] \mathbf{u}_k = [\mathbf{H}_k \mathbf{A}_k] \mathbf{x}_{k-1} + (\mathbf{H}_k \mathbf{v}_k + \mathbf{v}_k^*) \quad (6.5)$$

and for ease of notation we rename this equation to

$$\mathbf{z}_k = \mathbf{D}_k \mathbf{m}_k + \mathbf{e}_k. \quad (6.6)$$

Now the elements of the vector $\mathbf{m}_k = [x_{k-1}(1) \dots x_{k-1}(n)]^T$ are considered model parameters that have to be estimated. A regularization technique finds a pseudo-inverse of \mathbf{D}_k , by constraining the solution space of \mathbf{m}_k . In this study we use the weighted average of prior estimates and additional observations, not included in \mathbf{y}_k , as additional constraints (the weighted average is named \mathbf{m}_k^*). Adding these additional constraints to Equation 6.6, gives a problem that can be solved in a least-squares sense:

$$\begin{bmatrix} \mathbf{z}_k \\ r_k \mathbf{m}_k^* \end{bmatrix} = \begin{bmatrix} \mathbf{D}_k \\ r_k \mathbf{E} \end{bmatrix} \mathbf{m}_k + \begin{bmatrix} \mathbf{e}_k \\ \mathbf{e}_k^* \end{bmatrix} \quad (6.7)$$

where \mathbf{E} is a matrix with ones and zeros to select those elements from \mathbf{m}_k that correspond to \mathbf{m}_k^* , r_k is a so-called regularization parameter, which can be established on the basis of various criteria such as entropy, some sort of cross-validation or an l-curve [Tarantola, 1987; Hansen, 1992]. Here a technique known as generalized cross-validation is used to determine the desired value of r_k [Hansen, 1998]. In this study \mathbf{m}_k^* contains only soil moisture data, hence $\mathbf{m}_k^* = \mathbf{w}_k^*$. We will however continue to use \mathbf{m}_k^* to keep Equation 6.7 general. \mathbf{m}_k^* is defined as the weighted sum of the estimations at a previous time instant (\mathbf{m}_{k-1}) and additional observations ($\mathbf{m}_k^o = \mathbf{w}_k^o$), analogous to the geo-statistical model introduced in Section 6.2.3:

$$\mathbf{m}_k^* = (1 - o_k) \begin{bmatrix} 0 & & \\ & \mathbf{K} & \\ & & 0 \end{bmatrix} \mathbf{m}_{k-1} + (o_k) \begin{bmatrix} 0 & & \\ & \mathbf{L} & \\ & & 0 \end{bmatrix} \mathbf{m}_k^o \quad (6.8)$$

where the cross-correlation matrices \mathbf{K} and \mathbf{L} and the parameter o_k are the same as in Equation 6.2.

Whereas the values of \mathbf{z}_k , \mathbf{m}_k^o and \mathbf{m}_{k-1} are loose constraints on the values that \mathbf{m}_k may take, hard constraints can be imposed as well. In this study the mass balance at the catchment scale is imposed as a hard constraint on Equation 6.7. It is given by

$$\sum_{l=1}^L (w_{k,l} - w_{k-1,l} - q_{k,l} - d_{k,l} + p_{k,l} - ea_{k,l}) = 0 \quad (6.9)$$

where l ($1 < l < L$) is an index indicating each spatial element in the catchment and $ea_{k,l}$ is actual evapotranspiration ($ea_{k,l}$ is defined in Appendix 1). The precise mathematical operations involved to solve Equation 6.7, under the constraint given by 6.9 are given in Appendix 2.

It should be noted that in a model with time-variable parameters, state estimation is also possible via parameter estimation. This can be seen when the substitution of Equation 6.3 into 6.4 is expressed as

$$\mathbf{y}_k = [\mathbf{H}_k \mathbf{A}_k] \mathbf{x}_{k-1} + [\mathbf{H}_k \mathbf{B}_k] \mathbf{u}_k + \mathbf{e}_k, \quad (6.10)$$

whereafter the known vectors \mathbf{x}_{k-1} and \mathbf{u}_k are put into the data matrix \mathbf{D}_k , and all the unknown parameters (in \mathbf{A}_k and \mathbf{B}_k) are put in the parameter vector \mathbf{m}_k (for details see Appendix 3). This results in an equation similar to Eq. 6.6: $\mathbf{y}_k = \mathbf{D}_k \mathbf{m}_k + \mathbf{e}_k$, where the elements of the vector \mathbf{y} are observations, the matrix \mathbf{D} contains now the previous state as well as the inputs to the system and the elements in the vector \mathbf{m} are now the model parameters. After estimating the parameter vector \mathbf{m}_k , \mathbf{x}_k can be estimated through Equation 6.3.

The results of state estimation and parameter estimation are similar. However, the extension of the vector \mathbf{y}_k with additional observations \mathbf{m}_k^o is more elaborate with parameter estimation as it is in the case of state estimation. Moreover, with the parameter estimation approach the correlation matrices cannot readily be interpreted. For these reasons the state estimation approach is used here.

6.2.5. Comparing data assimilation with alternative estimation techniques

From Sections 3.2.2 to 6.2.4 it is clear that Equation 6.7, can be considered as a combination of a dynamic model (the upper row) and a statistical model (the lower row). By using each model separately and comparing the results with those obtained by the models in combination, the relative performance of the data assimilation approach can be assessed. This line of investigation is followed here, using different densities of observations in space as well as time, and considering different model resolutions. More specifically the following approach is applied.

1. The data set is divided into two parts: A) 20 July till 20 October, and B) 20 October till 21 December. From this last period nine sub-sets with soil moisture data are formed. These sub-sets differ with respect to data coverage in space and time, and a fixed portion of the observations (25%) is used as validation data (see Table 6.1). Observations in A are used for model identification and calibration, and those in B are used for prediction.
2. A dynamic model is identified and calibrated at each of the nine resolutions on the basis of rainfall, discharge, piezometer and soil moisture observations in the period A, as well as the soil and terrain information (see section 3.2.2). These models are called d-models.
3. Statistical models for soil moisture are derived on the basis of the observations in period A at each of the three spatial resolutions (the temporal resolution does not play a role in this model). The derivation is limited to the coefficients in the covariance matrices \mathbf{K} and \mathbf{L} , for which simple forms suffice (see Eq. 6.8), and the value of the weighting parameter σ_k . The resulting models are called s-models.
4. The d-models are applied in open-loop form (i.e. using only rainfall and potential evapotranspiration observations) to period B. Subsequently the performance is measured with the validation data.
5. The d-models are re-calibrated for each of the nine sub-sets of period B, using soil moisture, discharge and piezometer data for this period. The re-calibrated d-models are called dc-models. Then the nine resulting dc-models are applied in open-loop form and the performance is measured with the validation data.

6. The *s*-models are applied to each of the nine sub-sets in period B, and the performance is measured with the validation data.
7. Combinations of the *d*-models and *s*-models (defined at corresponding resolutions) are made according to Eq. 6.7. The combinations are called *da*-models. The *da*-models are applied to each of the nine sub-sets in period B, and the performance is measured with the validation data.

An alternative approach to make use of the additional soil moisture data in period B, could be to redo the identification and then compare the resulting models (the *d**-models) with the other models (*dc*- *da*- and *s*-models). Its relevance notwithstanding, this comparison is not used here for several reasons. In the first place because the parameterization of the *d**-model set could be very different from the *d*-model set (use different kernel functions) and therefore difficult to compare. The *d**-model set could actually perform worse than the *d*-model set because the existing identification techniques do not guarantee that the additional information is used optimally. Secondly, in many situations one has a ready-to-use dynamic hydrologic model, so that modifying the structure is not an issue anymore and it is exactly the tedious task of full calibration which one would like to bypass.

The different subsets for period B are established by a latin-hypercube sampling scheme, where 25, 50 and 75% of the observation times are combined with 25, 50 and 75% of the observation locations (thus yielding 9 combinations). The naming of the subsets is shown in Table 6.1.

Several aspects of the model performance are investigated. In the first place the deviation of the estimated from the observed soil moisture is considered. For this purpose the root mean squared error (RMSE) is used:

$$RMSE = \sqrt{\frac{\sum_{i=1}^I \sum_{j=1}^J (obs_{i,j} - pred_{i,j})^2}{I \cdot J}}$$

where the indices *i* and *j* refer to the *i*th time instant and *j*th location of the validation data, *obs*_{*i,j*} are observations, *pred*_{*i,j*} the model predictions and *I* and *J* the total number of time instants and locations for which validation data are available. Secondly, the spatial as well as temporal structure of the error of the soil moisture predictions by the three model sets is compared by error maps and time series respectively. Furthermore the spatial and temporal errors are related to various spatial and temporal attributes. The last aspect which is investigated, is the effect of soil moisture data assimilation on the prediction of discharge and ground water depths.

6.3. Results

6.3.1. Distribution of soil moisture in space

Soil moisture at our test site is not related very well to any particular soil or terrain feature. This is an important indication that a dynamic model is required to represent soil moisture at the desired accuracy. In addition, rather different soil moisture regimes exist within the catchment area. These moisture regimes are strongly related to soil classes defined by a combination of clay content and soil depth. Three soil classes can be identified: 1) soils with a depth less or equal than 0.2 m ('shallow'), 2) soils deeper than 0.2 m and

uncertainty. The parametric uncertainty is summarized by an index which is called 'total parameter range'. This value represents surface of the shaded area in Figure 6.5 relative to the entire plot area (averaged for all the kernel functions in a model), which implies that a value of zero refers to completely certain parameter values and a value of one to completely uncertain parameter values. The table also shows that the models perform quite similar in terms of RMSE for the calibration period (period A). With regard to the RMSE for the validation period (period B), the models show more differences. It is especially the parameter-rich models, which also appear to have a large total parameter range, that have a relatively low performance (high RMSE) for the validation set in comparison to the calibration set. Note that the number of parameters per kernel function in Table 6.2 refers to average size of the vector $\theta_{st}(i, j)$ in Equation 6.1 (there are in total seven vectors $\theta_{st}(i, j)$, one for each type of flux). For all nine models the RMSE of the validation and calibration periods is still quite close. In addition, the model resolution with maximum performance is also similar for the two periods (i.e. time instants of 15 minutes and grid-cells of 40×40 m). The inter-comparison of the d-, dc-, s-, da-models shows the significant added-value of the soil moisture observations. When compared to the d-models, the other models have a lower RMSE and have their optimal performance at a finer resolution. This implies that taking additional soil moisture observations into account during period B is useful for getting improved soil moisture predictions. Therefore we will concentrate on the dc-, s- and da-models in what follows.

The effect of varying the size of the set with additional data is shown in Figures 6.8 to 6.10. Note that the axes (or cells) in each of the nine sub-plots of the figures are defined by Table 6.1, i.e. the lowest observation density is in the upper left corner and the highest observation density is in the lower right corner. For the dc-models the effect of the size of the soil moisture data set appears to be quite large (i.e. lower RMSE when using more soil moisture data in the re-calibration), especially for models at fine resolutions (see Figure 6.8). The dc-model defined at the coarsest resolution (15 min and 60×60 m) is performing best. The s-models provide better predictions than the dc-models for all three resolutions and especially at finest resolution (see Figure 6.9). Here the effect of the size of the soil moisture data set is similar for the three resolutions. As noted in Section 6.2.3, the predictions by the s-models are independent of time-resolution since observations are only available once every four or two days (if observations were available at resolutions lower than 15 minutes, temporal resolution would also become a relevant factor for the s-models). The comparison of both the dc- and s-models with the results from the da-models learns that the average RMSE of the da-models is smaller than that of either the dc- or s-models. The spatial resolution where lowest RMSE is found, is equal to that for the s-models (40 m, see Figures 6.9 and 6.10). The temporal resolution with lowest RMSE for the da-model is much smaller than that for the dc-model (5 versus 15 minutes respectively, see Figures 6.8 and 6.10). In this context it is of relevance that the heterogeneity of soil moisture within a grid cell (also expressed as average RMSE in vol %; and using data set 3.3) is 2.1%, 2.7% and 2.9% for 20×20 , 40×40 and 60×60 m grid cells respectively. When these values are accepted as observation uncertainty, intrinsic to the TDR-technique used in this study, the differences between the dc-, s- and da-models are still quite significant. But the most accurate predictions by the s-models and dc-models are within the range of observation uncertainty, so that the differences between e. g. the predictions at 5 min. and 20×20 m or 40×40 m grid cells are irrelevant.

Also the comparison of the error-structure for the different models shows interesting patterns. Spatially, the errors in the dc-models display the most structure, followed by the da-models, while the s-models have a random error structure (see Figure 6.11). In con-

trast, the error is highly correlated in time for the s-models while it is much less correlated for the dc- and da-models (see Figure 6.12). Chi-square tests reveal that the spatial error of the s-models is close to Gaussian, and that the temporal error for the dc- and da-models is relatively close to Gaussian.

Relating the prediction error to other variables shows that in particular upstream area, soil depth, rainfall intensity and total water storage in the catchment can explain the spatial and temporal structure of the errors for the various models. The relation between the average prediction error and these variables is shown in Figure 6.13.

6.3.3. Predicting discharge and ground water level

Discharge and ground water observations have been taken into account in the dc-models as well as the da-models (via variable y_k in Eq. 6.4). In Figure 6.14 the error of the discharge prediction by the d-, dc- and da-models is shown (only for the most downstream location, the results for the upstream location were very similar). Here the errors for the 9 resolutions are averaged. In Figure 6.15 the predicted ground water table depth is shown for three locations (also for the d-, dc- and da-models, averaged over the nine resolutions). The locations of the piezometric observations are indicated in Figure 6.1. Apparently the discharge predictions are enhanced considerably by using the additional soil moisture observations in both the dc- and da-models. For all three model sets there is on average an under-prediction of discharge, and the structure of the error is quite comparable. This result is perhaps not surprising but nor is it self-evident, as illustrated by the prediction of ground water depth, which seems not to benefit from the additional soil moisture data. For the prediction of ground water depth all three models yield quite similar errors (especially the d- and dc-models) while the errors at the distinct locations seem to be unrelated and quite persistent in their bias. An explanation for this failure to enhance ground water predictions is that ground water depth only weakly relates to the local soil moisture content in the upper soil layer. Since such a relation is imposed in the models (see Figure 6.4) and the lateral connections in deep soil are in this research catchment relatively slow, little can be expected of the enhanced soil moisture prediction by the model. In other words, the model does not represent the ground water component of the system correctly, which causes the ground water predictions to be in error in spite of soil moisture data assimilation.

6.4. Discussion and Conclusions

The problem of characterizing and eventually estimating the spatio-temporal soil moisture distribution has received considerable interest over the past years (see Section 6.3.1). For cases where field observations are available, geo-statistical techniques have been much evaluated and tested for suitability. It has been shown by *Lehmann [1995]* and *Western et al. [1999]* that soil moisture dynamics can only partially be captured by statistical approaches. Another line of investigation, using almost exclusively remote sensing observations, has adopted the data assimilation approaches developed in the atmospheric and oceanographic disciplines [*Bennett, 1992; Daley, 1991; McLaughlin, 1995*]. Until now data assimilation techniques have only dealt with dynamic one-dimensional soil moisture models. As stated in the introduction, this study attempts to integrate elements from both approaches, in this way leading to a data assimilation technique which is adapted to

hydrological problems. The key-ingredient for this integration is regularization, i.e. the appropriate weighting of both model components in a structured and objective manner. In this study this has been achieved through generalized cross-validation [Hansen, 1992; Hanke and Hansen, 1993; Tikhonov and Arsenin, 1977]. Another way of seeing regularization is as a structured way of combining prior information (or assumptions) with an ill-posed hydrological problem to convert it into a well-posed problem. Until now this approach has not been used for data assimilation, and with this study we demonstrate that the regularization approach to data assimilation may indeed be suited for hydrological problems. It was our objective to get an idea about the performance of such an approach. As expected, it performed considerably better than the application of either the dynamic model or the statistical model, in terms of reducing predictive uncertainty (as measured by RMSE) as well as the structure of the error. It was however not *a priori* clear how the performance of the data assimilation approach would compare to that of a re-calibrated model. This study has shown that re-calibration leads to considerably higher RMSE and more structured prediction errors than data assimilation. A somewhat surprising result is that with a da-models the best predictions are generated at considerably finer spatial and temporal resolutions than with the dc-models.

However it is obvious that any distributed model shows structural errors, it is interesting to investigate these and to inter-compare different models. In this study it was *a priori* unknown where and how much the different models would over- or underestimate the time-averaged validation data. Furthermore it is still a question mark why the statistical model set does not show such a structure, and why the situation is reverse for the spatially averaged errors when analyzed over time (see Figures 6.11 and 6.12). We find it important to stress these aspects since we are not aware of any other study where a spatial error structure of a dynamic model is analyzed on the basis of validation with field data. In addition we did not come across a study where geo-statistical and distributed models of soil moisture dynamics are compared with regard to spatial or temporal error structures.

The effect of soil moisture assimilation on the prediction of other model variables like discharge and ground water depth was shown in Section 6.3.3. It was illustrated that assimilation might help in enhancing predictions, but not necessarily so. In this study discharge predictions did benefit from the soil moisture observations, whereas the predictions of ground water table depth did not. The explanation for this is that the models do reasonably well represent the runoff mechanism in the system, which is mainly related to the condition of the upper soil, but do not represent the ground water system properly. This last point is not surprising when considering that in the models the deep soil functions as a sink from where there is no outflow to either the upper soil or the catchment outlet (see Figure 6.4). In reality, fluxes, albeit small, do exist from deep soil to upper soil and the catchment outlet. The illustration that soil moisture assimilation might lead to enhanced discharge predictions is by itself valuable. It is to our knowledge the first realistic example which shows that soil moisture data assimilation might help to enhance the prediction of indirectly related hydrological variables.

The methodology employed in this study can be relatively easily adapted to different models or observations. If a hydrologic model can be linearized and written in state space form (according to Equations 6.3 and 6.4), the solution algorithm (Section 3.2.2 and Appendix 2) can be applied without any adjustments. However, it is important to notice that instabilities may arise when large time steps are used in combination with adjustments of state variables for which spatial interaction is possible in two directions. In addition, jumps may occur in the output in case of large deviations between predicted and observed

state variables. Both situations do not occur in this study: for soil moisture in the upper soil (the variable to which data assimilation has been applied) the spatial interaction is in the downslope direction only, and the deviation of predicted from observed soil moisture is never large due to the frequent soil moisture observations. The flexibility of the data assimilation technique is an important asset since it means that different dynamic models can be evaluated with the same or different conditioning data relatively easy, without the requirement to include exactly the same state variables in the model dynamics. This point has been demonstrated in this study by using a set of models instead of a single model in order to generate a prediction. Another characteristic of the technique used in this study is that it only requires the solution of (many) systems of linear equations of moderate size, for which there are numerous efficient solution algorithms available. The computational requirements of the da-model set in this study are about 12 times the time required to run the d-model set; the time requirement for the dc-model is about 20 times that of the d-model set.

As was already pointed out in section 6.2.4, parameter estimation may be quite close to the state estimation proposed in this study. This is also apparent when applying a (Kalman) filter, which allows to extend a state-vector with parameters, in case these have to be estimated. The major distinction between state and parameter estimation is that the variables estimated in the former are spatial functions while those estimated in the latter are (often) lumped model coefficients. Furthermore state estimation is a data-oriented activity, which uses models to constrain measurements; while parameter estimation is a model-oriented activity, which uses measurements to constrain models [McLaughlin, 1995]. Note that this does not imply that a (Kalman) filter would be a suitable tool for state or parameter estimation for the models in this study. Already for quite some time meteorologists and oceanographers have successfully used state estimation techniques for data assimilation [Courtier et al., 1993]. At first sight the key ingredients for an equally successful application of these techniques in hydrologic problems seem to be available: there is an abundance of different spatially distributed models which could provide a basis for data assimilation, and in addition there are also many papers which discuss the estimation of uncertain environmental variables from field data. All that is needed is the intersection of these two foci - estimation studies which rely on spatially distributed physically-based models. The reason why this development is not (yet) seen may partially be due to the complexity of the models, which often include thresholds and other nonlinearities that are not compatible with the data assimilation methods used in meteorology and oceanography, as noted by McLaughlin [1995]. But, in our view, a more serious problem is the fact that the spatial correlation lengths of hydrologic processes are much shorter than those of atmospheric/oceanographic processes, so that for many cases an impossible amount of observations would be required for successful data assimilation. Only remote sensing observations can effectively provide the required observational density. Presently these are not available at a sufficient frequency and cannot give reliable soil-moisture estimations under different conditions (e.g. different vegetation covers), to be used in an operational setting. The only remaining 'cure' against data-scarcity is regularization, which was the strategy chosen in this study. In other respects hydrologic data assimilation problems have favourable properties when compared to the atmospheric and oceanographic equivalents. The first is the dominance of terrain, which reduces in many cases the spatial domain to a 2D problem. Secondly, the boundary and initial conditions are in hydrologic problems relatively well known (also due to the dominance of terrain, but also due to the relatively long temporal correlation lengths), or at least more easily observed as in atmospheric/oceanographic problems. Both properties have been exploited in this study:

the three-dimensional problem has been reduced to a number of parallel one-dimensional problems, and by starting the study period from a dry condition the boundary and initial conditions were relatively certain.

Based on the results of this investigation, the next step will be to further develop the presented data assimilation technique by applying it to various other data sets and models and especially to test its feasibility in an operational setting with a much lower observational density. Another question is how the presented approach compares to the Kalman filter and smoother, where regularization can also be applied [Boutabyeb *et al.*, 1997; Reif *et al.*, 1998].

Table 6.1.: Naming of nine sub-sets that result from combining 25, 50 and 75% of the observation times with 25, 50 and 75% of the observation locations in the complete data set. Note that 1/4th of the data is always used for validation.

coverage in time	coverage in space		
	1/4	1/2	3/4
1/4	1.1	1.2	1.3
1/2	2.1	2.2	2.3
3/4	3.1	3.2	3.3

Table 6.2.: Characteristics of the d-models where the calibration period is from 20 July to 20 October, and the validation period from 20 October to 21 December, using data set 1.1 for both periods (see Table 6.1). The number of parameters refers to the number of breakpoint-coordinates in the kernel functions (7 in total), including the uncertainty information (i.e. the shaded areas in Figure 6.5). The total parameter range is an aggregate measure for the total parameter uncertainty of the model, 0 refers to complete certainty and 1 to complete uncertainty (see text).

model resolution :									
grid size (m)	20	40	60	20	40	60	20	40	60
temporal unit (min)	5	5	5	10	10	10	15	15	15
number of parameters (nr)	91	112	97	119	112	63	101	64	59
total parameter range (-)	0.17	0.18	0.22	0.29	0.27	0.17	0.24	0.08	0.11
RMSE calibration (vol %)	7.5	7.2	6.7	6.3	6.5	6.8	6.4	6.1	6.2
RMSE validation (vol %)	8.6	8.0	7.2	8.4	7.8	7.4	8.3	6.5	6.9

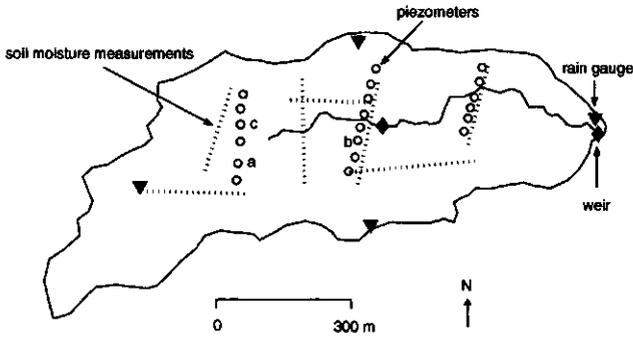


Figure 6.1.: Location of observations.

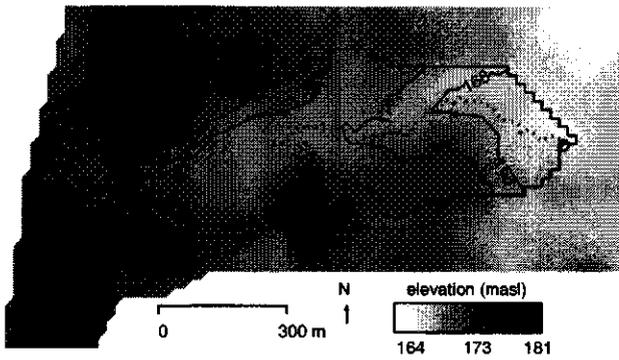


Figure 6.2.: Contour map of the study catchment.

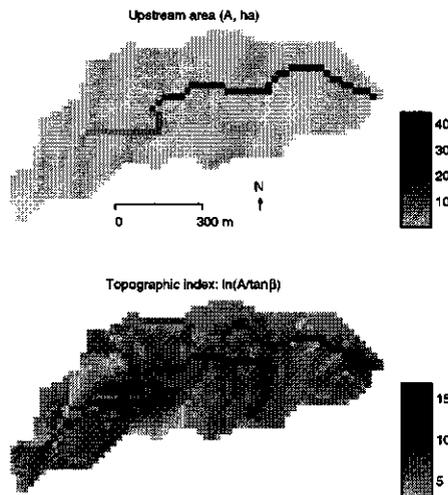


Figure 6.3.: Topographic features of the study catchment.

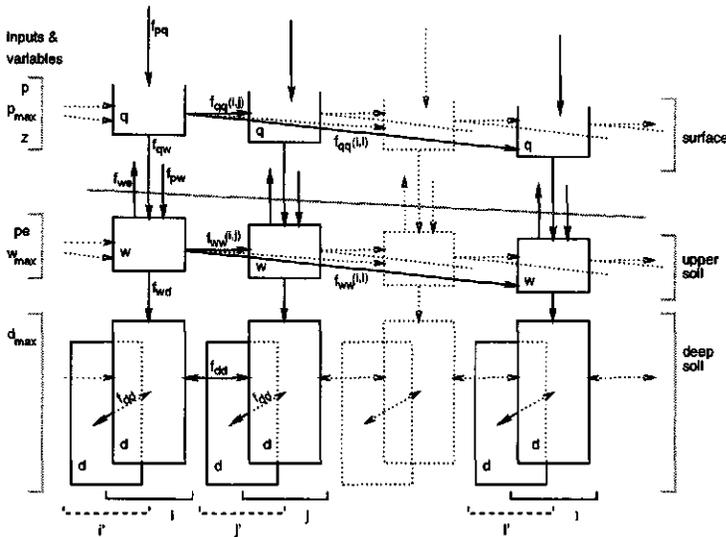


Figure 6.4.: Couplings between different model-components, where f_{pq} is the rain flux to surface water, f_{pw} the rain flux into the upper soil ($rain = f_{pq} + f_{pw}$), f_{qq} denotes the downslope surface water flux, f_{qw} the vertical flux from surface water to water in the upper soil, f_{ww} the downslope flux of water in the upper soil, f_{wd} the vertical water flux from upper soil to deeper soil, and f_{dd} the lateral flux in deeper soil. At the left the atmospheric forcings (p and pe) as well as the variables surface elevation (z), maximum infiltration capacity (p_{max}) and the storage capacity of upper and deep soil (w_{max} and d_{max} respectively) are indicated per layer. The atmospheric forcings and variables are defined for each grid cell. The figure illustrates that flow is possible from a grid cell i to the adjoining downslope cell j , and in the case of lateral fluxes in deeper soil also to several other adjoining cells (i'). In addition, surface flow and flow in the upper soil is possible between the cell i and cells further downslope (l). Note that w represents the unsaturated zone and d represents the saturated zone. There are seven independent types of fluxes in the model since $rain = f_{pq} + f_{qw}$ and $f_{qq}(i, j)$ is of the same type as $f_{qq}(i, l)$.

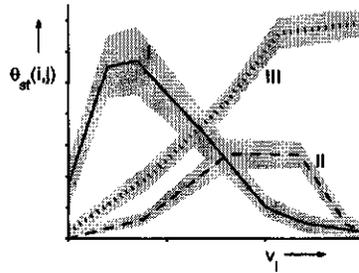


Figure 6.5.: Example of kernel functions, used to define the fluxes in Figure 6.4. The independent variable v_i is a scaled model variable, forcing or an average slope and $\theta_{st}(i, j)$ is the fraction of a state variable s_i assigned to a flux $f_{st}(i, j)$. The peaked shapes (I and II) are typical for the lateral fluxes (f_{qq} and f_{ww}), whereby the lateral distance is smaller for I as for II. The uni-modally increasing relation (III) is typical for the other fluxes (such as f_{qw} or f_{dd}).

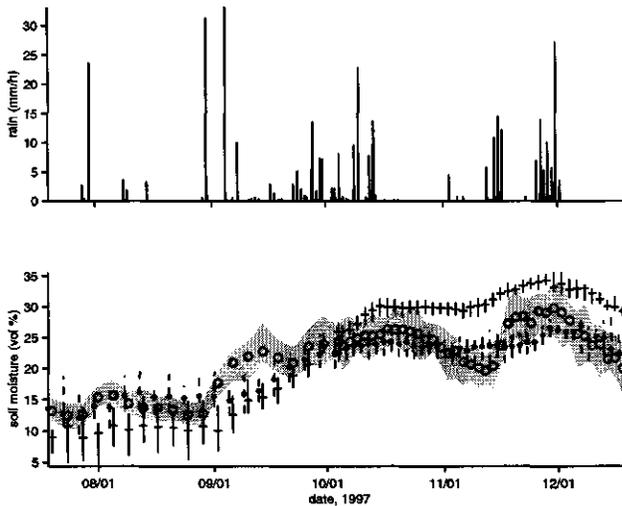


Figure 6.6.: Average rain and soil moisture for three different soil types, where symbols indicate the mean and the ranges the min/max measured soil moisture values within a soil unit. Values for vertisol are indicated with the solid line, for non-vertisol with the dotted line and points, and for shallow soil with the shaded area and circles.

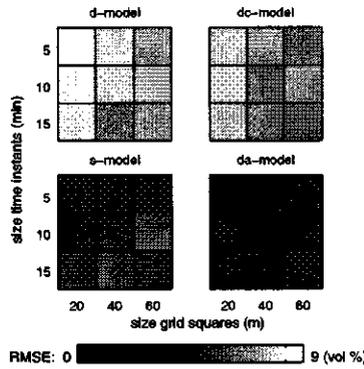


Figure 6.7.: RMSE of soil moisture predictions by the d-, dc-, s- and da-models. The latter three all use the smallest set of conditioning soil moisture data (1.1 according to Table 6.1). Note that the plot for the d-model contains the same information as the last row in Table 6.2.

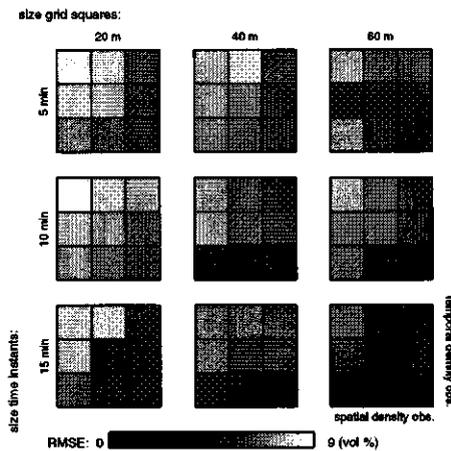


Figure 6.8.: RMSE of soil moisture predictions by the dc-model for different spatial and temporal model resolutions and observation densities (see Table 6.1 for the meaning of the observation densities).

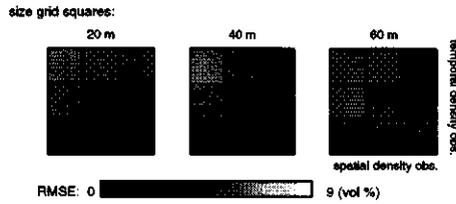


Figure 6.9.: RMSE of soil moisture predictions by the s-model for different spatial model resolutions and observation densities (see Table 6.1 for the meaning of the observation densities).

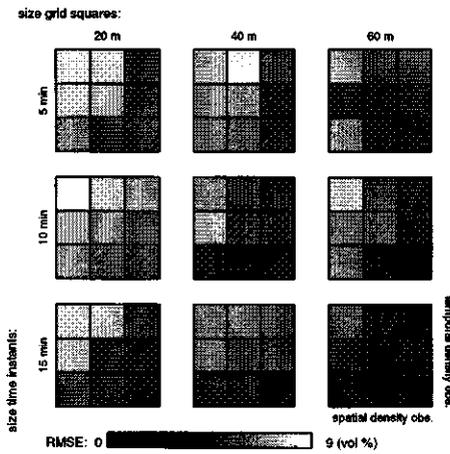


Figure 6.10.: RMSE of soil moisture predictions by the da-model for different spatial and temporal model resolutions and observation densities (see Table 6.1 for the meaning of the observation densities).

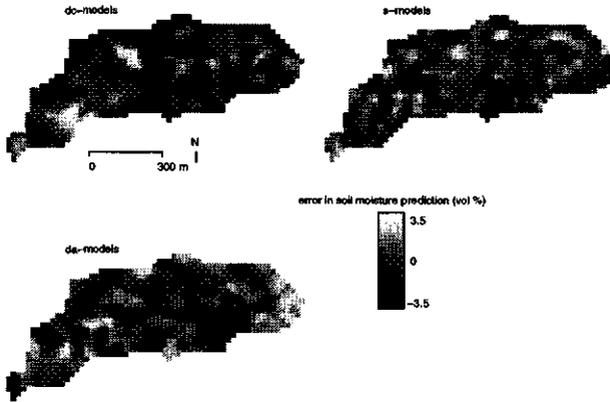


Figure 6.11.: Error maps for dc-, s- and da-models. The maps result from a distance weighted interpolation of RMSE of the soil moisture prediction at the various locations (see Figure 6.1) for each observation time, subsequently an averaging over time (the period 20 October - 21 December), and then an averaging over the different model resolutions (nine in the case of the dc- and da-models and three in the case of the s-model).

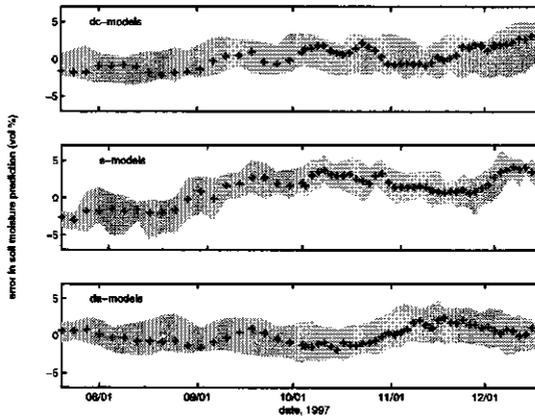


Figure 6.12.: Time series of error for dc-, s- and da-models, where the shaded areas indicate the minimum/maximum and + the mean deviation. At each time step the result is obtained by averaging the observed minus predicted soil moisture for all observation locations, and then an averaging over the different model resolutions (nine in the case of the dc- and da-models and three in the case of the s-model).

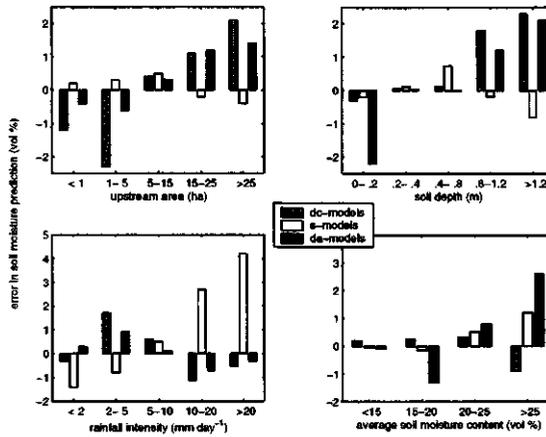


Figure 6.13.: Time-averaged error for different upstream areas and soil depth (top figures), and space-averaged error for different rainfall intensities and total water storage, for dc-, s- and da-models. The results are obtained by averaging over the respective class-intervals and subsequently over the different model resolutions (nine in the case of the dc- and da-models and three in the case of the s-model).

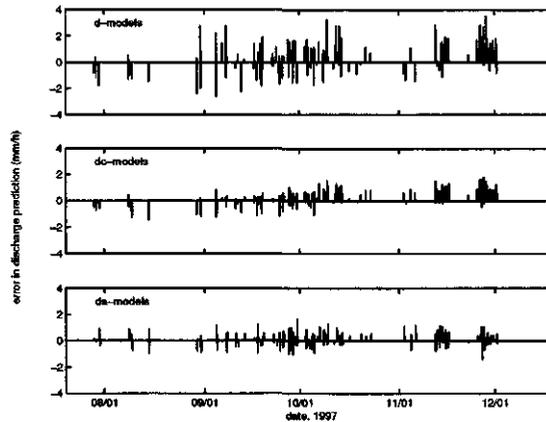


Figure 6.14.: Time series of the error in discharge prediction at the catchment outlet for for the d- dc- and da-models. The results are obtained by averaging the prediction errors (observed minus predicted hourly discharge) over the nine model resolutions.

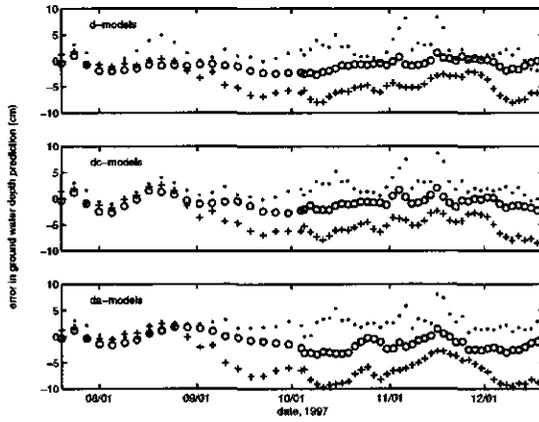


Figure 6.15.: Time series of the error in ground water depth at three locations for the d-, dc- and da-models. The locations are indicated in Figure 6.1. Location a is indicated by +, location b by ● and location c by ○. The results for each time step are obtained by averaging the observed minus predicted water table depths over the respective day, and subsequently averaging over the nine model resolutions.

Part II.

**Application of control models to
sub-catchments of the Meuse**

7. Automatic calibration of a groundwater model for the Geer catchment

Allain Dassargues

7.1. Introduction

In this particular study synthetic tests have been performed to improve automatic calibration of hydraulic conductivity values (K) of a 2D groundwater flow model by conditioning stochastic co-simulations of the K-field on geophysical data. The results are listed and the performance of the model and - linked to this - the calibration procedure is evaluated. The implications of this with regard to the focus new data assimilation studies are highlighted.

7.2. General structure of MOHISE

MOHISE is a deterministic, spatially distributed, physically-based model, composed of three interacting sub-models: a soil model, a groundwater model and a surface water model, which are linked dynamically and operated in a global structure. MOHISE stands for 'Modèle Hydrologique Intégré pour la Simulation du cycle de l'Eau' (Integrated hydrological model for the simulation of the water cycle). It was one of the results of the project "Integrated modeling of the Hydrological cycle in the Scope of Climatic Changes".

7.3. Improvement of the Geer groundwater model and presentation of results

Since the groundwater modeling code SUFT3D has been originally designed and developed to simulate very different hydro-geological situations (unsaturated zones, rainfall-seepage boundary conditions, Fourier boundary conditions, coupled flow and transport, ...), it can be considered as a complex and difficult to manage piece of software. Due to his powerful characteristics, this code has drawbacks in terms of computational time and total memory needed (RAM and disk space). Both can become very important when the number of finite elements is increased (e.g. when modeling basins larger than 100 km^2). Based on the experience gained previously (1997-2000), several improvements in terms of physical representation of the flow processes and in terms of numerical aspects have been brought to the SUFT3D code in order to obtain better results when applying in the Geer basin. Due to its ability to treat unsaturated/saturated groundwater flow, the SUFT3D code is

particularly convenient when high piezometric heads variations are observed or expected in function of time: it is the case in the Geer basin. New improvement consisted mainly in solving problems arising from the presence of galleries and withdrawal of groundwater in the Geer basin. A new one-dimensional finite element has been implemented to represent explicitly galleries and pumping wells. The Geer aquifer provides about $60.000 \text{ m}^3/\text{day}$ of drinking water for the city of Liège and its suburbs. The hydro-geological basin at the outlet Kanne does not correspond to the hydrological. The main aquifer is located in the Cretaceous chalks of Hesbaye. This chalky layer has a thickness varying from a few meters (in the south) to 100 meters (in the northeast). It is overlying a low permeable layer (the smectite of Herve) considered as an impervious bottom of the aquifer. This bottom layer slopes northwards with a gradient of 1% to 1.5%. The hydro-geology seems to be influenced by the fault of Horion-Hozémont which crosses the domain from the southwest to the northeast and by more fractured zones in the chalk corresponding, in the morphology, to dry valleys. 40 km of galleries, dug in the lower part of the chalk, also play a major role influencing the shape of the piezometric surface: groundwater is drained by them in most places but an important quantity of water can also be re-infiltrated in other places, depending on local piezometric conditions. Moreover, in low groundwater conditions, some parts of the galleries can become dry. In this basin, piezometric fluctuations in function of time can reach more than 15 meters. The SUFT3D code was thus chosen for its ability to consider a highly variable thickness of the unsaturated zone. The groundwater model is composed of seven layers of finite elements. From bottom to the top, these seven layers generally correspond to:

- three layers of chalk
- one layer of hardened chalk named hard-ground
- one layer of fractured chalk
- one layer of conglomerate
- one layer of silt.

Laterally, some of these layers include different geological units. The hard-ground layer is not present everywhere in the domain and the conglomerate layer disappears in the northern part of the Geer basin: there, it is replaced by a sandy layer (Tongrian, Landenian and Heersian). The thickness of the unsaturated zone can reach 40 meters. In some zones, groundwater levels fluctuate only in the chalky layers. The spatial discretization was carried out according to the following criteria: (1) a mean element size of about 700 meter; (2) local refinement in zones where important stresses are applied (fault, gallery and pumping well). The total mesh is made of 31423 finite elements (18680 nodes). The model calibration was conducted in two steps. First, a steady state calibration was performed, for two situations: one corresponding to high groundwater levels (1983-84), the second to low groundwater levels (1991-1992). In a second step, a calibration in transient state was performed. On basis of the set of available data, periods were chosen for the model calibration (1975-1988) and validation (1989-1995). For steady state calibration, annual averages of recharge and piezometric heads cannot be used because the delay of the effective recharge is highly variable in the domain (from 1 month to 18 months), depending on geology and thickness of the unsaturated zone. Average values, computed over several years, were used. Two steady state situations were selected, guided by the

amount of recent available data sets. Many data were available for the period 1983-1984 corresponding to the last extensive campaign for piezometric data collection. For the other period (1991-1992), about forty measured piezometric levels could be used. Working with these two opposite situations of steady state conditions is particularly indicative of the aquifer vertical heterogeneity. These calibrations allowed to determine zones where major heterogeneity is found and to set the values of hydraulic conductivity for each of them. In Figure 7.1, scatter plot diagrams are shown as a summary comparing observed (averaged) and computed piezometric head values. The averaged absolute errors were considered satisfactory in the two steady state situations. It must be noted that the simulations were carried out with a spatially uniform recharge, considering observed values averaged over two years. Results thus do not correspond to any actual state known at any given moment. Consequently, deviations equal to 2,5 meters are acceptable.

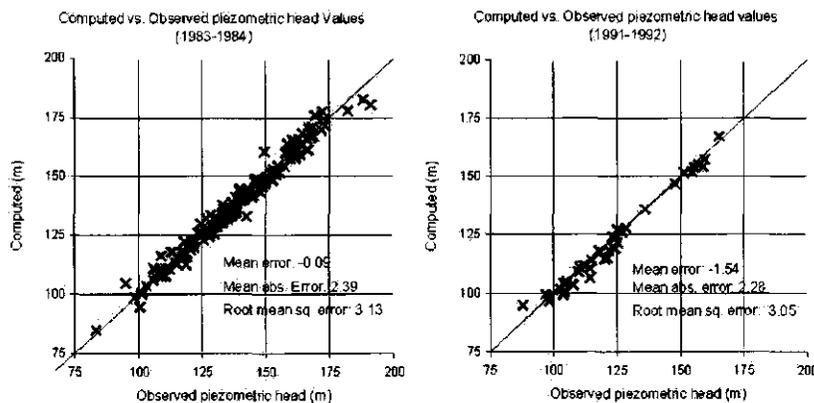


Figure 7.1.: Comparison of observed versus computed piezometric heads.

7.4. Transient state calibration and validation

The calibration in transient state was performed for the years 1975-1988, while the validation was performed for the period 1989-1995. Indeed, the transient calibration was especially useful to set intervals of values for effective porosity and specific storage coefficient and to improve those concerning hydraulic conductivity in this highly heterogeneous aquifer. The time evolution of piezometric heads measured in 34 piezometers was used to calibrate the groundwater model in transient state. Locations of these piezometers are distributed in the whole domain, except in the north of the Geer river. Data are not always available for the whole period of calibration or validation, but a minimum period of 5 years is available for each observation well. Horizontal heterogeneity can be also important in terms of hydraulic conductivity values to be entered for obtaining calibration. Main fractures and dry valleys are characterised by higher hydraulic conductivity values, of the order of 110^{-3} m/s. In the chalk, the hydraulic conductivity value varies from 110^{-5} to 610^{-4} m/s. In the upper chalk layer, hydraulic conductivity values are from 2 to 10 times higher than in the lower chalk layer. Between these two layers, the thin hard-ground seems to be characterised by relatively low hydraulic conductivity values (of the order of 110^{-5}

m/s), increasing the vertical heterogeneity of the aquifer. The calibration was difficult and time-consuming, but provided accuracy and reliability to the model for further use. In terms of simulated piezometric fluctuations, the relative quality of the calibration varies from one piezometer to another (Figures in 7.2). Generally speaking, it is observed that the computed piezometric fluctuations are not smoothed enough. This can be due to a simulated too short delay between precipitation and effective recharge for groundwater in some places of the domain, or to an inaccurate evaluation of the chalk storage capacity.

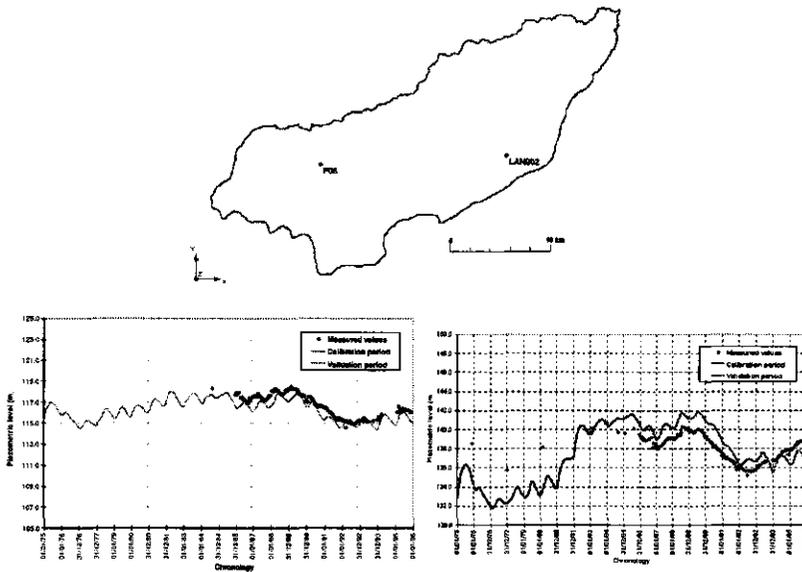


Figure 7.2.: The location of two observation wells in the Geer catchment (upper graph). Simulated and measured piezometric levels at the selected observation well F06 (bottom left) and well LAN00 (bottom right).

7.5. Conclusions

It is still an open question whether for a complex physically based and spatially distributed model like MOHISE it will be possible to effectively apply data assimilation technology, in the absence of a dense network of observations. Nonetheless this study illustrates that there is an urgent need of improved techniques for automatic model calibration (which is a sub-category of data assimilation), to obtain better calibration results and to make it easier to inter-compare different models or model parameterizations - in this study the structure of the sub-surface is e.g. uncertain and can be parameterized in various ways. For this study an automatic calibration procedure was already applied, but there is scope for much more automatisation.

Secondly, the study illustrates that, in spite of the great detail of the model and the considerable data set used for calibration, there is still much room for improvement by state-corrections. The calibration and validation periods namely show a similar degree

of model-to-data fit (see Figure 7.2). This suggest that parameter adjustments will most likely not result in in improved model predictions, so that state-adjustments offer the only alternative for improved model predictions.

8. Calibration of a flood prediction model for the Ourthe

Ad de Roo

8.1. Description of the LISFLOOD model

The physically-based LISFLOOD model has been developed for simulations of floods in large European drainage basins [*de Roo et al., 2000*]. Full basin-scale simulations can be carried out, such that influences of land use, spatial variations of soil properties and spatial precipitation differences are taken into account. LISFLOOD consists of a catchment-scale water balance model (LISFLOOD-WB), run with a daily time step, a catchment-scale flood simulation model (LISFLOOD-FS), run with an hourly time step, and a floodplain simulation model (LISFLOOD-FP) [*Bates and Roo, 2000*], run with a time step of several seconds (Figure 8.1). The water-balance model is started approximately one year or more before a flood, to simulate the initial conditions (discharge, soil moisture, snow cover, groundwater) before the flood event. The catchment flood simulation model starts just a few days before a flood. The main difference with the water-balance model is the time step, which is smaller to improve the river-routing. Typical model grid-sizes for the Meuse and Oder catchment are 1 km. Sub-basins of the Meuse and Oder are simulated using 100 – 300 m grids. The LISFLOOD floodplain model simulates with high spatial and temporal resolution a part of the floodplain of a river, using either observed discharges as boundary condition, or simulated discharge from the catchment LISFLOOD model.

The three LISFLOOD models are distributed models integrated in the PCRaster and/or ArcView GIS. In LISFLOOD-WB and LISFLOOD-FS, a simulated catchment consists of an overland flow grid and a separate channel grid, a topsoil layer, a subsoil layer, an upper groundwater zone, and a lower groundwater zone (Figure 8.2). Processes simulated are precipitation, interception, soil freezing, snowmelt, evapotranspiration, infiltration, percolation and capillary rise, groundwater flow and surface runoff. Overland flow and channel flows are simulated using a kinematic wave approximation. The user can choose both the spatial and temporal resolution of the model. The channel routing part contains a simple solution to account for floodplain storage and flow. A summary of the processes that are simulated is given below:

- Precipitation data from individual stations can be used in LISFLOOD, which are then interpolated using an inverse distance method of the 5 closest stations. Precipitation is corrected for altitude effects, based on precipitation-altitude relations found in the catchment to be simulated.

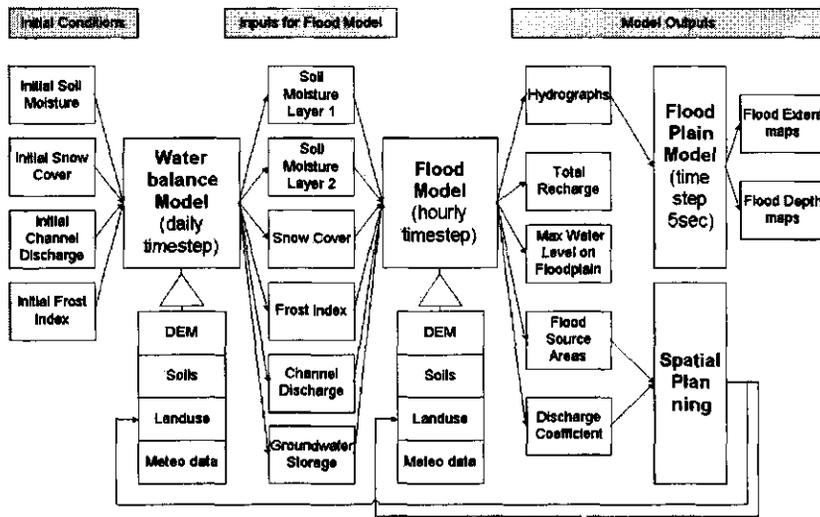


Figure 8.1.: Flowchart of the LISFLOOD model

- Snowfall is simulated when the average daily temperature is lower than 1.0 degree Celsius. Minimum and maximum daily temperature values from stations are interpolated using an inverse distance method of the 5 closest stations, and on each pixel are corrected for altitude.
- Interception of rainfall by the vegetation is simulated using the method of *von Hoyning-Huene [1981]* for all land use except forests, for which the approach of *Shuttleworth and Calder [1979]* is used. The equations are based on the Leaf Area Index of the vegetation. Seasonal changes of LAI are taken into account.
- Evapotranspiration is simulated using the Penman-Monteith method, as applied in the WOFOST model [*Supit et al., 1994*]. For forests, the Priestley-Taylor equation is used, as modified by *Shuttleworth and Calder [1979]*. Meteorological variables used are temperature, wind speed, sunshine duration, cloud cover and actual vapor pressure, which are all interpolated from station data using an inverse distance method and where appropriate corrected for altitude. The Leaf Area Index of each simulated pixel is used to calculate actual evapotranspiration from potential evapotranspiration.
- Snowmelt is simulated using a degree-day method [*Baumgartner et al., 1994*], when the average daily temperature is above 0 degrees Celsius.
- Infiltration is simulated using the Smith-Parlange equation [*Smith and Parlange, 1978*]. The capillary drive value is based on topsoil texture. Saturated hydraulic conductivity values are based on topsoil texture and land use. In city areas and on water bodies no infiltration takes place.
- Soil freezing is simulated using a degree-day method [*Molnau and Bissel, 1983*]. If the soil is frozen to a certain degree, infiltration is reduced to zero.

- Vertical transport of water in the two soil layers is simulated using a one-dimensional form of the Richards equation. Soil water retention and conductivity curves are described by the *van Genuchten [1980]* relationships. Pedotransfer-functions from the HYPRES project [*Wösten et al., 1998*] are used to calculate the water retention and conductivity curves from soil texture. Both soil texture and soil depth are derived from the European Soils Database [*Finke et al., 1998*] or local soil maps.
- Percolation to the groundwater store is calculated using the Darcy equation.
- Groundwater storage and transport to the channel system are simulated with an upper and a lower groundwater zone, and groundwater is then routed using a response function similar to the one adopted in the HBV model [*Lindström et al., 1997*].
- Overland flow and transport to the channel system is simulated using a four-point finite-difference solution of the kinematic wave [*Chow et al., 1988*] together with Manning's equation.
- Channel flow is also simulated using a four-point finite-difference solution of the kinematic wave [*Chow et al., 1988*] together with Manning's equation. The channel and floodplain dimensions (width and depth) are used to calculate the wetted perimeter. For rivers with low channel gradients, a dynamic wave algorithm can be selected.
- Special structures such as water reservoirs and retention areas can be simulated by giving their location, size and in- and outflow boundary conditions (maximum storage volume, minimum and maximum outflow, reservoir management parameters).

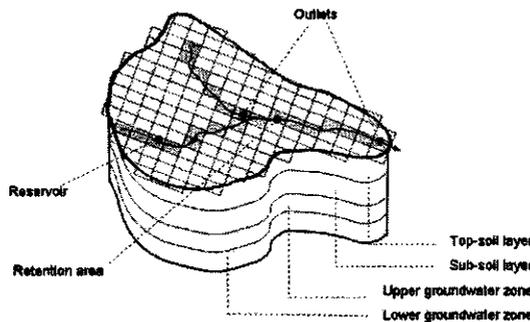


Figure 8.2.: Schematization of a catchment in LISFLOOD including soil and groundwater layers.

The input parameters for the LISFLOOD-WB and -FS model are maps of topography, land use type (Corine land-cover database), soil depth and soil texture (European Soils Database) (see Figure 8.3). Time-series of precipitation amounts and other meteorological parameters (minimum and maximum daily temperature, actual vapour pressure, sunshine duration, cloud cover, wind speed at 2 m) are needed for as many meteorological stations within the catchment as possible. Precipitation and temperature are corrected for altitude.

All meteorological parameters are spatially interpolated using an inverse distance method using the 5 closest stations. Seasonal NDVI profiles were derived from IRS-WIFS satellite images, showing the changes in vegetation cover during the year for each land use type. From the NDVI profiles Leaf Area Index values are derived for model parameterization. Antecedent soil moisture conditions are taken from the LISFLOOD water balance model, which is used as pre-processing for the flood simulation model.

The outputs of LISFLOOD (Figure 8.3) consist of hydrographs at user-defined locations in the catchment, usually the locations where also measured discharge is known. Furthermore, time-series of for example evapotranspiration, soil moisture content or snow depth can be created at selected locations, if validation data are available. The model produces a number of GIS maps, such as water source areas, discharge coefficient, total precipitation, total evapotranspiration, total groundwater recharge and soil moisture maps.

8.2. Finding optimum parameter sets for the Ourthe catchment

The Ourthe sub-catchment of the Meuse (Figure 8.4) (1618 km^2) has been simulated with a regular grid of $300 \times 300 \text{ m}$. The catchment is simulated upstream of Tabreux (coordinates: 233152, 126194). Three parameters were selected for determining optimum parameter sets for the LISFLOOD-WB model (daily time step) in the Ourthe catchment:

1. The parameter setting the threshold (in mm) for the total groundwater storage (hereafter referred to as ST). Above the value set for the ST the recharge is linearly reduced until it becomes zero when the recharge reaches the value of the threshold plus 10 mm. This parameter enables to simulate saturation excess overland flow.
2. Secondly the parameter the maximum daily amount of percolation (in mm) going from the upper to the lower response box (in the model GwPercValue, from now on GP). This parameter regulates the percolation velocity of the water, and therefore the response on the model to a rain event. Small values of GP result in a faster saturation of the upper response box, and consequently in a reduced infiltration capacity.
3. The last one is known in the model as GwK1. This is the recession coefficient used in the outflow equation of the groundwater lower response box. Changes in this parameter influence the model response in the descending phase of the flood. Smaller values of this coefficient determine a slower release of water from the groundwater system to the surface, implying a flattening of the discharge curve after the flood peak.

Several simulations have been run for the Ourthe river basin to obtain several the parameter sets of good performance. Beside this, other tests were made to evaluate which inverse distance weight for interpolation of the meteorological data was best performing in this catchment, considering as well other simulations carried out for the entire Meuse catchment. The results of the different runs have then been validated on the observed discharge data (acquired from RMI, Brussels) using the Nash-Sutcliffe goodness-of-fit index. In order to calibrate the model a yardstick was needed on which evaluate the goodness of the results. Several methods are available in the literature, mostly based on a goodness-of-fit index. Among these, the Nash-Sutcliffe method was chosen to evaluate the level of

agreement between measured and computed discharges with regard to both shape of the hydrograph and amplitude of the peaks.

In the next Sections the results are summarised in a table showing the goodness-of-fit for the different settings of the chosen calibration parameters.

8.3. Evaluation of optimal parameter values

The initial set of values for the selected parameters was $ST = 100$ mm, $GP = 0.1$, $GwK1 = 0.05$, with a degree for the polynomial interpolation for the rainfall profile (hereafter called RainPw) equal to 2. This has been referred to as the Standard set in the summary of the results starting from which different alternatives were simulated for the calibration.

Figure 8.5 shows the goodness-of-fit coefficients of all the parameter sets simulated at the outlet (Tabreux). The best fitting alternative is shown in Figures 8.6 and 8.7. (alt. 87: $ST = 30$ mm, $GP = 1.2$, $GwK1 = 0.05$, $RainPw = 1$). From the statistics it was concluded that the linear interpolation of the rain profile ($RainPw = 1$) is always preferable. Also, higher values of ST together with higher values of GP gave also good results. By comparing the results of the different simulations it appears that there is a range within which these calibration parameters produce good fitting results. Such a range varies between 5 and 35 for ST and between 0.1 and 1.5 for GP . Therefore, this is the optimum range of values for the calibration of the model.

To evaluate the uncertainty of input data, a sensitivity analysis has been executed for the Ourthe catchment on the optimum set of parameters. The following input data have been considered (source or calculation method between brackets):

- Rainfall (grid map calculated with an inverse distance polynomial interpolation from measured data at different stations locations);
- Leaf Area Index, LAI (grid map obtained from the combination of the Corine land use map and tabled values of NDVI for different land use types and time of the year);
- Saturated conductivity of top-soil, K_{sat1} , and sub-soil, K_{sat2} (created combining the soil texture map with tabled values);
- Reference potential evapotranspiration, PET (calculated in the model by means of the Penman-Monteith formula);
- Soil depth (input grid map);

Five stochastic maps have been created for each input data type:

- Rain, LAI and PET: stochastic maps with a normal distribution (average 1, standard deviation 0.1);
- K_{sat1} and K_{sat2} : stochastic maps with a normal distribution (average 0, standard deviation 0.25);
- Soil depth: stochastic maps with a normal distribution (average 1, standard deviation 30).

For Ksat the standard deviation was determined on the basis of the information contained in *Wösten et al. [1998]*. In Table 8.1 for each soil type the fits on the geometric averaged curves are reported together with the uncertainty bands. As one can notice, the range is about one order of 10 higher/smaller with respect to the average value. The standard deviation fixed for the generation of the stochastic maps permitted to reach this variance range.

Table 8.1.: Ksat1 and Ksat2 values for each soil type acquired from *Wösten et al. [1998]*.

	Soil type	Minimum	Maximum	Avg. curve fit
Topsoils (Ksat1)	Coarse	2.0	200	60.0
	Medium	2.0	90	12.1
	Medium fine	0.2	20	2.3
	Fine	2.0	300	24.8
	Very fine	0.4	300	15.0
Subsoils (Ksat2)	Coarse	3.0	200	70.0
	Medium	0.2	20	10.8
	Medium fine	0.5	18	4.0
	Fine	0.7	100	8.5
	Very fine	3.4	160	8.2
Histosols	No texture	1.2	100	8.0

The model has been run five times for each data type using the original grid map multiplied by one of the five generated stochastic maps (in the case of the soil depth the stochastic maps were added to the original input data). The results of these simulations were then compared with those obtained using the original data by means of the Nash-Sutcliffe coefficient. The sensitivity analysis performed on the optimum parameter set gave satisfactory results: the Nash-Sutcliffe coefficient obtained from the runs where a stochastic variance was introduced was always high and did not differ significantly from the value obtained from the base simulation (See Figure 8.14). Only the random variation of soil depth resulted in lower coefficients.

8.4. Summary of the results

Because several parameter sets yield similar satisfactory goodness-of-fits there is an uncertainty band around the simulation results. The two extreme results, the best fit and the observed discharges are plotted in Figure 8.15. Figure 8.16 shows that both peak discharges and low discharges are simulated well for Tabreux for the simulation period of 1978-1975. The agreement of observed and simulated discharges at the other three measuring stations Nisramont, Ortho and Mabompre are of similar quality. The bottom graph in Figure 8.16 shows that also the total discharge over a long period is corresponding with the observed discharge.

8.5. Conclusions

- The results obtained from LISFLOOD-WB with daily simulations, using more or less the standard parameter set (also used for the entire Meuse and Oder basins), were satisfactory.

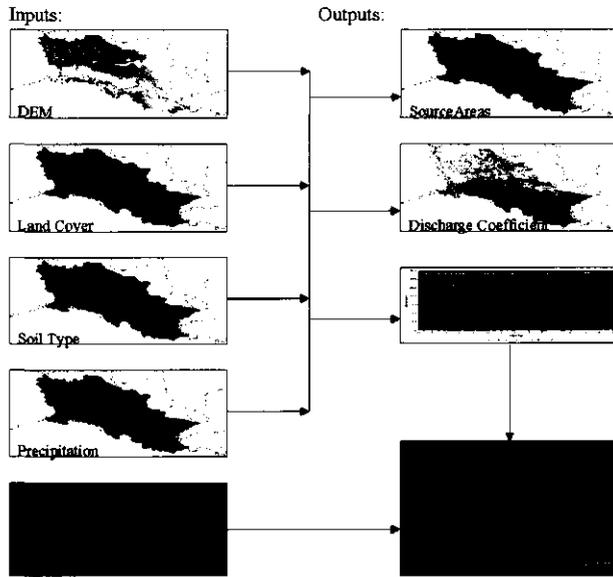


Figure 8.3.: Inputs and outputs of the LISFLOOD flood simulation model

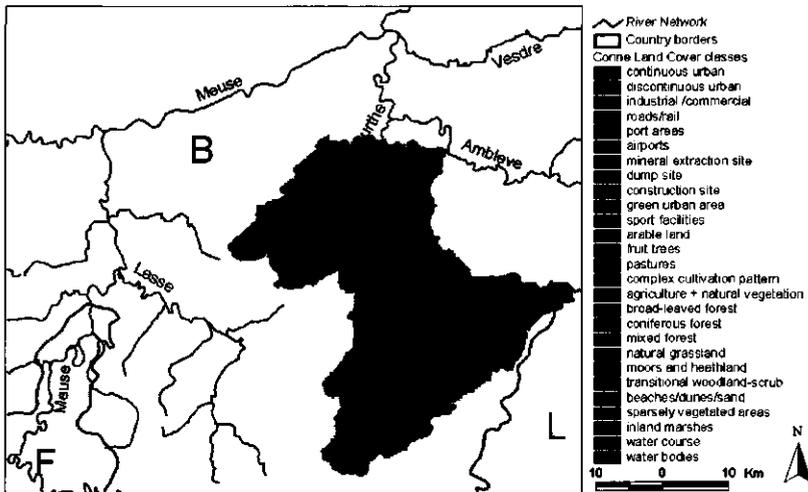


Figure 8.4.: Corine land use map of the Ourthe basin upstream of Tabreux.

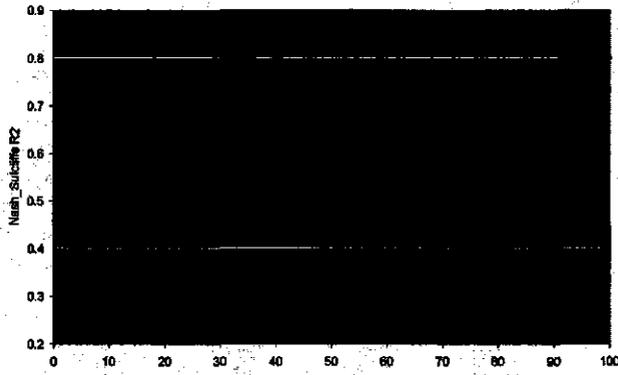


Figure 8.5.: Nash-Sutcliffe goodness-of-fit values for the 100 simulations with different parameter sets.

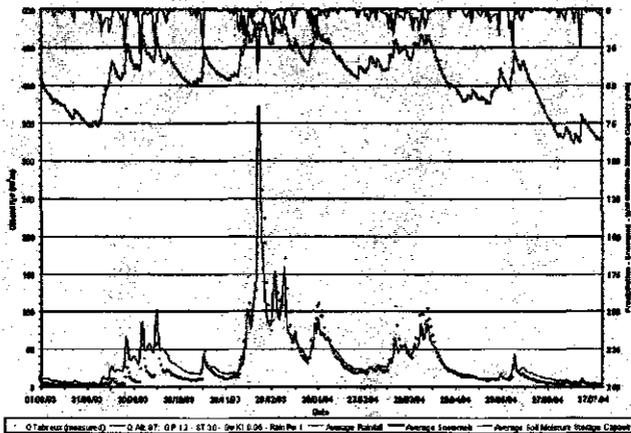


Figure 8.6.: Results of the best fitting parameters set (alt. 87) for Tabreux, period 1993-1994.

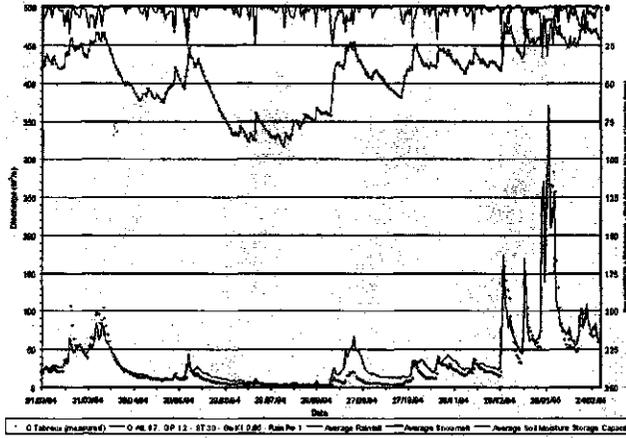


Figure 8.7.: Results of the best fitting parameters set (alt. 87) for Tabreux, period 1994-1995

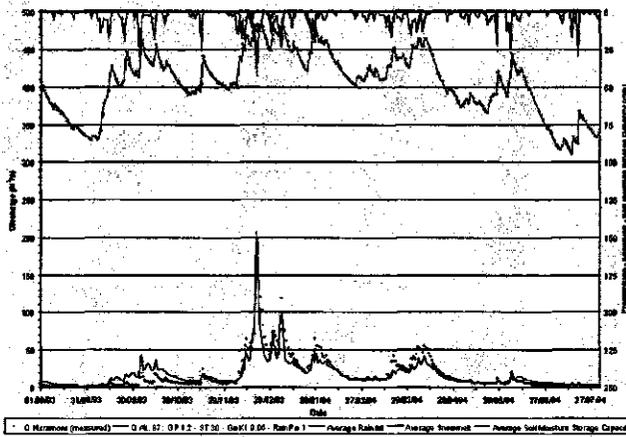


Figure 8.8.: Results of the best fitting parameters set (alt. 87) for the Nisramont gauge, period 1993-1994

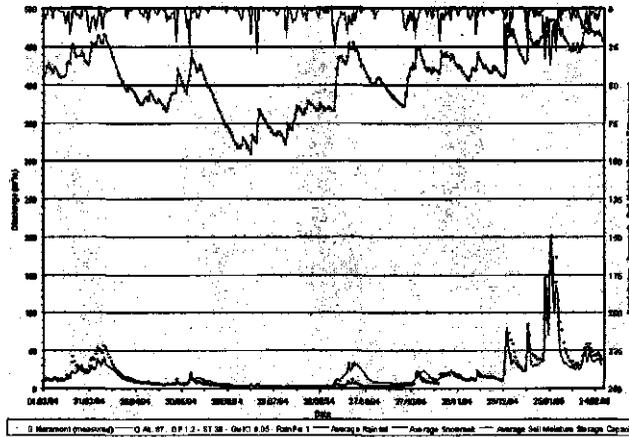


Figure 8.9.: Results of the best fitting parameters set (alt. 87) for the Nisramont gauge, period 1994-1995

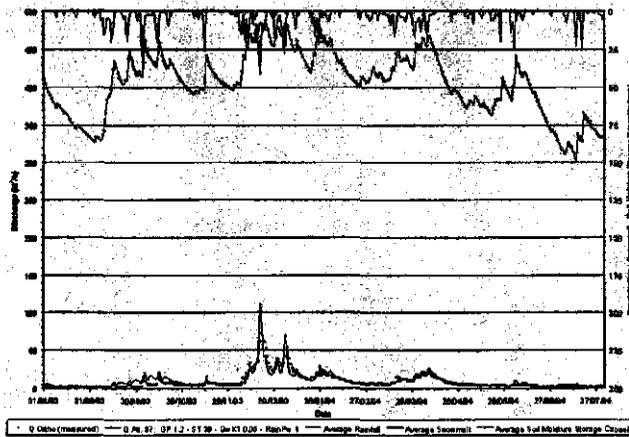


Figure 8.10.: Results of the best fitting parameters set (alt. 87) for Ortho gauge, period 1993-1994

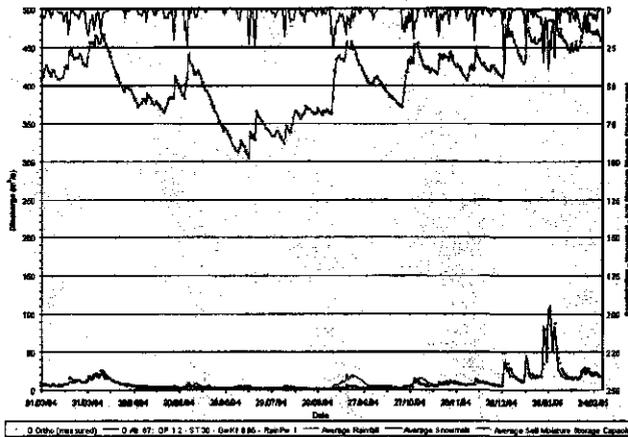


Figure 8.11.: Results of the best fitting parameters set (alt. 87) for the Ortho gauge, period 1994-1995

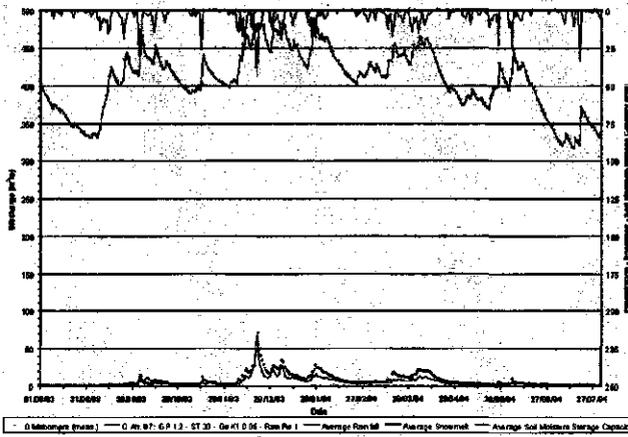


Figure 8.12.: Results of the best fitting parameters set (alt. 87) for Mabompre gauge, period 1993-1994

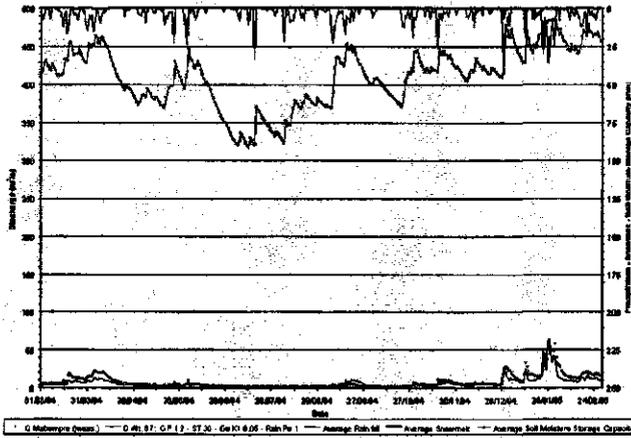


Figure 8.13.: Results of the best fitting parameters set (alt. 87) for Mabompre, period 1994-1995

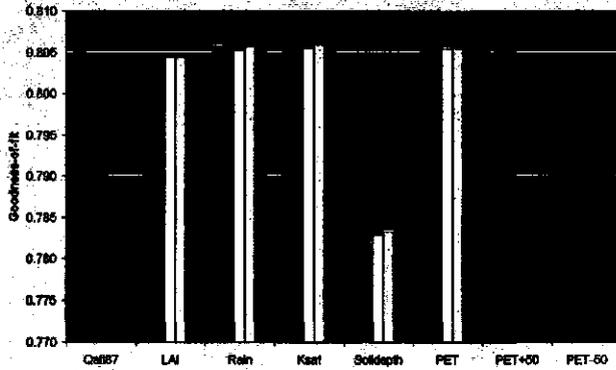


Figure 8.14.: Goodness of fit coefficients after adding a stochastic term to the baseline simulations for LAI, Rainfall, Soil depth, Saturated conductivity, and Potential Evapotranspiration.

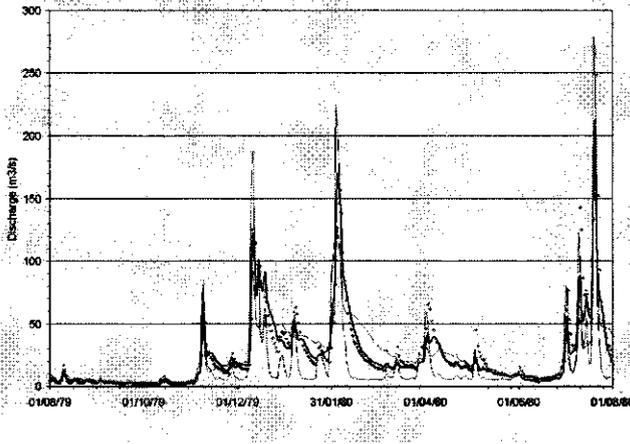


Figure 8.15.: Uncertainty bands around the simulation results for Tabreux.

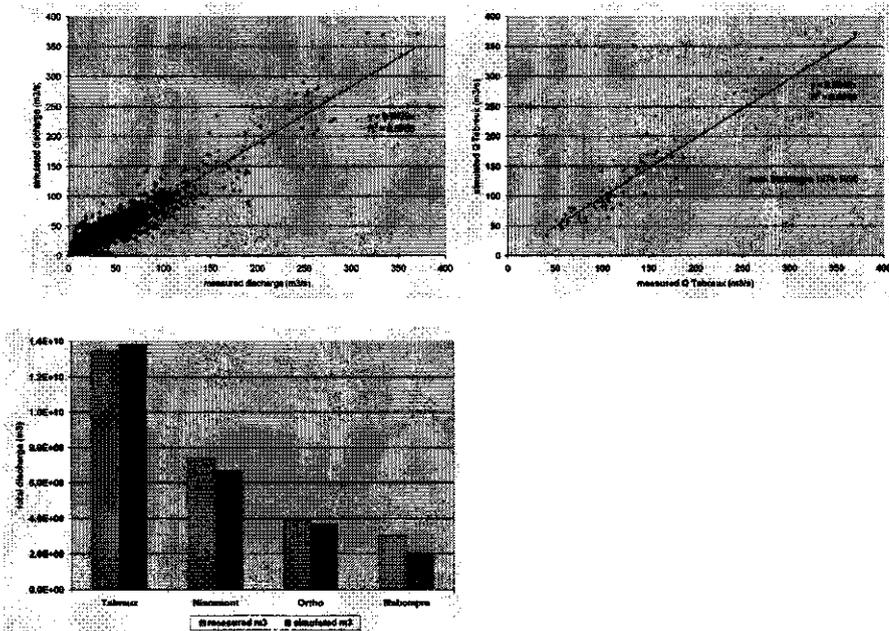


Figure 8.16.: Agreement between observed and simulated daily discharges (upper left) and daily peak discharges (upper right). The lower graph displays the observed and simulated total discharge. All for Tabreux over the period 1978-1995.

- The model performed well also for the 4 major flood events of the 20th century (February 1984, January 1993, December 1993, and January 1995) in the Meuse. There was good agreement between observed and predicted discharges.
- Improvement in the input data (soil maps and meteorological data) produced a significant improvement in the model results. Further knowledge on the soil properties will surely improve the results without need of further calibration.
- With respect to input data, LISFLOOD is a very demanding model and for this reason it is very sensitive to the quality of the available data.

9. Statistical correction applied to a water-balance model for the Meuse

Emmanuel Roulin

9.1. Introduction

Data assimilation in the SCHEME hydrological model has been investigated by using operationally available Meteosat data at the Royal Meteorological Institute of Belgium. These Meteosat data have already been processed to assess actual evapotranspiration. The use of weather radar images has been also briefly considered. Both data sources currently receive much attention due to their (unfulfilled) potential in enhancing hydrological predictions.

The structure of this chapter is as follows. First the methodology is presented. The accounting of soil moisture and actual evapotranspiration will be presented in detail in Section 9.2. The method developed to estimate actual evapotranspiration from satellite data is summarized in Section 9.3. The data assimilation method is described in Section 9.4. Soil moisture data from a field campaign by the Catholic University of Leuven (UCL) in 1995 (used as ground truth data in this study) is described in Section 9.5. The results of SCHEME with and without data assimilation are compared on the basis of soil moisture and water balance observations in Sections 9.6 to 9.8. Finally the conclusions and recommendations are given in Section 9.9.

9.2. Soil moisture and evapotranspiration in SCHEME

SCHEME is a comprehensive grid-based conceptual hydrological model that operates at coarse scales. The vegetation cover plays a central role in SCHEME. Seven vegetation covers are distinguished in the model. For each vegetation cover, three state variables describe respectively the water content of the vegetation itself and of two soil horizons. A water balance is computed successively for the three state variables by comparing the current status with the input, the potential evapotranspiration and a maximum value allowed to the water content. At each level, the potential evapotranspiration is decreased by an amount equal to the actual evapotranspiration at that level.

Referring to Figure 9.1, the water balance can be described as follows. In order to compute w_v , the water content at the level of the vegetation, the potential interception, IP , and the maximum water content, w_{vx} are calculated depending upon the vegetation cover, its leaf area index, the precipitation, P , the potential evapotranspiration, ETP , and the rainfall intensity if available. Three cases are described (time-indices are omitted for brevity).

Firstly, if potential evapotranspiration is large enough ($wv + IP < ETP$) then the water content of the vegetation is emptied ($wv = 0$), the evapotranspiration occurs at a rate $ETV = wv + IP$, and the potential evapotranspiration is reduced by the corresponding amount ($ETP = ETP - ETV$). Secondly, if ETP is not that large and if the resulting water content of the vegetation ($wv = wv - ETP + IP$), is lower than its maximum value (wvx), then evapotranspiration occurs at its potential value ($EVT = ETP$) so that no water will evaporate from the soil ($ETP = 0$). Thirdly, if the capacity to hold water by vegetation is exceeded, the water content is set to its maximum value ($wv = wx$), evapotranspiration occurs also at its potential rate ($ETV = ETP$; $ETP = 0$) and actual interception is re-calculated accordingly ($IP = wx + ETP - wv$). Throughfall is given by the precipitation reduced by the interception ($TH = P - IP$).

For the water balance of the upper soil horizon, two cases are expected. Firstly, if the throughfall is lower than the residual potential evapotranspiration, it is supposed to evaporate completely before filling the soil. Furthermore, the upper soil dries at a rate proportional to its saturation:

$$\Delta wu = (ETP - TH) \frac{wu}{wux}$$

The actual evapotranspiration is then $ETV = EGT + \Delta wv$, the upper soil moisture becomes $wv = wv - \Delta wv$ and the potential evapotranspiration is decreased accordingly $ETP = ETP + \Delta wv$. Secondly, if the throughfall is greater than or equal to the residual potential evapotranspiration, then the actual evapotranspiration is increased by $ETU = ETP$. If any, the remaining water fills in the upper soil. If the upper soil get saturated, part of the water infiltrates into the lower soil horizon (IN) and part reaches the surface water (Rc). The "runoff coefficient" that separates excess water to both paths is one of the parameters that is optimized. The capacity of the upper soil, wux is also optimized.

The lower soil reservoir is either dried or filled depending if throughfall evaporated completely or not. The lower soil dries at the rate:

$$\Delta wl = (ETP - IN) \frac{wl}{wlx}$$

In that case, evapotranspiration is $ETL = \Delta wl$. If water infiltrates in the lower soil horizon, the water contents rises up to a maximum value, wlx . If this value is exceeded, water percolates to the underground reservoir (Sc). The maximum water content of the soil ($wux + wlx$) is given a priori. For crops, it is made varying throughout the year following the depth of the root zone.

9.3. Evapotranspiration from Meteosat data

In this study latent heat flux data, based on Meteosat, are used. Meteosat data consist in half-hourly images in the infrared and in the visible range with pixels of about 5×9 km at the latitude of the Meuse study-catchments. Meteorological observations are available from radio-soundings and data from a micro-meteorological station measuring wind temperature and humidity profiles.

Satellite data were processed to estimate the surface net radiation [Roulin et al., 1995]. Data of cloud free pixels in the visible channel and in the infrared channel were corrected for the radiative transfer to provide the albedo and the infrared radiative exitance, respectively. The visible irradiance was estimated from images in the visible channel using

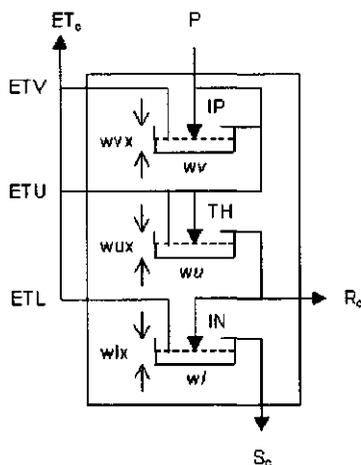


Figure 9.1.: Schematic view of the SCHEME model.

a empirical simple model. Based on histograms built from image series and based on cloudiness observed at the synoptic stations, the cloud cover was estimated. The infrared exitance of cloudy pixels as well as the infrared irradiance (atmospheric radiation) were estimated with empirical algorithms.

The turbulent sensible heat flux was estimated using the Monin-Obukhov similarity theory and data from the micro-meteorological station at Melle in the Scheldt basin. Latent heat flux was calculated from the energy balance. The "evaporative fraction" is the ratio of the latent heat flux, LE , to the available energy at the surface, $f = LE/(Rn - Qg)$ where Rn is the net radiation and Qg the heat flux into the ground [Gellens-Meulenberghs, 2000]. The daily evaporative fraction was computed with daily means of latent heat flux and available surface energy. The evaporative fraction at Melle is assumed to be representative for Belgium.

This procedure was developed on an operational basis. Nevertheless, due to the drift of Meteosat radiometers, the calibration coefficients has to be updated. For the visible channel, this has been done by other teams by the analysis of targets in the Sahara. Due to the lack of maintenance of the operational procedure, our analysis is limited to the year 1995. The methodology is currently further developed and adapted to Meteosat Second Generation (MSG) in the framework of the SAF-Land (Satellite Application Facility) network set up by Eumetsat.

Up to now, the SCHEME model has been applied using meteorological data from two Météo-France stations and the hydrometeorological stations. The hydrometeorological database is used for operational estimation of potential evapotranspiration over reference sub-catchments in Belgium. For the solar radiation, the network is shown in Figure 9.2. The interpolated field may be compared to the results of the Telsat project for the same day and the improvement in resolution with satellite information is obvious.

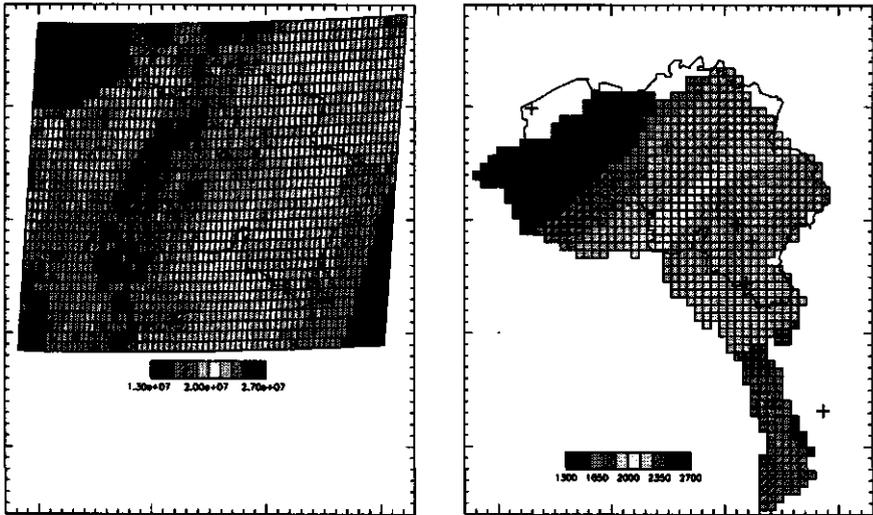


Figure 9.2.: Left: example of an archived Meteosat image, displaying daily solar radiation (Jm^{-2}) on 26th of May, 1995. Right: hydrometeorological stations with solar radiation data and interpolated fields (Jcm^{-2}) for the same day on the grid of the SCHEME model, applied to the Scheldt and Meuse basins.

9.4. Assimilation method

Evapotranspiration affects the water intercepted by vegetation as well as the water content of two soil horizons (the three state variables in SCHEME). Given the structure of the SCHEME, latent heat fluxes have been assimilated by adapting the potential evapotranspiration input variable for the interception and soil sub-models. In Figure 9.3, the flow diagram of the assimilation method is shown. The method starts to process a grid cell and ends before the underground and routing sub-models.

A simple statistical correction algorithm is used [Schuurmans *et al.*, 2002]:

$$ET_{new} = ET + \left(\frac{\sigma_{ET}^2}{\sigma_{ET}^2 + \sigma_{ET_{sat}}^2} \right) (ET_{sat} - ET)$$

where ET_{new} is the updated actual evapotranspiration, ET is the actual evapotranspiration estimated with SCHEME, ET_{sat} is the actual evapotranspiration estimated from satellite data and σ_{ET}^2 and $\sigma_{ET_{sat}}^2$ are the variances of the errors of actual evapotranspiration estimated with SCHEME and with the satellite data, respectively. The values of the variances are not known. We assume that both are related with $\sigma_{ET}^2 = \beta \sigma_{ET_{sat}}^2$, and therefore:

$$ET_{new} = ET + f(ET, ET_{sat}) = ET + \left(\frac{\beta}{\beta + 1} \right) (ET_{sat} - ET)$$

The updating is achieved through an iterative procedure that adapts the potential evapotranspiration, ET_{pot} , until ET converges to ET_{new} . Starting with the values $\delta ET_{pot} = \delta ET = f(ET, ET_{sat})$, rule of false position is:

$$\delta ET_{pot} = g(\delta ET_{pot}, \delta ET_0, \delta ET) = -\frac{\delta ET_{pot} \delta ET_0}{\delta ET - \delta ET_0}$$

The potential evapotranspiration of each of the vegetation classes are modified with this increment. Using a value of β as large as 10, convergence has been found to occur within 3 iterations for 96% of the cases. The percentage of failure is 2.4% among which only a marginal fraction could have been solved using a more robust procedure; the other cases of failure are due to a soil water content so low that the potential evapotranspiration is increased up to unrealistic values. Using $\beta = 1$, convergence occurs for 99% of cases within 3 iterations.

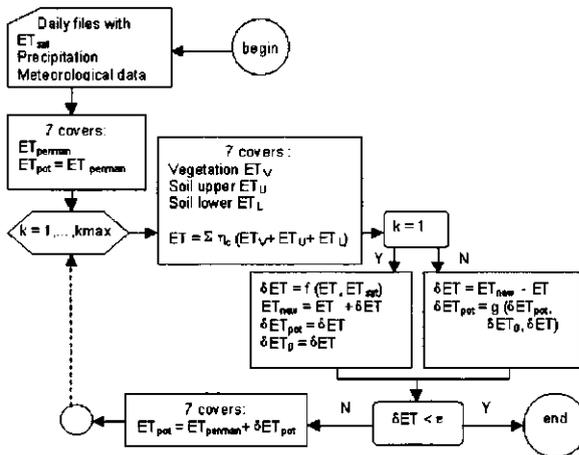


Figure 9.3.: Flow diagram of the assimilation of latent heat fluxes estimated with satellite data in the SCHEME model.

9.5. Soil moisture data and validation

Soil moisture data provide from a project study at the Université Catholique de Louvain [Auquière et al., 1997]. This project was devoted to maize growth monitoring using SAR images. Field surveys consisted in collecting ground truth data in a large area (Figure 9.4) at seven dates in 1995 corresponding to the satellite overpass. Soil cores of 5 cm depth have been also taken for gravimetric soil water measurements. The largest part of the study area lies in the Scheldt basin. Grid cells of the SCHEME model which contain at least 2 fields that were visited during these surveys have been included in this validation. Actually each field survey lasted 2 to 3 days. We have merged the data of each survey and kept the same date as in Auquière [2001].

In order to compare values measured on the top 5 cm of soil and values simulated for the "upper soil reservoir" of the SCHEME model, the parameter w_{ux} defining the capacity of this bucket has been fixed to the value of 18 mm. This value is estimated considering that the gravimetric soil moisture ranges between 0,02 and 0,29. Taking a dry bulk density of 1.3 g cm^{-3} , this range corresponds to a difference of 0.35 in volumetric moisture. The difference in volumetric moisture corresponds to a depth of $50 \text{ mm} \times 0.35 = 18 \text{ mm}$. The saturation of the upper soil reservoir simulated by the model is rescaled between 0.02 and 0.29 to be compared with measured gravimetric soil moisture.

Results are exhaustively presented on Figure 9.6 to 9.8 together with the assimilation of latent heat fluxes results. It can be seen that the SCHEME model performs well in following the trend of surface soil moisture. The regression line between the observed and the simulated soil moisture is $w_{sim} = 1.1w_{obs}$ ($n = 58$, $R^2 = 0.68$). A sub-sample corresponding to the data on the 18 July 1995 is plotted on Figure 9.5. This sub-sample presents the widest range in observed soil moisture and gives the opportunity to validate the pattern obtained with SCHEME (Figure 9.5). Comparing the raingauge network and the grid of the model, it is suggested that for three grid cells, the low values simulated with SCHEME compared to value measured on soil cores could be due to the interpolation of precipitation.

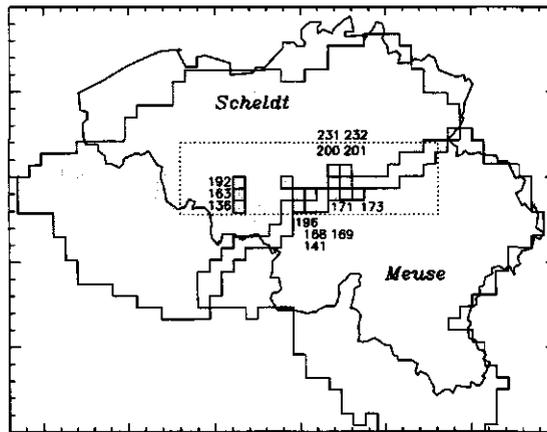


Figure 9.4.: Domain of the SCHEME model applied to the Scheldt and Meuse basins; study area and grid cells with soil moisture measured by *Auquière [2001]*

9.6. Actual evapotranspiration assimilation and soil moisture

On Figures 9.6 to 9.8, the time series of soil moisture simulated with SCHEME with and without assimilation of evapotranspiration are presented for the grid cells of the Scheldt basin. Mean and standard deviation of measured data are also shown. As mentioned above, the correlation coefficient between observed and measured soil moisture is 0.68. This value is decreased to 0.65 with data assimilation and $\beta = 1$ and drop to 0.53 with $\beta = 10$. With this latter value, the upper soil layer of the model is completely dried during long spells.

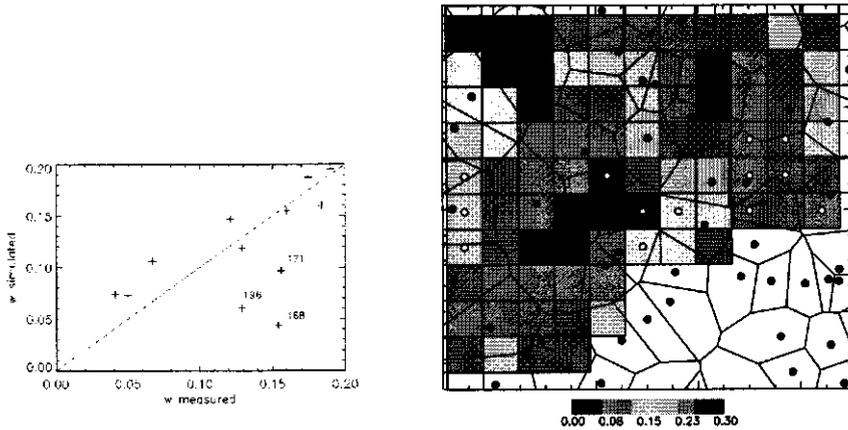


Figure 9.5.: Gravimetric soil moisture (w) of the upper 5 cm on July 18, 1995; left: observed versus simulated (with SCHEME); right: simulated soil moisture overlaid with the raingauge network and the corresponding Thiessen polygons. The labels and the white circles are explained in Figure 9.4.

Analysis of the data at the 18th of July and removing the three cells suspected of errors due to rainfall interpolation, the correlation coefficient is 0.72. This value is improved to 0.79 with assimilation of actual evapotranspiration and $\beta = 1$. The correlation is lost with $\beta = 10$. The slope of the regression line is 1.03 for the reference run and 0.84 with data assimilation and $\beta = 1$. It should be noted that this slope depends on the range used to rescale soil saturation to the gravimetric soil moisture.

A long drought spell occurred during August 1995 and the soil as simulated with SCHEME became saturated at the end of September. Discrepancies between measurements and simulation results may be due to the chosen minimum and maximum values for the simulated soil moisture. The comparison is also limited by the simplification in the conceptual model of replacing a set of soil profiles with two buckets. The differences are increased or remain unaffected with data assimilation.

The study area is not too far from the micro-meteorological station. The landscape is not too different so that the evaporative fraction estimated at this station may be used with more confidence than in the Ardennes. Improvement gained by assimilating actual evapotranspiration from Meteosat data in the SCHEME model has been found for the case of sufficient spatial variability. The importance of rainfall has been also stressed. All sources of spatial information have to be exploited optimally.

9.7. Comment on weather radar data

As mentioned in the previous section, the comparison between the soil moisture in maize fields measured on the 17-19 July 1995 and simulated with the SCHEME model (Figure 9.5) raises the question of the interpolation of rainfall data. A weather radar is operating at Zaventem (Brussel's airport). Images with 6 rain rates levels are sent every 15 minutes. Difficulties linked to the calibration of radar data [Gellens-Meulenberghs et al., 1992] is

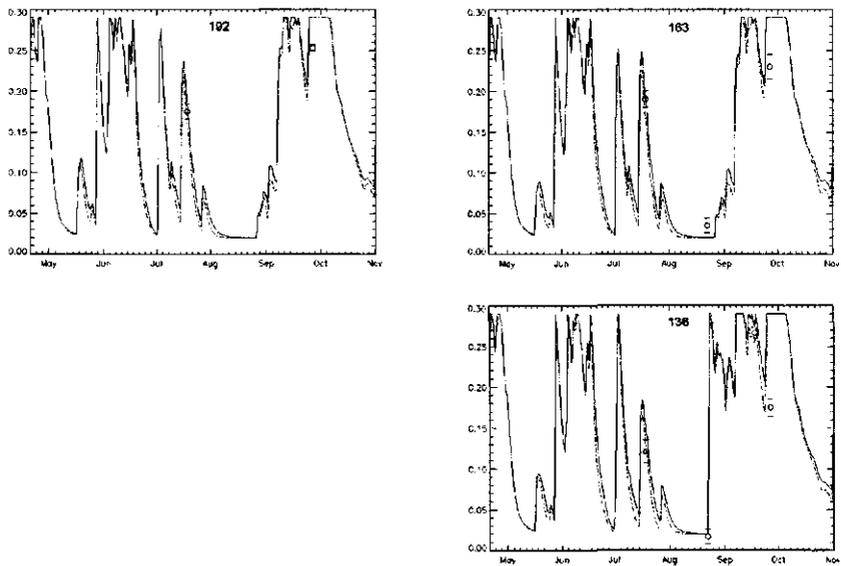


Figure 9.6.: Soil moisture of the upper of the upper 5 cm in maize fields during 1995: average and standard deviation of measured (gravimetrically) values (circles); simulated with SCHEME (solid line), simulated with assimilation evapotranspiration observations from Meteosat (dash-dotted line). The numbers refer to Figure 9.4.

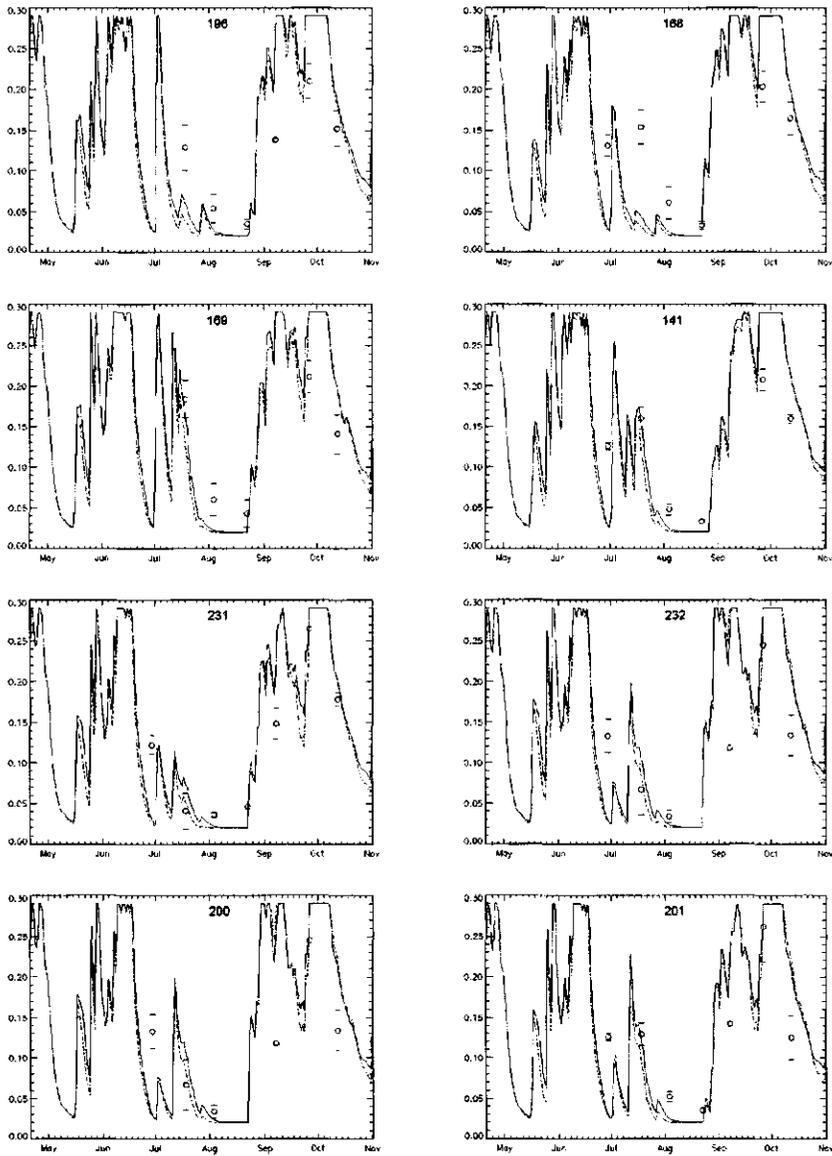


Figure 9.7.: Soil moisture of the upper of the upper 5 cm in maize fields during 1995: average and standard deviation of measured (gravimetrically) values (circles); simulated with SCHEME (solid line), simulated with assimilation evapotranspiration observations from Meteosat (dash-dotted line). The numbers refer to Figure 9.4.

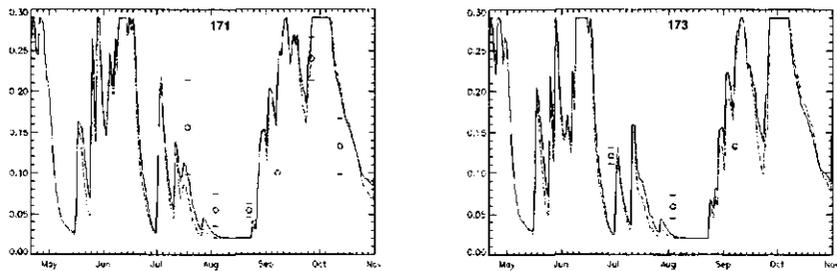


Figure 9.8.: Soil moisture of the upper of the upper 5 cm in maize fields during 1995: average and standard deviation of measured (gravimetrically) values (circles); simulated with SCHEME (solid line), simulated with assimilation evapotranspiration observations from Meteosat (dash-dotted line). The numbers refer to Figure 9.4.

one of the reasons why these data has not been used for hydrological purposes. We have examined the series of images of July 1995.

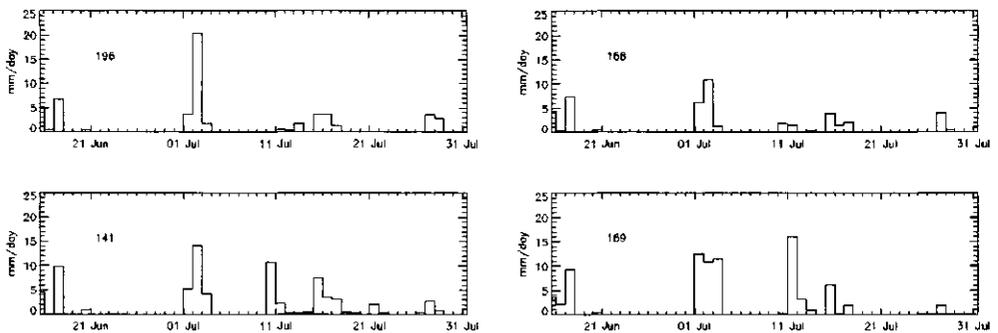


Figure 9.9.: Rainfall during July 1995, interpolated on 4 grid cells of the SCHEME model; the numbers refer to Figure 9.4.

An example is given on Figure 9.10 for the 15th of July 1995. The weather radar data have been processed in the following simple way. Echoes from surrounding buildings and from topographic features result in pixels with spurious data. These pixels have been removed using images taken during a clear day. Data from 18 rainfall stations, covering the area where also the previously described soil moisture measurements were taken, have been used to calibrate the radar data. For each station, histograms have been computed by summing the occurrence of the 6 levels over the 96 images (from 6 to 6 h U.T.) and over 3×3 pixels. The rain rates corresponding to the 6 levels have been adjusted to best fit data measured at the station. The root mean square of the difference between rainfall measured and estimated from radar is 5.5 mm (or $y = 1.01x$; $r^2 = 0.56$).

Assimilation of these data have not been attempted due to the poor calibration results and due to gaps in the radar data during the rainfall events on the 10 and 11 July 1995

(Figure 9.9) that could have helped to explain soil moisture pattern. Nevertheless, the comparison between both fields in Figure 9.10 - especially the different patterns of the area of higher rain intensities on the left part - argue for a closer look at the use of weather radar data. A new radar exclusively devoted to weather forecasting applications has been set up recently in the Ardennes. It is foreseen that these new data will be used in the hydrological modeling of the Meuse basin and in particular in the Ardennes catchments like the Ourthe.

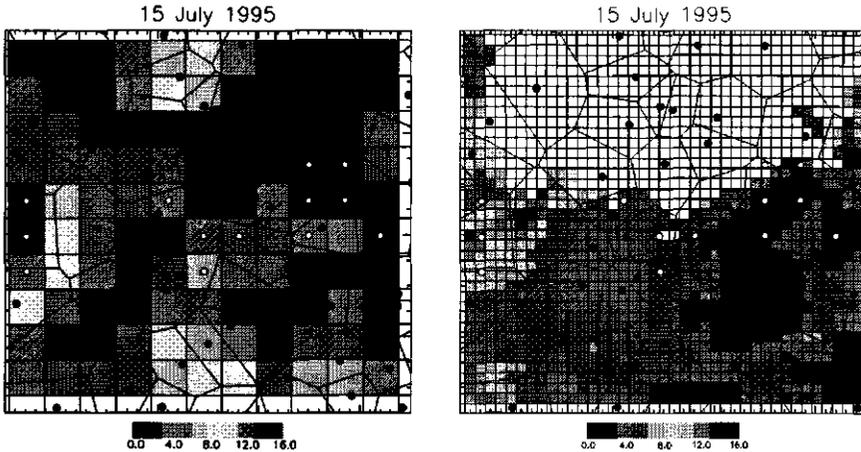


Figure 9.10.: Rainfall in mm/day; left: interpolated from the raingauge network data on the grid of the SCHEME model; right: derived from the weather radar.

9.8. Actual evapotranspiration and the water balance

The SCHEME model has been applied to the Meuse Basin with the capacity of the upper soil layer (*wux*) set to 35 mm, corresponding approximately to a layer of 10 cm thickness. Three runs have been carried out. For the reference run, input data consist in timeseries of interpolated precipitation and meteorological data of 1994 and 1995. For the second run, the values of solar radiation interpolated from the data of the hydrometeorological stations in 1995 are replaced with values estimated from Meteosat images. For the third run, values of actual evapotranspiration estimated from Meteosat images are assimilated using the method presented above and with a value of $\beta = 1$. Comparisons are made for the days with Meteosat images either as map of monthly averages or as daily values values integrated over four subcatchments (Figure 9.11).

Solar radiation is the driving component of the radiation balance. Global radiation can relatively easy be estimated from Meteosat data. It is free of the assumption of a homogeneous "evaporative fraction", which is a questionable assumption underlying the available actual evapotranspiration products, derived from Meteosat. Therefore, the use of global radiation inputs in SCHEME on the basis of Meteosat instead of interpolating radiation on the basis of observations at hydrometeorological stations is tested. as input of the SCHEME model has been tested for comparison with assimilation results. First, the daily values of global radiation, as interpolated from the data measured at the hydrometeorological stations and averaged over the four sub-catchments are compared with the

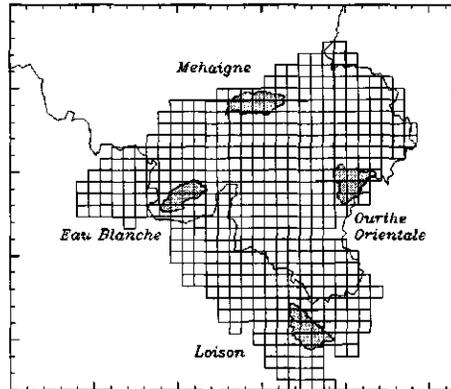


Figure 9.11.: Grid of the SCHEME model applied to the Meuse basin and the boundaries of sub-catchments selected for the analysis of data assimilation.

corresponding values estimated from Meteosat data. The results are shown in Table 9.1. In the Ourthe Orientale the two data-sources are correlated most, with a slope close to 1.

Table 9.1.: Regression between global radiation interpolated from hydrometeorological stations and values estimated from Meteosat data.

catchment	n	intercept	slope	R ²
Mehaigne	293	1	0.93	0.92
Ourthe Orientale	296	-11	0.99	0.94
Eau Blanche	295	4	0.95	0.92
Loison	297	-16	0.91	0.92

The root mean square of the difference between both is about 20 % of the average during 1995 and ranges from 9 to 24 % of the mean monthly values from March to October, 35 % in February and November and 55 % in January and December. This sub-catchment is very close to the radiation station of Saint-Hubert. For this station, *Gellens-Meulenberghs [2000]* report root mean square of the difference of hourly values ranging between 20 to 86 % from January to July 1995, stressing also the difficulty of validating values averaged on 5×3 pixels with point measurements.

The effect of replacing global radiation and of assimilating actual evapotranspiration is evaluated by first comparing the regression between daily actual evapotranspiration estimated with SCHEME and assessed from Meteosat data. Three regressions are made: a reference run where the results from SCHEME are derived in a straightforward manner, a run in which global radiation inputs in SCHEME are provided by Meteosat, and a run in which actual evapotranspiration data are assimilated. The results are shown in Table 9.2. For the first two runs, the best relationship is found for the Ourthe Orientale and the worst relationships are found for the Mehaigne and the Loison. Replacing the global solar radiation in SCHEME makes the actual evapotranspiration more similar to the values derived from Meteosat. As expected, the improvement is less important for the Ourthe Orientale than for the other three catchments. The effect of assimilating the actual evapotranspiration (using $\beta = 1$) is more important than just replacing the global solar

radiation: it results in lower intercepts and slopes and correlation coefficients even closer to 1. Differences among the results for the different sub-catchments are less pronounced when assimilating actual evapotranspiration.

Table 9.2.: Regression between daily actual evapotranspiration as calculated by SCHEME and by Meteosat data.

catchment	Reference			G substituted			ET assimilated		
	a	b	R ²	a	b	R ²	a	b	R ²
Mehaigne	0.29	0.66	0.72	0.30	0.69	0.77	0.14	0.81	0.94
Ourthe Orientale	0.17	0.78	0.81	0.20	0.80	0.83	0.10	0.88	0.96
Eau Blanche	0.26	0.73	0.74	0.28	0.76	0.83	0.14	0.85	0.94
Loison	0.25	0.61	0.72	0.29	0.73	0.81	0.12	0.82	0.97

The validation data consist of observed streamflow in the four catchments. In Table 9.3 the total annual streamflow and actual evapotranspiration are compared. Since the total streamflow during 1995 is overestimated for the Mehaigne, the Ourthe Orientale and the Eau Blanche, the total streamflow is improved with the use of Meteosat derived information. Both global solar radiation and actual evapotranspiration assessed from Meteosat data are namely larger than in SCHEME. When comparing the timeseries (Figure 9.12), the differences between observed and simulated streamflow turn out to vary for the different catchments. The large difference between observed and simulated total streamflow for the Mehaigne (104 mm) is partly due to water seepage towards adjacent catchments. In this respect, the Mehaigne catchment is similar to its neighboring Geer (Jeker) catchment. Adding one parameter for subsurface water exchange in the set of tunable parameters of the SCHEME model makes the model performances as high as for other catchments. Another example is given by the Ourthe Orientale for which the optimized and then regionalized values of the "runoff coefficient" in the autumn seems to be underestimated. As a consequence, no difference between the three runs could be found when comparing the regression between simulated and observed daily streamflow (R²: 0.85-0.90).

Table 9.3.: The validation results for streamflow ($mm\ year^{-1}$), for an open loop simulation and assimilation with G and ET, using data for 1995.

catchment	Observed	Reference		G substituted		ET assimilated	
	Q	Q	ET	Q	ET	Q	ET
Mehaigne	246	350	495	339	516	344	524
Ourthe Orientale	665	702	511	690	529	687	535
Eau Blanche	518	537	529	532	546	532	548
Loison	-	586	464	534	540	540	533

Due to the spatial and temporal resolution of geostationary satellite information, the actual evapotranspiration derived from Meteosat provides a potential tool to monitor droughts over large areas. On Figure 9.13 and 9.14, the actual evapotranspiration (based on Meteosat) is averaged for months April to September in 1995, covering the central part of the Meuse basin. The spatial variability of the actual evapotranspiration from Meteosat is lower than the spatial variability of the SCHEME results for the six months. This is probably due to the use of the single evaporative fraction value for Belgium as described previously. The effect of assimilating Meteosat-based actual evapotranspiration data (with $\beta = 1$) will therefore be a homogenization and a shift towards higher evapotranspiration

values, while the patterns obtained in the reference run are maintained.

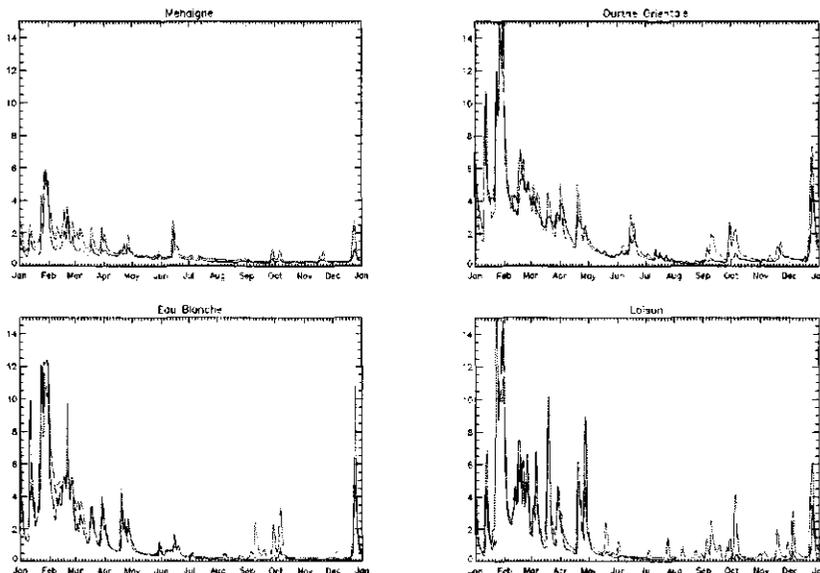


Figure 9.12.: Measured (blue) and simulated (red) streamflow($mm\ day^{-1}$) during 1995 at the outlet of the four sub-catchments.

9.9. Conclusions

Actual evapotranspiration estimated from Meteosat data has been assimilated in the SCHEME model by adapting the potential evapotranspiration inputs so that the actual evapotranspiration simulated by the model reaches a desirable value. This desired evapotranspiration value is an average that is weighted according to the ratio of the variances of the errors of the hydrological model and the satellite methodology.

The assimilation method has been applied to the loamy region in Belgium. The soil moisture values simulated with SCHEME have been compared with data from a field campaign during 1995 by UCL. The modeled soil moisture follows the measured sequence of soil drying and wetting. The impact of data assimilation is small, which probably is caused by the uncertainties associated to areal precipitation during thunderstorms. For such events, the use of weather radar images could be useful.

The assimilation of actual evapotranspiration has been evaluated by the analysis of the water balance in the Meuse basin. It has been shown how the actual evapotranspiration simulated with SCHEME is more variable spatially than the values assessed from Meteosat data. This is probably due to the use of a single value of the "evaporative fraction" for entire Belgium. Part of the differences are explained by differences between the global solar radiation interpolated from data from sparse stations and the values estimated from the Meteosat images in the visible channel. Assimilating actual evapotranspiration from Meteosat data reduces the spatial variability but the pattern remains closer to that of the SCHEME results than to the pattern of Meteosat results.

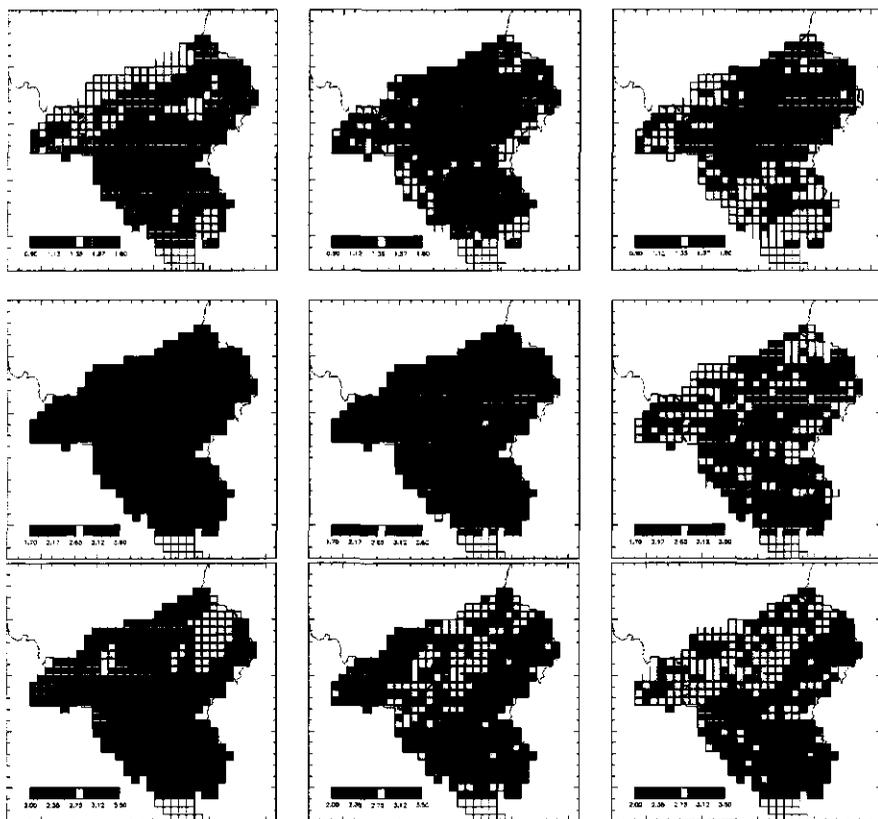


Figure 9.13.: Daily actual evapotranspiration ($mm\ day^{-1}$; averaged over the month). Left: based on Meteosat images. Middle: output from SCHEME. Right: output from SCHEME with data assimilation. Note that the scales are different for each month.

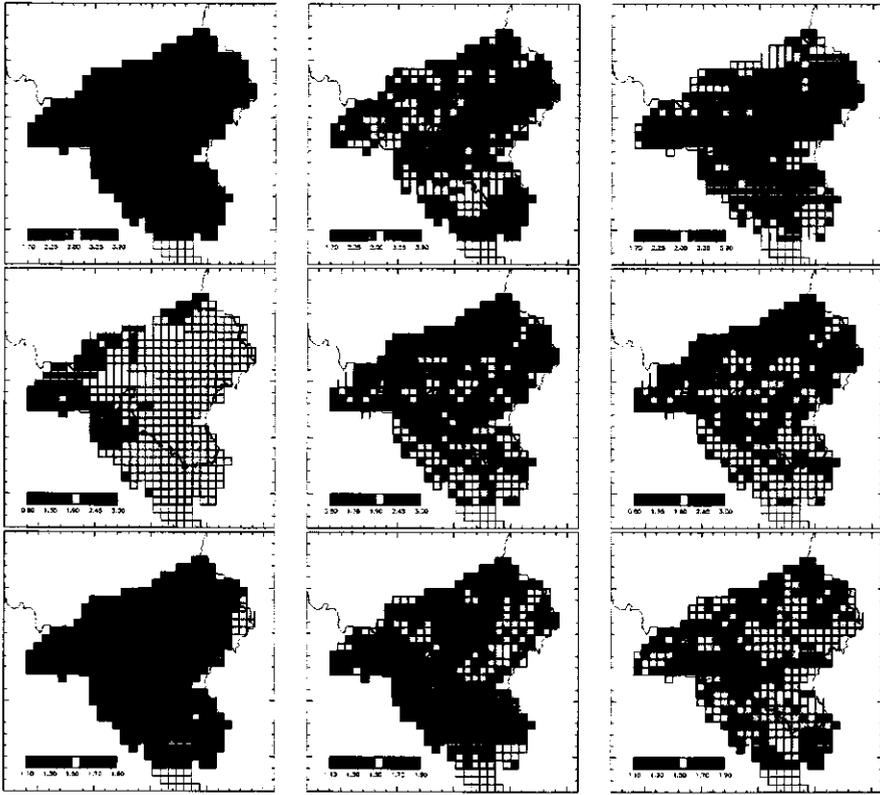


Figure 9.14.: Daily actual evapotranspiration ($mm\ day^{-1}$; averaged over the month). Left: based on Meteosat images. Middle: output from SCHEME. Right: output from SCHEME with data assimilation. Note that the scales are different for each month.

In this study the open-loop predictions by the SCHEME model were shown to be close to the Meteosat-derived values. In some respects the assimilation of Meteosat-derived evapotranspiration led to improved model predictions, but overall, the effect of using the Meteosat data is modest. Apart from that, the feasibility to derive useful actual evapotranspiration data from Meteosat images remains to be proven - especially since the errors associated with this estimation procedure are unclear and most likely correlated in space and time. This study also leads to the question which type(s) of data to assimilate: the use satellite radar images to update soil moisture or weather radar to enhance rainfall estimates may prove to be more effective than the use of Meteosat to update evapotranspiration data.

Part III.

Conclusions and Recommendations

10. Conclusions

10.1. Achievements by the DAUFIN project

The DAUFIN project succeeded in a 20-month period to achieve a great deal of the goals that were initially set:

- The REW model has been coded and tested on existing data;
- The REW model has partially been compared to existing distributed models (so-called control models);
- Data assimilation routines have been developed, tested and implemented in control models;
- Data assimilation for hydrologic modeling has been brought under the attention of the scientific community in Europe via an international scientific workshop on 'Catchment Scale Hydrological modeling and Data Assimilation' in Wageningen.
- The findings of the project have been disseminated to end-users via an end-user workshop in Delft, November 2001.

10.2. Project output

A lot of useful software was developed during the project. These comprise several hydrologic modeling codes: the REW-model, extensions to the CATHY model, extensions to the SCHEME and LISFLOOD models; and several data assimilation routines: a nudging routine, statistical correction, and cross-validation routines. These software products are meant to be used in further research, to be applied in case studies and in some cases made ready for operational use.

In addition to the computer codes, comprehensive data sets with long-term hourly meteorologic and hydrologic observations have been compiled for a large number of meteo-stations and a limited number of discharge stations in the Ourthe and Geer basins. Several data sets with terrain, soil and land use information have been compiled. A complete overview is given in *de Roo et al. [2000]* The data sets have proven to be valuable for the development and testing of the various models. These data sets will remain in tact for future use but were obtained under strict licences and are unfortunately not available for use outside the DAUFIN project.

In addition to this real data a comprehensive set of synthetic data has been generated with a full 3D Richards model (CATHY), while using real terrain data and meteorologic inputs (the Brisv catchment). This data set is described in *Oosterhoff and Paniconi [2001]*. Also

this data set has been of great value for the evaluation of several data assimilation and modeling techniques. This data set is freely available.

All the publications and presentations resulting from the DAUFIN project are listed in Appendix A.

10.3. Dissemination of results

The project results were disseminated in a variety of ways: through scientific papers in peer-reviewed journals, in poster and oral presentations at scientific conferences, in more informal workshops and on the DAUFIN website. Most of these are listed in Appendix A. Members of the DAUFIN team are actively engaged in other research projects involved with data assimilation and catchment scale hydrologic modeling and the results of the DAUFIN project will filter through to these projects.

Via the end-user workshop (November 2001), a larger group of interested parties were informed about the project results and at this workshop it was concluded that there is a large need for data assimilation techniques by the end-users. There was agreement among the end-users that the main constraint for the wide-scale application of data assimilation techniques is the ignorance about the topic and the absence of suitable software tools (not data).

11. Recommendations

The successful launch of Earth Observation (EO) satellite missions, such as ENVISAT and TERRA, the imminent launch of others (e.g. MSG, GPM), and the use of both airborne and in-situ instrumentation are expected to revolutionize hydrology by supplying the environmental drivers that have been lacking to date. The primary justification for these new EO satellite missions is to provide spatial and temporal information that when coupled to hydrological models allow the prediction of soil moisture and river flows that are vital to eco-systems, agricultural production and water resources. Remote sensing technology provides many types of data that are related to land surface variables of interest to hydrologists. However, very little of this information is available in a form that can be used directly for hydrological purposes. Therefore, data assimilation research in hydrology should focus on producing data products that are directly useful for water management. Such products need to be carefully designed to meet the needs of potential users, with due consideration to:

- resolution and spatial configuration of the data products;
- quantitative measures of data product reliability;
- quality control issues;
- sensitivity of data products to "hidden" model properties.

Classical hydrological models that have been optimized for use with sparse in situ observations are inadequate for extension to work with remote sensing data. There is a need for developing more appropriate distributed models at the catchment-scale. Only then we will be able to bridge the gap between continental-scale data sets (such as GLDAS) and catchment scale applications. The development of these appropriate catchment-scale models will involve downscaling and parameterization issues. In both, data assimilation can play a central role.

More research is needed to develop data assimilation algorithms that can handle the specific problems encountered in hydrological applications:

- subsurface processes are hard to "observe";
- systems are very heterogenous - especially in space;
- hydrologic systems function over a wide range of spatial and temporal scales.

Geostatistical techniques for describing multi-scale spatial heterogeneity need to be incorporated into algorithms that account for the multi-resolution nature of different but complementary hydrologic measurements.

Finally, all this calls for carefully selected case studies to introduce and demonstrate the potential of data assimilation in operational water resources management (e.g. improved flood predictions). Continued investments and coordination of data assimilation initiatives at the European level is needed, because:

- the wide range of research topics relevant to data assimilation;
- the strong need for innovation in each of these areas;
- the clear potential for water resources modeling and management;
- the transboundary nature of catchments and river basins;
- the need for common algorithms, models, tools and data standards;
- the leading role already demonstrated by European researchers.

Expertise from many disciplines will be needed to meet the challenge of data assimilation for improved river basin water resources management: hydrologists, meteorologists, remote sensing specialists, ecologists, mathematicians, information technology experts, water managers, etc.

Appendix A

DAUFIN project output: publications, presentations and data.

Scientific publications

- van Loon, E. E.; P. Torfs; P. A. Troch (2001) Notes on data assimilation algorithms suitable for hydrologic modeling. Internal report, Sub-Dept. of Water Resources, Wageningen UR.
- van Loon, E. E.; P. A. Troch (2001a) Directives for 4D soil moisture data assimilation for hydrologic modeling. In: Soil-vegetation-atmosphere transfer schemes and large scale hydrological models. Proceedings of the 6th Scientific Assembly of the IAHS, Maastricht 2001, IAHS publication nr. 270, pp 257-269.
- van Loon, E. E.; P. A. Troch (2001b) An intercomparison of cross-validation and Kalman filtering algorithms. Internal report, Sub-Dept. of Water Resources, Wageningen UR.
- van Loon, E. E.; P. A. Troch (2002) Tikhonov regularization as a tool for assimilating soil moisture data in distributed hydrological models. *Hydrologic Processes*, 16(2) 531-556.
- Marrocu, M. and C. Paniconi (2001) Assessment and formulation of data assimilation techniques for a 3D Richards equation-based hydrological model. Technical Report CRS4-TECH-REP-01/27. Center for Advanced Studies, Research and Development in Sardinia, Cagliari, Italy.
- Mirosław-Swiątek D., and A. Wita (2000) Mathematical model and computer programme of water movement through river embankments for changeable hydrological data. Proceedings of the conference on flood hazard in Sudety Mountains. Ładek Zdrój, October 12-13, 2000.
- Oosterhoff, R. T. and C. Paniconi (2001a) Implementation of a catchment hydrologic model for the Brisy subcatchment of the Ourthe watershed, and generation of a dataset for a 240-day storm-interstorm sequence. Technical Report CRS4-TECH-REP-01/28, Center for Advanced Studies, Research and Development in Sardinia, Cagliari, Italy.
- Reggiani, P., M. Sivapalan, S. M. Hassanizadeh, and W. G. Gray (2001) Coupled equations for mass and momentum balance in a bifurcating stream channel network: Theoretical derivation and numerical implementation. Proceedings of the Royal Society (London) Series A: Mathematical, Physical & Engineering Sciences, Vol. 456, 1-33.

- Reggiani, P., M. Sivapalan, and S. M. Hassanizadeh (2000) Conservation equations governing hillslope responses: Physical basis of water balance. *Water Resour. Res.*, 36, 1845-1864.
- de Roo, A.P.J., Wesseling, C.G., Van Deursen, W.P.A. (2000) Physically-based river basin modeling within a GIS: The LISFLOOD model. *Hydrological Processes*, Vol.14, 1981-1992.
- de Roo, A.P.J., Thielen, J, Schmuck, G., Somma, F, Somma, G., Price, D., Bates, P., Van Der Knijff, J., Gouweleeuw, B., Stam, M. (2001). The LISFLOOD water balance and flood simulation model for large drainage basins. In preparation. To be submitted to *Journal of Hydrology*.
- de Roo, A.P.J.; G. Somma (2001). Water balance and flood simulations in the Ourthe basin using the LISFLOOD model. In preparation. To be submitted to *Hydrological processes*.
- Somma, G.; A.P.J. De Roo (2001) Validating LISFLOOD in the Ourthe basin. JRC Internal report.

presentations

- Bixio, A. C., M. Putti and C. Paniconi (2001) Coupling of a surface routing model to a subsurface flow model, with an application to the Chernobyl watershed (abstract), *Geophysical Research Abstracts (European Geophysical Society 26th General Assembly)*.
- Botti, P; P. Pillai and R. Silvano (2000) Complex hydraulic systems management under hydrological uncertainty. *Proceedings of the International Conference "Siccità: Monitoraggio, Mitigazione, Effetti"*, Villasimius (Italy) September 2000.
- Dassargues, A., S. Brouyere and C. Rentier (2001) Integrating geophysical and tracer test data for accurate solute transport modeling in heterogeneous porous media, accepted for *GQ'2001, 3rd Int. Conf. on Groundwater Quality*, June 2001, Sheffield.
- Dassargues, A. Couplages de modèles d'eaux souterraines à des modèles agronomiques et d'eaux de surface en vue de la simulation intégrée du cycle de l'eau à l'échelle d'un sous-bassin. 29/5 EPFL Lausanne et 30/5 Université de Neuchâtel (Switzerland).
- Hassanizadeh M., P. Reggiani, M. de Neef, M. Sivapalan, T. Rientjes (2001) REW-Scale modeling of rainfall-runoff processes (abstract), *Geophysical Research Abstracts (European Geophysical Society 26th General Assembly)*.
- Hilberts, A. G. J., E. E. van Loon, C. Paniconi and P. A. Troch (2001) Validity of quasi-3D hillslope hydrological models (abstract), *Geophysical Research Abstracts (European Geophysical Society 26th General Assembly)*.
- Paniconi, C., P. A. Troch, M. Mancini and M. A. Dessena (2000) Soil moisture mapping from ASAR imagery for the Flumendosa and Meuse river basins}. In: *Proceedings European Space Agency ERS-Envisat Symposium: Looking Down to Earth in the New Millenium*. CD-ROM SP-461, ESA Publications Division, ESTEC, Noordwijk, The Netherlands.

- Paniconi, C. and M. Marrocu (2001) Nudging a Richards equation finite element model to soil moisture observations (abstract) Geophysical Research Abstracts (European Geophysical Society 26th General Assembly).
- van Loon, E. E. and P. A. Troch (2000) Data assimilatietechnieken in hydrologische modellering van stroomgebieden (Data assimilation techniques in catchment scale hydrological modeling). In: NCR publicatie 03-2001, NCR dagen 2000, het begin van een nieuwe reeks. ISSN 1568-234X, p. 54.
- van Loon, E. E. and P. A. Troch; E. E. van Loon (2001b) Directives for 4D data assimilation in hydrological modeling (abstract), Geophysical Research Abstracts (European Geophysical Society 26th General Assembly).
- de Neef, M., P. Reggiani, T.Rientjes, M.Hassanizadeh (2001) Application of the REW approach to a generic catchment in the Belgium Ardennes (abstract), Geophysical Research Abstracts (European Geophysical Society 26th General Assembly).
- Torfs, P.; R. Wojcik (2000) Local probabilistic networks in hydrology. ERB 2000 Conference on Monitoring and modeling catchment scale water quantity and quality. Ghent, Belgium September 27-29, 2000.
- Torfs, P. (2000) Technieken voor disaggregatie van hydrologische tijdreeksen (Statistical techniques for disaggregation of hydrological data). In: NCR publicatie 03-2001, NCR dagen 2000, het begin van een nieuwe reeks. ISSN 1568-234X p. 55.
- Troch, P. A. ; E. E. van Loon (2000a) New techniques for 4D soil moisture data assimilation. Presentation at the LASSA meeting in Utrecht, June 2000.
- Troch, P. A.; E. E. van Loon (2000b) Soil moisture data assimilation within a distributed hydrological model. ERB 2000 Conference on Monitoring and modeling catchment water quantity and quality. Ghent, Belgium September 27-29, 2000.
- Troch, P. A.; E. E. van Loon (2000c) Characteristic Subsurface Stormflow and Saturation Overland Flow Response Functions of Hillslopes. Eos Trans. AGU, 81 (48), Fall Meet. Suppl., Abstract H71E-09, 2000.

free data

- Oosterhoff, R. T. and C. Paniconi (2001b) Brisly subcatchment simulation dataset. CD-ROM, Center for Advanced Studies, Research and Development in Sardinia (CRS4) and Wageningen University and Research Center (WUR).

Summary

The aim of the DAUFIN project is to design and apply methodologies that are able to improve the predictive skill of a hydrological model by including, in the physical state of the simulation model, scattered measures of various kinds (satellite, in situ, indirect measures, etc.). The project aims to study the available data assimilation routines in depth, and apply these to existing distributed hydrological models. But in addition to this the DAUFIN project also aims at a unification of those different models through the REW concept. This unification is believed to assist much in developing better data assimilation routines, which will eventually lead to a better understanding of the hydrologic systems being studied and a higher predictability. The modest size and the short project period of 20 months is of course not sufficient to fully realize such an ambitious goal, therefore only initial steps are set towards such a unification. These are: the development of a fully operational REW modeling code, the inter-comparison of this code with an existing distributed model, and finally the design and application of data assimilation routines specific for the REW code.

The project was organised along four main activities that took place more or less in parallel. The first was the development of the REW computer code. The second project activity was the development of data assimilation routines. A third project activity was the application of existing distributed hydrological models (so-called control models) to sub-catchments of the Meuse. And a fourth activity was the inter-comparison of the REW computer model with one of the control models. In support of these activities there was a data collection effort to make a good set of meteorologic and hydrologic data available to all the partners, and several efforts to disseminate the project results in the scientific community and among end-users.

Chapter 2 gives an overview of the structure and performance of the REW model. The governing equations required to model a whole watershed underlain by an aquifer through the representative elementary watershed approach are introduced. The chapter lists the governing equations for mass and momentum. The momentum balance equations are vectorial and are therefore projected along the axes of an appropriate reference system. Subsequently the constitutive relationships are briefly described, which are used to close the unknown mass and momentum exchange terms found in the governing equations. Finally there are still a series of parameters appearing in the balance equations, which need to be expressed in terms of either the independent variables for which the governing equations are solved, or in terms of quantities which are directly measurable in the field. In addition a theoretical procedure has been pointed out, which allows to estimate the exchange coefficient for the partition of mass between a REW and its neighbours.

In the numerical modeling, the high potential of the REW-approach could not fully be explored. The model structure of the REW-model requires some minor modifications. First however, much more model tests on smaller, and preferably, shallow catchments must be executed in order to understand the model behavior in more detail. The overall

model behavior, however, has proven to be successful in modeling the runoff production mechanism of saturation overland flow and the routing of water in the channel network system. As such, the semi-distributed modeling approach with interconnected REWs and 5 model domains is a very promising approach. Future research also must focus on combining the REW approach to 4DDA and automated multi-objective model calibration to improve model simulation and forecasting. So far, due to the time consuming process of code development, debugging and testing, the REW model could not be combined with 4DDA techniques. In the near future, however, after the code has been further developed, the use of multi-objective model calibration and 4DDA is foreseen and regarded an absolute necessity to future model applications.

Chapter 3 provides a review to existing data assimilation technology. Data assimilation is a procedure to provide time-dependent spatially distributed estimates of a dynamic system using observations from various sources and physical constraints in an efficient way. Mathematically it can be seen as a state estimation procedure. The spatio-temporal prediction of soil moisture is pre-eminently a problem that can be dealt with by data assimilation techniques. Even though there is considerable literature on data assimilation in the environmental sciences, the application to soil moisture estimation is still new. To date only the problem of integrating remotely sensed data or field-observations over a small area with a one dimensional soil moisture model of the root-zone has been studied. In all cases the assimilation technique comprises simplifications or variants of the weak-constraint variational technique (equivalent to the Kalman smoother). Problems involving more spatial dimensions and more diverse information sources have not yet been dealt with. This is partly because the dimensionality of the problem prohibits a straightforward extension of the common algorithms, i.e. new data assimilation algorithms need to be developed. The lack of appropriate data sets and regularization tools are also serious impediments. As a first step towards applying soil moisture data assimilation to problems of a higher dimension, this study outlines the desirable structure of 4D hydrological data assimilation algorithms, and discusses the consequences for data handling and the specification of hydrologic models.

In Chapter 4 the implementation and testing of a Newtonian nudging (or relaxation) technique for data assimilation is described. This technique has been selected as a suitable scheme for conducting some initial investigations into the applicability and effectiveness of data assimilation for a three-dimensional finite element Richards equation-based distributed hydrological model. In the nudging procedure an additional forcing term is added to the model, so the implementation requires significant, though not complicated, code modifications. Nonetheless the overheads associated with nudging, including computation of the new forcing term, add negligible cost to the model simulation. The simplicity and efficiency of the Newtonian relaxation scheme are the two main advantages of this approach over more sophisticated data assimilation methods. Another appealing feature of the algorithm, and our implementation of it for the CATHY hydrological model, is its flexibility. It is always forward-marching in time (we assume prior knowledge of the observation times), spatial and temporal variability in the nudging parameters can be handled (to account for seasonal and geomorphological effects on the state variables, for instance), and concurrent observation datasets distributed over the entire grid or at scattered points can be assimilated (allowing utilization of information from different measurement sources). In this chapter the relative merits and tradeoffs between factors such as computational efficiency, flexibility, and hydrological performance are investigated. The test cases presented here show that the nudging technique is successful in decreasing the error between the computed and observed values, the latter are derived from a numerically generated

surrogate true integration. Surprisingly, nudging was also found to generally improve the numerical performance of the model in comparison to prior integrations, in terms of convergence, time stepping, and computational efficiency. This can perhaps be attributed to a smoothing influence of the weighting functions in the nudging relaxation term. A contrasting effect of deterioration in the numerical solution was shown for relatively high values of an empirical nudging parameter (G), suggesting that care is required in selecting G to ensure that the nudging term does not become dominant in Richards equation. The limited sensitivity trials indicated, for our test case, an appropriate value of around 0.1 for this parameter. The effectiveness of adapting G to the model time step size throughout the simulation in order to satisfy the stability criterion $G \leq 1/\Delta t$ can easily be tested. More effort will be needed to investigate model time scales in relation to the nudging term. More rigorous sensitivity analysis will be needed to assess not only the impact of G , but also of the weighting function parameters τ , R_{xy} , and R_z . Of these three parameters, τ has perhaps the subtlest effect because of its close interaction with the relaxation difference term, its relationship with time step size in an adaptive time evolution scheme, the possibility of overlapping time windows of different lengths, and the evidence of numerical under and overshoot (as seen with the centered time influence windows used in our tests). Increasing the horizontal radius of influence leads to a considerable decrease in the mean spatial difference between observed and computed surface soil moisture values. Though not investigated, similar results are expected for the vertical radius of influence parameter R_z , with an important bearing on soil moisture profile and water table estimation. These parameters, however, must in some way be consistent with or reflect the correlation scales of the state variable being assimilated. Indeed the four-dimensional weighting functions used in the nudging approach provide a simple and intuitive way of incorporating any available knowledge regarding characteristic length scales and the spatio-temporal variability of state variables. As our understanding of this variability and how to represent it improves, the weighting functions will acquire more physical relevance, taking on shapes that conceivably will be quite different from the Cressman functions used in this and previous studies, and the nudging technique will become a little less heuristic in the process. In addition to inter-comparison studies, more detailed sensitivity analyses, and improved weighting function formulations, other topics for more immediate future research include: implementation of the nudging algorithm for the Newton linearization option already available in the CATHY code; implementation of a nudging term better suited to gridded and/or high-density data such as remotely sensing observations of surface state variables; application of the model to an actual catchment with assimilation of real observation data; and evolution to more advanced data assimilation methods for the Richards equation-based coupled surface-subsurface model, using the nudging implementation as a basic framework.

In Chapter 5 the nudging technique is explored by application to the estimation of 2D soil moisture profiles in the Brisby catchment. The nudging technique applied is the same as that described in Chapter 4. The forcing term is proportional to the difference between simulation and observation (relaxation component) and contains four-dimensional weighting functions that can incorporate prior knowledge about the spatial and temporal variability and characteristic scales of the state variable(s) that are being assimilated. The numerical model couples a three-dimensional finite element Richards equation solver for variably saturated porous media and a finite difference diffusion wave approximation based on digital elevation data for surface water dynamics. The implementation of the data assimilation algorithm for the coupled model is described in detail. Similar to the results reported in Chapter 4, the properties of the assimilation scheme in terms of computational

cost and improvement of simulation results are favourable. Nudging in some cases even improves numerical performance compared to model runs without nudging. We also examine the sensitivity of the model to nudging term parameters including the spatio-temporal influence coefficients in the weighting functions. Overall the nudging algorithm is quite flexible, for instance in dealing with concurrent observation datasets, gridded or scattered data, and different state variables. The implementation presented here can be readily extended to any of these features not already incorporated. Moreover the nudging code and tests can serve as a basis for implementation of more sophisticated data assimilation techniques in a Richards equation-based hydrological model.

Chapter 6 investigates whether a data-based data assimilation technique can be used for catchment scale soil moisture modeling. On the basis of a set of observations two distinct soil moisture models are identified and calibrated: a linear stochastic time-varying state-space model, and a geo-statistical model. Both models are defined at various spatial and temporal resolutions. Next, four different ways to predict the soil moisture dynamics are compared: 1) the application of the dynamic models in open-loop form; 2) a re-calibration of the dynamic models with soil moisture data, and subsequent prediction in open-loop form; 3) prediction with the geo-statistical models, using the soil moisture data in period B; 4) prediction by combining the outcomes of 1 and 3 via generalized cross-validation. The last method, which is a form of data assimilation, compares favourably to the three alternatives. Over a range of resolutions, the predictions by data assimilation have overall uncertainties that are approximately half that of the other prediction methods and have a favourable error structure (i.e. close to Gaussian) over space as well as time. In addition, data assimilation gives optimal predictions at finer resolutions as compared to the other methods. Compared to prediction with the models in open-loop form, both re-calibration with soil moisture observations and data assimilation result in enhanced discharge predictions, whereas the prediction of ground water depths is not improved.

In Chapter 7 the possibilities to apply automatic calibration routines to the SUFT3D code (within the more general MOHISE hydrological model). The model is applied to the Geer catchment. For this purpose the existing model has been intensively reworked and modified in order to improve calibration and validation on historical data in the studied catchment. Synthetic tests have been performed to improve automatic calibration of hydraulic conductivity values (K) of a 2D groundwater flow model by conditioning stochastic co-simulations of the K -field on geophysical data.

Chapter 8 shows the results from an automatic calibration procedure applied to the LISFLOOD-WB model, with daily basis simulations, using more or less the standard parameter set (also used for the entire Meuse and Oder basins). It turns out that model performance is satisfactory in the sense that the model produced acceptable results, especially during the 4 major flood events of the 20th century (February 1984, January 1993, December 1993, and January 1995) in the Meuse. The study shows that improvement in the input data (soil maps and meteorological data) produces a significant improvement in the model results. Further knowledge on the soil properties will improve the results without need of further calibration. With respect to input data, LISFLOOD is a very demanding model and for this reason it is very sensitive to the quality of the available data. It is still unclear how data assimilation can be implemented in LISFLOOD. The most logical start, considering the model's focus on flood-prediction, would be the implementation of state correction on the basis of observed discharges.

Chapter 9 shows how actual evapotranspiration estimated from Meteosat data can be assimilated in the SCHEME model by adapting model inputs (potential evapotranspiration)

via a technique known as statistical correction. This method has been applied to the loamy region in Belgium. The soil moisture simulated with SCHEME has been compared with data from a field campaign during 1995 by UCL. The modeled soil moisture follows the measured sequence of soil drying and wetting. The impact of data assimilation may be hidden by the uncertainties associated with precipitation (especially during thunderstorms). For such events, the use of weather radar images would be useful. The assimilation of actual evapotranspiration has been evaluated by the analysis of the water balance in the Meuse basin. It is shown how the actual evapotranspiration simulated with SCHEME is more variable spatially than the values assessed from Meteosat data. This is probably due to the use of a single value of the "evaporative fraction" for entire Belgium. Part of the differences are explained with differences between the global solar radiation, interpolated on the basis of sparse stations and the values estimated from the Meteosat images in the visible channel. Assimilating actual evapotranspiration from Meteosat data reduces the spatial variability but the pattern remains closer to that of the SCHEME results than to the pattern of Meteosat results. In this study the open-loop predictions by the SCHEME model were shown to be close to the Meteosat-derived values. In some respects the assimilation of Meteosat-derived evapotranspiration led to improved model predictions, but overall, the effect of using the Meteosat data is modest. Apart from that, the feasibility to derive useful actual evapotranspiration data from Meteosat images remains to be proven - especially since the errors associated with this estimation procedure are unclear and most likely correlated in space and time. This study also leads to the question which type(s) of data to assimilate: the use satellite radar images to update soil moisture or weather radar to enhance rainfall estimates may prove to be more effective than the use of Meteosat to update evapotranspiration data.

The report ends in Chapter 10 with a list of project achievements and recommendations. The main project achievements are:

- The REW model has been coded and tested on existing data;
- The REW model has partially been compared to existing distributed models (so-called control models);
- Data assimilation routines have been developed, tested and implemented in control models;
- Data assimilation for hydrologic modeling has been brought under the attention of the scientific community in Europe via an international scientific workshop on 'Catchment Scale Hydrological modeling and Data Assimilation' in Wageningen, September 2001.
- The findings of the project have been disseminated to end-users via an end-user workshop in Delft, November 2001.

Central in the recommendations is the statement that data assimilation research in hydrology should focus on producing data products that are directly useful for water management. Such products need to be carefully designed to meet the needs of potential users, with due consideration to:

- Resolution and spatial configuration of the data products;

- Quantitative measures of data product reliability;
- Quality control issues;
- Sensitivity of data products to "hidden" model properties.

To achieve this, more research is needed to develop data assimilation algorithms that can handle the specific problems encountered in hydrological applications:

- Subsurface processes are hard to "observe";
- High degree of heterogeneity of the physical system;
- Hydrologic systems function over a wide range of spatial and temporal scales.

Geostatistical techniques for describing multi-scale spatial heterogeneity need to be incorporated into algorithms that account for the multi-resolution nature of different but complementary hydrologic measurements.

Finally, all this calls for carefully selected case studies to introduce and demonstrate the potential of data assimilation in operational water resources management (e.g. improved flood predictions). Continued investments and coordination of data assimilation initiatives at the European level is needed, because:

- the wide range of research topics relevant to data assimilation;
- the strong need for innovation in each of these areas;
- the clear potential for water resources modeling and management;
- the transboundary nature of catchments and river basins;
- the need for common algorithms, models, tools and data standards;
- the leading role already demonstrated by European researchers.

Expertise from many disciplines will be needed to meet the challenge of data assimilation for improved river basin water resources management.

Bibliography

- M. B. Abbott, J. C. Bathurst, J. A. Cunge, P. E. O'Connell, and J. Rasmussen. An introduction to the European Hydrological System — Système Hydrologique Européen, "SHE", 1: History and philosophy of a physically-based, distributed modelling system. *J. Hydrol.*, 87:45-59, 1986a.
- M. B. Abbott, J. C. Bathurst, J. A. Cunge, P. E. O'Connell, and J. Rasmussen. An introduction to the European Hydrological System — Système Hydrologique Européen, "SHE", 2: Structure of a physically-based, distributed modelling system. *J. Hydrol.*, 87: 61-77, 1986b.
- F. Ács, D. T. Mihailović, and B. Rajković. A coupled soil moisture and surface temperature prediction model. *J. Appl. Met.*, 30:812-822, 1991.
- B. D. O. Anderson and J. B. Moore. *Optimal Filtering*. Prentice-Hall, Englewood Cliffs, NJ, 1979.
- E. Auquièrè. *SAR temporal series interpretation and backscattering modeling for maize growth monitoring*. PhD thesis, Université Catholique de Louvain, Louvain-la-Neuve, Belgium, 2001.
- E. Auquièrè, P. Defourny, V. Baltazart, and A. Guissart. Multitemporal SAR data for crop monitoring - signal modeling and analysis of biophysical variables dynamic. final report of the ostc project T3/10/29. Technical report, Université Catholique de Louvain, Louvain-la-Neuve, Belgium, 1997.
- P.D. Bates and A.P.J. De Roo. A simple raster-based model for flood inundation simulation. *Hydrol. Proc.*, 236:54-77, 2000.
- J.C. Bathurst and J.M. Wicks. *Recent Advances in the Modeling of Hydrologic Systems*, chapter Framework for erosion and sediment yield modelling, pages 269-288. Kluwer, 1991.
- M.F. Baumgartner, J. Martinec, A. Rango, and R. Roberts. Snowmelt runoff model (srn) user's manual. Technical report, University of Bern, Dept. of Geography, 1994.
- A. F. Bennett. *Inverse methods in physical oceanography*. cambridge, USA, 1992.
- K. Beven. Changing ideas in hydrology — the case of physically-based modelling. *J. Hydrol.*, 105:157-172, 1989.
- K. Beven and F. Wood. Flow routing and the hydrological response of channel network. In K. Beven and M. J. Kirkby, editors, *Channel Network Hydrology*, pages 99-128. John Wiley & Sons, 1993.

- K. J. Beven. A discussion of distributed hydrological modelling. In M. B. Abbott and J. C. Refsgaard, editors, *Distributed Hydrological Modelling*, pages 255–278, Kluwer Academic, Dordrecht, The Netherlands, 1996.
- K.J Beven. Towards a new paradigm in hydrology. In N.C. Matalas and J.C. Rodda, editors, *Water for the Future: Hydrology in Perspective (proceedings of the Rome symposium April 1987, IAHS publication no. 164)*, pages 393–403. Int. Assoc. of Hydrol. Sci., 1987.
- K.J. Beven and M.J. Kirkby. A physically-based variable contributing area model of basin hydrology. *Hydrological Sciences Bulletin*, 24(1):43–69, 1979.
- A. Binley and K. Beven. Three-dimensional modelling of hillslope hydrology. *Hydrol. Proc.*, 6:347–359, 1992.
- A. C. Bixio, S. Orlandini, C. Paniconi, and M. Putti. Physically-based distributed model for coupled surface runoff and subsurface flow simulation at the catchment scale. In L. R. Bentley, editor, *Computational Methods in Water Resources*, pages 1115–1122, Balkema, Rotterdam, The Netherlands, 2000.
- M. Boutabyeb, H. Rafaralahy, and M. Darouach. Convergence analysis of the extended kalman filter used as an observer for nonlinear deterministic discrete-time systems. *IEEE Transactions on Automatic Control*, 42(4):581–586, 1997.
- F. Bouttier, J.-F. Mahfouf, and J. Noilhan. Sequential assimilation of soil moisture from atmospheric low-level parameters. part i: Sensitivity and calibration studies. *J. Appl. Meteorol.*, 32(8):1335–1351, 1993a.
- F. Bouttier, J.-F. Mahfouf, and J. Noilhan. Sequential assimilation of soil moisture from atmospheric low-level parameters. part ii: Implementation in a mesoscale model. *J. Appl. Meteorol.*, 32(8):1352–1364, 1993b.
- F. Bouyssel, V. Cassé, and J. Pailleux. Variational surface analysis from screen level atmospheric parameters. *Tellus*, 51A:453–468, 1999.
- R.H. Brooks and A.T. Corey. Properties of porous media affecting fluid flow. *J. Irrig. Drain. Div Am. Soc. Civil Eng.*, 92(IR2):61–88, 1966.
- S. C. Brubaker, A. J. Jones, D. T. Lewis, and K. Frank. Soil properties associated with landscape position. *ssaj*, 57:235–239, 1993.
- A. T. Cahill, D. R. Nielsen, F. Ungaro, M. B. Parlange, and M. Mata. Combined spatial and kalman filter estimation of optimal soil hydraulic properties. *Wat. Resour. Res.*, 35(4):1079–1088, 1999.
- U. Callies, A. Rhodin, and D. P. Eppel. A case study on variational soil moisture analysis from atmospheric observations. *J. Hydrol.*, 212–213:95–108, 1998.
- J. C. Calvet, J. Noilhan, and P. Bessemoulin. Retrieving the root-zone soil moisture from surface soil moisture or temperature estimates: A feasibility study based on field measurements. *J. Appl. Meteorol.*, 37(4):371–386, 1998.
- W. J. Capehart and T. N. Carlson. Decoupling of surface and near-surface soil water content: A remote sensing perspective. *Water Resour. Res.*, 33(6):1383–1395, 1997.

- V.T. Chow, D.R. Maidment, and L.W. Mays. *Applied Hydrology*. McGraw-Hill, 1988.
- P. Courtier, J. Derber, R. Errico, J.-F. Louis, and T. Vukčević. Important literature on the use of adjoint, variational methods and the kalman filter in meteorology. *tellus*, 45A: 342–357, 1993.
- N.H. Crawford and R.K. Linsley. Digital simulation in hydrology: Stanford watershed model IV. Technical Report 39, Department of Civil Engineering, Stanford University, Stanford, California, 1966.
- R. Daley. *Atmospheric Data Analysis*. Cambridge University Press, Cambridge, UK, 1991.
- R. Daley. *Atmospheric data analysis*. Cambridge University Press, Cambridge, UK, 1991.
- H. C. Davies and R. E. Turner. Updating prediction models by dynamical relaxation: An examination of the technique. *Quart. J. R. Met. Soc.*, 103(436):225–245, 1977.
- A.P.J. de Roo, C.G. Wesseling, and W.P.A. van Deursen. Physically-based river basin modelling within a gis: The lisflood model. *Hydrol. Proc.*, 14:1981–1992, 2000.
- J. Dozier. Opportunities to improve hydrologic data. *Rev. Geophys.*, 30(4):315–331, 1992.
- T. Dunne. *Field studies of hillslope flow processes*, pages 227–293. John Wiley & Sons, 1978.
- P. S. Eagleson. *Dynamic Hydrology*. mcgraw, 1970.
- D. Entekhabi, H. Nakamura, and E. G. Njoku. Solving the inverse problem for soil moisture and temperature profiles by sequential assimilation of multifrequency remotely sensed observations. *IEEE Trans. Geosci. Remote Sensing*, 32(2):438–447, 1994.
- H. Nakamura Entekhabi, D. and E. G. Njoku. Solving the inverse problem for soil moisture and temperature profiles by sequential assimilation of multifrequency remotely sensed observations. *IEEE Trans. Geosci. Remote Sensing*, 32:438–448, 1994.
- G. Evensen. Using the extended Kalman filter with a multilayer quasi-geostrophic ocean model. *J. Geophys. Res.*, 97(C11):17,905–17,924, 1992.
- D. Entekhabi F. Castelli and E. Caporali. Estimation of surface heat flux and an index of soil moisture using adjoint-state surface energy balance. *Wat. Resour. Res.*, 35(10): 3115–3125, 1999.
- J. S. Famiglietti and E. F. Wood. Multiscale modeling of spatially variable water and energy balance processes. *Wat. Resour. Res.*, 30(11):3061–3078, 1994a.
- J.S. Famiglietti and E.F. Wood. Multiscale modeling of spatially variable water and energy balance processes. *Wat. Resour. Res.*, 30(11):3061–3078, 1994b.
- P. Finke, R. Hartwich, R. Dudal, J. Ibà nez, M. Jamagne, D. King, L. Montanarella, and N. Yassoglou, editors. *Georeferenced Soil Database For Europe - Manual of procedures. Version 1. European Soil Bureau Scientific Committee*. Number EUR 18092 EN. Office for Official Publications of the European Communities, Luxembourg, 1998.
- R.A. Freeze. *Hillslope hydrology*. Wiley-Interscience, 1978.

- J. F. Galantowicz, D. Entekhabi, and E. G. Njoku. Tests of sequential data assimilation for retrieving profile soil moisture and temperature from observed l-band radiobrightness. *Ieee Transactions On Geoscience and Remote Sensing*, 37(4):1860-1870, 1999.
- F. Gellens-Meulenberghs. Evapotranspiration and surface heat fluxes over belgium: outcome and perspectives. *Agronomie*, 20:857-868, 2000.
- F. Gellens-Meulenberghs, D. Gellens, R. Heylen, A. Manhout, and S. Vannitsem. Le radar météorologique: des échos mesurés à l'estimation des intensités des précipitations. *Ciel et Terre*, 108:56-59, 1992.
- K. P. Georgakakos and O. W. Baumer. Measurement and utilization of on-site soil moisture data. *J. Hydrol.*, 184:131-152, 1996a.
- K. P. Georgakakos and O. W. Baumer. Measurement and utilization of on-site soil moisture data. *J. Hydrol.*, 184:131-152, 1996b.
- M. Ghil and P. Malanotte-Rizzoli. Data assimilation in meteorology and oceanography. *Advances in Geophysics*, 33:141-266, 1991. Academic Press, San Diego, Calif.
- R. Grayson and G. Blöschl, editors. *Spatial Patterns in Catchment Hydrology: Observations and Modelling*. Cambridge University Press, Cambridge, UK, 2001.
- R. B. Grayson, A. W. Western, F. H. S. Chiew, and G. Blöschl. Preferred states in spatial soil moisture patterns: Local and nonlocal controls. *Water Resour. Res.*, 33(12): 2897-2908, 1997.
- P. D. Greminger, Y. K. Sud, and D. R. Nielsen. Spatial variability of field-measured soil-water characteristics. *sssaj*, 49:1075-1082, 1985.
- H. V. Gupta, S. Sorooshian, and P. O. Yapo. Toward improved calibration of hydrologic models: Multiple and noncommensurable measures of information. *Water Resour. Res.*, 34(4):751-763, 1998a.
- H. V. Gupta, S. Sorooshian, and P. O. Yapo. Toward improved calibration of hydrological models: Multiple and noncommensurable measures of information. *Wat. Resour. Res.*, 34(4):751-763, 1998b.
- A. B. Hairston and D.F. Grigal. Topographic influences on soils and trees within single mapping units on a sandy outwash landscape. *For. Ecol. Man.*, 43:35-45, 1991.
- M. Hanke and P. C. Hansen. Regularization methods for large-scale problems. *Surv. Math. Ind.*, 3:253-315, 1993.
- A. Y. Hanna, P. W. Harlan, and D. T. Lewis. Soil available water as influenced by landsape position and aspect. *Agron. J.*, 74:999-1004, 1982.
- P. C. Hansen. Analysis of discrete ill-posed problems by means of the l-curve. *SIAM Review*, 34:561-580, 1992.
- P. C. Hansen. *Rank-deficient and discrete ill-posed problems: numerical aspects of linear inversion*. SIAM, Philadelphia, 1998.
- R. Hoeben and P. A. Troch. Assimilation of active microwave observation data for soil moisture profile estimation. *Wat. Resour. Res.*, 36(10):2805-2819, 2000a.

- R. Hoeben and P. A. Troch. Assimilation of active microwave observation data for soil moisture profile estimation. *Wat. Resour. Res.*, 36(10):2805–2819, 2000b.
- R. Hoeben and P. A. Troch. Assimilation of active microwave observation data for soil moisture profile estimation. *Water Resour. Res.*, 36(10):2805–2819, 2000c.
- G. M. Hornberger and E. W. Boyer. Recent advances in watershed modelling. In R. A. Pielke Sr. and R. M. Vogel, editors, *U.S. National Report to the International Union of Geodesy and Geophysics 1991–1994: Contributions in Hydrology*, pages 949–957. American Geophysical Union, Washington, DC, 1995.
- P. R. Houser, W. J. Shuttleworth, J. S. Famiglietti, H. V. Gupta, K. H. Syed, and D. C. Goodrich. Integration of soil moisture remote sensing and hydrologic modeling using data assimilation. *Water Resour. Res.*, 34(12):3405–3420, 1998a.
- P.R. Houser. *Remote-Sensing Soil Moisture Using Four-Dimensional Data Assimilation*. PhD thesis, University of Arizona, Dept. of Hydrology and Water Resources, 1996.
- P.R. Houser, W.J. Shuttleworth, J.S. Famiglietti, H.V. Gupta, K.H. Syed, and D.C. Goodrich. Integration of soil moisture remote sensing and hydrological modeling using data assimilation. *Wat. Resour. Res.*, 34(12):3405–3420, 1998b.
- P. L. Houtekamer and H. L. Mitchell. Data assimilation using an ensemble kalman filter technique. *Monthly Weather Review*, 126(3):796–811, 1997.
- Y. Hu, X. Gao, W. J. Shuttleworth, H. Gupta, J.-F. Mahfouf, and P. Viterbo. Soil-moisture nudging experiments with a single-column version of the ecmwf model. *Q. J. R. Meteorol. Soc.*, 125:1879–1902, 1999.
- Z. Hu and S. Islam. Prediction of ground temperature and soil moisture content by the force-restore method. *Wat. Resour. Res.*, 31:2531–2540, 1995.
- P. S. Huyakorn, S. D. Thomas, and B. M. Thompson. Techniques for making finite elements competitive in modeling flow in variably saturated porous media. *Water Resour. Res.*, 20(8):1099–1115, 1984.
- S. Islam and E. T. Engman. Why bother for 0.0001% of Earth's water? Challenges for soil moisture research. *Eos, Trans., Amer. Geophys. Union*, 77(43):420, 1996.
- T. J. Jackson. Soil moisture estimation using special satellite microwave/imager satellite data over a grassland region. *Water Resour. Res.*, 33(6):1475–1484, 1997.
- T. J. Jackson, T. J. Schmugge, and E. T. Engman. Remote sensing applications to hydrology: soil moisture. *hsj*, 41:517–530, 1996.
- T. A. Johansen. On tikhonov regularization, bias and variance in nonlinear system identification. *Automatica*, 33(3):441–446, 1997.
- G. G. Katul, O. Wendroth, M. B. Parlange, C. E. Puente, M. V. Folegatti, and D. R. Nielsen. Estimation of in situ hydraulic conductivity function from nonlinear filtering theory. *Wat. Resour. Res.*, 29(4):1063–1070, 1993.
- V. Klemeš. Dilettantism in hydrology: Transition or destiny? *Wat. Resour. Res.*, 22(9): 177s–188s, 1986.

- O. Klepper and D. I. Rouse. A procedure to reduce parameter uncertainty for complex models by comparison with a real system output illustrated on a potato growth model. *Agricultural Systems*, 36:375-395, 1991.
- K. G. Kostov and T. J. Jackson. Estimating profile soil moisture from surface layer measurements - A review. In *Proc. of the International Society of Optical Engineering (SPIE)*, volume 1941, pages 125-136, Orlando, USA, 1993.
- W. R. Kreznor, K.R. Olson, W.L. Banwart, and D.L. Johnson. Soil, landscape and erosion relationships in a northwest illinois watershed. *ssaj*, 53:1763-1771, 1989.
- V. Lakshmi. A simple surface temperature assimilation scheme for use in land surface models. *Wat. Resour. Res.*, 36(12):3687-3700, 2000.
- W. Lehmann. *Anwendung geostatistischer verfahren auf die bodenfeuchte in ländlichen einzugsgebieten*. PhD thesis, Institut für Hydrologie und Wasserwirtschaft, Universität Karlsruhe, Karlsruhe95, 1995.
- L. B. Leopold and T. Maddock Jr. The hydraulic geometry of stream channels and some physiographic implications. Prof. Pap. 252, U.S. Geol. Surv., Washington, DC, 1953.
- G. Lindström, B. Johansson, M. Persson, M. Gardelin, and S. Bergström. Development and test of the distributed hbv-96 model. *J. Hydrol.*, 201:272-288, 1997.
- D. S. Mackay and L. E. Band. Extraction and representation of nested catchment areas from digital elevation models in lake-dominated topography. *Water Resour. Res.*, 34(4):897-901, 1998.
- J.-F. Mahfouf. Analysis of soil moisture from near-surface parameters: A feasibility study. *J. Appl. Meteorol.*, 30:1354-1365, 1991.
- D. Marks, J. Dozier, and J. Frew. Automated basin delineation from digital elevation data. *Geo-Processing*, 2:299-311, 1984.
- D. McLaughlin. Recent developments in hydrologic data assimilation. In R. A. Pielke Sr. and R. M. Vogel, editors, *U.S. National Report to International Union of Geodesy and Geophysics 1991-1994: Contributions in Hydrology*, pages 977-984. American Geophysical Union, Washington, DC, 1995.
- W. Menke. *Geophysical data analysis: discrete inverse theory*, volume 45 of *International Geophysics Series*. Academic Press, revised edition, 1989.
- P. C. D. Milly. Integrated remote sensing modeling of soil moisture: sampling frequency, response time, and accuracy of estimates. In M. E. Moss, editor, *Integrated Design of Hydrological Networks. Proceedings of the Budapest Symposium, July. 1986*. IAHS Publ. no. 158, pages 201-211, Wallingford, UK, 1986. IAHS.
- P. C. D. Milly and Z. J. Kabala. Integrated modeling and remote sensing of soil moisture. In A.I. Johnson, editor, *Hydrologic Applications of Space Technology. Proceedings of the Cocoa Beach Workshop, Florida, Aug. 1985*. IAHS Publ. no. 160, pages 331-339, Wallingford, UK, 1986a. IAHS.
- P. C. D. Milly and Z. J. Kabala. Integrated modeling and remote sensing of soil moisture. In *Hydrologic Applications of Space Technology (Proceedings of the Cocoa Beach Workshop, Florida, August 1985)*, pages 331-339. IAHS Publication No. 160, 1986b.

- H. L. Mitchell and P. L. Houtekamer. An adaptive ensemble kalman filter. *Mon. Wea. Rev.*, 128:416–433, 2000.
- M. Molnau and V.C. Bissel. A continuous frozen ground index for flood forecasting. In *Proc. 51st Ann. Meet. West. Snow Conf.*, pages 109–119, 1983.
- D. R. Montgomery and E. Foufoula-Georgiou. Channel network source representation using digital elevation models. *Water Resour. Res.*, 29(12):3925–3934, 1993.
- I. D. Moore, P. E. Gessler, G. A. Nielsen, and G. A. Peterson. Soil attribute prediction using terrain analysis. *ssaaj*, 57:443–452, 1993.
- I. D. Moore, R. B. Grayson, and A. R. Ladson. Digital terrain modeling: A review of hydrological, geomorphological, and biological applications. *hp*, 5:3–30, 1991.
- P. Naden, P. Broadhurst, N. Tauveron, and A. Walker. River routing at the continental scale: use of globally-available data and an a priori method of parameter estimation. *Hydrology and Earth System Sciences*, 3:109–124, 1999.
- S.P. Neuman. Saturated-unsaturated seepage by finite elements. *J. Hydraul. Div. Amer. Soc. Civil Eng.*, 99(HY12), 1973.
- R. W. Newton, J. L. Heilman, and C. H. M. Van Bavel. Integrating passive microwave measurements with a soil moisture/heat flow model. *Agricultural Water Management*, 7:379–389, 1983.
- R. T. Oosterhoff and C. Paniconi. Implementation of a catchment hydrologic model for the brisy subcatchment of the ourthe watershed, and generation of a dataset for a 240-day storm-interstorm sequence. technical report crs4-tech-rep-01/28. Technical report, Center for Advanced Studies, Research and Development in Sardinia, Cagliari, Italy, 2001.
- S. Orlandini and R. Rosso. Diffusion wave modeling of distributed catchment dynamics. *J. Hydrol. Engrg., ASCE*, 1(3):103–113, 1996.
- C. Paniconi, M. Marrocu, M. Putti, and M. Verbunt. Newtonian nudging for a Richards equation-based distributed hydrological model. *Advances in Water Resources*, 26(2): 161–178, 2003.
- M. B. Parlange, G. G. Katul, M. V. Folegatti, and D. R. Nielsen. Evaporation and the field scale soil water diffusivity function. *Wat. Resour. Res.*, 29(4):1279–1286, 1993.
- V. R. N. Pauwels, R. Hoeben, N. E. C. Verhoest, and F. P. De Troch. The importance of the spatial patterns of remotely sensed soil moisture in the improvement of discharge predictions for small-scale basins through data assimilation. *J. Hydrol.*, 251:88–102, 2001.
- C. Perrin, C. Michel, and V. Andréassian. Does a large number of parameters enhance model performance? Comparative assessment of common catchment model structures on 429 catchments. *J. Hydrol.*, 242:275–301, 2001.
- W. H. Press, S. A. Teukolsky, W. T. Vettering, and B. P. Flannery. *Numerical Recipes in C - The Art of Scientific Computing*. Cambridge University Press, Cambridge, 1992.

- M. Putti, C. Paniconi, and A. C. Bixio. A process-based numerical model of surface-subsurface water interactions. *Water Resour. Res.*, submitted, 2002.
- S. Sorooshian Q. Duan and V. K. Gupta. Effective and efficient global optimization for conceptual rainfall-runoff models. *Wat. Resour. Res.*, 28(4):1015-1031, 1992.
- J. Refsgaard and J. Knudsen. Operational validation and intercomparison of different types of hydrological models. *Wat. Resour. Res.*, 32:2189-2202, 1996.
- P. Reggiani, S. M. Hassanizadeh, M. Sivapalan, and W. G. Gray. A unifying framework for watershed thermodynamics: constitutive relationships. *Advances in Water Resources*, 23:15-39, 1999.
- P. Reggiani, M. Sivapalan, and S. M. Hassanizadeh. A unifying framework for watershed thermodynamics: balance equations for mass, momentum, energy and entropy, and the second law of thermodynamics. *Advances in Water Resources*, 22(4):367-398, 1998.
- P. Reggiani, M. Sivapalan, and S. M. Hassanizadeh. Conservation equations governing hillslope responses: Exploring the physical basis of water balance. *Wat. Resour. Res.*, 36(7):1845-1863, 2000.
- P. Reggiani, M. Sivapalan, S. M. Hassanizadeh, and W.G.Gray. Coupled equations for mass and momentum balance in a stream network: Theoretical derivation and computational experiments. *Proceedings of the Royal Society, Series A*, 45, no 2005:157-189, 2001.
- R. H. Reichle, D. Entekhabi, and D. B. McLaughlin. Downscaling of radio brightness measurements for soil moisture estimation: A four-dimensional variational data assimilation approach. *Water Resour. Res.*, 37(9):2353-2364, 2001.
- R. H. Reichle, D. Entekhabi, and D.B. McLaughlin. Downscaling of radiobrightness measurements for soil moisture estimation: A four-dimensional variational data assimilation approach. *Wat. Resour. Res.*, page submitted., 2000.
- R. H. Reichle, D. McLaughlin, and D. Entekhabi. Variational data assimilation of soil moisture and temperature from remote sensing observations. In F. Stauffer, W. Kinzelbach, K. Kovar, and E. Hoehn, editors, *ModelCARE 99 Proceedings of the International Conference on Calibration and Reliability in Groundwater Modeling: Coping with Uncertainty. IAHS Publication no. 265*, pages 353-359, Wallingford, UK, 1999. IAHS.
- K. Reif, F. Sonnemann, and R. Unbehauen. An EKF-Based nonlinear observer with a prescribed degree of stability. *Automatica*, 34(9):1119-1123, 1998.
- A. Rhodin, F. Kucharski, U. Callies, D. P. Eppel, and W. Wergen. Variational analysis of effective soil moisture from screen-level atmospheric parameters: Applications to a short-range weather forecast model. *Q. J. R. Meteorol. Soc.*, 125:2427-2448, 1999.
- A. Robock, K. Y. Vinnikov, G. Srinivasan, J. K. Entin, S. E. Hollinger, N. A. Speranskaya, S. Liu, and A. Namkhai. The Global Soil Moisture Data Bank. *Bull. Amer. Meteor. Soc.*, 81(6):1281-1299, 2000.
- E. Roulin, F. Gellens-Meulenberghs, and J. Gosset. Operational assessment of surface radiative fluxes over belgium by means of meteosat pdus and meteorological data. In *Proceedings of the Symposium " Progress in Environmental Research and Applications and workshop on Satellite Meteorology and Hydrology "*, Bâle, 4-7 September 1995, pages 409-416, Rotterdam, The Netherlands, 1995. Balkema.

- G. R. Schiffler and A. Bárdossy. Geostatistical assessment of space-time distributed data: application to soil moisture measurements in an experimental catchment. In P. Nachthebel and K. Kovar, editors, *Hydrological Basis for Ecologically Sound Management of Soil and Groundwater. Proceedings of the Vienna Symposium, August 1991*. IAHS Publ. no. 202, pages 279–287, Wallingford, UK, 1991. IAHS.
- T. J. Schmugge. Remote sensing of soil moisture. In M. G. Anderson and T. P. Burt, editors, *Hydrological Forecasting*, chapter 5, pages 101–124. John Wiley & Sons, 1985.
- J.M. Schuurmans, P.A. Troch, A.A. Veldhuizen, W. Bastiaansen, and M. Bierkens. Assimilating remotely sensed latent heat fluxes in a distributed hydrological model. *Advances in Water Resources*, 26(2):151–161, 2002.
- W.S. Shuttleworth and I.R. Calder. Has the priestley-taylor equation any relevance to forest evaporation? *J. Appl. Meteor.*, 18:639–646, 1979.
- R.E. Smith and J.Y. Parlange. A parameter-efficient hydrologic infiltration model. *Wat. Resour. Res.*, 14(3):533–538, 1978.
- J.D. Snell and M. Sivapalan. On the application of the meta-channel concept: Construction of the meta-channel hydraulic geometry for a natural catchment. *Hydrol. Proc.*, 9(5/6): 485–505, 1995.
- D. R. Stauffer and N. L. Seaman. Use of four-dimensional data assimilation in a limited-area mesoscale model. Part I: Experiments with synoptic-scale data. *Monthly Weather Review*, 118(6):1250–1277, 1990.
- D. R. Stauffer and N. L. Seaman. Multiscale four-dimensional data assimilation. *Journal of Applied Meteorology*, 33(3):416–434, 1994.
- W. Stol, D. I. Rouse, D. W. G. van Kraalingen, and O. Klepper. FSEOPT, a FORTRAN program for calibration and uncertainty analysis of simulation models. CABO-TT 24, CABO-DLO, Wageningen, the Netherlands, 1992.
- I. Supit, A.A. Hooijer, and C.A. van Diepen. System description of the wofost 6.0 crop simulation model implemented in cgms. volume 1: Theory and algorithms. Technical Report EUR 15956 EN, Joint Research Centre, 1994.
- A. Tarantola. *Inverse problem theory : methods for data fitting and model parameter estimation*. Elsevier, Amsterdam, 1987.
- D.G. Tarboton. A new method for the delineation of flow directions and upslope areas in grid digital elevation models. *Water Resour. Res.*, 33(2):309–319, 1997.
- A. N. Tikhonov and V. Y. Arsenin. *Solutions of Ill-Posed Problems*. Winston & Sons, Washington, D.C., 1977.
- U. S. Congress Office of Technology Assessment. *Remotely Sensed Data: Technology, Management, and Markets*. Number OTA-ISS-604. U.S. Government Printing Office, Washington, DC, 1994.
- U. S. National Research Council. *Four-Dimensional Model Assimilation of Data: A Strategy for the Earth System Sciences*. National Academy Press, Washington, DC, 1991.

- K. Unlu, D. R. Nielsen, J. W. Biggar, and F. Moroc. Statistical parameters characterizing the spatial variability of selected soil hydraulic properties. *ssaj*, 54:1537-1547, 1990.
- J. C. van Dam. *Field-scale water flow and solute transport. SWAP model concepts, parameter estimation and case studies*. PhD thesis, Wageningen University, 2000.
- M. Th. van Genuchten. A closed-form equation for predicting the hydraulic conductivity of unsaturated soils. *ssaj*, 44:892-898, 1980.
- E. E. van Loon and K. J. Keesman. Investigating scale-dependent models: the case of overland flow at the hillslope scale. *Wat. Resour. Res.*, 36(1):243-255, 2000.
- E. E. van Loon, K. J. Keesman, and L. Stroosnijder. Overland flow: identification and calibration of distributed models. *submitted to Water Resour. Res.*, 2001.
- E. E. van Loon and P. A. Troch. Tikhonov regularization as a tool for assimilating soil moisture data in distributed hydrological models. *Hydrol. Proc.*, 16(2):531-556, 2001.
- E. E. van Loon and P. A. Troch. Soil moisture data assimilation with a distributed hydrological model: a case study. *Hydrol. Proc.*, *in press.*, 2002.
- L. van Loon, P. J. H. Bultjes, and A. J. Segers. Data assimilation of ozone in the atmospheric transport chemistry model lotos. *Environmental Modelling and Software*, 15: 603-609, 2000.
- M.S. Viney and M. Sivapalan. A conceptual model of sediment transport: application to the avon river basin in western australia. *Hydrol. Proc.*, 13:727-743, 1999.
- J. von Hoyning-Huene. Die interzeption des niederschlags in landwirtschaftlichen pflanzenbeständen. Technical report, Arbeitsbericht Deutscher Verband für Wasserwirtschaft und Kulturbau DVWK, Braunschweig, 1981.
- J. P. Walker and P. R. Houser. A methodology for initializing soil moisture in a global climate model: Assimilation of near-surface soil moisture observations. *J. Geophys. Res.*, 106(D11):11761-11774, 2001.
- A. W. Western, G. Blöschl, and R. B. Grayson. Geostatistical characterisation of soil moisture patterns in the tarrawarra catchment. *J. Hydrol.*, 205:20-37, 1998.
- A. W. Western, G. Blöschl, and R. B. Grayson. Toward capturing hydrologically significant connectivity in spatial patterns. *Water Resour. Res.*, 37(1):83-97, 2001.
- A. W. Western and R. B. Grayson. The Tarrawarra data set: Soil moisture patterns, soil characteristics, and hydrological flux measurements. *Water Resour. Res.*, 34(10): 2765-2768, 1998.
- A. W. Western, R. B. Grayson, G. Blöschl, G. R. Willgoose, and T. A. McMahon. Observed spatial organization of soil moisture and its relation to terrain indices. *Wat. Resour. Res.*, 35(3):797-810, 1999.
- J.H.M. Wösten, A. Lilly, A. Nemes, and C. Le Bas. Using existing soil data to derive hydraulic parameters for simulation models in environmental studies and in land use planning. Report 156, Winand Staring Centre, Wageningen, The Netherlands, 1998.

- C. Wunsch and J. F. Minster. Methods for box models and ocean circulation tracers: Mathematical programming and non-linear inverse theory. *Journal of Geophysical Research*, 87:5647-5662, 1982.
- Y. Xue, P. J. Sellers, J. L. Kinter, and J. Shukla. A simplified biosphere model for global climate studies. *J. Climate*, 4:345-364, 1991.
- G.T. Yeh. On computation of darcian velocity and mass balance in the finite element modeling of groundwater flow. *Wat. Resour. Res.*, 17(5), 1981.
- T. C. J. Yeh, L. W. Gelhar, and P. J. Wierenga. Observations of spatial variability of soil water pressure in a field soil. *Soil Sci.*, 142:7-12, 1986.
- P.C. Young and K.J. Beven. Data-based mechanistic modelling and the rainfall-flow non-linearity. *Environmetrics*, 5:335-363, 1994.
- Z. Yu, T. N. Carlson, E. J. Barron, and F. W. Schwartz. On evaluating the spatial-temporal variation of soil moisture in the Susquehanna River Basin. *Water Resour. Res.*, 37(5):1313-1326, 2001.
- H. Zaradny. *Groundwater Flow in Saturated and Unsaturated Soil*. A.A. Balkema, Rotterdam, Netherlands, 1993.
- Y. B. Zecharias and W. Brutsaert. Recession characteristics of groundwater outflow and base flow from mountainous watersheds. *Water Resour. Res.*, 24(10):1651-1658, 1988.
- O.C. Zienkiewicz. *The finite element method in engineering science*. McGraw-Hill, London, 1971.

**Inlichtingen zijn verkrijgbaar bij het
secretariaat:**

Wageningen Universiteit
Departement Omgevingswetenschappen
Sectie Waterhuishouding
Nieuwe Kanaal 11
6709 PA Wageningen
telefoon : 0317 - 482778
telefax : 0317 - 484885

**For information please contact the
secretariat:**

Wageningen University
Department of Environmental Sciences
Sub-department Water Resources
Nieuwe Kanaal 11
6709 PA Wageningen
The Netherlands
telephone : +31 - (317) - 482778
telefax : +31 - (317) - 484885

Internet: www.dow.wau.nl/whh

EERDER VERSCHENEN RAPPORTEN/PREVIOUS REPORTS

Nr	Auteur(s) + titel/author(s) + title	Prijs/Price (€)
51.	Boiten, W., A. Dommerholt en M. Soet, 1995. Handboek debietmeten in open waterlopen.	*
52.	Boiten, W., 1995. Het opstellen van de afvoerkrommen van klepstuwen.	8,00
53.	Boiten, W., 1995. Afvoerrelaties klepstuwen op de sarnenvloeiing van de Veengoot en de Van Heeckerenbeek.	*
54.	Benning, R.G., 1995. Towards a new lumped parameterization at catchment scale.	11,00
55.	Bastiaansen, C.J.M., 1995. Lui River Valley Model, and some of its applications.	10,00
56.	He, Q., K.D.W. Nandalal, J.J.K.M. Bogardi and D. Milutin, 1995. Application of stochastic dynamic programming models in optimization of reservoir operations: A study of algorithmic aspects.	15,00
57.	Lanen, H.A.J. van, B. van de Weerd, R. Dijkma, H.J. ten Dam en G. Bier, 1995. Hydrogeologie van het stroomgebied van de Noor en de effecten van grondwateronttrekkingen aan de westrand van het Plateau van Margraten. Basisrapport.	*
58.	Lanen, H.A.J. van, R. Dijkma en B. van de Weerd, 1995. De effecten van grondwateronttrekkingen aan de westrand van het Plateau van Margraten op de hydrogeologie van het Noordal. Samenvattend rapport.	*
59.	Bogardi, J.J.K.M., B.A.H.V. Brorens, M.D.U.P. Kularathna, D. Milutin en K.D.W. Nandalal, 1995. Long-term assessment of a multi-unit reservoir system operation: the <i>ShellDP</i> programme package manual.	25,50
60.	Dommerholt, A., 1995. Afvoerrelatie meetoverlaat Stermerdinkbrug.	*
61.	Dijkma, R., J.H. Bouma en H.F. Gertsen, 1995. Proefproject verdroging Duurswouderheide. Hoe effectief zijn de beheersmaatregelen?	*
62.	Verburg, P.H., 1995. De relatie tussen de vochttoestand van de bodem en de vochtindicatie van de vegetatie. Een nadere bepaling van de grens tussen 'vochtig' en 'droog' binnen het ecotopensysteem.	12,00

* niet te koop bij de sectie, alleen ter inzage/not for sale at the department, for inspection only.

Nr	Auteur(s) + titel/author(s) + title	Prijs/Price (€)
63.	Denecke, H.W., 1995. Voorlopige evaluatie drainagesysteem militaire oefenterrein Marnewaard.	*
64.	Arts, M.P.T. en W. Boiten, 1995. Meetnet voor afvoermetingen in de Renkumse Beken.	*
65.	Arts, M.P.T. en R. Dijkma, 1995. Morra park, voorlopige resultaten grond- en oppervlaktewater meetnet.	*
66.	Akker, M.F.A. van den en B.J.H. van de Wiel, 1996. Hoogwatervoor- spellingen voor de Rijn bij Lobith met hybride methoden. De hydrologische Muskingum methode uitgebreid met lineaire en niet- lineaire updatingstechnieken.	18,00
67.	Soet, M., P. Petrovic, J.N.M. Stricker, W. Meijninger, A. van Schaik en T. Lapsansky, 1996. Water budget of maize on heavy clay in a continental climate: field experiment and model simulation. Final report of the project "Evaporation estimation comparison".	*
68.	Torfs, P.J.J.F. en H. Middelkoop, 1996. Analysis of discharge series from the Rhine basin at different levels of aggregation.	*
69.	Dijkma, R. (ed.), 1996. International excursion Hydrogeology - SLOVAKIA, september 8-15, 1996. Excursion report.	,-
70.	Oevelen, P.J. van and I.H. Woodhouse, 1996. NOPEX/FOREST-DYNA- MO ground data collection and data analysis report.	11,50
71.	Dam, J.C. van, J. Huygen, J.G. Wesseling, R.A. Feddes, P. Kabat, P.E.V. van Walsum, P. Groenendijk and C.A. van Diepen, 1997. Theory of SWAP version 2.0. Simulation of water flow, solute transport and plant growth in the Soil-Water-Atmosphere-Plant environment.	20,50
72.	Warmerdam, P.M.M. (ed.) en M. Quist (ed.), 1997. Verslag van de buitenlandse excursie Hydrologie naar België, 1 t/m 5 september 1997.	,-
73.	Moors, E.J., J.N.M. Stricker and G.D. van den Abeele, 1998. Evapo- transpiration of cut over bog covered by <i>Molinia Caerulea</i> .	8,50
74.	Eertwegh, G.A.P.H. van den en C.R. Meinardi, 1998. Water- en nutriën- tenhuishouding van het stroomgebied van de Hupselse beek.	*
75.	Eertwegh, G.A.P.H. van den, J.R. Hoekstra en C.R. Meinardi, 1998. Praktijkproef nutriëntenbalans: nutriëntenbelasting in oppervlakte- water via drainage van akkerbouwpercelen op zeeklei.	*
76.	Boiten, W. 1998. Levering drie Rossum-stuwen Schouwen West.	*
77.	Oevelen, P.J. van, M.A.M. Vissers and I.H. Woodhouse, 1998. RESMEDES Spain 1996. Ground data collection and analysis report.	12,50
78.	Dijkma, R., H.A.J. van Lanen, W.J. Ackerman en H.F. Gertsen, 1998. De afvoer van de Noor (Zuid-Limburg). Periode 1992 - 1997.	*
79.	Soet, M., J.N.M. Stricker, P. Droogers and J. Esenbrink, 1998. EFEDA- Spain and HAPEX-Sahel. A further analysis of data.	*
80.	Wojcik, R., P.J.J.F. Torfs and S. Ignar, 1998. Sensitivity of output covariance to non-second order properties of the input in the case of unsaturated zone models.	11,50

Nr	Auteur(s) + titel/author(s) + title	Prijs/Price (€)
81.	Kroes, J.G., J.C. van Dam, J. Huygen en R.W. Vervoort, 1998. Users Guide of SWAP 2.0. Simulation of water flow, solute transport and plant growth in the Soil-Water-Atmosphere-Plant environment.	16,00
82.	Boiten, W., 1998. Levering Rossum-stuw. Polder de Noordplas.	*
83.	Boiten, W., 1998. IJking Friese spuisluisen. Friese Sluis te Zoutkamp, Suatiesluis Nes, Ameland.	*
84.	Dijksma, R. en J.F. Borsje, 1999. Het Morra Park. Water als lust of last.	*
85.	Klok, E.J. and K.P. Roelofsma, 1999. Modelling of glacier and snow melt processes within the hydrological catchment model WaSiM-ETH.	12,50
86.	Boiten, W., 1999. Ontwerp Hobrad-overlaat. Project Evertsekoog, Texel.	*
87.	Warmerdam, P.M.M. en P.J.J.F. Torfs, 1998. Verslag van de buitenlandse excursie Hydrologie naar de landendriehoek Frankrijk, Duitsland en Luxemburg. 30 augustus tot 5 september 1998.	-,-
88.	Wójcik, R., 1999. Non-linear stochastic methods for discharge prediction.	27,00
89.	Kalma, J.D., R.A. Feddes, G. Boulet, M.F. McCabe and S.W. Franks, 1999. Towards effective land surface parameters for use with SVAT models: the use of similarity scaling and inversion techniques.	6,50
90.	Kalma, J.D., S.W. Franks, B.J.J.M. van den Hurk, M.F. McCabe and R.A. Feddes, 1999. Estimating large scale land surface fluxes: the use of remote sensing data with SVAT and NWP models	7,00
91.	Ovaa, B.P.S.A., 1999. Gebiedscontracten. Een nieuwe kijk op sturing van regionaal landgebruik en waterbeheer.	11,50
92.	Boiten, W., 2000. Report Overseas Study Tour Rhine Basin, 7-21 November 1999. Hydrology Project India.	*
93.	Boiten, W., 2000. Debietmeetstations in het gebied Halkenbroek.	*
94.	Schaaf, S. van der en P.M.M. Warmerdam, 2000. Herstel van het watersysteem in het bebouwde gebied van Wageningen. Haalbaarheidsonderzoek.	14,00
95.	Dommerholt, A., 2000. Meetopstelling voor aan- en afvoer van water in de Vlietpolder bij Hoogmade. Onderdeel project "Water- en Nutriëntenhuishouding Veenweideproject".	*
96.	Boiten, W. 2000. Debietmeetstations op de Schipbeek; Kloosterstuw te Deventer, inlaatgemaal Twentekanaal te Markelo.	*
97.	Dijksma, R. (ed.), 2000. International excursion Hydrogeology - Poland. 3-10 September 2000.	-,-
98.	Dijksma, R. en H.A.J. van Lanen, 2001. De afvoer van de Noor (Zuid-Limburg). Periode 1992-2000.	*
99.	Boiten, W., 2001. Rehabilitation of the Busongo Reservoir near Sirigu, Bolgatanga, Ghana.	8,00
100.	Uijlenhoet, R., M.J.M. de Wit, P.M.M. Warmerdam and P.J.J.F. Torfs, 2001. Statistical Analysis of Daily Discharge Data of the River Meuse and its Tributaries (1968-1998): Assessment of Drought Sensitivity.	7,00

Nr	Auteur(s) + titel/author(s) + title	Prijs/Price (€)
101.	Loon, E.E. van and P.A. Troch, 2001. Book of Abstracts. International Workshop on Catchment scale Hydrologic Modeling and Data Assimilation. Wageningen, September 3-5, 2001.	14,00
102.	Boiten, W., 2001. Inrichting debietmeetnet Vallei en Eem. Vooronderzoek, richtingsadvies en ontwerp van een lange overlaat achter de Grebbesluis.	*
103.	Dommerholt, A., W. Boiten en M.R. Hoffmann, 2001. Meestopstelling voor hoogwater peilregistratie. Afwijking tussen gemeten en werkelijke waterhoogte.	*
104.	Wit, M.J.M. de (ed.), P. Warmerdam, P. Torfs, R. Uijlenhoet, E. Roulin, A. Cheymol, W. van Deursen, P. van Walsum, M. Ververs, J. Kwadijk and H. Buiteveld, 2001. Effect of climate change on the hydrology of the river Meuse.	18,00
105.	Loon, A. van (ed.) en G. v.d. Meijden (ed.), 2001. Hydrologie excursie – Duitsland. 2 – 7 september 2001.	,-
106.	Dommerholt, A., W. Boiten en R.T. Oosterhoff, 2002. Modelonderzoek van de Venturi-meetdoorlaat. Een vispasseerbare debietmeet-inrichting.	*
107.	Breukels, M., B.P.S.A. Ovaa en P.M.M. Warmerdam, 2002. Water als ordenend principe. Interviews.	10,00
108.	Boiten, W., 2002. Inlaatduiker Oosterhout. Herijking van de afvoerrelatie	*
109.	Uijlenhoet, R., 2002. Parameterization of Rainfall Microstructure for Radar Meteorology and Hydrology.	20,00
110.	Uijlenhoet, R., 2002. Development of a Stochastic Model of Rainfall for Radar Hydrology.	8,00
111.	Boiten, W., R. Velnor, R. Dijkma en J.W. Kole, 2002. Waterbalansen voor de polders Atsjetille, Edens, Rodenburg en Grouster Laagland. Rapportage over de gegevensverwerking en de waterbalansen van de meetplichtige gebieden binnen Wetterskip Marne-Middelsee voor 1997, 1998 en 1999.	*
112.	Käfer, J., 2002. Het gewicht van bedekking. Optimalisatie van de berekening van gewogen gemiddelde indicatiewaarden voor vegetatieopnamen.	10,00
113.	Boiten, W., 2002. Visgeleiding Stedelijke Roer te Roermond.	*
114.	Boiten, W. en A. Dommerholt, 2003. IJking Aflaatwerk naar het Twenthekanaal in Lochem.	*
115.	Loon, E.E. van and P.A. Troch (eds.), 2003. Final Report of the DAUFIN project.	23,00