Swedish University of Agricultural Sciences

Improvement of the Watbal model for till soil hydrology with application to heavy metal transport

Master's thesis, 30 credits

by

Claudia Teutschbein

Supervisors:
Prof. Lars Lundin, SLU Uppsala, Sweden
Lage Bringmark, SLU Uppsala, Sweden
Mike Starr, University Helsinki, Finland

Department of Environmental Assessment, SLU
P.O. Box 7050, S-750 07 Uppsala
Report 2008:9

Improvement of the WATBAL model for till soil hydrology with application to heavy metal transport

Master's thesis, 30 credits

by

Claudia Teutschbein

May 2008

ISSN 1403-977X

Department of Environmental Assessment, SLU
P.O. Box 7050
S-750 07 Uppsala
www.ma.slu.se

Printed: 25 copies, Department of Environmental Assessment, SLU, 2008

Abstract

In environmental research, the movement, accumulation and dissolving of pollutants are often of great interest. An important field of investigation is, for instance, the transport of heavy metals - such as cadmium (Cd), lead (Pb) and mercury (Hg) - through the soil. To understand the behaviour and estimate the quantity of those substances in certain subterranean parts, the chemical concentrations as well as hydrological fluxes need to be known. Unfortunately, at this point modern science approaches its limits. It is complicated to accomplish accurate and reliable measurements of hydrological fluxes at a small scale, i.e. percolating soil water.

This problem can be solved by applying a soil hydrological model. WATBAL is such a soil water balance model that can help to estimate the monthly soil water flux values within the rooting zone. The model was developed by Mike Starr (University Helsinki, Finland), but was not adequately tested and verified in the past, especially not applied to Swedish till soils.

In this study, the WATBAL model was intensely examined and slightly improved by changing a few parameters. It was applied to several Swedish sites to investigate its grade of accuracy and plausibility. The results showed that WATBAL is definitely a reasonable soil water model. The numerical statistical analysis resulted in a moderate grade of accuracy with R^2 values between 0.13 and 0.58 combined with MAPE values from 0.02 to 0.21. However, the graphical analysis of the model results revealed a good model fit, so that WATBAL in general is rated as acceptable. The overall average percentage error was estimated as $8.91\% \pm 1.08\%$ (95% confidence interval).

Furthermore, the obtained results from the WATBAL simulations were used to investigate the heavy metal transport within the soil. With available chemical data and the calculated small scale water fluxes, the mass balance of cadmium, lead, mercury, zinc and copper could be estimated. Even though chemical values were rather incomplete, results were obtained and showed that between 68.37 % (cadmium, Kindla) and 97.28 % (copper, Kindla) of the incoming heavy metals remained in the soil. Litterfall, throughfall and corresponding absolute accumulation values were much higher in Kindla than in Gammtratten. Besides, a seasonal dependency of monthly absolute accumulation was detected, with much greater values for April to October than for November to March. Based on long-term considerations, the amount of accumulated heavy metals spread over the whole catchment can add up to several kilograms.

Keywords: WATBAL, soil water model, till soils, hydrological fluxes, rooting zone, heavy metal transport, integrated monitoring, Sweden, Kindla, Gammtratten, Kloten.

Acknowledgements

This master thesis could not have been written without my main supervisor Prof. Lars Lundin (SLU Uppsala, Sweden) as well as my assistant supervisors Lage Bringmark (SLU Uppsala, Sweden) and Mike Starr (University Helsinki, Finland). They supported and challenged me, helped me with organisation, administration and guided me through the thesis process, never accepting less than my best efforts. I thank them all.

I would also like to thank Prof. Kevin Bishop for introducing me to his research group and letting me participate in great scientific discussions.

Furthermore, I would like to acknowledge and extend my gratitude to the following persons who have made the completion of this master thesis possible: María for always cheering me up and having so much patience with me. Thank you for some sunshine in the middle of a dark Swedish winter! Juliane, Michael and Charlotte for giving me strength to go on and making life in Uppsala even more pleasant. My proofreaders for reviewing the manuscript.

And last but not least, many thanks to my family: My parents Doris and Gerald for supporting me mentally and financially in every life situation. My sisters Janka and Kristina with their families for always believing in me and for their encouragement. My grandparents for keeping me informed about news from home. There are no words to say thank you for how much you mean to me!

Contents

List of Figures and Tables

List of Figures

List of Tables

List of Symbols

Symbol	Unit	Description
AET	[mm]	actual evapotranspiration
C	[%]	cloudiness
c	$[g \cdot L^{-1}]$	concentration
$clay$	[%]	clay content
$Cov(x, y)$	[-]	covariance
D	[mm]	drainage
ΔSM	[mm]	changes in soil moisture content
d_{snow}	[-]	no. of snow-covered days
ΔSOG	[mm]	changes in snow on the ground
ET	[mm]	evapotranspiration
ET_0	[mm]	reference potential evapotranspiration
ET_c	[mm]	forest potential evapotranspiration
$E(x)$	[-]	expected value
F	[mm]	infiltration during snowmelt
h	[kPa]	absolute capillary pressure
i	[-]	control variable
$infil\_coeff$	[-]	infiltration coefficient
LH_vap	[mm]	latent heat of evapotranspiration from soil
LH_vap_Ep	[mm]	latent heat of evapotranspiration from reference crop
m	[g]	mass
$\dot{m}$	$[g \cdot month^{-1}]$	mass rate flow
MAE	[-]	mean absolute error
$MAPE$	[-]	mean absolute percentage error
n	[-]	samlpe size
P	[mm]	precipitation
PE	[%]	percentage error
PET	[mm]	potential evapotranspiration

Symbol	Unit	Description
PS_{melt}	[mm]	potential snowmelt
PWL	[mm]	potential soil water loss
Q	$[L \cdot month^{-1}]$	volumetric flow rate
R	[mm water]	runoff
R^2	[-]	R square value
R_g	$[MJ \cdot m^{-2} \cdot month^{-1}]$	global radiation
$\rho_{x,y}$	[-]	correlation
$RMSE$	[-]	root mean squared error
R_{surf}	[mm]	surface runoff
R_{unsat}	[mm]	unsaturated flow
$sand$	[%]	sand content
SE	[-]	standard error
σ	[-]	deviation
SM	[%]	soil moisture
S_{melt}	[mm]	snowmelt
SOG	[mm water]	snow on the ground
T	[°C]	temperature
$T_{max,long}$	[°C]	long-term (30-years average) air maximum temperatures for the warmest month of the year (July)
$T_{min,long}$	[°C]	long-term (30-years average) air minimum temperatures for the warmest month of the year (July)
θ	[%]	soil moisture
θ_a	[mm]	available water capacity (AWC)
θ_{FC}	[%]	soil moisture at field capacity (equals water holding capacity WHC)
θ_{PWP}	[%]	soil moisture at permanent wilting point
V	[L]	volume
y_i	[-]	observed value
$\hat{y}_i$	[-]	modelled value
$\bar{y}_i$	[-]	mean of observed values

List of Acronyms

Symbol	Description
AWC	available water capacity
FC	field capacitiy
GA	Gammtratten site
ICP	International Cooperative Programme
ICP IM	International Cooperative Programme on Integrated Monitoring of Air Pollution Effects on Ecosystems
ID	identification
IM	Integrated Monitoring
IVL	Svenska Miljöinstitut IVL (Swedish Environmental Research Institute)
KI	Kindla site
KL	Kloten site
LRTAP	Convention on Long-Range Transboundary Air Pollution
PMK	Programme for Monitoring of Environmental Quality
PWP	permanent wilting point
SGU	Sveriges Geologiska Undersökning (Geological Survey of Sweden)
SLU	Sveriges Lantbruksuniversitet (Swedish University of Agricultural Sciences)
SM	soil moisture
SOG	snow on the ground
TDR	time domain reflectometer (electronic device for soil moisture measurements)
UN ECE	UN Economic Commission for Europe
WGE	Working Group on Effects of the Convention on Long-range Transboundary Air Pollution
WHC	water holding capacity (equals FC)

Chapter 1

Introduction

Integrated Monitoring (IM) is a useful control method applied in natural sciences and environmental protection that makes it possible to observe positive and negative changes in ecosystems. Basically, IM implies continuous observation of certain sites, including measurements of biological, chemical and physical parameters over time. According to the *Working Group on Effects of the Convention on Long-range Transboundary Air Pollution* (WGE) [Kleemola and Forsius, 2006], these investigations allow to obtain an overall picture of a certain ecosystem, to understand the interdisciplinary connections of its functionality, to quickly recognize changes of specific parameters and to react or go against these changes, if necessary. Thus, IM undertakes the task of assisting communities or countries with coming to legislative and executive decisions in terms of anticipatory environmental precautions, pollution control and sustainable treatment of ecosystems. Practically, IM is accomplished by creating several subprogrammes according to different subject areas. In general, these subprogrammes are either linked by using the same parameters or the same gauging stations [Kleemola and Forsius, 2006].

Within the framework of the *UN Economic Commission for Europe* (UN ECE), in 1979 the *Convention on Long-range Transboundary Air Pollution* (LRTAP) was adopted [SLU Uppsala, 2007]. As a direct result, several *International Cooperative Programmes* (ICPs) were set up for the purpose of indicating the effects of emission reductions on different biotopes and man-made structures.

In Sweden IM has been applied since 1981 as part of the national *Programme for Monitoring of Environmental Quality* (PMK). For many years, various data - for instance parameters of soil chemistry, deposition, runoff and soil water, groundwater, soil biology, vegetation and fauna - have been collected from more than fifteen sites, representing different physical geographical regions of Sweden. However, a new national programme was started in 1995 and four new, smaller sites were established. In contrast to the former sites of great diversity, these sites are hydrologically well defined, homogeneous, dominated by coniferous forest, free of lakes or large mires and they represent different positions in the air pollution gradient of the country. Since that time, the new IM programme has focused more on the hydrological transport of pollutants through the catchment and the biological impact on vegetation and soil.

Due to the early implementation of IM, Sweden became a leading country of the *International Cooperative Programme on Integrated Monitoring* (ICP IM) *of Air Pollution Effects on Ecosystems* in Europe and throughout the world along with Finland [SLU Uppsala, 2007]. Nowadays, the ICP IM network includes 50 sites in 19, mainly European, countries [Kleemola and Forsius, 2006]. Two of the four Swedish IM sites are discussed in this paper in terms of their geological, hydrological and chemical characteristics.

Research within the ICP IM network includes the use of certain computer-based models. From various sets of parameters and initial conditions obtained from earlier IM investigations, different scenarios can be simulated to predict the behaviour of ecosystems, such as groundwater flow or biogeochemical processes.

Since one of the main interests of the IM programme is the hydrological transport of pollutants, fluxes of chemical substances need to be known. Those can be determined with the help of hydrological fluxes and concentrations of the substances. Most of the required chemical data is available within ICP IM subprogrammes, but unfortunately the hydrological fluxes are only identified as precipitation and runoff, because measurements of water flow at a smaller scale inside the soil are much more complicated [Starr, 1999]. However, this problem can be solved by applying a soil hydrological model that makes the estimation of soil water fluxes possible.

Most of the water movement processes in soils are well-known but very intricate so that process models usually require complex input parameters, which might be difficult to obtain. Thus, any applied soil water model should be simple to ensure that the required input data is generally available. WATBAL (developed by Mike Starr, University Helsinki) is such a relatively simple soil water balance model that can help to provide monthly soil water information. It is a physical profile model that simulates variables like monthly evapotranspiration, drainage outflow and change in soil water storage with a set of input parameters defined from commonly available data.

Therefore, WATBAL was used to study monthly soil water balances of several Swedish sites presented in this paper. The obtained results were used to investigate heavy metal transport in glacial till soils. With the gained information about hydrologic fluxes and the existing chemical data, time-dependent transport and accumulation of heavy metals within the soil were ascertained. Such an analysis is important for environmental protection issues, because heavy metals, e.g. cadmium, lead and mercury, can not only cause hazardous conditions for flora and fauna in an ecosystem, but are a danger for human health as well.

Chapter 2

Methodology

2.1 Study sites

This project first intended to investigate the four Swedish IM sites Aneboda, Gammtratten, Gårdsjön and Kindla [SLU Uppsala, 2007]. However, required soil hydrological data were not available for Gårdsjön (SE04) and Aneboda (SE14). Therefore, this study focused on the remaining sites **Gammtratten** (SE16) and **Kindla** (SE15).

To compensate the lack of data of the neglected sites and to receive more extensive as well as significant results, **Kloten**, which is another site close to Kindla, was selected for further research. It is not part of the IM programme but had earlier been intensively studied by Lars Lundin [Lundin, 1982] so that adequate data already existed.

Thus, a total of three measurement sites - Gammtratten, Kindla and Kloten site - were investigated (Fig. 2.1).

Fig. 2.1. Map of the three study sites in Sweden

2.1.1 Gammtratten (GA)

Gammtratten is situated in northern Sweden, 100 km straight to the west of Umeå and ca. 50 km southeast of Åsele in the Västernorrland County. The geographic coordinates are N63°51' and E18°07', the total catchment size mounts up to 0.45 km^2 and the biome is stated as middle boreal [SLU Uppsala, 2007]. The vegetation period is about 145 days and the dominant vegetation type is defined as Norway spruce (*Picea abies*) - blueberry (*Vaccinium myrtillus*) forest.

2.1.2 Kindla (KI)

Kindla is a site in the Örebro County in central Sweden, between Hällefors, Kopparberg and Lindesberg. Longitude and latitude are N59°45' and E14°54'. The catchment size is 0.191 km^2 and the biome is defined as southern boreal [SLU Uppsala, 2007]. In this region the vegetation period lasts 180 days and prevalent type of vegetation is Norway spruce (*Picea abies*) - blueberry (*Vaccinium myrtillus*) forest.

2.1.3 Kloten (KL)

Kloten is located relatively close to Kindla in central Sweden, in the northern part of the Örebro County. The geographic coordinates are N59°54' and E15°15', the catchment has a total size of 0.186 km^2 [Lundin, 1982] and the biome is given as southern boreal. The duration of vegetation period is 185 days and main vegetation types are alternating Norway spruce (*Picea abies*) and pine (*Pinus sylvestris*) forest.

For all investigation areas, Podsol is the dominant soil type and the bedrock is mainly composed of granite.

Long-term climate parameters are similar for sites Kindla and Kloten [Raab and Vedin, 1995], but are noticeably different for Gammtratten site (Table 2.1).

Table 2.1. Long-term climate parameters for Gammtratten (GA), Kindla (KI) and Kloten (KL)

Parameter	Symbol	Unit	GA site	KI site	KL site
annual mean temperature	T	$[^\circ C]$	1.2	4.0	4.0
annual precipitation	P	$[mm]$	750	850	850
annual evapotranspiration	ET	$[mm]$	340	440	430
annual mean cloudiness	C	[%]	66	65	65
no. of snow-covered days	d_{snow}	$[-]$	175	130	140
annual global radiation	R_g	$[\frac{MJ}{m^2}]$	3,132	3,474	3,456
max. temperature July	$T_{max,long}$	$[^\circ C]$	18.6	20.7	20.7
min. temperature July	$T_{max,long}$	$[^\circ C]$	8.0	9.8	9.8

2.2 Specification of Sampling Points

2.2.1 Arrangement of Test Points

At each IM site, measurements of soil parameters and water content were carried out at three different locations - with sample number 1 belonging to the test point with highest elevation, number 2 to a medium elevation and number 3 to the lowest point of the slope (Fig. 2.2). Available data sets were not complete, so that only test points 1 and 2 were used for sites Kindla and Gammtratten. In the Kloten catchment the installation of measurement stations was not slope depending and thus, just one test point was used.

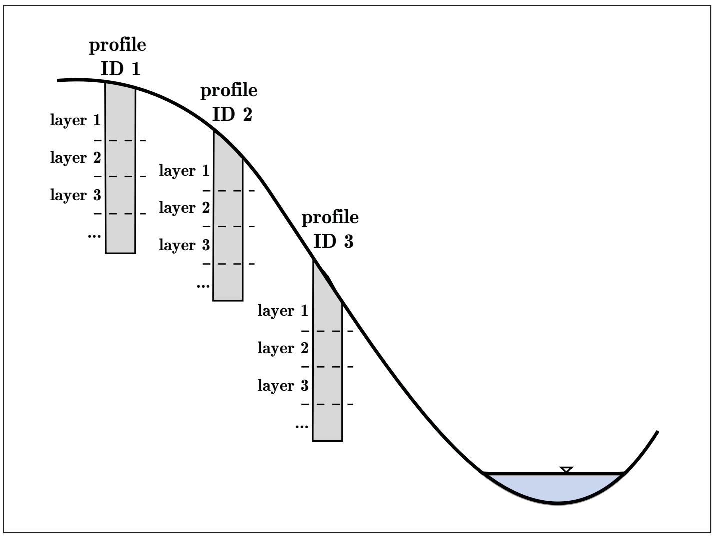


Fig. 2.2. Layout of the sampling positions at each site.

Complete data were available for depths up to $55cm$ below the humus layer. In sites Gammtratten and Kloten, soil moisture studies were carried out for five layers, in Kindla site only four layers were classified (Table 2.2).

Table 2.2. Site-specific layer arrangement (depth below humus layer) for Gammtratten (GA), Kindla (KI) and Kloten (KL)

Layer	GA site		KI site		KL site
	ID 1	**ID 2**	**ID 1**	**ID 2**	**ID 1**
layer 1	5 cm	5 cm	5 cm	5 cm	5 cm
layer 2	20 cm	13 cm	20 cm	10 cm	10 cm
layer 3	30 cm	29 cm	30 cm	30 cm	20 cm
layer 4	40 cm	40 cm	55 cm	55 cm	30 cm
layer 5	55 cm	55 cm	-	-	55 cm

2.2.2 Soil Information

The soil profile development is similar for all five test points (Fig. 2.3) with an organic humus (mor) layer on top of a podsolic mineral soil profile.

The **humus layer** consists primarily of organic material formed from accumulation of only partially decomposed organic material at the soil surface [FAO, 1998]. The thickness varies between $3cm$ in site Gammtratten (ID 1 and 2) and $25cm$ in Kindla site ID 2 (Fig. 2.3).

In the top mineral soil, eluviation of metals and base cations - principally caused by organic acids - had occurred. The outcome of this is a bleached **E horizon**. The metals had precipitated in the underlying **B horizon**, characterised by illuvial concentrations of iron and aluminium. In the investigated sites, the B horizon reaches depths of 0.5 to $0.7m$, partially merging into a transition zone below - labelled as **B/C**. In the Kloten site a hard pan had developed in the B horizon (Fig. 2.3). Beneath, a mostly unchanged **C horizon** forms undisturbed parent material.

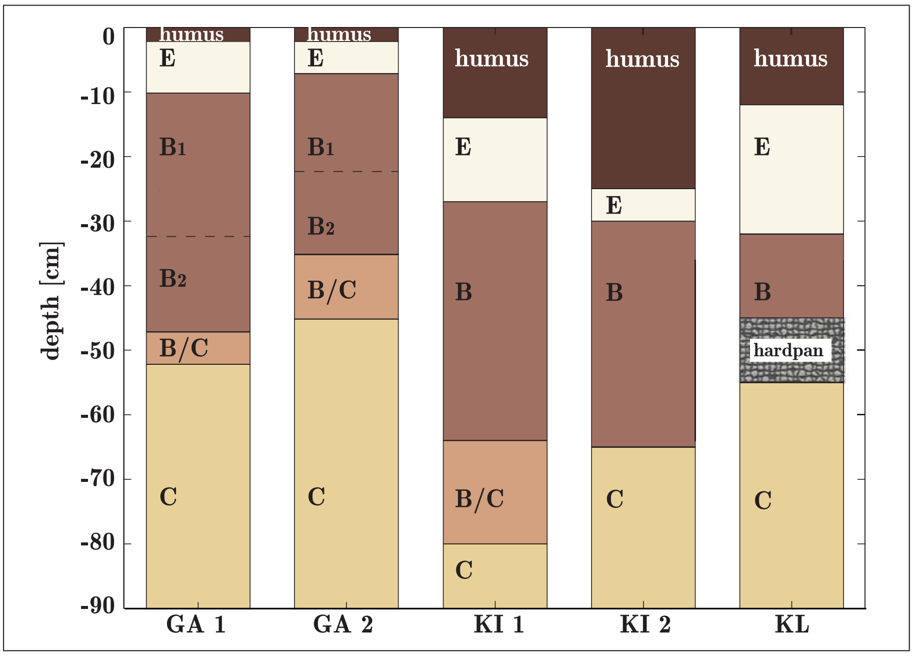


Fig. 2.3. Soil profile diagnostic horizons for all study sites

In upslope locations (recharge areas) the soil profiles were fairly well drained with a dominating vertical percolation. Groundwater tables in the upper soil and organic humus layer occurred rather seldom. The profiles can be considered to be ferric Podsols.

However, in downslope locations close to the stream (discharge areas), the groundwater tables were often found in the upper soil layers, partly close to the soil surface. Thus, higher contents of organic matter in the upper soil layers were provided. Developed soil types vary from humic Podsols to Gleysols or Regosols. Water movement was dominated by lateral flow with only occasionally occurring vertical percolation (in low rainfall periods).

2.3 The WATBAL Model

2.3.1 Model Basics

WATBAL is a soil water balance model developed by Mike Starr at the University of Helsinki. It simply realises an accounting of the components of the water balance for any stand or plot at the end of every month [Starr, 1999]. It was created based on the standard water balance equation:

$$P = ET + R \pm \Delta SM \tag{2.1}$$

P ... precipitation

ET ... evapotranspiration

R ... stream runoff

Δ *SM* ... changes in soil moisture storage

Runoff is usually the sum of surface runoff and interflow, whereas surface runoff is the part that does not enter the soil but directly moves over the land surface. In contrast, interflow (subterranean constituent of the runoff R) is infiltrated water, which travels through the macropores laterally within the vadoze zone of the ground (above groundwater table and below soil surface) under saturated conditions. The interflow component is not particularly taken into account in WATBAL.

In addition to the basic components of the general water balance, WATBAL also considers snow on the ground and two further runoff components: unsaturated flow and drainage. Thereby, more accurate results can be obtained for a smaller scale (Fig. 2.4). Thus, the runoff can be represented as follows:

$$R = R_{surf} + D + R_{unsat} \tag{2.2}$$

D ... drainage

$\boldsymbol{R_{surf}}$ ... surface runoff

$\boldsymbol{R_{unsat}}$ ... unsaturated flow

By combining equations (2.1) and (2.2) as well as considering snow on the ground, the resulting water balance equation used in WATBAL is:

$$P = ET + R_{surf} + D + R_{unsat} \pm \Delta SM \pm \Delta SOG \tag{2.3}$$

Δ *SOG* ... changes in snow on the ground

The drainage component D, also called percolation, describes the amount of water moving downward through the soil. In WATBAL, it occurs when the soil moisture is at field capacity, which means that no more water can be hold against gravity. Unsaturated flow R_{unsat} is the soil water loss under unsaturated conditions due to micropores, intricated flow paths and the matric potential in the capillary system.

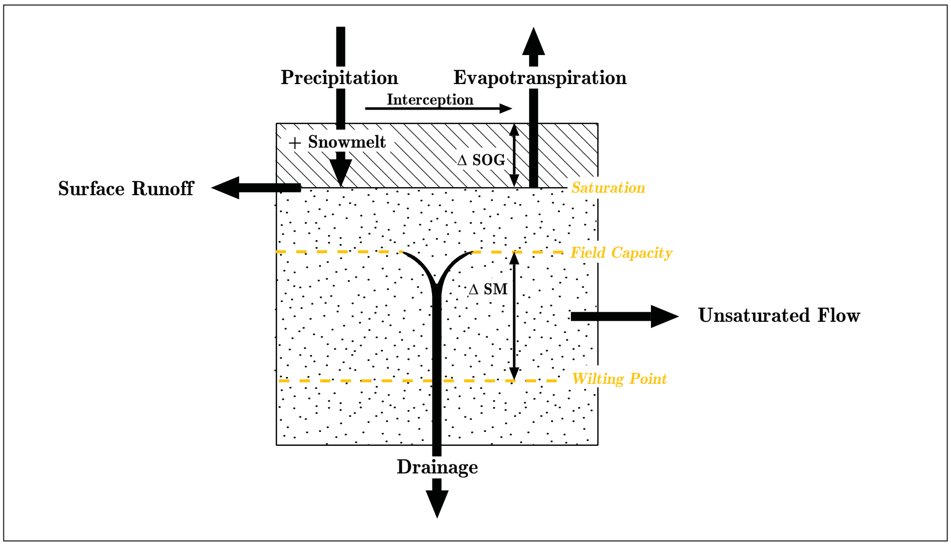


Fig. 2.4. WATBAL main principle. Modified from Starr, M. (unpublished)

WATBAL is a relatively simple model, because it is based on generally known approaches and equations. Furthermore it uses input data that are normally on-hand or can simply be estimated from other available parameters. With the aid of WATBAL it is possible to model the water balance for the entire rooting zone or just for a certain layer therein. The site surface can be of any slope (horizontal or sloping) and aspect. All simulation results are presented in graphical and tabular form; they can be saved and imported into other programmes for editing or analysing purposes. Some of the output parameters are:

- drainage (D)
- evapotranspiration, including monthly ET deficit, potential and actual ET (PET and AET respectively)
- infiltration (F)
- potential soil water loss (PWL)
- snow on the ground (SOG)
- snowmelt (S_{melt}) over a month and potential S_{melt} (PS_{melt})
- soil moisture (SM), e.g. monthly ΔSM, SM at the end of every month and monthly SM deficit
- surface runoff (R_{surf})
- unsaturated flow (R_{unsat})

2.3.2 Required Data

Although WATBAL has a comparatively easy structure, it requires a couple of essential parameters in order to provide appropriate results. These input data give information about the geology of a site, the climate and some other parameters, for instance geographical and layer characteristics.

Geological Data

To model the water flow in a soil, it takes information about the soil itself, such as texture, soil type and infiltration coefficient. Besides, two very important parameters are the soil moisture at field capacity[1] (θ_{FC}) and permanent wilting point[2] (θ_{PWP}). With those parameters it is possible to estimate a couple of other values such as infiltration coefficient and critical soil moisture. It is essential to define the initial SM value (SM of the previous month) before starting the simulation.

Required geological information for Gammtratten, Kindla and Kloten - the three sites investigated in this study - were obtained from *Swedish Geological Survey* (SGU) and from literature [Olsson et al., 1985].

Meteorological Data

The needed climate input data include monthly values of air temperature, precipitation and global radiation (or cloudiness, alternatively). Additionally, the long-term minimum and maximum temperature of the warmest month of the year are required, i.e. the 30-years-average of the warmest and coldest day in July - the warmest month of the year in Sweden. Before running the model, the initial SOG value (SOG of the preceding month) needs to be estimated.

Meteorological data for this project were obtained from literature [Raab and Vedin, 1995] and from the IM database of *Swedish University of Agricultural Sciences* (SLU).

Other Necessary Information

In addition to geological and meteorological values, other parameters, such as aspect, canopy (fraction of forest cover), crop coefficient, elevation, latitude, fraction of roots in the layer and slope (tilt angle), are indispensable to model the water balance. These parameters are partly measured and partly estimated with help of the given geological and meteorological data.

[1] θ_{FC} ... water content at -16 ... -33kPa suction pressure (= pF1.8 ... 2.5)
[2] θ_{PWP} ... water content at -1,500kPa suction pressure (= pF4.2)

2.3.3 Model Parameters

The WATBAL model uses a wide range of parameters with different units and abbreviations, including all necessary climate and geological data as well as other factors (Appendix A). Certain variables need to be defined explicitly by the user (Table 2.3).

Table 2.3. List of input variables to be specified by model user

Input parameters
canopy
critical ratio SM/WHC
crop coefficient
diffuse radiation tilt factor
elevation
file names of calibration, climate, output and runoff file
fraction of roots
infiltration coefficient
initial SM and initial SOG
long-term average max. and min. temperature in warmest month of the year
matrix loss
multiplier for SM/AWC ratio
run-time
soil moisture at field capacity and permanent wilting point

2.3.4 Model Implementation

WATBAL was originally implemented in *ModelMaker Run-time Version 3.0.3*©, but within this project the entire model was rewritten and realised in *MATLAB*® - a programming language for technical computing. The model's equations and conditions were carefully analysed and step by step transferred into a *.m-file, a *MATLAB*® specific file type that contains the programme code with all commands that have to be executed.

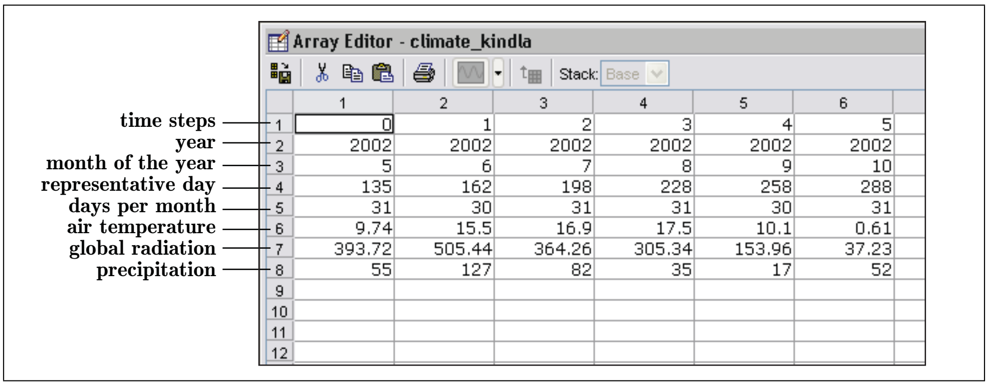
Altogether, this step was intended to simplify model handling and analysis. Editing of the model, its equations and parameters as well as saving certain results, e.g. tables and graphs, can be accomplished much easier and less time consuming in the *MATLAB*® environment compared to *ModelMaker*©. Furthermore, improved plotting functions were added to allow advanced interpretation and comparison of simulation results.

The structograms (Appendix B) visualise the programme structure, divided into radiation routine, snow routine and a routine with main calculations comprising further subdivision.

The programme source code (Appendix C) shows the model design as implemented in *MATLAB*®.

2.3.5 Model Setup and Calibration

Before running WATBAL, the model needs to be set up. The first step is to define the mentioned input parameters (Table 2.3). The specified climate input file must match certain requirements: The file (Fig. 2.5) is a simple *.mat-file - a data file format that is used by *MATLAB*® to save variables and parameters. It contains a matrix with a certain climate parameter (e.g. temperature, precipitation, etc.) per row, whereas every column represents one time step (equals one month). Such a *.mat-file can, for instance, be created by copying cells from an Microsoft®Excel or text file into *MATLAB*®'s array editor and saving this as a *.mat-file, e.g. *climate_kindla.mat*.



		1	2	3	4	5	6
time steps	1	0	1	2	3	4	5
year	2	2002	2002	2002	2002	2002	2002
month of the year	3	5	6	7	8	9	10
representative day	4	135	162	198	228	258	288
days per month	5	31	30	31	31	30	31
air temperature	6	9.74	15.5	16.9	17.5	10.1	0.61
global radiation	7	393.72	505.44	364.26	305.34	153.96	37.23
precipitation	8	55	127	82	35	17	52
	9						
	10						
	11						
	12						

Fig. 2.5. Screenshot of the climate input file. Columns equal time-steps (months), rows contain certain climate parameters.

As soon as all necessary parameters are stated by the user, the model can be run for the first time. The programme plots the simulated soil moisture per month for the predefined time period. It also provides the function to upload another *.mat-file containing time domain reflectometer (TDR) measurements for calibration purpose. This file should be called *SMcali.mat* and can be generated the same way as the climate input file.

Accordingly, the main plot contains the modelled and the really measured data in one chart so that the user is able to compare them and decide whether there is a good fit or not. To improve matching, a calibration can be performed by changing some parameters until the model fit is satisfactory. To simplify the decision making, some calculated statistical parameters and graphics, for instance scatter plots comparing modelled and simulated soil moisture, are presented to the user as well.

In this paper, available TDR measurements of short time periods (approximately one year) - provided by SGU - were used for calibration.

2.4 Testing and Improving the Model

After calibrating the model for each site and all layers within, the soil moisture was modelled over time (subsection 3.1.1 and Appendix F.1). Ideas for improving the model, such as changing and adding certain parameters or rearranging some equations, were tested and statistically evaluated by means of graphical analyses and numerical methods. That way, a final version of WATBAL, providing the best fit for all sites and layers, was obtained.

It turned out, that - without changing the basic principals and algorithms - only a few modifications could be made (subsection 2.4.1) to obtain slightly better results concerning model fit.

2.4.1 Model Modifications

Within the snow routine, snow on the ground (SOG) was earlier calculated by adding precipitation (P) to snow on the ground of the previous month (SOG_{i-1}) and subtracting snowmelt (S_{melt}), in case the temperature is less than $1°C$. The snowmelt component can be neglected since it does not appear as long as the temperature is below the freezing point. This reduction has no effect on the model outcome, but might save some computing time. Thus, the equation was changed as follows:

$$\begin{aligned} SOG &= SOG_{i-1} - \mathbf{S_{melt}} + P \\ &\Longrightarrow SOG_{i-1} + P \end{aligned} \tag{2.4}$$

Furthermore, the equation for the reference potential evaporation (ET_0) was modified. The constant value 1.11 - multiplying the fraction of latent heat of evaporation from a reference crop (LH_vap_Ep) and latent heat of evaporation from soil (LH_vap) - is a correction factor for the known underestimation of lysimeter measured alfalfa ET_0 values. In the original WATBAL model, a value of 1.42 was recommended, but 1.11 was already used for calculation. However, during model test period a value of 1.13 gave the best results for all test sites, so it was changed:

$$\begin{aligned} ET_0 &= \frac{LH\_vap\_Ep}{LH\_vap} \cdot 1.11 \\ &\Longrightarrow \frac{LH\_vap\_Ep}{LH\_vap} \cdot \mathbf{1.13} \end{aligned} \tag{2.5}$$

The formula for calculating the infiltration (F) into the soil was slightly adjusted, too. In case the snowmelt equals zero and temperature (T) as well as the difference of precipitation and forest potential evapotranspiration ($P - ET_c$) are not less than zero, it had been assumed in the original version that all surplus water ($P - ET_c$) infiltrates the soil. But this is not consistent with reality; especially during heavy rain events some surface runoff will always occur. Therefore, the original equation for infiltration determination needed to be multiplied by an infiltration coefficient - a value between zero and one - that reduces the amount of water entering the soil:

$$\begin{aligned} F &= P - ET_c \\ &\Longrightarrow (P - ET_c) \cdot \mathbf{infil\_coeff} \end{aligned} \tag{2.6}$$

Due to the insertion of the infiltration coefficient in equation (2.6), there will always be some surface runoff as long as $P - ET_c$ is greater than zero, irrespective of any other condition. But, in the original model, one condition for calculating the surface runoff is that the snowmelt needs to be greater than zero. Since this requirement is not given any longer, the conditions for runoff calculation were consequently shortened as follows:

$$\begin{aligned} \Longrightarrow\ & \cancel{\mathbf{S}_{melt} > \mathbf{0}} \quad \text{and} \\ & (P - ET_c + S_{melt}) \cdot (1 - \mathit{infil\_coeff}) > 0 \end{aligned} \tag{2.7}$$

Moreover, the formula for computing actual evapotranspiration (AET) needed to be changed as well, in that case by subtracting the surface runoff component:

$$\begin{aligned} AET &= P + S_{melt} - \Delta SM - R_{unsat} \\ &\Longrightarrow P + S_{melt} - \Delta SM - R_{unsat} - \mathbf{R_{surf}} \end{aligned} \tag{2.8}$$

To avoid drainage values higher then infiltration, the way of calculating drainage for each layer was changed as well (equation (2.9)). This has no influence on the soil moisture estimation, but gives slightly different drainage values from the original equation:

$$\begin{aligned} D &= P - \Delta SOG - R_{surf} - ET - |\Delta SM| - R_{unsat} \\ &\Longrightarrow \mathbf{F} - \mathbf{\Delta SM} - \mathbf{R_{unsat}} \end{aligned} \tag{2.9}$$

2.4.2 Application of Pedotransfer Functions

In the course of this thesis it became apparent that some of the given soil property data were not satisfactory and had negative effects on the model fit. Hence, instead of using the given soil moisture at field capacity and wilting point, for some sites these values were calculated with help of a pedotransfer function. The following equations developed by Saxton et al. (1986) were used [Guber et al., 2006]:

$$h = A \cdot \theta^{B} \tag{2.10}$$

$$A = 100 \cdot e^{(-4.396 - 0.071 \cdot clay - 0.000488 \cdot sand^2 - 0.00004285 \cdot sand^2 \cdot clay)} \tag{2.11}$$

$$B = -3.140 - 0.00222 \cdot clay^2 - 0.00003484 \cdot sand^2 \cdot clay \tag{2.12}$$

$\boldsymbol{h}$ ... absolute capillary pressure [kPa]

clay ... clay content [%]

sand ... sand content [%]

$\boldsymbol{\theta}$ ... soil moisture [-]

Equation (2.10) can be rearranged as follows:

$$\theta = \left(\frac{h}{A}\right)^{\frac{1}{B}} \tag{2.13}$$

Assuming that h equals $33kPa$ at field capacity (upper limit, same as pF= 2.5) and $1,500kPa$ at permanent wilting point, equation (2.13) gives soil moisture contents of:

$$\theta_{FC} = \left(\frac{33}{A}\right)^{\frac{1}{B}} \cdot 100 \tag{2.14}$$

$$\theta_{PWP} = \left(\frac{1500}{A}\right)^{\frac{1}{B}} \cdot 100 \tag{2.15}$$

$\boldsymbol{\theta}_{FC}$ ... soil moisture at field capacity [%]

$\boldsymbol{\theta}_{PWP}$ ... soil moisture at permanent wilting point [%]

Aside from the equations above, many different pedotransfer functions - established within other research projects - can be found in literature, each giving unique results for a certain soil type or climate zone. Swedish soils are characterised by high silt contents. Saxton's equations (2.10) - (2.12) do not consider silt as a separate variable, but include only sand and clay content in the calculations. Unexpectedly, those formulas (equations (2.14) and (2.15)) gave the best results out of all tested approaches.

The pedotransfer functions (equations (2.14) and (2.15)) were applied for the three locations Kindla ID 1, Gammtratten ID 1 and ID 2. For all other sites (Kindla ID 2 and Kloten) the given FC and PWP values were used.

2.5 Runoff Estimation

In order to check WATBAL's further application possibilities, the catchment runoff was estimated by means of simulated components of the hydrological cycle. That way, it was also possible to test the consistency of these modelled parameters. Assuming that surface runoff, drainage and unsaturated runoff contribute to the total amount of water in the stream, the runoff was calculated based on equation (2.2).

2.6 Heavy Metal Transport

The final step after the fully completed simulation of soil moistures and associated water fluxes was the determination of heavy metal loads for the two IM sites Kindla and Gammtratten. Heavy metal mass balances are important for risk assessment concerning environmental protection and human health. In general, such a mass balance can be done by comparing input and output values of certain critical elements. With help of water fluxes and chemical concentrations it is possible to calculate the monthly or yearly mass transfer (equations (2.16) and (2.17)). In this paper, the elements cadmium (Cd), lead (Pb), mercury (Hg), zinc (Zn) and copper (Cu) were analysed.

$$m = V \cdot c \tag{2.16}$$

$$\dot{m} = Q \cdot c \tag{2.17}$$

$\boldsymbol{c}$ ... concentration $[\frac{g}{L}]$

$\boldsymbol{m}$ ... mass $[g]$

$\boldsymbol{\dot{m}}$ ... mass flow rate $[\frac{g}{month}]$

$\boldsymbol{Q}$ ... volumetric flow rate $[\frac{L}{month}]$

$\boldsymbol{V}$ ... Volume [L]

The scheme of heavy metal transport into, through and out of the subsurface (Fig. 2.6) illustrates that the input is represented by heavy metals reaching the soil surface by (i) throughfall, which is the precipitation hitting the ground, including direct precipitation, canopy drop and stemflow [Klaassen et al., 1996], and (ii) litterfall, which is accumulated biomass on the soil surface, i.e. leaves and needles that have fallen from trees. One part of the input is directly removed by surface runoff or interflow, both reaching the stream quite fast. The remaining part is entering the soil and moving from soil layer to layer. During this process, heavy metals can accumulate to a greater or lesser extent, depending on certain parameters (e.g. soil type and grain size). However, some soil water fluxes get to the stream directly, others move towards groundwater, which also reaches the stream after some time. The stream can be assumed to be the only discharge of heavy metals, if measured at the catchment outlet.

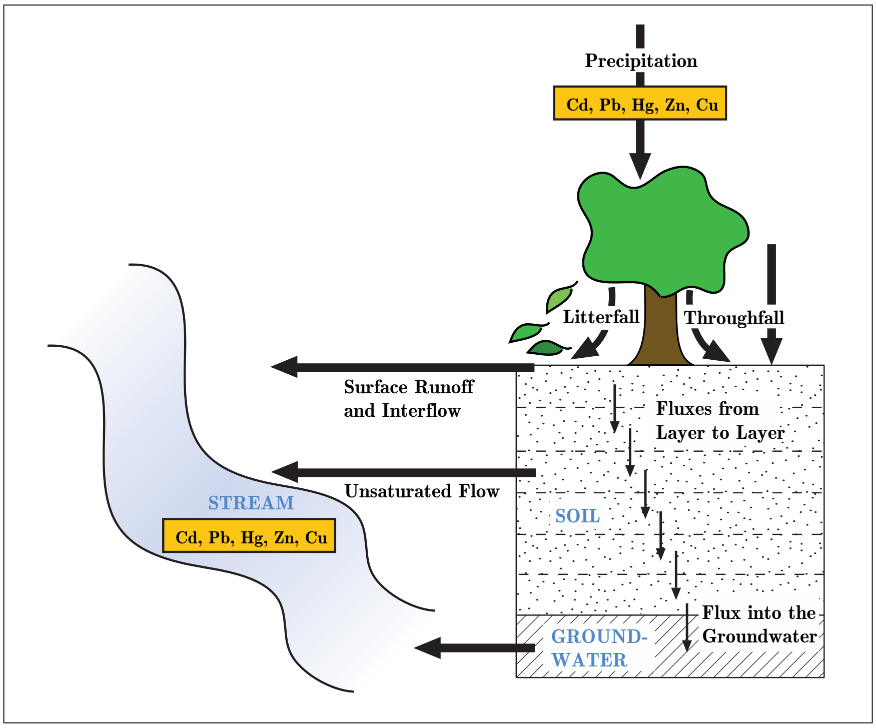


Fig. 2.6. Principle of heavy metal fluxes

2.6.1 Identification of Incoming Mass Rates

To obtain the total input, a couple of calculation steps were necessary. Throughfall data $[g \cdot L^{-1}]$ were given for years 1998-1999 (Kindla site) and 1999-2000 (Gammtratten site). From the respective years, monthly average values were estimated. It was assumed that these values do not vary a lot from year to year and can therefore also be applied to other time periods. Monthly precipitation data was on hand for the simulation period 2002-2006 and was multiplied by the calculated throughfall averages to determine the mass of the mentioned heavy metals that reach the soil surface with the precipitation per month.

Litterfall data $[g]$ were given for years 1998-2002 (Kindla site) and 2000-2002 (Gammtratten site) respectively. Unfortunately, there were only three values measured per year. Thus, it was required to assess the missing months by interpolation (cubic spline). Afterwards, the monthly means were calculated from these years to obtain mass rates $[g \cdot month^{-1}]$ of heavy metals that reach the ground by falling biomass from trees.

In this study, plant uptake was neglected, because (i) it is only a relatively small component compared to litterfall, at least in case of lead [Klaminder et al., 2005] and mercury [Bishop et al., 1998], and (ii) no information on tree uptake was available. Therefore, the total input was obtained by adding the mass values of throughfall and litterfall together.

2.6.2 Determination of Mass Rates at Catchment Outlet

Monthly chemical concentrations $[g \cdot L^{-1}]$ in the stream water were given for years 1998-2002 (Kindla site) and 1999-2004 (Gammtratten site). Based on that, averages for every month were calculated. The amount of runoff $[L \cdot month^{-1}]$ was available for the simulated time period 2002-2006 and was multiplied by the averaged concentrations to estimate the mass transport of the mentioned heavy metals at catchment outlet.

2.6.3 Accumulation in the Soil

With the ascertained data, the mass difference of in- and output could be calculated to get an idea of the amount of cadmium, lead, mercury, zinc and copper that is accumulating in the soil zone per month.

Aim of this study was not only to get information about the quantity of heavy metal enrichment in the entire soil, but also to find out more details about the accumulation in certain soil layers. Soil water heavy metal data $[g \cdot L^{-1}]$ were only existing for a depth of $30cm$ below humus layer for both Kindla ID 1 as well as Gammtratten ID 1, and for time intervals 1998-1999 (Kindla site) and 2000-2003 (Gammtratten site). Only three to four measurements were performed per year, so that an interpolation had to be done to obtain values for the missing months. These values were assumed to represent the concentrations in the output (drainage and unsaturad flow) of the respective layer at $30cm$ below humus layer. On this basis, accumulation was devided into two soil zones I and II (Fig. 2.7).

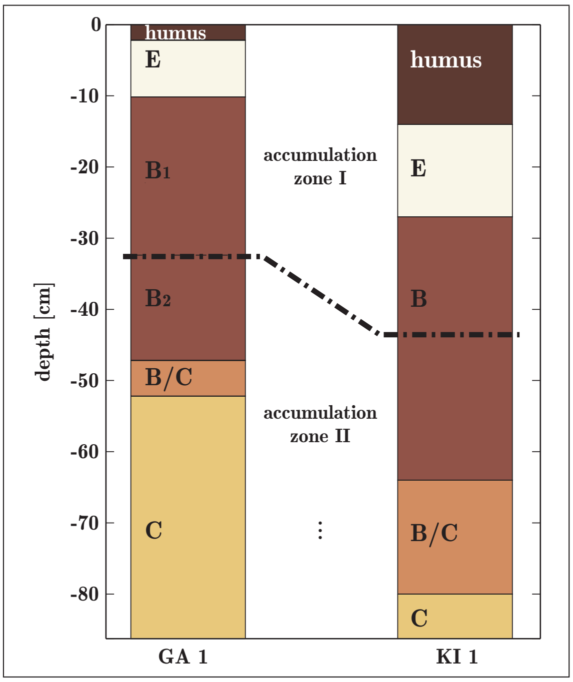


Fig. 2.7. Devision of accumulation zones

For Gammtratten, accumulation zone I is the upper soil part until a depth of $33cm$ (equals $3cm$ humus layer plus $30cm$ of soil horizons E and B) and zone II is the remaining part below that. In Kindla ID 1, accumulation zone I equals the first $42cm$ of the soil (i.e. $12cm$ humus layer and $30cm$ of soil horizons E and B) and zone II is the part below $42cm$.

Heavy metal concentrations in throughfall, litterfall and stream water were provided by SLU and the *Swedish Environmental Research Institute* (IVL).

2.7 Statistical Analysis

2.7.1 Basic Statistics

To understand and test the basic WATBAL model functions as well as relationships of certain variables, a couple of investigations were made, including calculations of expected value $E(x)$, covariance $Cov(x, y)$, correlation $\rho_{x,y}$, deviation σ, standard error SE and so forth. Equations are not given in this paper, but can be found in every standard statistics textbook, for instance *Schaum's Outline of Theory and Problems of Probability, Random Variables, and Random Processes* [Hsu, 1997] or *Statistical Design and Analysis of Experiments* [Mason et al., 2003].

2.7.2 Model Fit

The statistical values R^2, mean absolute error (MAE), mean absolute percentage error ($MAPE$) and root mean squared error ($RMSE$), as well obtained from standard textbooks, were used to analyse the model fit for a given time period (i.e. calibration period):

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2} \tag{2.18}$$

$\boldsymbol{i}$ ... control variable

$\boldsymbol{n}$ ... sample size

$\boldsymbol{R}^2$ ... R square value

$\boldsymbol{y}_i$ ... observed values

$\hat{\boldsymbol{y}}_i$ ... modelled values

$\bar{\boldsymbol{y}}$ ... mean of observed values

$(\boldsymbol{y}_i - \hat{\boldsymbol{y}}_i)$ ... residual

$$MAE = \frac{1}{n} \cdot \sum_{i=1}^{n} |(y_i - \hat{y}_i)| \tag{2.19}$$

$$MAPE = \frac{1}{n} \cdot \sum_{i=1}^{n} \left| \frac{y_i - \hat{y}_i}{y_i} \right| \tag{2.20}$$

$$RMSE = \sqrt{\frac{1}{n} \cdot \sum_{i=1}^{n} (y_i - \hat{y}_i)^2} \tag{2.21}$$

MAE ... mean absolute error

MAPE ... mean absolute percentage error

RMSE ... root mean squared error

For equations (2.19) - (2.21), the model fit is, in general, the better the lower the value is. $MAPE$ is a useful parameter and easy to understand. It can have a value from zero to $+\infty$, whereas zero implies a perfect fit and every value greater than one a more than 100 percent deviation from the original model. Thus, the closer the value is to zero, the better is the fit. R^2 should have values between zero and one, whereas one is a perfect fit (the modelled curve comes close to the observed points) and zero means that a horizontal line going through the mean of all observed y values fits the data better than the modelled values.

R^2 is not a reliable parameter (at least not for use as main criterion), because high R^2 values imply that the modelled curve comes close to the observed values, which does not necessarily mean that it is a good fit in other ways (possibly high confidence interval).

The other two parameters, MAE and $RMSE$, can vary within a range of zero and $+\infty$, indicating a good model fit by low values.

For model accuracy studies, the overall average percentage error was calculated. It gives information about the percentage deviation of the modelled from the observed values and is calculated similar to the $MAPE$ value:

$$PE = \frac{|(y_i - \hat{y}_i)|}{y_i} \cdot 100\% \tag{2.22}$$

PE ... percentage error

The respective deviations $(y_i - \hat{y}_i)$ could not be calculated for other years than those used for calibration, because there were no further observed values available. To estimate the model's total grade of correctness for all predictions nevertheless, even for those without available $\hat{y}_i$ values (equals statistical population), the *average* percentage error was calculated for the available $(y_i - \hat{y}_i)$ data pairs (equivalent to statistical sample) in consideration of corresponding mean and standard error.

Acceptable benchmarks that seemed to be suitable for the above mentioned evaluation parameters were assumed to be:

- $R^2 > 0.6$
- $MAPE < 0.1$
- MAE in general low (e.g. < 5)
- $PE < 10\%$
- $RMSE$ in general low (e.g. < 10)

Chapter 3

Results and Discussion

3.1 Model Outcome

3.1.1 Simulated Soil Moisture

WATBAL was run for all layers within each site. Soil moisture simulations were obtained for Kindla site ID 1 (Fig. 3.1) and all other test points (Appendix F.1).

The results indicated connections between both the climate parameters and runoff (Fig. 3.1 (a)) as well as climate parameters and soil moisture (Fig. 3.1 (b)).

Runoff and soil moisture curves follow the seasonal patterns of temperature, global radiation and precipitation relatively well. Especially during winter time (January and February), runoff and soil moisture are a bit lower, because most of the precipitation is accumulating at the soil surface as snow due to low temperatures (Fig. 3.1 (a)). In spring (March and April), as the weather gets warmer and the snow starts to melt, water content in the soil as well as runoff show high values. In the middle of the year (June-August), the runoff and soil moisture curves show a low point as a result of high temperature and global radiation (hence greater evapotranspiration). During autumn (October and November), which is characterised by heavy rain falls as well as lower temperature and global radiation (consequently less evapotranspiration), soil water content and runoff reach their maximum.

The following consequential relationship between soil depth and soil moisture could be observed: The thicker the soil layer from ground surface to respective depth, the higher is the water content (Fig. 3.1 (b)).

On the basis of these coherences, this analysis confirms that WATBAL is able to comprehend and reproduce natural hydrological processes, which can be seen in the curve progression of soil moisture content (Fig. 3.1 (b)). Furthermore, soil water content is not only dependent on climate parameters, but also on various other influencing factors, for instance certain soil characteristics (e.g. grain size, porosity and soil type), and can therefore vary from layer to layer and site to site.

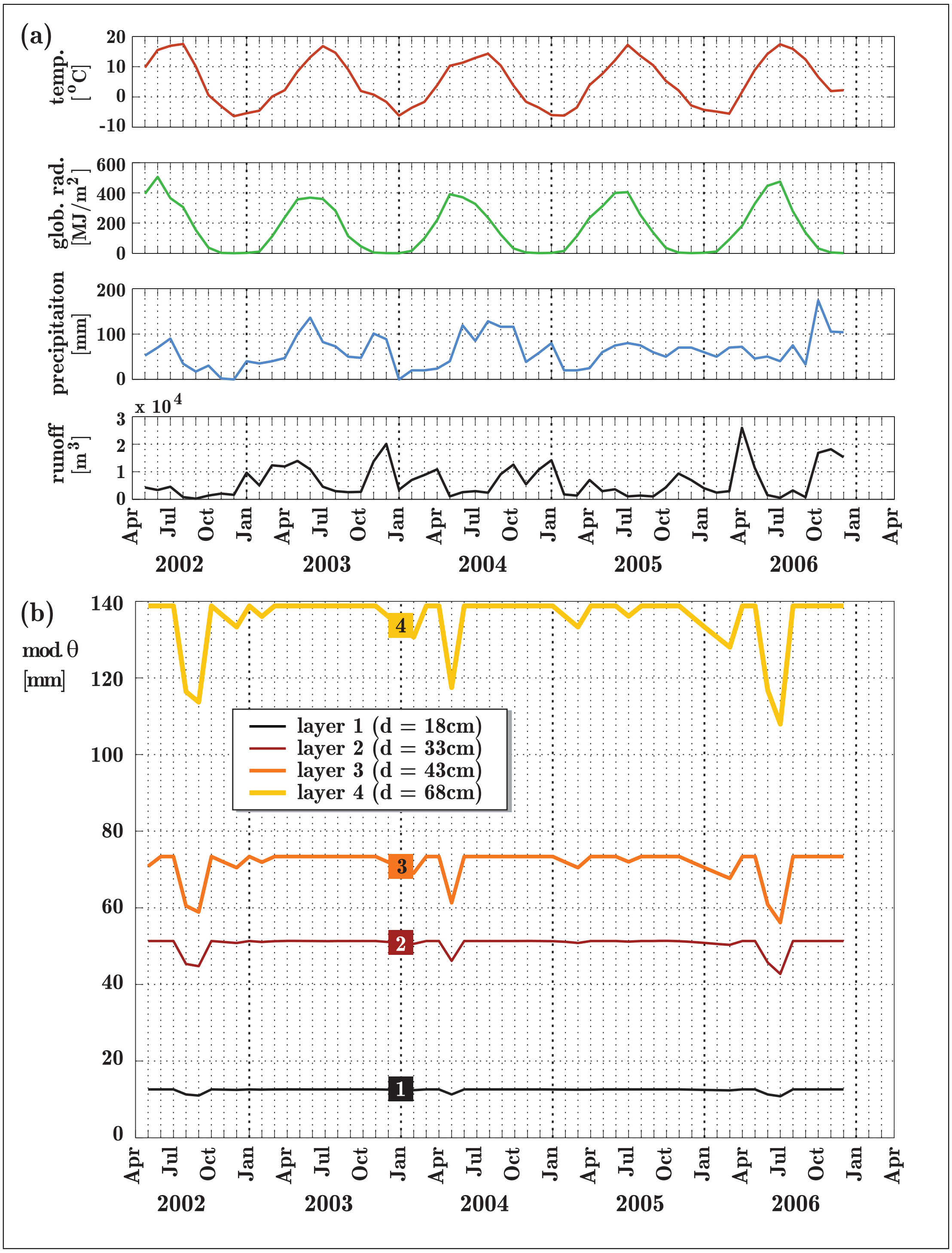


Fig. 3.1. Kindla site: (a) Monthly climate data - temperature, global radiation, precipitation and runoff. (b) Monthly modelled SM for all layers (from ground surface to respective depth) with thickness d.

3.1.2 Dependency on Climate Parameters

In order to check the impact of certain climate parameters on the simulation outcome of WATBAL, Pearson product-moment correlation coefficients of modelled soil moisture with precipitation, temperature and global radiation, respectively, were determined.

Since higher precipitation causes a higher amount of water reaching the soil, a positive correlation was expected in this case. Contrarily, increased temperature as well as global radiation values result in greater evapotranspiration and consequently in less water entering the soil. Thus, negative correlations were expected in these both cases.

The correlation coefficients for Kindla site ID 1 (Table 3.1) confirm the previous assumptions. Similar correlations of varying intensity were found for all layers in all sites.

Table 3.1. Pearson correlation coefficients for SM and climate factors, Kindla ID 1

Layer no.	$\rho_{temperature,SM}$	$\rho_{precipitation,SM}$	$\rho_{radiation,SM}$
Layer 1	-0.331	0.354	-0.278
Layer 2	-0.342	0.348	-0.293
Layer 3	-0.210	0.400	-0.085
Layer 4	-0.280	0.377	-0.185

There is a positive relationship between precipitation and simulated soil moisture (Fig. 3.2) and a negative relationship for the other two parameters (Fig. 3.3 and 3.4). In general, the strongest correlation can be found for precipitation and soil moisture. On the contrary, radiation and soil moisture feature the lowest degree of linear dependency.

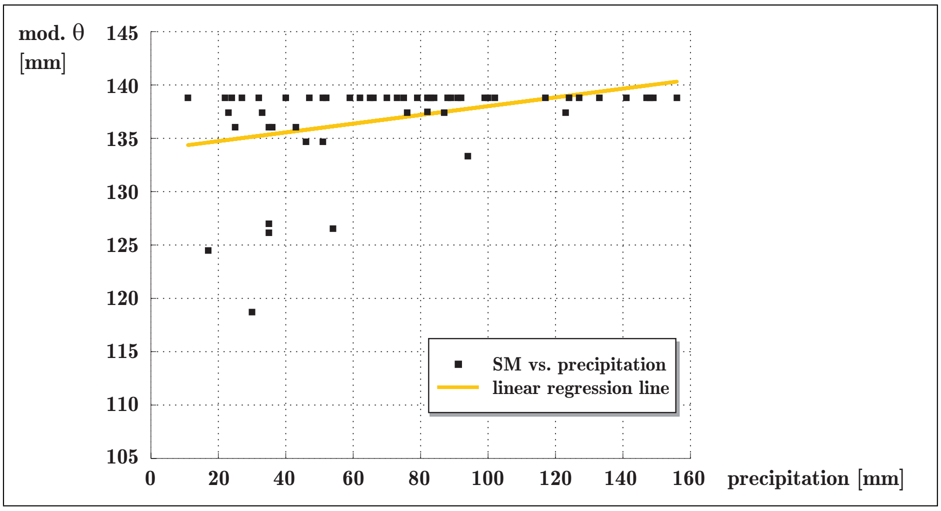


Fig. 3.2. Modelled soil moisture versus precipitation, Kindla site ID 1, layer 4 (until 55cm below humus layer). The grey trend line indicates a positive relationship.

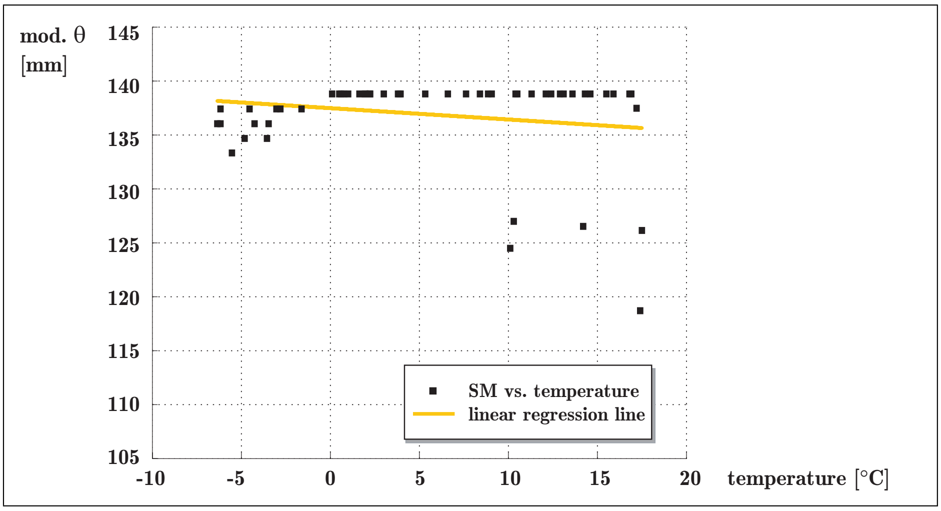


Fig. 3.3. Modelled soil moisture versus temperature, Kindla site ID 1, layer 4 (until 55cm below humus layer). The grey line's slope shows a negative relationship.

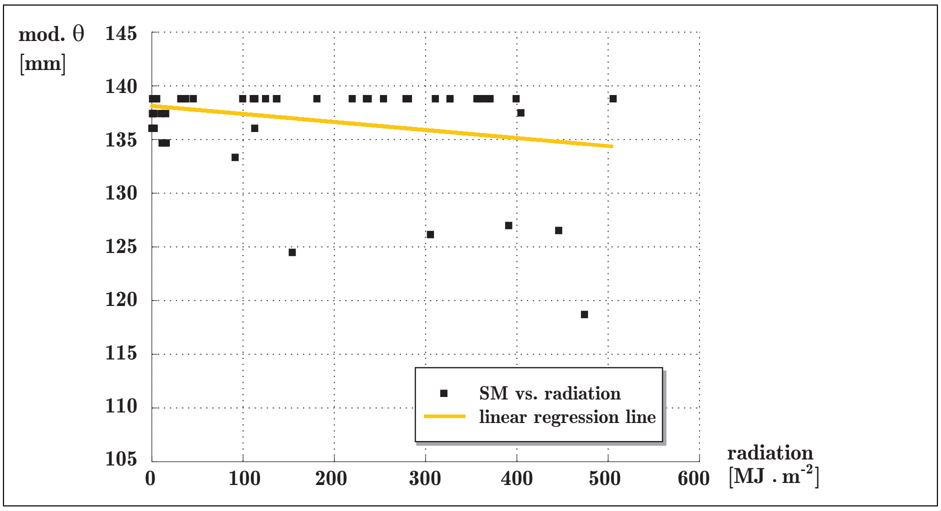


Fig. 3.4. Modelled soil moisture versus glob. radiation, Kindla site ID 1, layer 4 (until 55cm below humus layer). The grey trend line is a sign of a negative relationship.

A gathering of points at field capacity (138.8mm) becomes apparent (Fig. 3.2 - 3.4), irrespective of the amount of precipitation, temperature or radiation. This is a first sign of limitations and a poor model fit. These points abate the slope of the regression line, which would otherwise be steeper and show stronger correlations.

3.2 Model Fit and Accuracy

3.2.1 Numerical Analysis of the Prior WATBAL Version

Computations of the numerical parameters R^2, $MAPE$, $RMSE$ and MAE resulted in a poor model fit after calibration in almost all cases. Just for a few layers the values met the demands mentioned in subsection 2.7.2.

Kindla site ID 1 (Table 3.2) showed the best model fit compared to the other sites (Appendix D.1). The best values, indicating an acceptable fit, occur in Kindla site ID 1, layer 1 ($R^2 = 0.5843$, $MAPE = 0.0152$, $RMSE = 0.3488$ and $MAE = 0.1805$).

Table 3.2. Statistics, Kindla site, ID 1 (original model)

Profile-ID	Layer	R^2	MAPE	RMSE	MAE
ID 1	*Layer 1*	0.5843	0.0152	0.3488	0.1805
	Layer 2	0.4621	0.0167	1.6326	0.8341
	Layer 3	0.2361	0.0740	6.8604	4.5673
	Layer 4	0.1293	0.0345	6.9171	4.4669

3.2.2 Numerical Analysis of the Revised WATBAL Model

Since the model fit was poor with the original model, a couple of changes were made in WATBAL (subsection 2.4.1), which resulted in slightly enhanced values. The improvements were not outstanding but significant anyway. Considering all calculated evaluation parameters, in 75 % of all cases an improvement was obtained. The attained enhancements were more relevant and of larger extend than the partly occurred declines.

The results of the improved model regarding Kindla site ID 1 (Table 3.3) as well as all other sites (Appendix D.2) point out a much better model fit compared to the results of the previous model version. But nevertheless, R^2 is still relatively low and only $MAPE$ shows acceptable values many times.

Table 3.3. Statistics, Kindla site, ID 1 (improved model)

Profile-ID	Layer	R^2	MAPE	RMSE	MAE
ID 1	*Layer 1*	0.5843	0.0152	0.3488	0.1805
	Layer 2	0.4661	0.0168	1.6416	0.8372
	Layer 3	0.2567	0.0728	6.7255	4.4973
	Layer 4	0.1464	0.0338	6.8001	4.3928

These results give a first idea of the model fit. But to make a more precise statement about the match of WATBAL, graphical analyses need to be considered, too.

3.2.3 Graphical Analysis of the Model Results

Additional to numerical methods, the model fit can also be estimated by examining statistical graphics, e.g. diagrams of modelled and observed data as well as plots that show their deviation and faultiness. Site Kindla was chosen to demonstrate the results obtained with the improved model exemplarily.

Modelled and Observed Data over Time

On the basis of plots, simulated and observed soil moisture content over time had a passable match (Fig. 3.5 - 3.8). Most of the actually measured peaks are also visible in the modelled curve so that the overall impression of WATBAL's model suitability is acceptable.

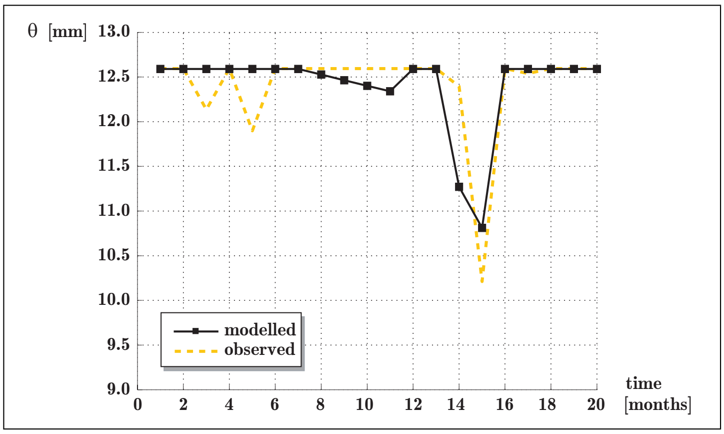


Fig. 3.5. SM over time, Kindla site ID 1, Layer 1 (5cm below humus layer).

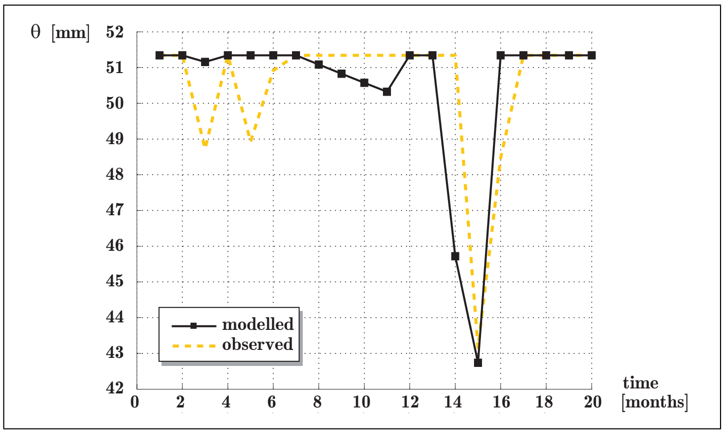


Fig. 3.6. SM over time, Kindla site ID 1, Layer 2 (20cm below humus layer).

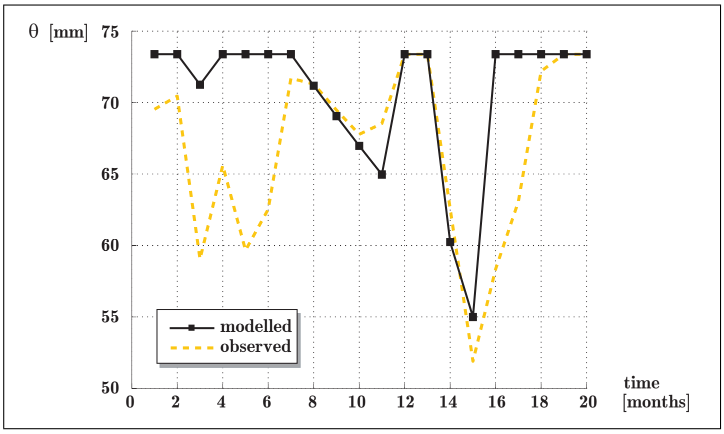


Fig. 3.7. SM over time, Kindla site ID 1, Layer 3 (30*cm* below humus layer).

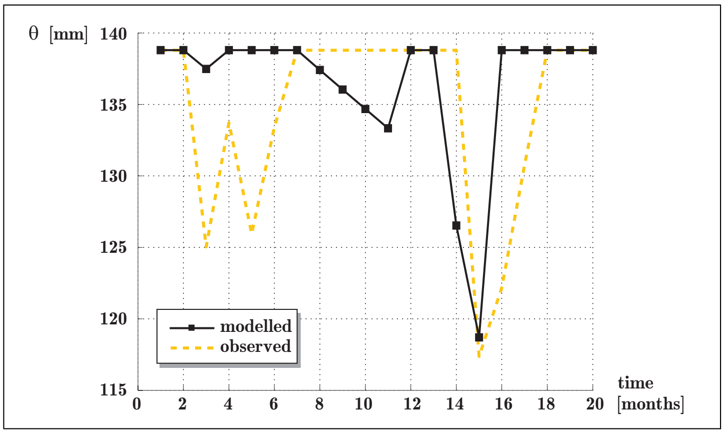


Fig. 3.8. SM over time, Kindla site ID 1, Layer 4 (55*cm* below humus layer).

The model fit is much better for later months of calibration period for all layers within Kindla site. This could be a seasonal problem or be caused by a certain set of input parameters. This phenomenon did not occur for the other study sites.

Residual Analysis

For further model fit investigation, a couple of statistical plots were created and residuals ($y_i - \hat{y}_i$), in form of absolute as well as percentage values, were analysed.

In general, a linear relationship between modelled and observed data can be observed (Fig. 3.9), which is usually a sign of good model fit. But there are some differences between modelled and observed data, indicated by residuals between $-53.58mm$ and $+46.94mm$ (Fig. 3.10) as well as -84.40% and $+56.77\%$ (Fig. 3.11), respectively. Absolute residuals average out to $-1.04mm$, the mean of percentage residuals is -3.41%.

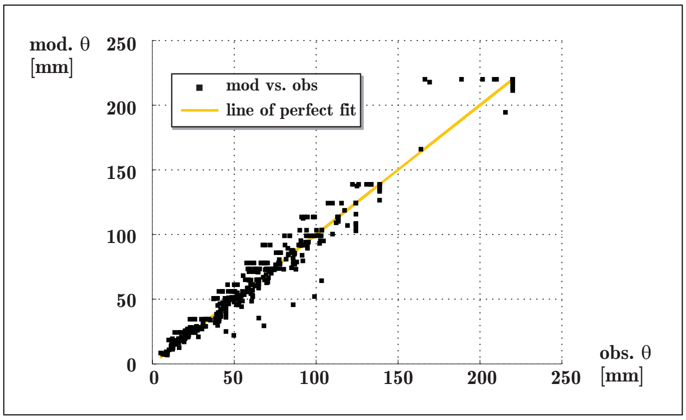


Fig. 3.9. Modelled versus observed soil moisture. Values from all study sites.

Another indication for good model fit is the equal variance of model residuals, which could not be verified for the existing data: For high observed soil moistures, the absolute residuals [mm] tend to have large variation (Fig. 3.10). Percentage residuals instead show a tendency to greater variation for low observed soil water contents (Fig. 3.11). Either way, variation of the residuals is irregular and can not proof a good model fit.

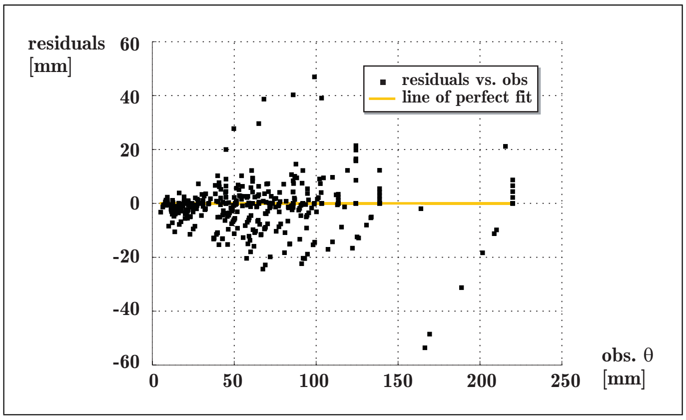


Fig. 3.10. Absolute residuals over observed SM.

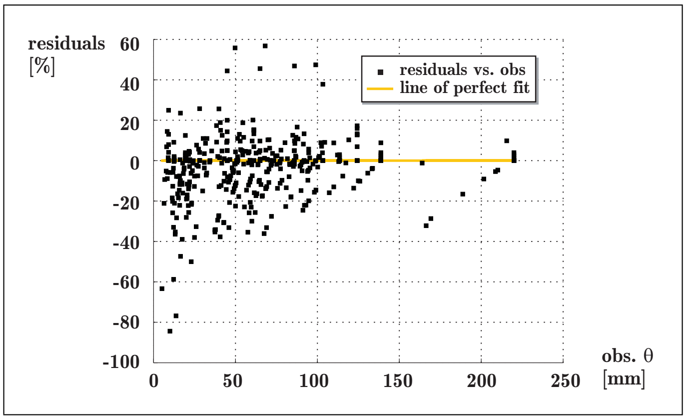


Fig. 3.11. Precentage residuals over observed SM.

Furthermore, a common sign for a good model fit is the normal distribution of the variation. In order to test this, histograms for both absolute as well as percentage residuals were produced. At first sight, the residuals seem to be normal distributed (Fig. 3.12), but the bar around zero is suspiciously large and does not fit with the normal distribution curve.

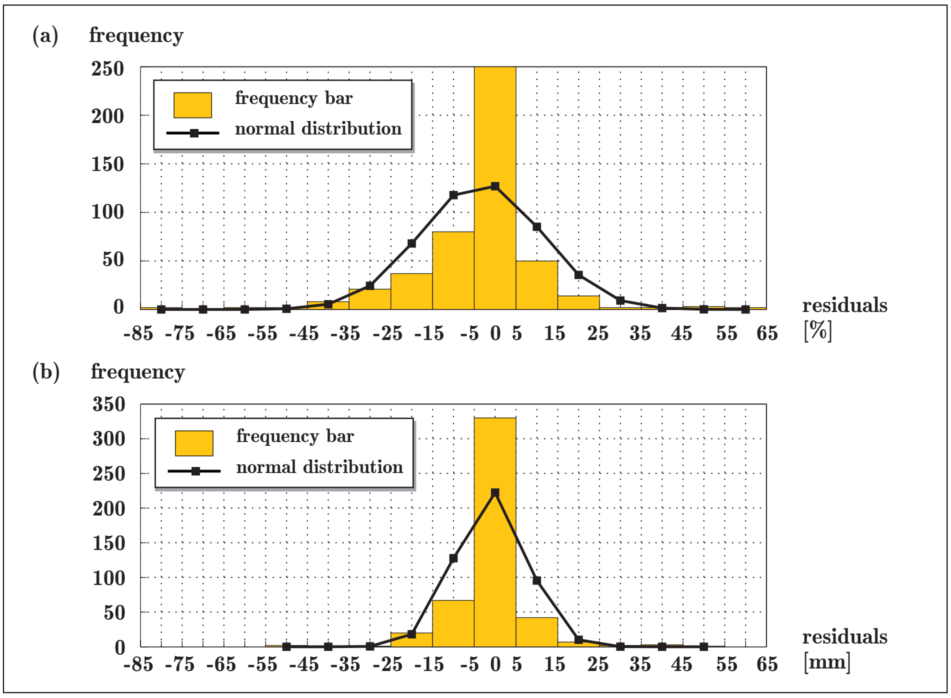


Fig. 3.12. Histograms for (a) percentage as well as (b) absolute residuals, combined with normal distribution curves

To examine this more detailed, probabilities of the residuals were plotted in a normal probability plot (Fig. 3.13). In the ideal case of normal distribution, the points would form a straight line. Unfortunately, the probability plot does not follow a straight line (Fig. 3.13) but is rather S-shaped and seems to follow a Student's distribution.

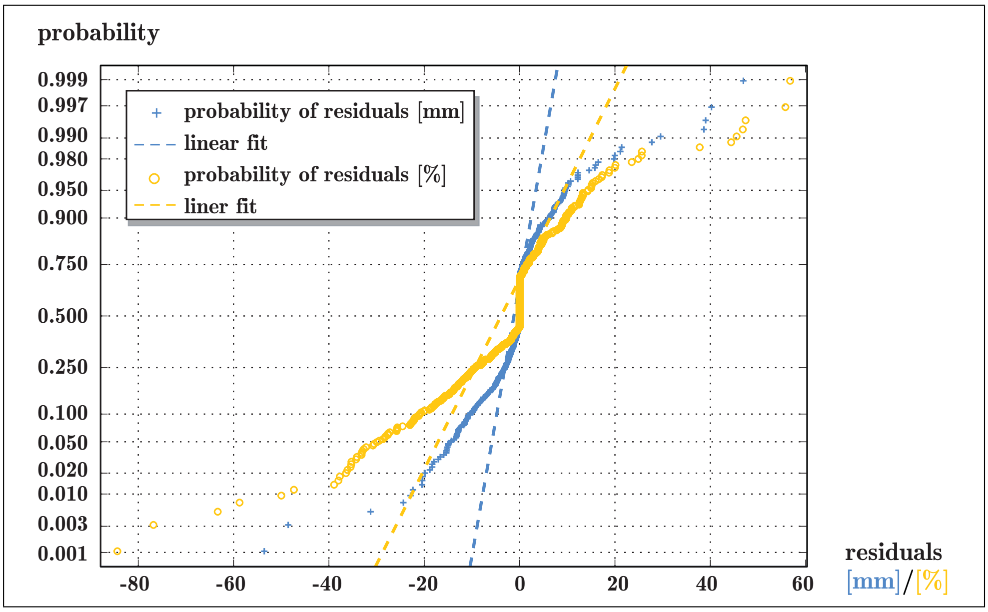


Fig. 3.13. Normal probability plot for absolute and percentage residuals. Straight lines represent normal distribution.

For that reason, the residual analysis can not confirm a good model fit regarding the whole calibration period. The two procedures trying to prove equality as well as normal distribution of residual variation failed, although the scatterplot (Fig. 3.9) shows a straight line. The model accuracy might result in much better values considering just the last months of calibration period that show a very good fit (Fig. 3.5 -3.8).

3.2.4 WATBAL's Accuracy

Aside from the parameters R^2, $MAPE$, MAE and $RMSE$ that were used to investigate the model fit (subsection 3.2.2), percentage error was used to describe WATBAL's total accuracy, which also includes the reliability of prospective predictions. The overall absolute percentage error of WATBAL - resulting from this study - was estimated as $8.91\% \pm 1.08\%$ with a 95% confidence interval.

But, this percentage error is site depending, e.g. for Kindla, ID 1 - the site with the best model fit - the average percentage error was determined as only $3.47\% \pm 1.24\%$ with a 95% confidence interval. In contrast, the average percentage error for Kloten, which is the site with the worst model fit, was $12.20\% \pm 2.95\%$ with a 95% confidence interval. However, the overall average percentage error was less than 10% and is therefore considered as satisfactory.

3.3 Cause of Errors and Weaknessess

3.3.1 Inaccuracy Causes

The WATBAL model is influenced by various factors, such as a range of measured climate parameters, estimated geological values and other adjusted variables. Thus, the occurring errors and deviations from the expected model outcome can have manifold causes. For instance, WATBAL parameters that are adjusted during calibration can cause inexact results if they are not tuned in the right way. Furthermore, certain soil properties need to be applied for the right layer (depth) and there might be a need of slight adjustments (e.g. pedotransfer functions).

The model is based on a large amount of measured data, so that even a small uncertainty or faultiness in measurements can lead to a bad model outcome. Especially for the Kloten site, a major part of inaccuracy was caused by a lack of appropriate data. Since there were no climate and runoff measurements done directly in the Kloten catchment, the observed climate and runoff values were obtained from nearby stations, which had a not negligible negative impact on the modelling results.

Furthermore, errors can also occur due to error-prone calculation approaches. For example, pedotransfer functions were applied for three sites, because the measured data did not seem to fit well. But as mentioned earlier, the used functions might not be perfect for the soil types in Sweden and therefore be an additional error source.

3.3.2 Model Weakness

Considering the performed WATBAL simulations for all sites Gammtratten, Kindla and Kloten, the gained results were generally acceptable. Anyway, there might still be some need for improvement since the calculated statistical parameters as well as the graphics did not show an outstanding model fit. But unfortunately, in order not to go beyond the scope of this thesis, limits needed to be set and further improvement possibilities - for instance reconsidering some equations and finding alternative formulas and parameters in literature - are not discussed in this paper.

A mentionable weak point of the model is that the upper limit of possible soil moisture is assumed to be equal to the soil moisture content at field capacity. Any water above that goes directly to drainage. In reality, the upper limit is soil saturation instead, which is greater than field capacity. Moistures above field capacity are also included in measurements, i.e. available measurements usually show much higher soil moisture values than the simulation. Thus, moisture contents between field capacity and saturation need to be excluded, in order to compare simulated and measured values. This causes a reduction of the soil moisture curve's amplitude.

To clarify this further, original and corrected observed soil moisture for layer 1 in Kindla ID 1 were plotted over time (Fig. 3.14). Especially between October 2005 and May 2006 the effect of amplitude reduction can be seen. The original curve shows a typical moisture increase in October and November due to heavy rain falls in autumn. In winter, the water content decreases a little bit and finally in spring the soil moisture takes higher values again as a result of snowmelt. However, this seasonal pattern can not be recognized in the corrected - by field capacity limited - curve at all.

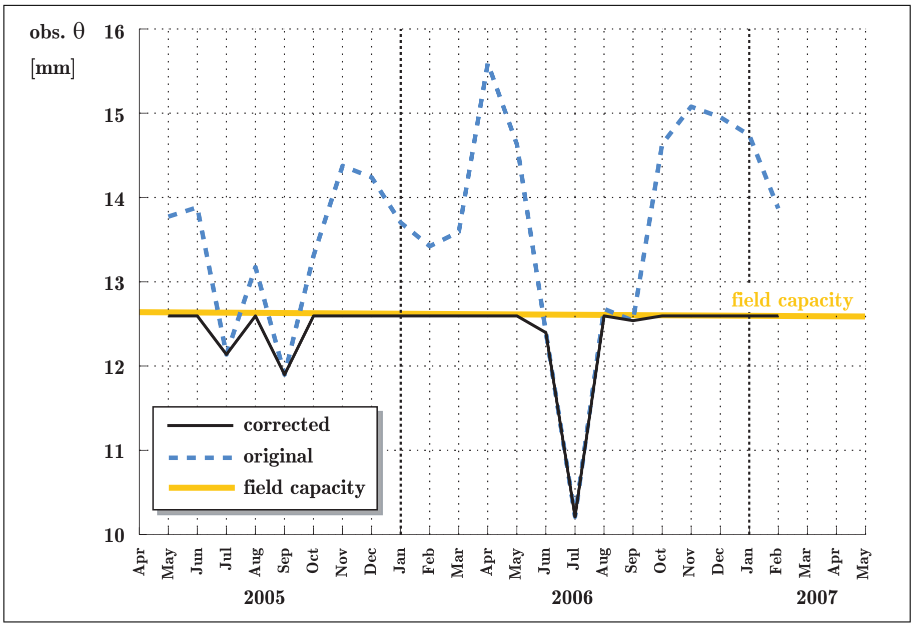


Fig. 3.14. Original observed SM (distinct seasonal scheme) and corrected, by field capacity limited, SM (abated seasonal pattern), Kindla site ID 1, layer 1.

This inconsistency in the WATBAL model is the main cause of imprecise simulation results, because the investigated Swedish soils mainly have soil moisture contents above field capacity.

Unfortunately, this problem could not be overcome in the course of this work, because the idea of field capacity being the maximum possible value and every surplus going to drainage is the basis for all calculations in WATBAL (capacity approach) and therefore could not be changed.

3.4 Hydrological Fluxes

3.4.1 Runoff Estimation at Catchment Outlet

With the aid of given and modelled hydrological constituents, for instance precipitation, evapotranspiration and surface runoff, it was also possible to calculate the water flow in the stream at catchment outlet. The fit of calculated and measured data was not expected to be very good due to the use of a simple water balance approach (equation (2.2)) for runoff estimation. In example, water retaining in surface ponds, time delayed runoff and spatial variability within the catchment were disregarded.

Table 3.4. Runoff statistics for sites Gammtratten, Kindla and Kloten.

Site	R^2	MAPE	RMSE	MAE
Gammtratten	0.5464	0.5979	33.7755	20.3742
Kindla	0.4188	0.9258	40.9389	27.5960
Kloten	0.4403	0.8526	48.6563	31.8465

However, the match of calculated and measured values was good, especially for Gammtratten site, which shows the best numerical parameters (Table 3.4) as well as best fit of modelled and observed runoff curves (Fig. 3.15). For the catchments Kindla and Kloten, the fit was a bit worse (Table 3.4), but the monthly measured runoff characteristics can still be recognised in the modelled runoff progression (Fig. 3.16 and 3.17, respectively). The deviating simulated runoff might also be originated in the difference of scale. The WATBAL model is used for small scale water transport in the soil and, by trying to calculate runoff at a catchment scale, inconsistencies are likely to occur.

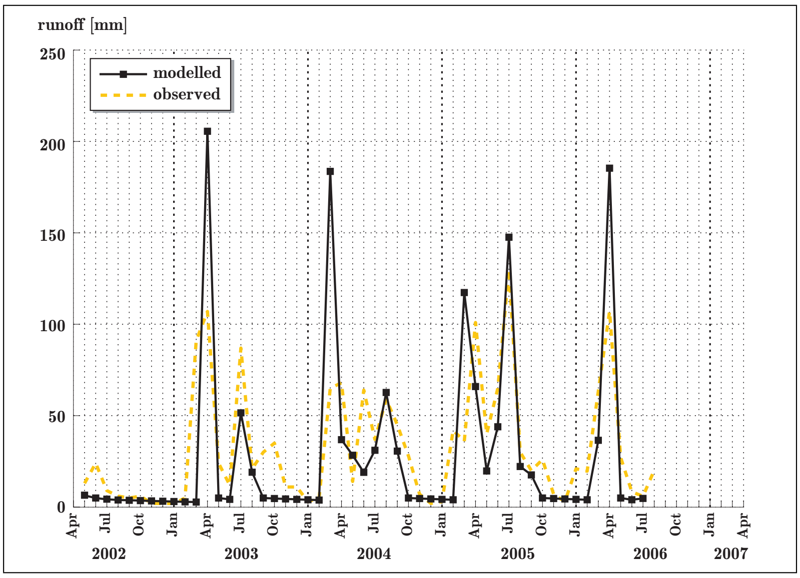


Fig. 3.15. Calculated and measured stream runoff, Gammtratten.

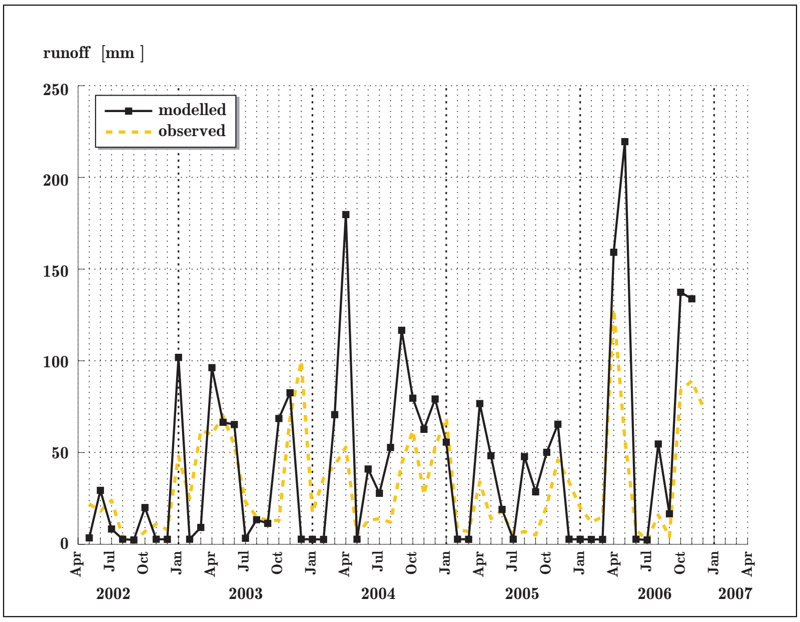


Fig. 3.16. Calculated and measured stream runoff, Kindla.

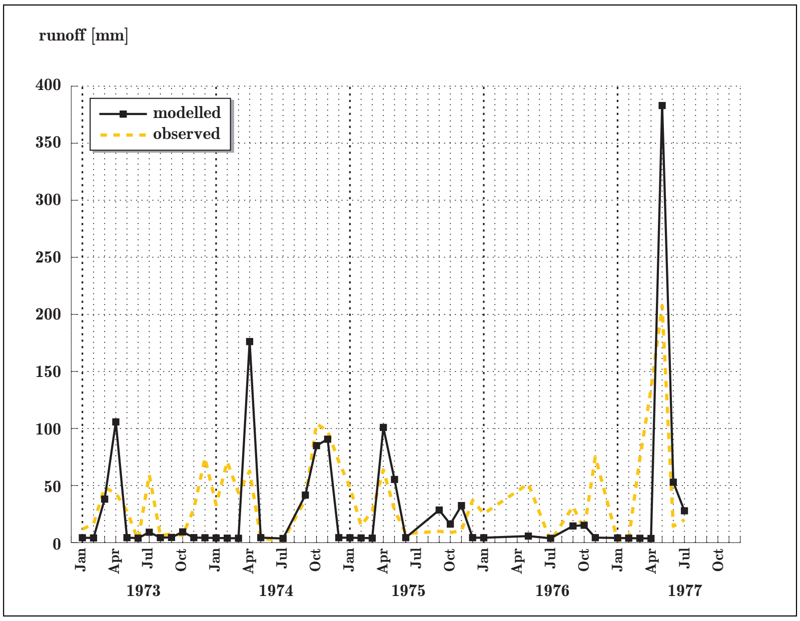


Fig. 3.17. Calculated and measured stream runoff, Kloten.

3.4.2 Comparison of Water Fluxes

Annual average values for the main hydrological fluxes, i.e. precipitation, evapotranspiration, surface runoff, infiltration and runoff, as well as for drainage and unsaturated flow within the soil were calculated. This allows a long-term balance as well as the comparison of observed climate parameters with modelled soil water fluxes.

The resultant values (Table 3.5) reveal an overestimation of stream runoff in sites Kindla ID 1 and 2, allowing for the conclusion that infiltration and consequently drainage as well as unsaturated flow are overrated. In contrast, runoff is underestimated for the other three locations, which is a sign of undervalued soil water fluxes.

Unsaturated flow values are much higher for sites Gammtratten and Kloten than for Kindla, whereas drainage behaves the other way around with greater values in Kindla site.

Comparing the drainage and unsaturated flow values of all layers within one site, it becomes clear that drainage is higher for upper than for lower layers, whereas the cumulative unsaturated flow component becomes larger for deeper layers (due to greater layer thickness and consequently higher water contents).

Table 3.5. Hydrological fluxes [*mm*], annual averages over years 2002-2006

	Flux	**GA 1**	**GA 2**	**KI 1**	**KI 2**	**KL**
surface	precipitation	710	710	872	872	717
	evapotranspiration	311	311	296	296	459
	surface runoff	23	23	20	61	22
	infiltration	384	384	546	424	311
layer 1	drainage	354	365	542	416	288
	unsaturated flow	29	19	3	7	20
layer 2	drainage	334	355	535	406	283
	unsaturated flow	43	28	6	10	24
layer 3	drainage	321	339	509	390	269
	unsaturated flow	54	40	28	18	35
layer 4	drainage	307	327	489	373	251
	unsaturated flow	65	49	43	30	46
layer 5	drainage	299	314	-	-	237
	unsaturated flow	72	60	-	-	59
stream	runoff (measured)	445	445	387	387	358
	runoff (simulated)	394	398	525	443	318

More detailed values were determined for years 2003 to 2006 for Kindla ID 1 (Table 3.6) as well as the other sites Gammtratten ID 1 and 2, Kindla ID 2 and Kloten site (Appendix E).

The above mentioned soil water characteristics apply here, too. Apart from year 2003, the runoff for Kindla site ID 1 is overrated for each year (Table 3.6), probably due to a too high estimated infiltration coefficient, resulting in high infiltration values and consequently overestimated soil water fluxes. The high amount of precipitation in 2004 and 2006 is reflected in the higher drainage and unsaturated flow values for those years, resulting in higher simulated runoff.

For the Gammtratten catchment (Appendix E.1), precipitation values were particularly high in 2005, leading to the same effect of higher soil water fluxes and stream runoff, which could also be observed for the wet years 1974 and 1977 in Kloten (Appendix E.3).

Table 3.6. Annual mean hydrological fluxes [mm], Kindla site ID 1.

	Flux	**2003**	**2004**	**2005**	**2006**
surface	precipitation	835	897	775	1,083
	evapotranspiration	263	246	264	305
	surface runoff	22	26	8	39
	infiltration	505	714	392	851
layer 1	drainage	502	708	389	844
	unsaturated flow	3	4	3	5
layer 2	drainage	499	702	387	835
	unsaturated flow	6	7	5	7
layer 3	drainage	476	670	368	801
	unsaturated flow	28	31	25	34
layer 4	drainage	459	646	354	773
	unsaturated flow	43	47	41	49
stream	runoff (measured)	552	379	268	510
	runoff (simulated)	524	718	403	737

3.5 Application to Heavy Metal Transport

3.5.1 Estimation of Mass Flow Rates

With the available data, heavy metal investigations could be done for ID 1 in sites Kindla (Table 3.7) and Gammtratten (Table 3.8).

Table 3.7. Heavy metal mass rates [$\mu g \cdot year^{-1} \cdot m^{-2}$] for Kindla, ID 1

Kindla, ID 1	**Cd**	**Pb**	**Hg**	**Zn**	**Cu**
Input (litter-, throughfall)	98	3,223	31	42,696	2,903
Drainage (30cm below humus)	68	369	-	6,831	428
Output (runoff)	31	158	1	3,168	79
Accumulation (entire soil)	67	3,065	30	39,528	2,824
Ratio output/input [%]	31.63	4.90	3.23	7.42	2.72
Ratio accumulation/input [%]	68.37	95.10	96.77	92.58	97.28
Ratio drainage/input [%]	69.39	11.45	-	16.00	14.74

Table 3.8. Heavy metal mass rates [$\mu g \cdot year^{-1} \cdot m^{-2}$] for Gammtratten, ID 1

Gammtratten, ID 1	**Cd**	**Pb**	**Hg**	**Zn**	**Cu**
Input (litter-,throughfall)	57	1,604	14	24,976	1,435
Drainage (30 cm below humus)	12	263	-	9,352	385
Output (runoff)	6	79	1	847	120
Accumulation (entire soil)	51	1,525	13	24,129	1,315
Ratio output/input [%]	10.53	4.93	7.14	3.39	8.36
Ratio accumulation/input [%]	89.47	95.07	92.86	96.61	91.64
Ratio drainage/input [%]	21.05	16.40	-	37.44	26.83

The values clarify a high mean accumulation of heavy metals per year. On an average, between 68.37% (cadmium) and 97.28% (copper) of the incoming metals were accumulating in the entire soil zone per year in Kindla site (Table 3.7), in Gammtratten (Table 3.8) it were between 89.47% (copper) and 96.61% (zinc). Unfortunately, only the total amount of input (litterfall and throughfall) was given (Fig. 2.6), but there were no adequate information about the chemical concentrations in the identified water fluxes (e.g. surface runoff and infiltration). Thus, it was impossible to estimate the exact amount of heavy metals that is actually entering the soil and to distinguish between accumulation in the upper soil zone (humus layer + 30cm) and in the part below. To calculate these values, detailed information about concentrations in surface runoff, infiltration as well as unsaturated flow would be required.

Based on the average values for sites Kindla (Table 3.7) and Gammtratten (Table 3.8), total heavy metal throughfall and litterfall per square meter and year in Kindla were more than twice as much as in Gammtratten. This confirms that there exists a pattern of lower deposition in northern and higher deposition in southern Sweden [Rühling and Tyler, 1973]. A gradient in the runoff output can not be recognised.

Considering the absolute amount of accumulation [$\mu g \cdot year^{-1} \cdot m^{-2}$], Kindla retains more metals in the soil compared to Gammtratten, which is in all likelihood caused by the higher deposition. But allowing for the percentage accumulation based on input, the values vary from metal to metal and site to site. Thus, it cannot be generalised that one of the investigated sites has a better or worse heavy metals retention. For instance, in Kindla the average accumulation of copper was 97.28% of the input and therefore much higher than in Gammtratten. Contrarily, in Gammtratten 89.47% cadmium were kept in the soil, which was approximately 21% more than in Kindla site.

Apart from a few exceptions considering cadmium, heavy metal litterfall and throughfall were constantly much higher than the stream output, causing high accumulation.

During the time period 2002 - 2006, circa $310\mu g \cdot m^{-2}$ cadmium (Fig. 3.18), about $14,305\mu g \cdot m^{-2}$ lead (Fig. 3.19) and $140\mu g \cdot m^{-2}$ mercury (Fig. 3.20) accumulated in the soil in Kindla site. Furthermore, approximately $184,466\mu g \cdot m^{-2}$ zinc as well as $13,179\mu g \cdot m^{-2}$ copper (Appendix F.2.1) retained in the soil during this time. The absolute amount of accumulating heavy metals for the same time period in site Gammtratten (Appendix F.2.2) is roughly half as much compared to Kindla site.

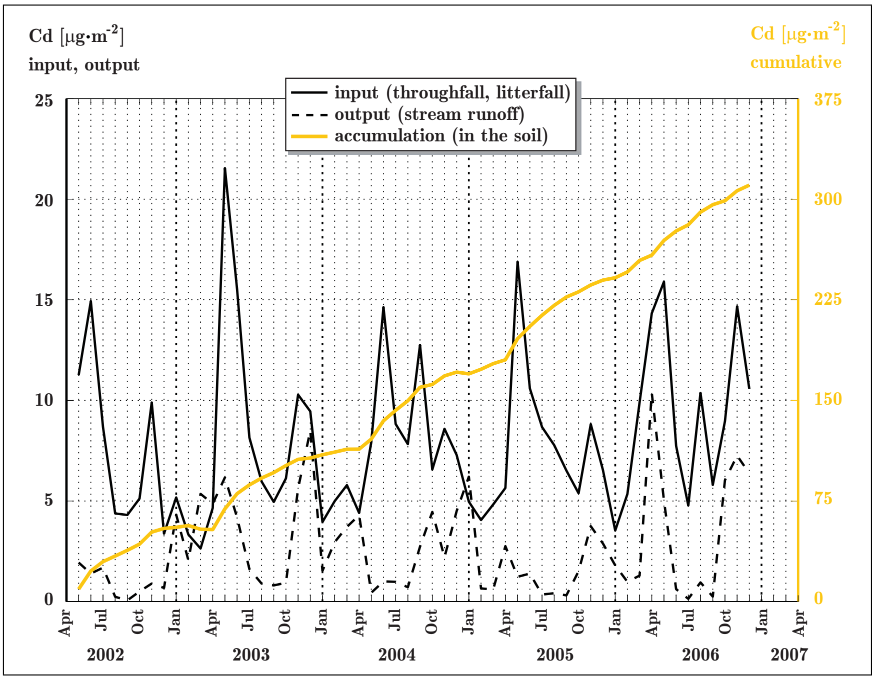


Fig. 3.18. Kindla: Cd input, output and accumulation.

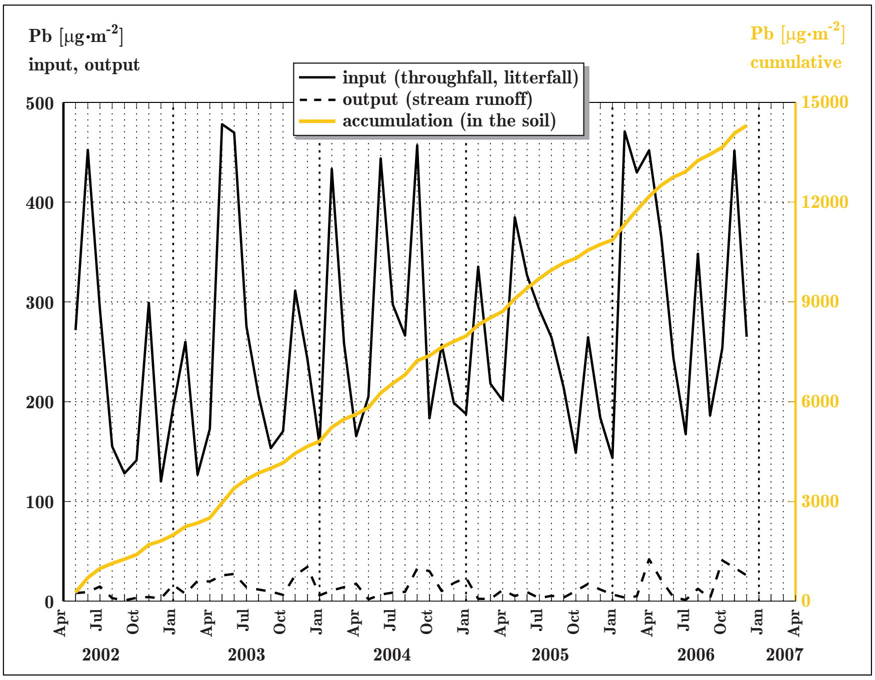


Fig. 3.19. Kindla: Pb input, output and accumulation.

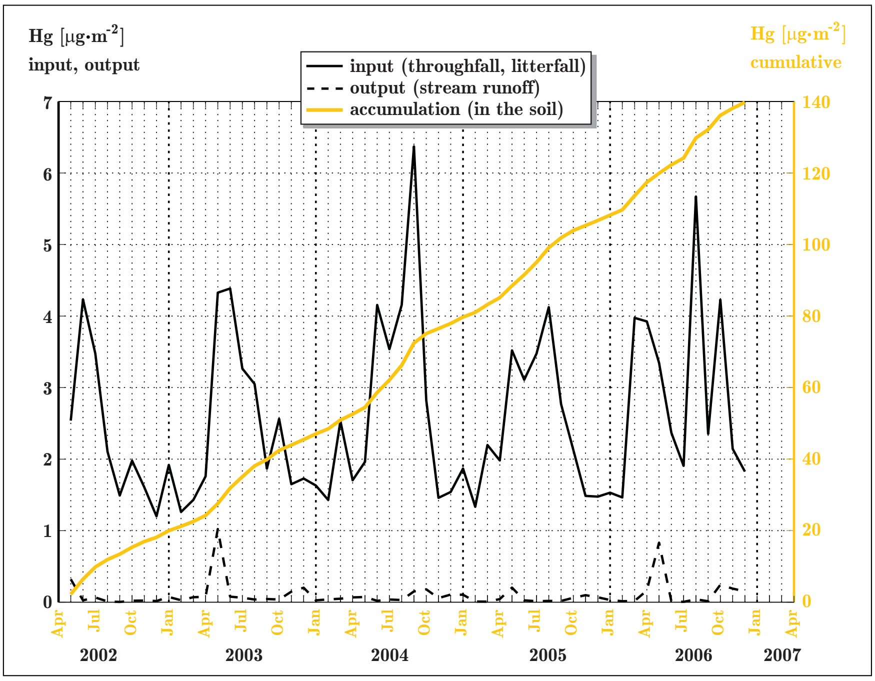


Fig. 3.20. Kindla: Hg input, output and accumulation.

There is a seasonal pattern in the absolute metal accumulation (Fig. 3.18 - 3.20). The absolute difference of input to the catchment and output through the stream is much higher for months April to October than it is for the time period November to March, which might possibly be a result of different climate as well as vegetation characteristics.

Accumulation values per square meter for the time interval 2002 - 2006 are already remarkable. Moreover, average accumulations per year of all investigated metals were calculated for the entire catchment area (Table 3.9) for Kindla ($0.191km^2$) as well as Gammtratten ($0.45km^2$).

Especially zinc values are very high and an extrapolation for a time period of 20 years would sum up to approximately $151kg$ (Kindla) and $217kg$ (Gammtratten) zinc spread over the entire catchment area. But also the other substances add up to high values based on long-term considerations, which will result in disturbed soil chemistry and biology processes.

Table 3.9. Heavy metal accumulation in the whole catchment

Site	**Cd**	**Pb**	**Hg**	**Zn**	**Cu**
Kindla $[g \cdot year^{-1}]$	12.71	585.47	5.72	7,549.94	539.39
Gammtratten $[g \cdot year^{-1}]$	23.05	686.42	5.82	10,857.84	591.72
Kindla $[kg \cdot (20years)^{-1}]$	0.25	11.71	0.11	151.00	10.73
Gammtratten $[kg \cdot (20years)^{-1}]$	0.46	13.73	0.12	217.16	10.79

3.5.2 Problems and Inaccurate Approaches

It is important to keep in mind that all heavy metal values and calculations are based on rather incomplete chemical data. It was difficult to find a proper way to combine the given values that were, for instance, measured with varying time lag or obtained for totally different time periods. The best solution was to work with averages and interpolated values that are definitely a source of impreciseness. In addition, not only the obtained chemical values were subject to error, but also the modelled soil water fluxes should be considered skeptically, because WATBAL cannot represent reality free of errors - as discussed in section 3.3. Thereby, all calculation outcomes based on the combination of faulty values cannot be trusted implicitly, but further heavy metal soil studies are suggested in order to validate these results.

Chapter 4

Summary and Conclusion

4.1 WATBAL Evaluation

The objectives of this study were to analyse and improve the given soil water model WATBAL as well as to investigate its accuracy. Furthermore, WATBAL was used to model soil moisture over time for different depths within the rooting zone of three different Swedish sites. The last challenge was to calculate chemical fluxes of certain heavy metals (cadmium, lead, mercury, zinc and copper) on basis of the model results.

The model outcome of the original WATBAL version was not very good, so that some improvements and test runs were accomplished in order to enhance the model fit. With the available data from central and northern Swedish till soils, the soil moisture was simulated with the corrected model for different layers within the rooting zone for a period of approximately five years.

As expected, there were negative correlations detected for soil moisture and temperature as well as soil moisture and global radiation. Correlation between soil moisture and precipitation was positive.

After application of numerical and graphical methods to determine the accuracy of the corrected model, WATBAL was rated as satisfactory and a perfectly suitable soil water model. The modelling results were not outstanding with R^2 values between 0.13 and 0.58 (on average 0.23) as well as MAPE values from 0.02 to 0.21 (averaged 0.09), but they showed that WATBAL simulations in general cover main seasonal patterns in soil moisture as well as noticeable peaks. The overall average percentage error was estimated to be $8.91\% \pm 1.08\%$.

For that reason, WATBAL is recommended for future studies in order to simulate soil moisture and soil hydrological fluxes. Since the model fit varied a lot from site to site, it is suggested to carry out further testing and model validation with additional measurement sites and soil types to check WATBAL's behaviour and to be able to give a more accurate statement about the model precision. It might also be beneficial to investigate WATBAL's accuracy based on different seasons.

4.2 Heavy Metal Analysis

Heavy metal investigations were carried out to get an overview of transport and accumulation behaviour of cadmium, lead, mercury, zinc and copper.

Many difficulties with incomplete data had to be overcome and inaccuracies in the approaches had to be accepted. But the final values and graphs revealed a couple of interesting facts, for instance, information about the scale of enrichment values. Regarding the accumulation in microgram per year and square meter, the ascending order of metal accumulation in the soil was mercury, cadmium, copper, lead and zinc, with zinc values more than 1000 times higher than mercury. Concerning percentage accumulation based on the amount of heavy metal input, cadmium was the metal with the lowest retention (between 68.37% and 89.47%), whereas the other elements had varying but generally high accumulation values between 91.64% and 97.28%.

Furthermore, the results showed greater litterfall, throughfall and corresponding absolute accumulation values for Kindla than for Gammtratten. Based on this study, the preliminary conclusion would be: the more heavy metals are introduced to the system, the more are remained in the soil, which is the reason for higher absolute heavy metal contents in the soil in Kindla. But nevertheless, the accumulation/input ratio did not indicate a comprehensible connection between percentage accumulation rate and site location nor amount of input.

Another discovery was the relation of monthly absolute accumulation in the soil and time of the year. Absolute accumulation had much greater values for April to October than for November to March.

In addition, it was detected that a huge amount of metals remained in the soil during the investigated time interval 2002 - 2006. Based on long-term considerations, the amount of accumulated heavy metals spread over the whole catchment can even add up to several kilograms. It was estimated that up to $11.71kg$ ($13.73kg$) lead and $110g$ ($120g$) mercury would accumulate in the entire catchment within 20 years in site Kindla (Gammtratten).

All things considered, the results of the accomplished heavy metal analysis are a good starting point for future investigations. Basic information were created and it is recommended to perform prospective studies in order to check, improve and specify the current findings.

Appendix A

Important WATBAL Variables

Variable	Unit	Description
a	[-]	temperature coefficient
awc	[mm]	availale water capacity
b	[-]	temperature coefficient
bypassflow	[mm]	surface runoff
c		matrix with modelled and simulated SM
cali_filename		filename of the file with calibration data
canopy	[-]	fraction of forest cover
corr_coeff	[-]	matrix that contains the correlation coefficients
delta_sm	[mm]	difference between *sm* and *smt_1*
delta_sog	[mm]	difference between *sog* and *sog_1*
doy	[days]	day of year that represents a specific month
drainage	[mm]	soil water drainage flux
e1	[kPa]	saturation vapor pressure at *july_tmin*
e2	[kPa]	saturation vapor pressure at *july_tmax*
elevation	[m a.s.l.]	elevation of the site
equation		number of used soil moisture equation
et_0	[mm]	reference potential evapotranspiration
et_c	[mm]	forest potential evapotranspiration
et_c_adj	[mm]	actual evapotranspiration
et_deficit	[mm]	monthly evapotranspiration deficit
g		matrix with modelled and simulated runoff
infil_coeff	[-]	infiltration coefficient
infiltration	[mm]	amount of infiltration

Variable	Unit	Description
initial_sm	[mm]	initial SM for the month before simulation
initial_sog	[mm]	initial SOG for the month before simulation
input_filename		filename of climate input file
j	[-]	number of current layer
july_tmax	[°C]	long-term mean maxi. air temperature for July
july_tmin	[°C]	long-term mean min. air temperature for July
kc	[-]	crop coefficient
latitude	[°]	decimal latitude of site
layer_number	[-]	number of layers
lh_vap	$[\mathrm{MJ} \cdot \mathrm{kg}^{-1}]$	latent heat of vaporisation
lh_vap_ep	$[\mathrm{MJ} \cdot \mathrm{m}^{-2} \cdot \mathrm{month}^{-1}]$	modified Jensen-Haise ET_0 (alfalfa)
lin_reg		slope and y-intercept of regression line
mape	[-]	MAPE (mean absolute percentage error)
matric_losses	[mm]	unsaturated flow
matrix_loss	[-]	factor to determine *matric_losses*
meanabserror		MAE (mean absolute error)
melt_coeff	$[\mathrm{mm} \cdot {}^{\circ -1} \cdot \mathrm{day}^{-1}]$	melt coefficient
numdays	[days]	number of days for a specific month
p_et_c	[mm]	difference between precipitation and *et_c*
pot_snowmelt	[mm]	potential amount of snowmelt
pot_wl	[mm]	potential soil water loss
precipitation	[mm]	precipitation (e.g. rain and snow)
radiation	$[\mathrm{MJ} \cdot \mathrm{m}^{-2} \cdot \mathrm{month}^{-1}]$	global radiation
rain_plus_snowmelt	[mm]	sum of precipitation and snowmelt
rd	[-]	diffuse radiation tilt factor
results_help		help file
rmse		RMSE (root mean square error)
root	[-]	fraction of roots in layer
rsquare	[-]	R^2 value
runoff_filename		filename of runoff input file
runoff_real	[mm]	measured runoff

Variable	**Unit**	**Description**
runoff_sim	[mm]	simulated runoff
runtime	[months]	model run-time
save_filename		filename for saving simulation results
season	[-]	seasonal coefficient
sm	[mm]	soil moisture
sm_crit	[-]	critical SM/AWC ratio controlling AET
sm_deficit	[mm]	soil moisture deficit
sm_fc	[mm]	SM content at field capacity
sm_pwp	[mm]	SM content at permanent wilting point
sm_rate	[-]	multiplier for SM/AWC
*smt_*1	[mm]	soil moisture at the end of previous month
snowmelt	[mm]	amount of snowmelt over month
sog	[mm]	snow on the ground
*sogt_*1	[mm]	SOG at the end of previous month
statisticsmatrix		matrix with statistical parameters
temperature	[°C]	air temperature
total_sm		matrix that contains soil moistures for all layers

Appendix B

Structograms

To establish a better understanding of the WATBAL model and its implementation, structograms (Nassi-Shneiderman diagrams) of the programme are listed below for explanatory purpose. They are intended solely to show the program sequence and to clarify the model's structure and its equations in a comprehensible way.

The main programme is devided into three subsystems: snow routine, radiation routine and main routine, containing further subsystems.

Main Programme

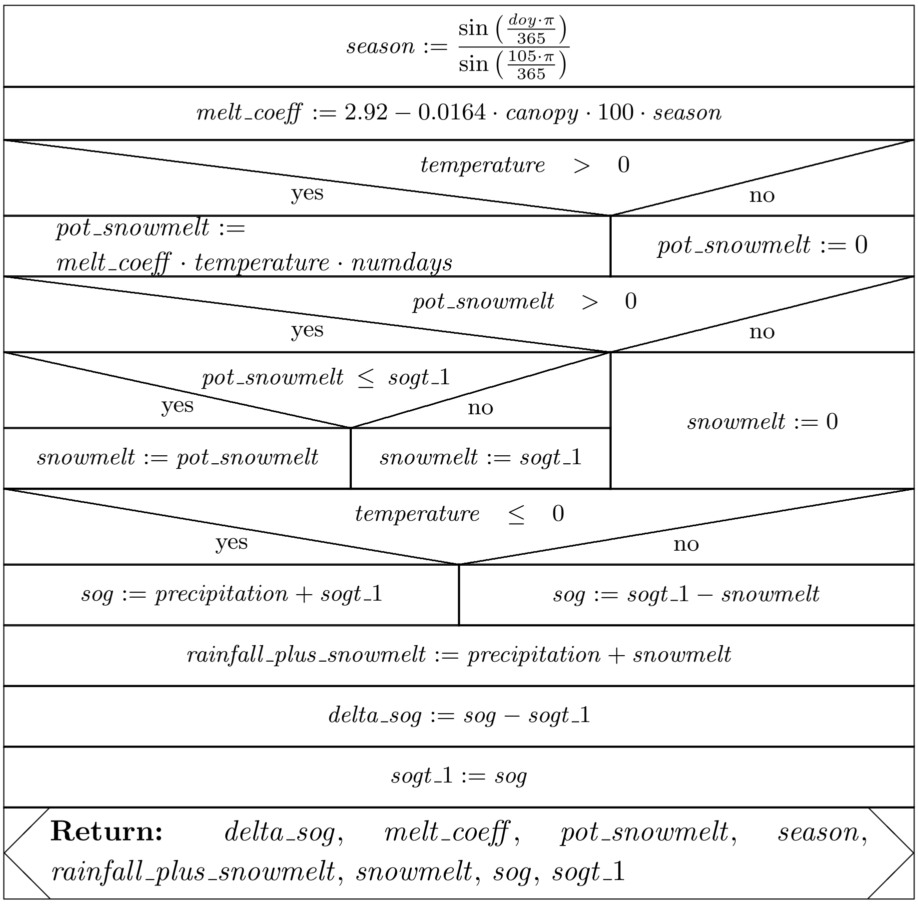
Load file: Climate data

Input: *canopy*, *elevation*, *infil_coeff*, initial *smt_1*, initial *sogt_1*, *july_tmax*, *july_tmin*, *kc*, *latitude*, *matrix_loss*, *rd*, *root*, *sm_fc*, *sm_pwp*, *sm_crit*, *sm_rate* and more

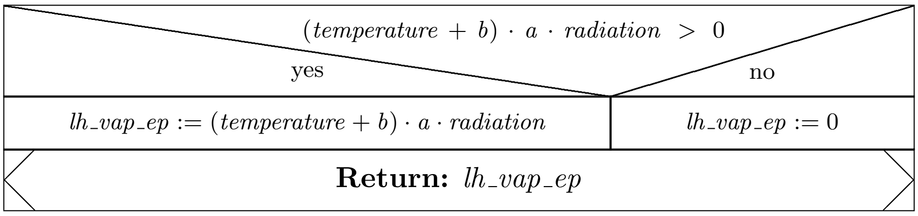
for $t = 0$ to n by $+1$ do

- **Read:** *radiation*(n), *doy*(n), *numdays*(n), *precipitation*(n), *temperature*(n)
- Subsystem Snow Routine
- Subsystem Radiation Routine
- Subsystem Main Routine

Subsystem Snow Routine



Subsystem Radiation Routine



Subsystem Main Routine

$awc := sm\_fc - sm\_pwp$

If $temperature > 0$

- yes:
  - $lh\_vap := 2.501 - 0.002361 \cdot temperature$
  - $et\_0 := \left(\frac{lh\_vap\_ep}{lh\_vap}\right) \cdot 1.13$
  - $et\_c := kc \cdot et\_0$
- no:
  - $lh\_vap := 0$
  - $et\_0 := 0$
  - $et\_c := 0$

$p\_et\_c := precipitation - et\_c$

If $(p\_et\_c + snowmelt) \cdot (1 - infil\_coeff) > 0$ **AND** $snowmelt > 0$

- yes: $bypassflow := (p\_et\_c + snowmelt) \cdot (1 - infil\_coeff)$
- no: $bypassflow := 0$

SUBSUBSYSTEM POT_WL

SUBSUBSYSTEM Infiltration

SUBSUBSYSTEM Soil Moisture

If $sm \geq sm\_pwp$

- yes: $matrix\_losses := (smt\_1 + infiltration) \cdot (1 - matrix\_loss)$
- no: $matrix\_losses := 0$

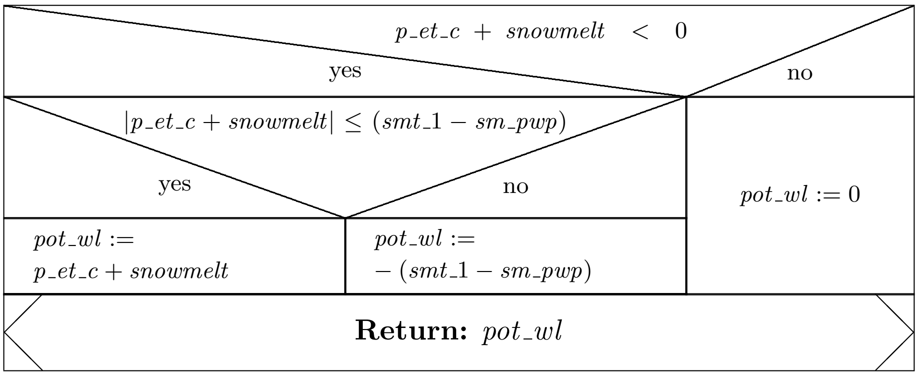
SUBSUBSYSTEM et_c_adj

If $et\_c - et\_c\_adj \geq 0$

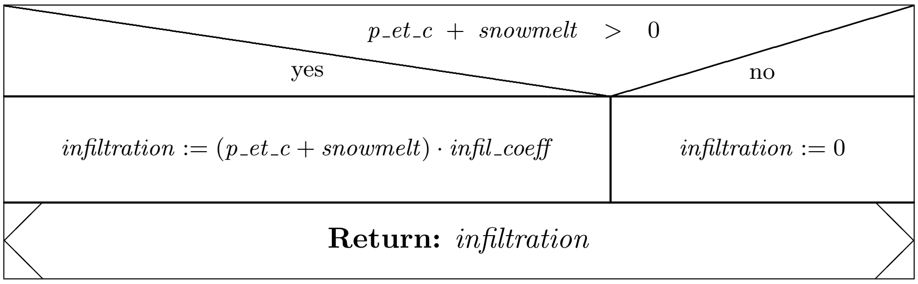
- yes: $et\_deficit := -(et\_c - et\_c\_adj)$
- no: $et\_deficit := 0$

Write: *lh_vap*, *et_0*, *et_c*, *p_et_c*, *pot_wl*, *infiltration*, *bypassflow*, *sm*, *delta_sm*, *sm_deficit*, *smt_1*, *matric_losses*, *et_c_adj*, *et_deficit*, *drainage* and more

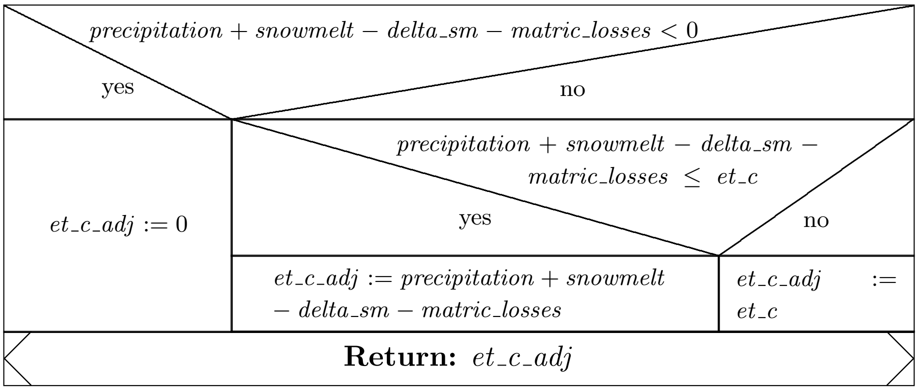
Subsubsystem POT_WL



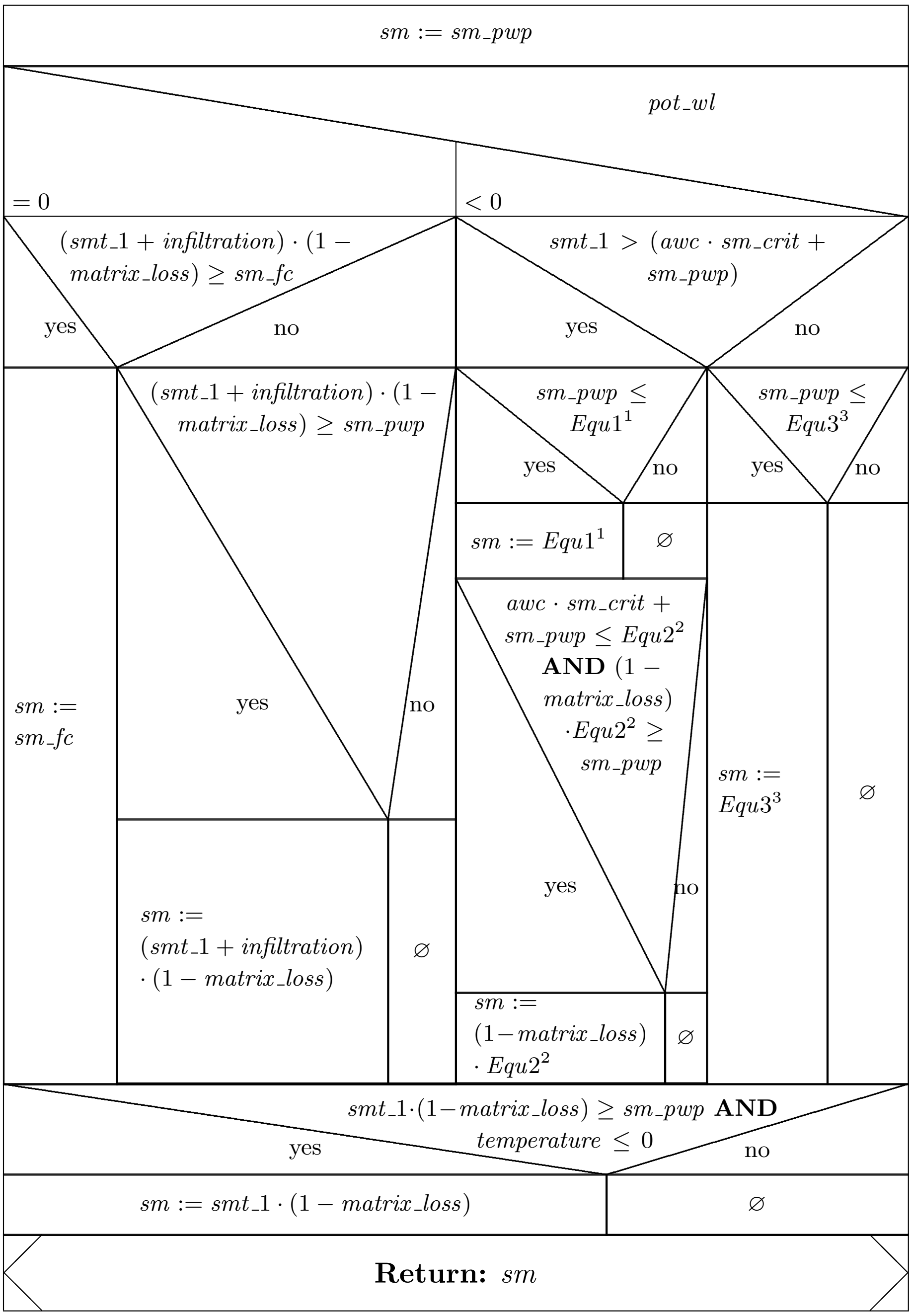
Subsubsystem Infiltration



Subsubsystem et_c_adj



Subsubsystem Soil Moisture



[1] $Equ1 = (smt\_1 - ((smt\_1 - (awc \cdot sm\_crit + sm\_pwp)) + (((awc \cdot sm\_crit + sm\_pwp)$
$-(smt\_1 - |pot\_wl| \cdot root)) \cdot sm\_rate))) \cdot (1 - matrix\_loss)$

[2] $Equ2 = smt\_1 - |pot\_wl| \cdot root$

[3] $Equ3 = (smt\_1 - (sm\_rate \cdot |pot\_wl| \cdot root)) \cdot (1 - matrix\_loss)$

Appendix C

MATLAB Source Codes

Based on structures of the original WATBAL implementation in *ModelMaker*©, the programme is controlled in a similar way in *MATLAB*®. One main file, called watbal.m is used, to define main parameters that are valid for all different layers within one measurement site. The main programme calls subprogrammes, so-called functions, that are saved in separate *.m-files. There are six functions, named as follows: The first files are snowroutine.m, radiationroutine.m and maincalculations; as the names say they contain the routines that are the main parts of the WATBAL model. The other files are runoff.m, parameter_def.m and graphics.m, estimating the runoff of the catchment, defining the layerspecific parameters and including the code for several plots. In the following section, the first part of the source code is listed (file watbal.m). Codes of the other *.m-files are not shown in order to protect data privacy.

```
%% WATBALMODEL (c) 2007/2008
%  - developed by Mike Starr, University Helsinki, Finland
%  - written in Matlab by Claudia Teutschbein, SLU Uppsala, Sweden

%% BEGIN
clear   clc         % clear workspace and command window

%% INPUT PARAMETERS (Kindla)
% define the site input parameters
canopy=0.8;         % CANOPY: fraction of forest cover (0=no trees, 1=totally covered)
elevation=376;      % ELEVATION: elevation of the site [meters above sea level]
july_tmax=20.7;     % MAX. TEMPERATURE JULY: Long-term mean maximum air
                    % temperature for July [°C]
july_tmin=9.8;      % MIN. TEMPERATURE JULY: Long-term mean minnimum air
                    % temperature for July [°C]
kc=1.3;             % CROP COEFFICIENT: crop specific evapotranspiration
                    % values, derived from research
latitude=59.75;     % LATITUDE: Decimal latitude of site
rd=1;               % DIFFUSE RADIATION TILT FACTOR: Rd = 1 for flat sites
                    % (Rd is only used for cloudiness data)
initial_sog=0;      % INITIAL SNOW ON THE GROUND: SOG for the month before
                    % simulation start
runtime=56;         % model run-time [months]: specifies how many months
                    % the model will run
input_filename='climate_kindla.mat';
                    % define filename of climate input file (with required
                    % climate values, such as temperature, radiation, etc.)
runoff_filename='kindla_runoff.mat';
                    % define filename of runoff input file, that contains
                    % the measured runoff data for comparison purposes
```

```
%% LOADING CLIMATE PARAMETERS
% load the climate input file and save as variable "input"
load(input_filename); input=ans;
% preparation: splitting the input climate data into several matrices (different variables)
doy(1,:)=input(1,:); doy(2,:)=input(4,:);
                     % DOY is the "day of year" that represents a specific
                     % month (e.g. 17 for January and 344 for December)
numdays(1,:)=input(1,:); numdays(2,:)=input(5,:);
                     % number of days for a specific month (e.g. 31 for
                     % January, 28 (29) for February and 30 for November)
temperature(1,:)=input(1,:); temperature(2,:)=input(6,:);
precipitation(1,:)=input(1,:); precipitation(2,:)=input(8,:);
radiation(1,:)=input(1,:); radiation(2,:)=input(7,:);
sog(1,:)=input(1,:); sogt_1(2,1)=initial_sog; sm(1,:)=input(1,:);

%% WATBALMODEL - BEGIN OF SIMULATION
% calculation of basic parameters required for snow and radiation routine
e1=exp((16.78*july_tmin -116.7)/(july_tmin+237.3));
e2=exp((16.78*july_tmax -116.7)/(july_tmax+237.3));
a=1/(38-(2*elevation/305)+36.5/(e2-e1));
b=2.5+1.4*(e2-e1)+elevation/550;
% CALL SNOW ROUTINE
[season,melt_coeff,pot_snowmelt,snowmelt,sog,delta_sog,sogt_1,...
                     rain_plus_snowmelt]=snowroutine(runtime,doy,canopy,...
                     temperature,numdays,sogt_1,precipitation);
% CALL RADIATION ROUTINE
[lh_vap_et]=radiationroutine(a,temperature,runtime,b,radiation);
for j=1:4        % run simulation for layer 1 to 4
% define the layer specific input parameters, call parameter_def.m
    [infil_coeff,output_filename,save_filename,cali_filename,...
                     matrix_loss,root,sm_fc,sm_pwp, sm_crit,...
                     sm_rate,initial_sm,layer_number]=parameter_def(j);
    smt_1(2,1)=initial_sm;    % initial soil moisture of month
                              % before simulation start
    awc=sm_fc-sm_pwp;         %availale water capacity
% MAIN CALCULATIONS
    [lh_vap,et_0,et_c,p_et_c,pot_wl,infiltration,bypassflow,sm,...
                     equation,delta_sm,sm_deficit,smt_1,matric_losses,...
                     et_c_adj,et_deficit,drainage]...
                     =maincalculations(temperature,infil_coeff,runtime,...
                     lh_vap_et,kc,precipitation,delta_sog,snowmelt,smt_1,...
                     sm_pwp,matrix_loss, sm_fc,awc,sm_crit,root,sm_rate);
    total_sm(1,:)=sm(1,:);    % creates a matrix that contains modelled soil
    total_sm(j+1,:)=sm(2,:); % moisture for each layer; row 1 = control variable,
                              % row 2 = layer 1, row 3 = layer 2 ...
%STATISTICAL ANALYSIS
[F,y2,R,c,meanabserror,residualmeansquare,mape,cali,rsquare,mce,rmsr]=...
                      statistics(runtime,total_sm,cali_filename,j);
% call graphics.m that creates certain data plots
    [cali]=graphics(y2,c,j,total_sm,cali_filename,runtime);
% call runoff.m, that generates the simulated runoff and compares it with
% real measurements in the catchment outlet
    [runoff_meanabserror,runoff_residualmeansquare,runoff_mape,g,...
                     runoff_real,runoff_sim]...
                     = runoff(j,runoff_filename,runtime,precipitation,...
                     et_c,delta_sog,p_et_c,snowmelt,infil_coeff,delta_sm,...
                     drainage,bypassflow,matric_losses);
save(save_filename,'p_et_c','snowmelt','F','R','rsquare','runoff_sim',...
                     'g','delta_sog','smt_1','bypassflow','et_c_adj',...
                     'mape','sm','cali','c','meanabserror',...
                     'residualmeansquare','canopy','drainage','elevation',...
                     'infiltration','infil_coeff','july_tmax','july_tmin',...
                     'kc','matric_losses','matrix_loss','root','sm_fc',...
                     'sm_pwp','sm_crit','sm_rate','initial_sm',...
                     'initial_sog','input_filename','output_filename',...
                     'delta_sm','sm','total_sm','save_filename', ...
                     'latitude','total_sm','mce','rmsr')
                              % save file -> contains all necessary data that
                              % can be important for later analysis
statisticsmatrix(j,1)=rsquare;                % R2
statisticsmatrix(j,2)=mape;                   % MAPE
statisticsmatrix(j,3)=residualmeansquare;     % MSE
statisticsmatrix(j,4)=meanabserror;           % MAE
save('statistics_layer.mat','statisticsmatrix')          %saving statistic parameters
end
```

Appendix D

WATBAL Testing Results

D.1 Prior WATBAL Version

Table D.1. Statistics Gammtratten (original model)

Profile-ID	**Layer**	**R^2**	**MAPE**	**RMSE**	**MAE**
ID 1	*Layer 1*	0.3039	0.1639	2.6492	2.2865
	Layer 2	0.3920	0.0979	6.8515	4.2293
	Layer 3	0.2084	0.0852	8.4595	5.0859
	Layer 4	0.1960	0.0842	11.346	6.6892
	Layer 5	0.1410	0.0773	13.119	6.8949
ID 2	*Layer 1*	0.1121	0.0893	1.0647	0.6037
	Layer 2	0.2071	0.0926	2.9764	2.1158
	Layer 3	0.1863	0.2011	9.4153	8.7114
	Layer 4	0.1648	0.1918	12.5510	11.7580
	Layer 5	0.1554	0.0538	7.7882	4.9411

Table D.2. Statistics Kindla (original model)

Profile-ID	**Layer**	**R^2**	**MAPE**	**RMSE**	**MAE**
ID 1	*Layer 1*	0.5843	0.0152	0.3488	0.1805
	Layer 2	0.4621	0.0167	1.6326	0.8341
	Layer 3	0.2361	0.0740	6.8604	4.5673
	Layer 4	0.1293	0.0345	6.9171	4.4669
ID 2	*Layer 1*	0.1398	0.0359	1.0371	0.6536
	Layer 2	0.1477	0.1287	7.0967	5.1761
	Layer 3	0.1612	0.0766	11.1130	7.2880
	Layer 4	0.1678	0.0592	19.3010	10.9400

Table D.3. Statistics Kloten (original model)

Profile-ID	Layer	R^2	MAPE	RMSE	MAE
ID 1	*Layer 1*	0.2347	0.2226	4.5848	3.4682
	Layer 2	0.2340	0.1321	4.5891	3.4396
	Layer 3	0.1939	0.0846	6.5703	4.5272
	Layer 4	0.2029	0.1063	10.7200	7.8060
	Layer 5	0.1486	0.0673	10.7130	7.9478

D.2 Improved WATBAL Version

Table D.4. Statistics Gammtratten (improved model)

Profile-ID	Layer	R^2	MAPE	RMSE	MAE
ID 1	*Layer 1*	0.3075	0.1771	2.7241	2.4388
	Layer 2	0.3743	0.0854	6.1533	3.6958
	Layer 3	0.1956	0.0723	7.6131	4.3455
	Layer 4	0.1846	0.0706	10.1610	5.6497
	Layer 5	0.1283	0.0724	12.2530	6.4620
ID 2	*Layer 1*	0.1596	0.0891	1.0457	0.6393
	Layer 2	0.2339	0.0773	2.7942	1.8429
	Layer 3	0.1688	0.1776	8.5493	7.8384
	Layer 4	0.1737	0.1683	11.4400	10.4920
	Layer 5	0.1980	0.0474	9.6501	4.6022

Table D.5. Statistics Kindla (improved model)

Profile-ID	**Layer**	$\mathbf{R}^2$	**MAPE**	**RMSE**	**MAE**
ID 1	*Layer 1*	0.5843	0.0152	0.3488	0.1805
	Layer 2	0.4661	0.0168	1.6416	0.8372
	Layer 3	0.2567	0.0728	6.7255	4.4973
	Layer 4	0.1464	0.0338	6.8001	4.3928
ID 2	*Layer 1*	0.1500	0.0355	1.0240	0.6467
	Layer 2	0.1535	0.1288	7.0621	5.1822
	Layer 3	0.1693	0.0767	11.0830	7.2976
	Layer 4	0.1757	0.0588	19.1810	10.8890

Table D.6. Statistics Kloten (improved model)

Profile-ID	**Layer**	$\mathbf{R}^2$	**MAPE**	**RMSE**	**MAE**
ID 1	*Layer 1*	0.2519	0.2166	4.4914	3.3708
	Layer 2	0.2374	0.1258	4.5114	3.2663
	Layer 3	0.1840	0.0880	6.7266	4.7183
	Layer 4	0.1968	0.1052	10.7470	7.7345
	Layer 5	0.1431	0.0700	11.0520	8.2579

Appendix E

Hydrological Fluxes

E.1 Gammtratten Site

Table E.1. Annual mean hydrological fluxes [mm], Gammtratten site ID 1.

	Flux	**2003**	**2004**	**2005**	**2006**
surface	precipitation	610	695	825	575
	evapotranspiration	320	294	318	508
	surface runoff	25	25	21	37
	infiltration	323	394	436	372
layer 1	drainage	295	363	404	337
	unsaturated flow	25	30	32	28
layer 2	drainage	267	347	389	305
	unsaturated flow	37	45	47	41
layer 3	drainage	251	335	378	286
	unsaturated flow	47	56	59	52
layer 4	drainage	234	322	366	273
	unsaturated flow	57	68	70	62
layer 5	drainage	224	315	358	267
	unsaturated flow	63	75	78	69
stream	runoff (measured)	427	406	502	383
	runoff (simulated)	311	414	457	372

Table E.2. Annual mean hydrological fluxes [mm], Gammtratten site ID 2.

	Flux	**2003**	**2004**	**2005**	**2006**
surface	precipitation	610	695	825	575
	evapotranspiration	320	294	318	508
	surface runoff	25	25	21	37
	infiltration	323	394	436	372
layer 1	drainage	305	374	415	350
	unsaturated flow	17	20	22	19
layer 2	drainage	294	364	406	336
	unsaturated flow	25	28	30	27
layer 3	drainage	274	351	393	313
	unsaturated flow	36	41	43	39
layer 4	drainage	259	340	383	296
	unsaturated flow	44	51	53	48
layer 5	drainage	242	328	371	277
	unsaturated flow	54	62	65	59
stream	runoff (measured)	427	406	502	383
	runoff (simulated)	321	415	457	373

E.2 Kindla Site

Table E.3. Annual mean hydrological fluxes [mm], Kindla site ID 2.

	Flux	**2003**	**2004**	**2005**	**2006**
surface	precipitation	835	897	775	1,083
	evapotranspiration	263	246	264	305
	surface runoff	65	78	25	117
	infiltration	393	555	305	662
layer 1	drainage	386	545	299	650
	unsaturated flow	6	8	5	9
layer 2	drainage	382	534	296	633
	unsaturated flow	10	12	9	12
layer 3	drainage	374	517	289	603
	unsaturated flow	18	19	17	20
layer 4	drainage	358	503	277	578
	unsaturated flow	30	32	29	32
stream	runoff (measured)	552	379	268	510
	runoff (simulated)	453	612	331	628

E.3 Kloten Site

Table E.4. Annual mean hydrological fluxes [mm], Kloten.

	Flux	**1973**	**1974**	**1975**	**1976**	**1977**
surface	precipitation	642	738	593	705	905
	evapotranspiration	490	453	505	435	411
	surface runoff	16	20	17	14	46
	infiltration	200	424	262	178	492
layer 1	drainage	180	396	240	160	463
	unsaturated flow	17	24	19	16	27
layer 2	drainage	178	389	234	154	458
	unsaturated flow	21	28	22	20	31
layer 3	drainage	166	374	219	141	444
	unsaturated flow	32	38	33	30	41
layer 4	drainage	140	359	202	127	430
	unsaturated flow	43	49	45	42	53
layer 5	drainage	124	343	187	114	415
	unsaturated flow	56	62	57	55	66
stream	runoff (measured)	333	530	262	202	462
	runoff (simulated)	196	425	261	183	523

Appendix F

Graphics

F.1 Simulated Soil Moisture

This section contains the plots of the WATBAL simulated soil moistures for all sites other the Kindla ID 1, which is already presented in subsection 3.1.1. Each plot includes the soil moistures for all layers within the respective site.

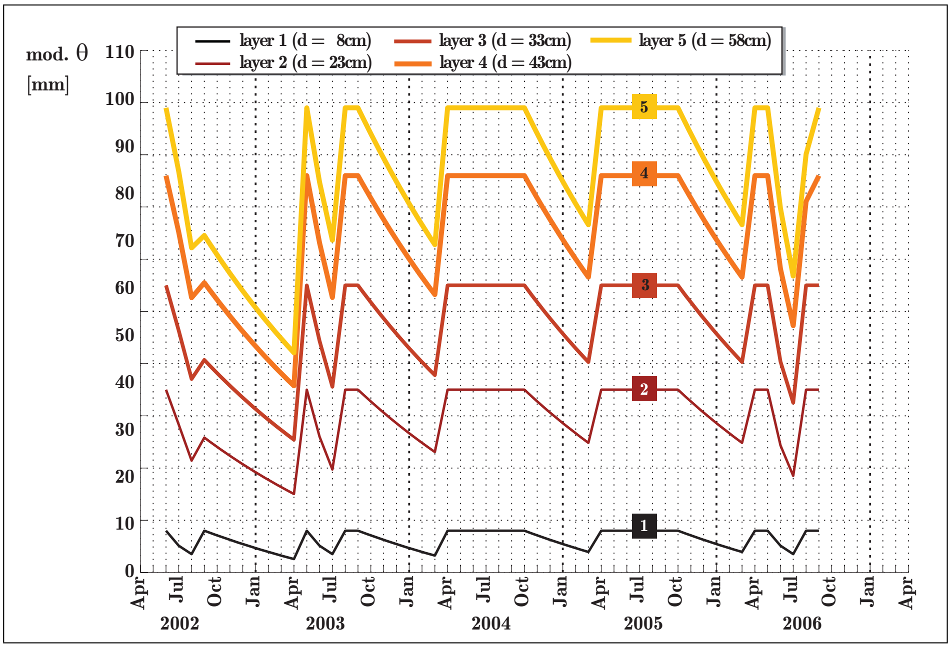


Fig. F.1. Modelled soil moisture for layers 1-5 in Gammtratten ID 1.

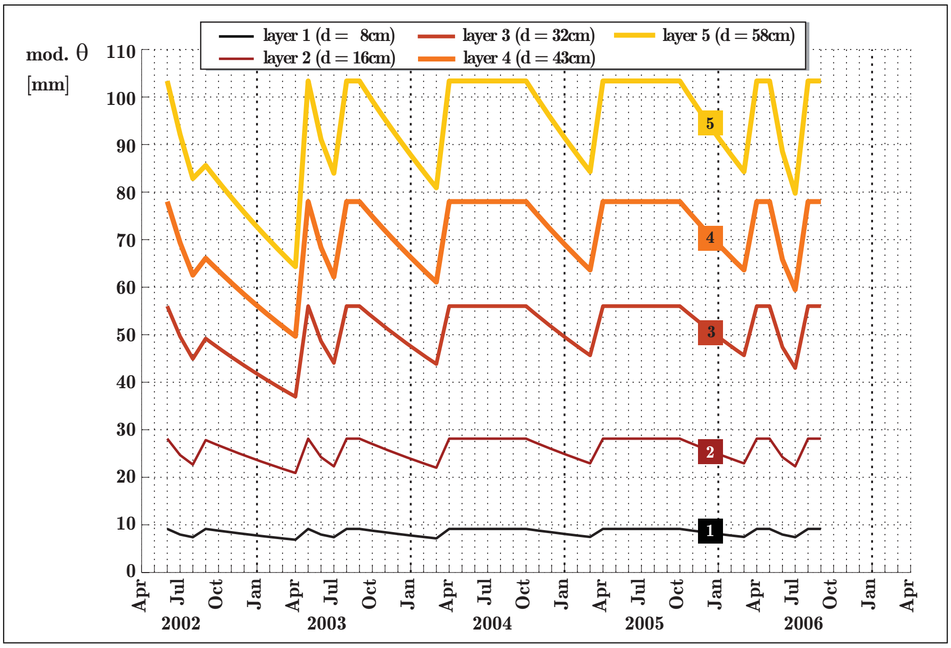


Fig. F.2. Modelled soil moisture for layers 1-5 in Gammtratten ID 2.

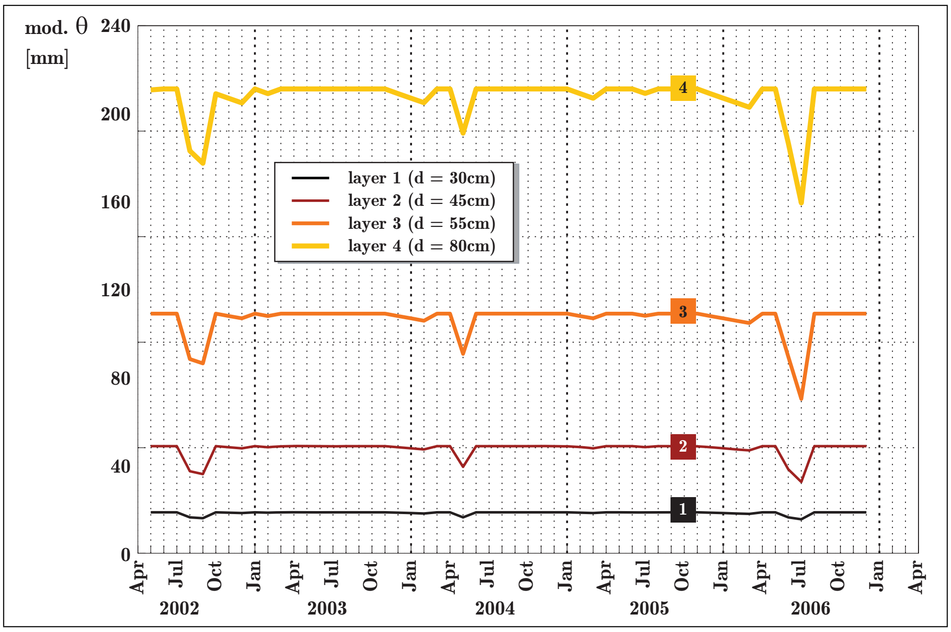


Fig. F.3. Modelled soil moisture for layers 1-4 in Kindla ID 2.

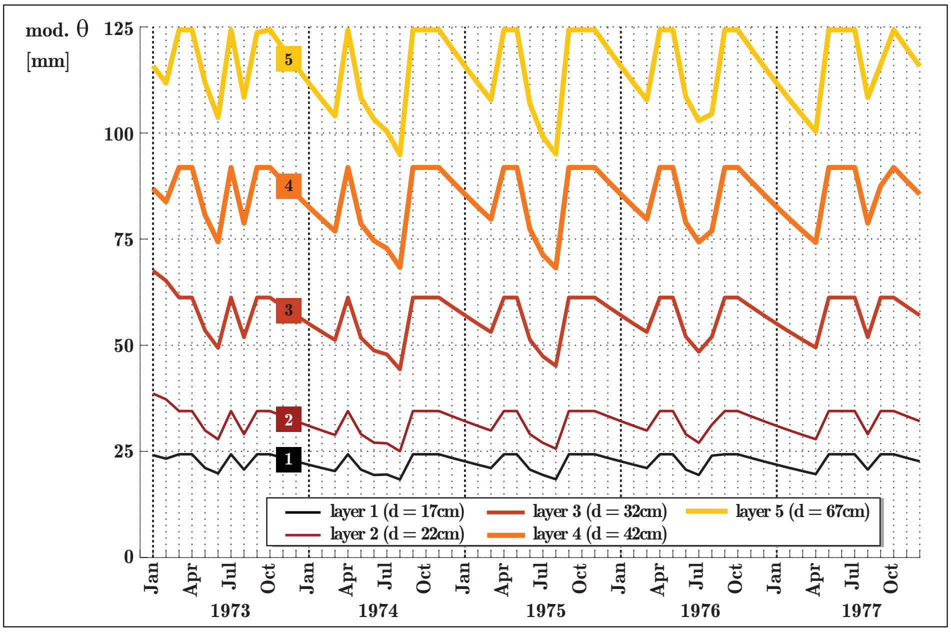


Fig. F.4. Modelled soil moisture for layers 1-5 in Kloten.

F.2 Heavy Metal Balance

F.2.1 Kindla

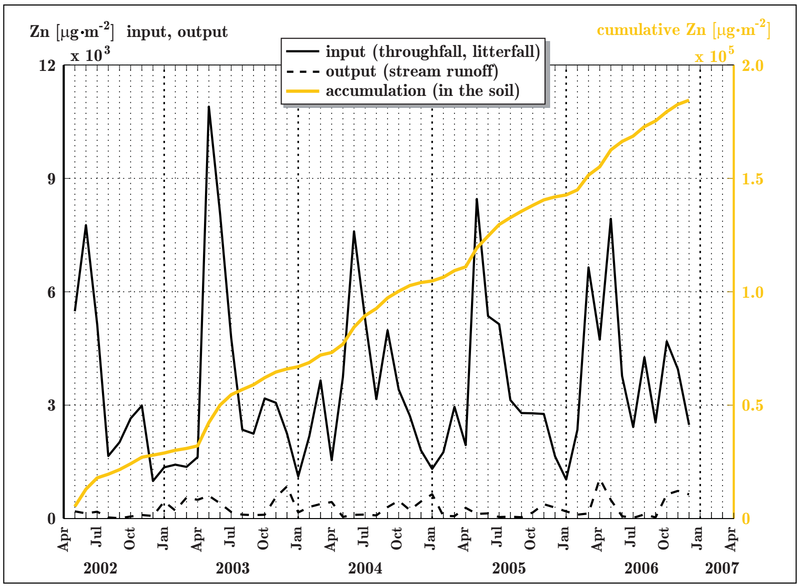


Fig. F.5. Kindla: Zn input, output and accumulation.

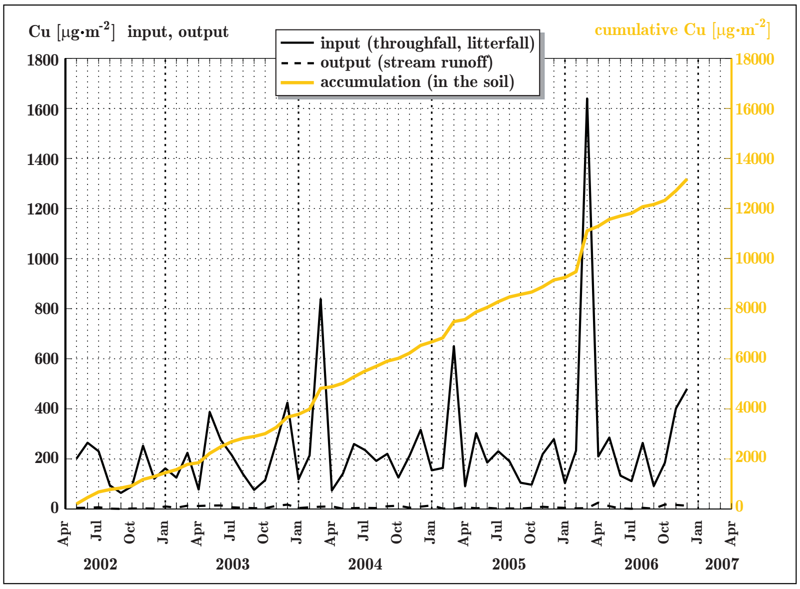


Fig. F.6. Kindla: Cu input, output and accumulation.

F.2.2 Gammtratten

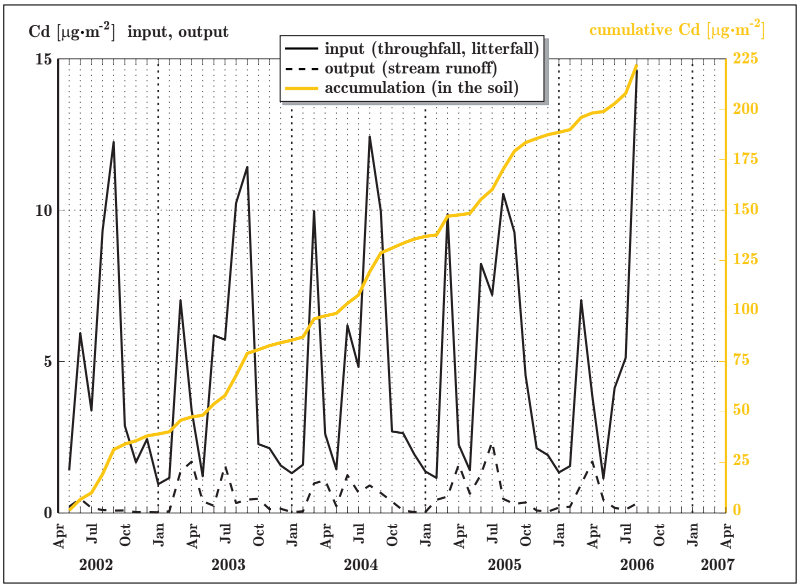


Fig. F.7. Gammtratten: Cd input, output, accumulation.

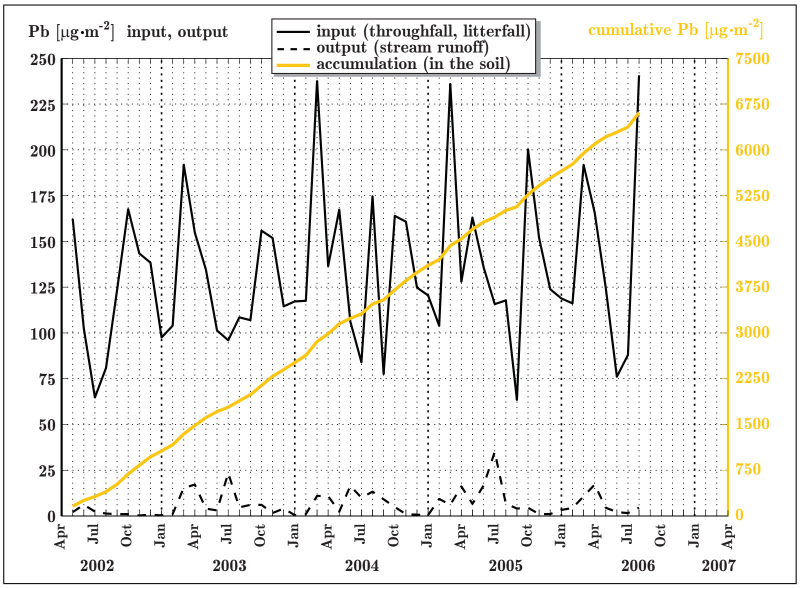


Fig. F.8. Gammtratten: Pb input, output, accumulation.

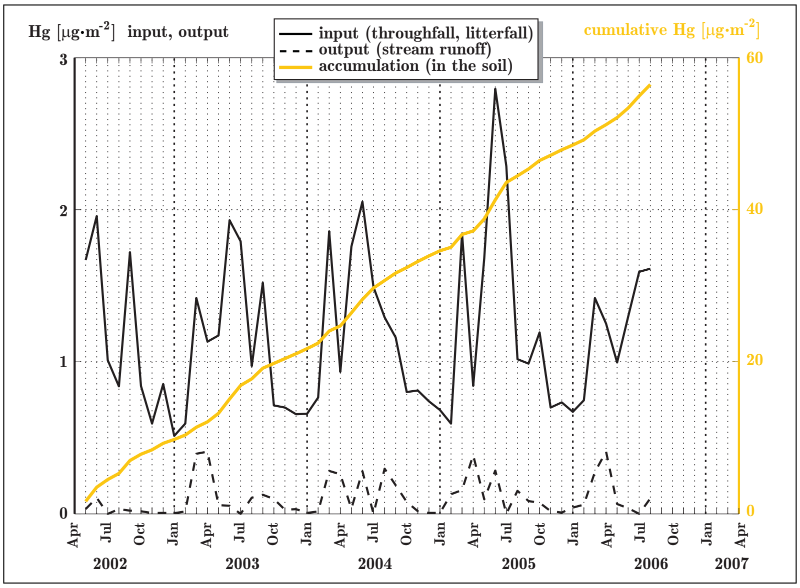


Fig. F.9. Gammtratten: Hg input, output, accumulation.

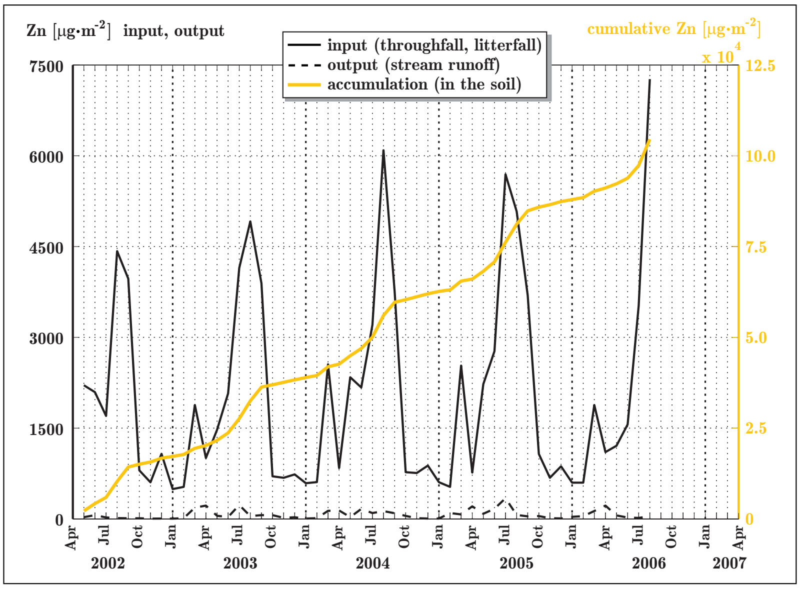


Fig. F.10. Gammtratten: Zn input, output, accumulation.

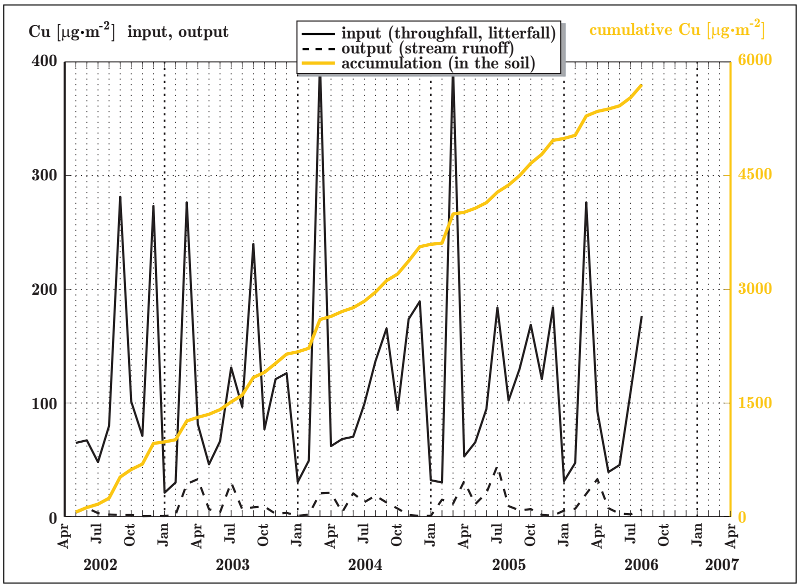


Fig. F.11. Gammtratten: Cu input, output, accumulation.

Bibliography

Bishop, K. H., Lee, Y.-H., Munthe, J., and Dambrine, E. Xylem sap as a pathway for total mercury and methylmercury transport from soils to tree canopy in the boreal forest. *Biogeochemistry*, 40(2-3):101–113, 1998.

FAO. World reference base for soil resources. Appendix 1: Soil horizon designations. In *World Soil Resources Reports*, volume 84. Food and Agricultural Organization of the United Nations, International Society of Soil Science, International Soil Reference and Information Centre, Rome, 1998.

Guber, A. K., Pachepsky, Y. A., van Genuchten, M. T., Rawls, W. J., Simunek, J., Jacques, D., Nicholson, T. J., and Cady, R. E. Field-Scale Water Flow Simulations Using Ensembles of Pedotransfer Functions for Soil Water Retention. *Vadose Zone J*, 5(1):234–247, 2006.

Hsu, H. P. Schaum's outline of theory and problems of probability, random variables, and random processes. In *Schaum's Outline Series*. McGRAW-HILL, New York, USA, 1997.

Klaassen, W., Lankreijer, H. J., and Veen, A. W. Rainfall interception near a forest edge. *Journal of Hydrology*, 185(1-4):349–361, 1996.

Klaminder, J., Bindler, R., Emteryd, O., and Renberg, I. Uptake and recycling of lead by boreal forest plants: Quantitative estimates from a site in northern sweden. *Geochimica et Cosmochimica Acta*, 69(10):2485–2496, 2005.

Kleemola, S. and Forsius, M. Background and objectives of ICP IM. In Kleemola, S. and Forsius, M., editors, *15th Annual Report 2006. ICP Integrated Monitoring*, volume 30/2006, pages 12–20. Finnish Environment Institute, Helsinki, Finland, 2006.

Lundin, L. *Mark- och grundvatten i moränmark och marktypens betydelse för avrinningen.* PhD thesis, Uppsala University, Department of Physical Geography, Division of Hydrology, Uppsala, Sweden, 1982.

Mason, R. L., Gunst, R. F., and Hess, J. L. *Statistical Design and Analysis of Experiments. With Applications to Engineering and Science.* John Wiley & Sons, Inc., Hoboken, New Jersey, USA, second edition, 2003.

Olsson, B., Hallbäcken, L., Hohansson, S., Melkerud, P.-A., Niölsson, S. I., and Nilsson, T. The lake gårdsjön area - physiographical and biological features. In *ECOLOGICAL BULLETINS 37: Lake Gårdsgön - An acid forest lake and its catchment*. Publishing House of the Swedish Research Councils, Stockholm, Sweden, 1985.

Raab, B. and Vedin, H. Klimat, sjöar och vattendrag. In *SVERIGES NATIONAL ATLAS*. SNA Förlag, Stockholm, Sweden, 1995.

Rühling, A. and Tyler, G. Heavy metal deposition in scandinavia. *Water, Air, and Soil Pollution*, 2(4):445–455, 1973.

SLU Uppsala. Integrerad Miljöövervakningn i Sverige/Integrated Monitoring in Sweden. http://info1.ma.slu.se/IM/program/IMomr.html and http://info1.ma.slu.se/IM/IMeng.html, 2007.

Starr, M. WATBAL: A model for estimating monthly water balance components, including soil water fluxes. In Kleemola, S. and Forsius, M., editors, *8th Annual Report 1999. ICP Integrated Monitoring*, volume 325, pages 31–35. Finnish Environment Institute, Helsinki, Finland, 1999.