

Regionalised time series models for water table depths

Promotor: prof. dr. ir. P.A. Troch
Hoogleraar in de hydrologie en kwantitatief waterbeheer
Copromotor: dr. ir. M.F.P. Bierkens
Senior wetenschappelijk onderzoeker Alterra

Martin Knotters

Regionalised time series models
for water table depths

Proefschrift

ter verkrijging van de graad van doctor
op gezag van de rector magnificus
van Wageningen Universiteit,
prof. dr. ir. L. Speelman,
in het openbaar te verdedigen
op woensdag 3 oktober 2001
des namiddags te vier uur in de Aula

Nederlandse vertaling titel:

Geregionaliseerde tijdreeksmodellen voor grondwaterstanden

CIP-DATA KONINKLIJKE BIBLIOTHEEK, DEN HAAG

Knotters, Martin

Regionalised time series models for water table depths. / M. Knotters
Doctoral Thesis Wageningen University. – With ref. – With summary in Dutch.
ISBN 90-5808-478-7

Knotters, M., 2001. *Regionalised time series models for water table depths*. Ph.D. thesis, Wageningen University, Wageningen, The Netherlands, 167 p., summary in Dutch.

Because of its shallow depths, the water table is of significant importance for agriculture and nature conservation in the Netherlands. Water management therefore requires accurate information on the spatial and temporal variations of the water table depth. This information is preferably expressed in terms of probabilities, in order to enable risk assessment. Furthermore, to support strategic decisions in water policy, the information on the water table dynamics should reflect the prevailing climatic conditions (say, the average weather over a 30-year period). Since the number of observation wells and the lengths of the time series are limited for regional studies, spatio-temporal prediction methods should be able to incorporate additional measurements and additional information related to the water table depth.

Stochastic methods are devised for estimating fluctuation characteristics representing the prevailing climatic and hydrologic conditions. These methods are based on various models for the dynamic relationship between precipitation surplus and water table depth: a physical descriptive, one-dimensional model, SWATRE, supplemented with a univariate time series model for the noise (SWATRE+ARMA), linear transfer function-noise models (TFN), dynamic regression models (DR) or autoregressive exogenous variable models (ARX), and non-linear threshold autoregressive models (TARSO). These models are applied to extrapolate observed time series of water table depths, by using observed input series on the precipitation surplus having a length of 30 years. Uncertainty is accounted for by generating a large number of realisations using the stochastic model component. The models perform only slightly differently in simulating water table depths, despite their clearly different theoretical starting points. It is shown that a first-order ARX model can easily be expressed in terms of a water balance for a soil column. Moreover, the physically based ARX model can be applied in predicting the effects of human interventions in the hydrological regime on the water table dynamics.

The ARX model is regionalised to a RARX model, by making its parameters dependent of the spatial co-ordinates. Because of their physical basis, the RARX model parameters can be guessed from auxiliary information such as a digital elevation model (DEM), digital topographic maps and digitally stored soil profile descriptions. Next, the guessed RARX parameters are used to transform a precipitation surplus series into a series of water table depths. Predictions obtained by this 'direct' method are compared with observed water table depths. The observed errors are used to correct the final predictions for systematic errors, and to perform stochastic simulations ('indirect' method). The RARX model is incorporated into a space-time Kalman filter algorithm, which enables predictions conditional to observed water table depths. A cross-validation experiment shows that Kalman filter approaches predict the temporal variation of the water table depths relatively precise, whereas the 'indirect' method yields relatively accurate estimates of expected water table depths, since systematic errors are small. The uncertainty about the temporal variation of the water table depth is underestimated by all methods evaluated. Given the sampling design, the accuracy of the uncertainty about the mean water table depth could not be assessed. Besides efforts to reduce uncertainty, it would be interesting to optimise sampling designs in order to obtain accurate estimates of uncertainty.

Additional index words: groundwater head, time series analysis, physical interpretation, resampling, stochastic simulation, accuracy, quantified uncertainty

Contents

Voorwoord	xi
1 General Introduction	1
1.1 The water table depth and its significance	1
1.2 Problem definition	4
1.3 Objectives and scope	5
1.4 Thesis outline	5
1.5 Notation	7
2 Estimating fluctuation statistics of water table depths	9
Abstract	9
2.1 Introduction	10
2.2 Modelling, extrapolation and stochastic simulation	11
2.2.1 Methodological principles	11
2.2.2 Modelling time series of water table depths using SWATRE . .	12
2.2.3 Modelling time series of water table depths using TFN models	14
2.2.4 Simulating time series of water table depths with extensive length	15
2.2.5 Estimating <i>MHW</i> and <i>MLW</i> from simulated extensive time series of water table depths	16
2.3 Time series of water table depths and precipitation surplus	16
2.4 Validation of estimates	18
2.5 Results and discussion	19
2.5.1 Modelling results	19
2.5.2 Estimated <i>MHW</i> and <i>MLW</i> values	22
2.5.3 Results of validation	22
2.5.4 Influence of weather on estimated <i>MHW</i> and <i>MLW</i> values . .	25
2.6 Conclusions	28
Appendix A2.1 Minimum length of calibration period	29
A2.1.1 Introduction	29
A2.1.2 Materials and methods	29
A2.1.3 Results	30
A2.1.4 Discussion	31
A2.1.5 Conclusion	35
Appendix A2.2 The effect of extrapolating time series on <i>MHW</i> and <i>MLW</i> estimates	36
	vii

CONTENTS

3	TARSO modelling of water table depths	39
	Abstract	39
3.1	Introduction	39
3.2	Data	40
3.3	Models	41
	3.3.1 TARSO and Subset-TARSO Model	41
	3.3.2 TFN Model	42
	3.3.3 SWATRE	43
3.4	Selection of SSTARSO models	43
3.5	Design of the simulation experiment	45
3.6	Results	47
	3.6.1 Modelling results	47
	3.6.2 Results of the simulation experiment	52
3.7	Some concluding remarks	55
4	Physical basis of time series models for water table depths	59
	Abstract	59
4.1	Introduction	59
4.2	Time series models	60
4.3	A water balance for the phreatic groundwater zone	62
4.4	Application: predicting the effect of a hydrological intervention	66
	4.4.1 The well sites	67
	4.4.2 The interventions	68
	4.4.3 The effect of the interventions	70
4.5	Some concluding remarks	74
5	Predicting water table depths in space and time using a regionalised time series model	75
	Abstract	75
5.1	Introduction	75
5.2	Data	77
5.3	A regionalised ARX model (RARX)	80
5.4	Guessing RARX parameters from physical information	81
	5.4.1 Guessing drainage resistance	82
	5.4.2 Guessing effective porosity	83
	5.4.3 Guessing $\mu(\mathbf{u})$	83
5.5	Spatial prediction of time series and validation	84
	5.5.1 A direct method	84
	5.5.2 An indirect method	85
5.6	Results	86
	5.6.1 Regionalised RARX parameters	86
	5.6.2 Validation results for the direct method	89
	5.6.3 Validation results for the indirect method	90
5.7	Sources of uncertainty	92
5.8	Mapping risks	94
5.9	Summary and discussion	97

CONTENTS

6 Accuracy of spatio-temporal RARX model predictions of water table depths	99
Abstract	99
6.1 Introduction	100
6.2 Study area; data set	101
6.3 Spatio-temporal prediction methods	103
6.3.1 The RARX model	103
6.3.2 The ‘direct’ method (DM) and the ‘indirect’ method (IM) . . .	104
6.3.3 The space-time Kalman filter (KF)	105
6.3.4 Summary of prediction methods	108
6.4 Set up of the validation of spatio-temporal predictions	110
6.5 Results and Discussion	111
6.5.1 Predicting water table depths	111
6.5.2 Estimating expected water table depths	112
6.6 Conclusions	114
Appendix A6.1 Accuracy of quantified uncertainty about water table depths	115
A6.1.1 Accuracy of quantified uncertainty about the temporal variation	116
A6.1.2 Accuracy of quantified uncertainty about the mean depth . . .	120
7 General discussion and conclusions	123
7.1 Introduction	123
7.2 Estimation of fluctuation statistics	123
7.3 Time series models for water table depths	124
7.4 Finding an appropriate model	125
7.5 The physically based ARX(1,0) model	127
7.6 Spatio-temporal predictions of water table depths	128
7.7 The accuracy of quantified uncertainty	130
7.8 Some topics for further research	131
8 Samenvatting	137
References	147
Curriculum vitae	153

Voorwoord

Als het meeste werk aan een proefschrift achter de rug lijkt te zijn vind je de tijd om een voorwoord te schrijven. Voor mijn persoonlijke aanleiding om dit proefschrift te schrijven moet ik terug naar het jaar 1993, maar het kan ook 1994 zijn geweest, waarin ik mij voorzichtig afvroeg of een promotieonderzoek misschien iets voor mij zou zijn. Ik werkte destijds met veel plezier bij de afdeling Landinventarisatiemethoden van DLO-Staring Centrum. In die prettige club van gedreven onderzoekers werd mijn idee om een promotieonderzoek te verrichten enthousiast begroet. Er was belangstelling voor methodieken om de grondwaterstand op een kwantitatieve manier te beschrijven. Ik vond het een uitdaging om mij gedurende langere tijd op de ontwikkeling van zulke methodieken te concentreren. In 1995 schreef ik een projectplan en in 1997 verscheen het eerste artikel dat zou bijdragen aan dit proefschrift.

Tijdens het promotieonderzoek veranderde er veel op allerlei vlak, zozeer zelfs dat ik mij enkele keren heb afgevraagd waar ik aan begonnen was en, vooral, waar het moest eindigen. Door de steun van een aantal mensen om mij heen mag ik nu dit voorwoord schrijven en hoop ik straks te promoveren. Mijn promotor prof. dr. ir. Peter A. Troch ben ik zeer erkentelijk voor de wijze waarop hij mij gestimuleerd heeft om dit proefschrift af te ronden en voor zijn adviezen die mij daarbij hebben geholpen. Peter, bedankt en ik hoop op een voortzetting van de prettige contacten na mijn promotie. Veel dank ben ik verschuldigd aan mijn copromotor dr. ir. Marc F.P. Bierkens, die naast copromotor ook collega, tegenwoordig zelfs teamleider, en coauteur van een aantal artikelen is. We verschillen nogal en dan heb ik het nog niet eens over muziek gehad. Toch heeft dit een goede samenwerking nooit in de weg gestaan, integendeel zelfs. Marc, bedankt voor je enthousiasme, de vele prikkelende discussies, je deskundige inbreng en je humor.

Bijzondere dank ben ik verschuldigd aan Jaap de Gruijter. Eerst als hoofd van de afdeling Landinventarisatiemethoden van DLO-Staring Centrum en later als collega bij het GIST-team van Alterra heb je me gesteund, ook op cruciale momenten. Ik ben je er erg dankbaar voor en ik beschouw het als een groot voorrecht om een wetenschapper van jouw kaliber te kennen. Bedankt voor je waardevolle adviezen en ik hoop nog vaak met je van gedachten te wisselen onder de koffie of wat dan ook. Dick Brus, je was geloof ik ooit de eerste aan wie ik voorzichtig mijn gedachten over een promotieonderzoek toevertrouwde, bedankt voor alles wat je sindsdien voor mij hebt betekend, inclusief je commentaar bij hoofdstuk 1. Wim te Riele was niet alleen mijn collega maar ook lange tijd mijn kamergenoot en daardoor oog- en vooral oorgetuige van allerlei wel en wee rond mijn promotieonderzoek. Er vloog nog weleens een muis door de kamer en jij wist die dan telkens met grote souplesse te ontwijken. Bedankt dat je dit alles geduldig hebt doorstaan. Bovendien bedankt voor je hulp en inbreng bij het onderzoek. Met je gedrevenheid en vasthoudendheid ben je een groot voorbeeld voor mij. Na Wim werden achtereenvolgens Kim Trouwborst en Patrick Hommel mijn ‘slapie’: ook jullie bedankt voor jullie solidariteit.

Mirjam Hack-ten Broeke, Peter Finke, Willy de Groot, Kees Hendriks, Tom Hoogland, Jack van der Horst, Ellis Leeters, Reind Visschers, Folkert de Vries en Dennis Walvoort zijn naast Marc Bierkens, Dick Brus en Jaap de Gruijter mijn collega's binnen het GIST-team. Bedankt voor jullie collegialiteit, ook als de ansichtkaarten uit Moddergat begonnen te vervelen. Mirjam, jou bedank ik bovendien voor de steun en

belangstelling die je getoond hebt toen je teamleidster was en die je nog steeds toont. Peter, bedankt voor de stimulerende rol die je gespeeld hebt en speelt als programmaleider van het LNV-programma 328, Geodata Groene Ruimte. Het grootste deel van het onderzoek vond plaats in het kader van dit programma.

Ben van der Pouw, bedankt voor de ruimte die je me hebt gegeven en de wijze waarop je me gestimuleerd hebt als voormalig hoofdafdelingshoofd bij DLO-Staring Centrum en als programmaleider. Heidi Hamers, bedankt voor je steun en de mogelijkheden die je me biedt als hoofd van de afdeling Bodem en Landgebruik van Alterra, vooral op die momenten dat mijn gezondheid het een beetje liet afweten. Betty Wennekes, Inge Koning en Liesbeth van der Lippe bieden belangrijke ondersteuning binnen de afdeling, bedankt dat ik daarvan mag meegenieten.

John Mulder en Dennis Walvoort zijn mijn paranimfen en dat zijn ze niet voor niets. John Mulder ken ik al een hele tijd en Liesbeth Ruyten niet geheel toevallig een dito periode. Het zijn goede vrienden van mij en daarbij ook collega's. Met John een dag de rivierstreek in en dan maar zien waar je uitkomt, 's avonds genieten van Liesbeth's onvolprezen kookkunst en nog lekker lang natafelen, waarbij menige Stiboka-anekdote nog eens uitvoerig wordt naverteld, prachtig toch? Dennis Walvoort is een enthousiaste onderzoeker die nogal veel weet, 'gewoon' erg handig is met de computer en niet geheel is ontbloot van humor. Een woord van dank aan mijn paranimfen stel ik liever nog even uit tot na de plechtigheid, maar Liesbeth, jou durf ik nu reeds zwart op wit te bedanken voor alle betrokkenheid die je steeds toont.

Paul van Walsum, coauteur van het eerste artikel dat heeft bijgedragen aan dit proefschrift, bedankt voor de prettige samenwerking toentertijd en de prettige contacten sindsdien. Bij datzelfde artikel gaven Michiel Jansen (Biometris) en Frans van Geer (TNO-NITG) adviezen, waarvoor dank. Rini Schuiling bedank ik voor het goede werk op GIS-gebied. Arie van Kekem wil ik graag bedanken voor zijn hulp bij de bodemclassificatie. Prof. dr. ir. Jan G. de Gooijer (Universiteit van Amsterdam) bedank ik voor de samenwerking die geleid heeft tot het artikel dat ten grondslag ligt aan hoofdstuk 3 van dit proefschrift.

Graag bedank ik Martin Jansen, die als vormgever prachtig werk heeft geleverd aan de omslag en de figuren. Herman Eijsackers, directeur Onderzoek van Alterra, bedank ik voor zijn waardevolle opmerkingen bij het concept van dit proefschrift. Bert Jansen bedank ik voor zijn adviezen bij de uitgave van Alterra Scientific Contribution 3.

Jacqueline Smits en Dik Engelfriet hielpen mij met de Engelse taal in delen van het proefschrift. Jacqueline omlijstte deze hulp vaak met lekker eten, mooie muziek en gezellige gesprekken bij een goed glas wijn. Jacqueline en Dik, heel veel dank. Tja, de omslag doet het al vermoeden, muziek is erg belangrijk voor mij. De inspirerende lessen traverso van Nancy Latour-Possman, het blokfluitkwartet met Anneke Bruin, Marijke Piepers en Dik Engelfriet, de avonden duetten spelen met Cees van Woerkum, de workshops traverso in Bunnik door Frédérique Chauvet en fantastisch georganiseerd door Miep Blomberg, al die gezellige Bunnik-musici, andere mensen met wie ik minder frequent maar niet met minder plezier samenspeel ... alle muziekvrienden ben ik dankbaar voor de ontspanning en afleiding die ze mij bieden.

De rij dankjes zou nog veel langer kunnen worden, maar staat u mij toe dat ik tenslotte mijn ouders bedank. Pa en moe, wat ik aan u te danken heb is teveel om op te noemen.

Bij de omslag

De omslag vertoont een grondwaterstandsverloop vanaf 1959 dat is gesimuleerd met het transfer-ruismodel dat is gegeven voor ‘Well B’ in tabel 2.5:

$$\begin{aligned}h_t^* &= h_{F,t} + n_t^* \\h_{F,t} &= 0.9223h_{F,t-1} + 8.138p_t - 1.982p_{t-1} \\n_t^* - 207.2 &= 0.5262(n_{t-1}^* - 207.2) + \epsilon_t^*\end{aligned}$$

waarin h_t^* een grondwaterstand is (cm) die mogelijk op tijdstip t is opgetreden, $h_{F,t}$ de component van de grondwaterstand is die kan worden verklaard uit het gemiddelde potentiële neerslagoverschot p_t (mm/dag) dat tussen $t - 1$ en t is opgetreden, n_t^* de component van de grondwaterstand is die niet uit het neerslagoverschot kan worden verklaard en ϵ_t^* een ‘foutje’ is dat aselekt met teruglegging is getrokken uit de foutenreeks die resteerde na kalibratie van het model op waarnemingen uit de periode 1984-1991 (*resampling*).

Elke achtste noot vertegenwoordigt een halfmaandelijke grondwaterstand, zodat elke driekwartsmaat een kwartaal bestrijkt. Het eerste kwartaal is gebaseerd op een G-akkoord, het tweede op een D-akkoord, het derde op een a-akkoord en het vierde op een D7-akkoord. Op die manier ontstaat er een *basso ostinato* die het seizoensverloop tot uiting brengt. De grondwaterstandsreeks begint op de voorzijde van het proefschrift in het jaar 1959 en loopt door op de achterzijde. Een slotmaat is er niet. Ondiepe grondwaterstanden klinken hoog, diepe grondwaterstanden laag. De droge zomer van 1976 bevindt zich in de tweede maat van het vijfde systeem op de achterzijde van het proefschrift.

Dennis Walvoort programmeerde het bovenstaande model in R (<http://www.r-project.org>) en wist er behulp van LilyPond (<http://www.lilypond.org>) nootjes van te maken zodat de grondwaterstand tot klinken kan worden gebracht. Wij hebben de gesimuleerde grondwaterstanden zodanig geschaald dat zij vallen tussen een D^1 en een A^3 , het bereik van een traverso, de dwarsfluit uit de achttiende eeuw. Als *seed* voor de *resampling* van ϵ_t^* is het geboortjaar van J.S. Bach gekozen, vermenigvuldigd met -1 . Ik wens u veel speel- en leesplezier. Er zit muziek in de grondwaterstand!

Chapter 1

General Introduction

1.1 The water table depth and its significance

When making a borehole in the soil anywhere in the countryside of the Netherlands, it may happen that, after some time, a depth is reached at which the soil is saturated with water. Unbounded and unconfined, free, or ‘phreatic’ water flows into the borehole, until a certain level is reached. At this level the pressure head of the phreatic water equals the atmospheric pressure. Next, a second borehole is made a bit farther away, say some meters. Most probably phreatic water is found at almost the same depth as in the first borehole. Driven by curiosity, for a commercial reason or for an educational purpose we continue making boreholes so that finally a large area is surveyed. After some time, depending on the conductivity of the soil, in all these boreholes the depth to the phreatic water level can be measured. This is often done by using a tape measure at which a little bell is attached. The bell is let down into the borehole little by little. When the bell contacts the water surface it causes a floplike sound; this is the correct moment to read the depth to the water surface from the tape measure. The observations in boreholes can be considered as point observations on a continuous surface of phreatic water. This surface is referred to as the water table. The depth of the water table below the ground surface is called the water table depth. Suppose that the water table depth could be measured instantaneously in all the boreholes, then we have an interesting data set of observations on water table depth at many locations in an area and at a certain time. A first analysis of the data most likely indicates that the water table depth at a certain time varies in space. Next, tubes with permeable filters are installed in the boreholes, to make future measurements at the borehole locations possible. Time and again we measure the water table depth in these so called observation wells, for instance with intervals of some days length. A further analysis of the data will indicate that the water table depth not only varies in space, but also fluctuates in time following a seasonal pattern with a top in the winter and a trough in the summer. The water table depth is determined at points and time steps at which the water table depth was measured in observation wells or boreholes. For unvisited locations and for time steps at which no observations were made water table depths need to be predicted. These predictions

General Introduction

may deviate from the true water table depth. Therefore, uncertainty may rise about the true water table depth at locations and time steps for which predictions were made.

With the exception of areas like dunes, ice-pushed ridges, the hills of Limburg in the very south of the Netherlands, and areas with excessive groundwater withdrawal, the water table is already found within two meters below the ground surface. Because of its shallow depths, the water table is of significant importance for agriculture and nature conservation in the Netherlands. Many conditions for agricultural land use, such as soil temperature and soil trafficability, are related to the water table depth. The conditions for plant growth and the potential for development of vegetation types are also related to the water table depth. It is important to notice that most of these relationships are indirect in some way, however. For instance, a relationship between water table depth and plant growth exists because the water table depth has an effect on variables like temperature, degree of aeration, concentrations of nutrients, degree of acidity, and thickness of the root zone. However, in general these variables can not be observed against low costs, in contrast to the water table depth which is relatively easy to measure. Moreover, the water table depth is only one variable explaining many others. Therefore, the water table depth is frequently observed in a network of thousands of observation wells, all over the Netherlands. The data are stored in the National Groundwater Archive (Van Bracht, 1989).

The water table depth is monitored systematically since 1948, with the purpose of making a groundwater map of the Netherlands (Krul, 1952). The extremely dry summer of 1947 and its impact on food production confirmed the need for detailed information on water table depths. The Committee on Agro-hydrological Research (COLN-TNO) started an extensive survey (Visser, 1958). During the COLN survey the water table depth was observed quarterly in some 23,000 wells in the period 1952-1955, semi-monthly in some 2,000 wells and in 65 wells the water table depth was observed every day. In 1958 the survey came to a close with the publication of two maps at a scale 1 : 300,000 for each province in the Netherlands, representing the mean water table depths in winter and summer, respectively. Together, these maps can be considered as the first space-time model of the water table depth in the Netherlands. Apart from these maps, maps of drought losses and salinisation were published. Note that this immense job was done in a short period without the aid of computers and geographical information systems!

In 1961 the Dutch Soil Survey Staff started to present seasonal fluctuation of the water table on soil maps. The fluctuations in the water table depth were described by water table classes which combine the average top and trough of the seasonal fluctuation. Van Heesen (1970) describes how water table classes are mapped by combining observations on hydromorphic soil profile characteristics such as gley horizons, landscape characteristics and observed water table depths. In 1992 the soil map of the Netherlands on a scale 1 : 50,000, including a map of water table classes, was completed. Apart from this national soil map, maps on larger scales (1 : 25,000 or 1 : 10,000) are made for specific areas like forests, re-allotment areas and areas where groundwater is withdrawn. In the last mentioned areas the purpose of the surveys is to estimate the lowering of the water table and the ensuing agricultural production losses (e.g. Bannink *et al.*, 1985). In general, detailed soil surveys concern areas of

1.1 The water table depth and its significance

about 2,000 to 15,000 hectares, in which only a limited number of observation wells is present. During the survey, which takes generally no more than one year, additional time series of the water table depth are collected in a network of temporarily installed observation wells. When the water table is approximately at the highest or at the lowest level of the seasonal fluctuation, the water table depth is measured instantaneously in a large number of boreholes. Ten Cate *et al.* (1995) describe the details of mapping water table classes on large scales. Like the COLN maps, water table class maps can also be considered as space-time models of water table depths.

During the second half of the twentieth century, the water table depth was lowered by human intervention in large areas. For instance, because of increasing agricultural mechanisation the trafficability of the soil needed to be improved, which was often achieved by intensifying the drainage and lowering the surface water levels. Groundwater was withdrawn increasingly because of the growing human population, changing life-styles and expanding economy. Consequently, the 1 : 300,000 COLN maps and the 1 : 50,000 Dutch water table class map do not represent the actual hydrological conditions, at least in parts of the country. Therefore projects to update the 1 : 50,000 Dutch water table class map were initiated in the last few years (Finke, 2000).

Were food production the main interest when the COLN survey started in 1952, nowadays ‘verdroging’ (literally: ‘drying out’, ‘desiccation’ or ‘withering’) in nature conservation areas is one of the main issues in research projects on the water table depth (Runhaar, 1999). The man-induced lowering of the water table in the Veluwe, a large area of forests and moors, was investigated by Gehrels (1999). Information on the water table depth is also used to assess the risk that fertilizers leach into the groundwater and the surface water system (e.g. Brus, 1994). Moreover, information on water table depths is needed for house-building, road-building and open-air recreation. Summarising, information on water table depths is needed for interests which are often incompatible. Conflicting demands are often made on the water table depth and its dynamics.

Water table depths are not only significant in the Netherlands. The following examples may illustrate the diversity in recent studies on the water table depth outside the Netherlands. For instance, Boucneau *et al.* (1996a) described a methodology to quantify trends in the water table depth due to changes in water management in a part of Belgium. To this purpose they used a regression model for the relationship between water table depth, precipitation and evapotranspiration. In another interesting paper the same authors linked the parameters of this model to soil data and used these relationships in predicting a time series of water table depths (Boucneau *et al.*, 1996b). Mew Jr. *et al.* (1997) evaluated monitoring strategies to estimate the mean water table depth in the eastern U.S.A. by comparing means, estimated from weekly up to annually observations, to the means calculated from hourly observations in 17 observation wells. Sun *et al.* (1995, 2000) investigated the spatio-temporal variation of water table depths in Northern Central Florida, because of the significance of the water table depth to wetland and upland ecosystems. Muñoz-Reinoso (2001) investigated the impact of groundwater abstraction on water table fluctuation and vegetation in the Doñana National Park in Spain. Coulibaly *et al.* (2001) devised artificial neural network (ANN) models to simulate water table fluctuations in the Gondo aquifer in Burkina Faso. The results of this study are very promising. The

authors showed that ANN models can capture nonlinear relationships and the time-varying dynamic behaviour of the hydrologic system. The models were successfully applied in forecasting deep water table depths during the dry season. Interestingly, ANN models do not require costly data on the physical system, just like the time series models applied in this thesis. The studies quoted here indicate the international interest of accurate information on the water table depth.

1.2 Problem definition

An observed time series of water table depths reflects the water table dynamics given the meteorological and hydrological conditions during the monitoring period which is generally restricted to a limited number of years, say 2 to 10 years. However, to support strategic decisions in water policy, the information on water table dynamics should reflect the prevailing climatic conditions, i.e., the average weather over, say, 30 years, rather than the meteorological circumstances during the monitoring period. Thus methodology is needed to extrapolate observed time series of water table depths to series of 30 years length from which characteristics can be calculated.

Apart from its restricted length, an observed time series represents the water table depth for the well site only. In order to support water management, information is needed for any location in an area. Information on the water table dynamics is needed at a national or a regional scale. At a national scale a dense network of observation wells exists, in regional studies the number of observation wells is often very limited. Regional studies may for instance concern re-allotment areas, rural development projects, district water boards, or nature conservation areas, the sizes of which roughly vary from 2,000 to 15,000 hectares. Given a density of one suitable observation well per 750 to 1,250 hectares (Finke, 2000), the number of observation wells in regional studies varies roughly from 2 to 15. In nature conservation areas and in areas where groundwater is withdrawn for drinking water or industrial water supply, the monitoring network may be denser. However, in general the number of observed time series in regional studies will be too small to make reliable spatio-temporal predictions of the water table depth by using observed time series only. Therefore, prediction methods are needed which incorporate additional measurements of the water table depth and additional information related to water table depth.

In fact, the water table class maps mentioned in the previous section are spatio-temporal models of water table depth, which incorporate additional measurements and additional information such as hydromorphic soil profile characteristics and landscape characteristics. These maps have proved their practical value, but they have some disadvantages. Firstly, although the map patterns are based on the knowledge and insights of experienced soil surveyors, the methodology is not reproducible. Basically a water table class map is a conclusion rather than a result. Secondly, the information on fluctuation is restricted to classes defined in terms of the highest and lowest level of seasonal fluctuation. These water table classes have proved to be useful in agricultural land evaluation. In other applications, for instance nature conservation, characteristics such as duration curves and regime curves are needed (e.g. De Haan, 1992; Bierkens, 1998). Finally, without further efforts the reliability of the water table class maps is unknown, which makes them less useful in risk assessment.

Risk assessment is particularly needed to find optimal solutions for water management in areas where conflicting interests on the water table depth are present. Therefore, there is a demand for stochastic methods which enable to describe the water table dynamics in terms of probabilities. Strictly spoken risk assessment can be performed by counting observed or deterministically simulated water table depths. However, observed time series of water table depths are generally not long enough to represent the prevailing climatic conditions, thus models are applied to extrapolate the observed time series. By using stochastic methods the probability distribution of the water table depth can be estimated more accurately than by applying deterministic methods, because the unexplained part or noise component is taken into account. This is particularly relevant if one is interested in the probability of extreme water table depths. The uncertainty about the true water table depth can be accounted for in stochastic methods, for instance by generating large numbers of possible realisations.

1.3 Objectives and scope

The main objectives of this thesis are:

1. to develop stochastic methods for estimating fluctuation characteristics of the water table depth representing the prevailing climatic and hydrologic conditions;
2. to develop methods to predict water table depths in space and time in terms of probabilities, for application in regions where suitable observation wells are scarce;
3. to evaluate the accuracy of the stochastic methods developed with respect to objectives 1 and 2.

The study is restricted to hydrologic situations where the local precipitation surplus is the main driving force behind the fluctuation of the water table. Situations where the water table is strongly influenced by open water levels, groundwater withdrawal etc. are not considered. The study mainly focuses on methodology to support decision making in *long term* water policy. To this purpose information on the water table dynamics (i.e., statistics) under given hydrologic and climatic conditions is necessary.

1.4 Thesis outline

In chapters 2 to 4 only the time dimension of water table depths is considered. In chapter 2 two methods are given to estimate fluctuation quantities of water table depths. These represent the prevailing climatic and hydrologic conditions, despite the limited length of the available time series of water table depths (4-10 years). In the first method the physical descriptive model SWATRE (Belmans *et al.*, 1983) is used to describe the relationship between precipitation surplus and water table depth. The temporal variation of the unexplained part is described with an additional univariate time series model (SWATRE+ARMA). In the second method a transfer-function model with added noise (TFN) is used to describe the relationship between precipitation surplus and water table depth (Box and Jenkins, 1976). In chapter 3 a time series

model that accounts for threshold nonlinearities in the relationship between precipitation surplus and water table depth is applied. The performance of this so called TARSO model (Tong, 1990) is compared to the performance of the SWATRE+ARMA model, the TFN model and a special form of the TFN model, namely the dynamic regression model (DR). The DR model is equivalent to the autoregressive-exogenous variable model (ARX). In chapter 4 the physical basis of a simple ARX model of order one (ARX(1,0) model) is discussed. The effects of human interventions in the hydrologic regime on the water table dynamics are predicted by using the physically based ARX model.

A thread in chapters 2 to 4 is shown by the different ways to find a time series model that describes the relationship between precipitation surplus and water table depth adequately. In chapter 2 an iterative procedure for model selection, proposed by Box and Jenkins (1976) is applied. This procedure starts with postulating a general class of models. Next, a subclass of models is identified by employing the data and knowledge of the system. A tentative model is fitted to the data and its parameters are estimated. Next diagnostic checks are applied to find possible model inadequacies. This procedure of identification, estimation (or calibration) and diagnostic checking (or verification) is repeated until an adequate time series model is found. A disadvantage of this procedure is that its results are generally not reproducible. Automatic model selection criteria enable us to select an appropriate model from a large set of candidate models in a reproducible way (see De Gooijer *et al.* (1985) for a review). In chapter 3 an automatic model selection criterion (Bayes information criterion, BIC) is applied to select dynamic regression models and time series models which account for threshold nonlinearities (TARSO). In chapter 4 an ARX(1,0) model is *chosen* that can easily be expressed in terms of a soil water balance. The physical interpretation of time series models was discussed among others by Salas and Smith (1981). The so called data-based mechanistic modelling (DBM) approach (see for instance Young and Beven, 1994) is also based on physical interpretation of time series models. Price *et al.* (2000) describe a recent application of DBM. Time series models which are based on knowledge of the physical system, besides data, are also referred to as *grey box models* (Bohlin, 1991; Hangos, 1995). In *black box modelling* knowledge of the physical system is not used or simply not available; only data are used. Models which are based on knowledge of the physical system only are referred to as *white box models*. Because of its physical basis the ARX(1,0) model can easily be linked to additional information. The physically based time series model forms the basis of the spatio-temporal prediction methods described in chapters 5 and 6, because in areas where only a limited number of observed time series is present it makes sense to incorporate additional information into the prediction methods.

In chapters 5 and 6 spatio-temporal prediction methods for water table depths are presented and their prediction performance is evaluated in a study area of about 1,200 hectares in the northern part of the Netherlands. These spatio-temporal prediction methods are all based on a regionalised ARX(1,0) model, which is referred to as the RARX(\mathbf{u} ;1,0) model with \mathbf{u} representing the spatial co-ordinates. In chapter 5 two methods are presented: the ‘direct’ and the ‘indirect’ method. In the direct method additional physical information is used to ‘guess’ the RARX(\mathbf{u} ;1,0) model parameters at any point in space. These guessed parameters are used to transform precipitation

surplus series into time series of water table depths. In the indirect method the preliminary predictions made by the direct method are compared to observed time series of water table depths. Next the observed mean errors are interpolated to correct the preliminary predictions. The standard deviations of the errors are interpolated to enable risk assessment. In chapter 6 the accuracy of five prediction methods is evaluated: the direct method, the indirect method and three alternative applications of the space-time Kalman filter algorithm described by Bierkens *et al.* (2001). Distinction is made between prediction of actual water table depths and estimation of statistics. Apart from the accuracy of the predicted water table depths and estimates of statistics, the accuracy of the quantified uncertainty is assessed.

In chapter 7 the main results and conclusions are summarised and discussed. The thesis ends with a summary in Dutch in chapter 8.

1.5 Notation

In this thesis the customs on notation of three disciplines are found, namely time series analysis, geostatistics and hydrological modelling. For this reason some overlap in the symbols used could not be avoided, unfortunately. For instance, γ is used for drainage resistance as well as the variogram, h is used for an observed water table depth as well as the lag distance in a variogram, etcetera. Therefore, formulas and symbols will be defined where they appear.

Uniformity is pursued at some points. For water table depth and precipitation surplus, distinction is made between stochastic variables, denoted by H and P , respectively, and single realisations or deterministic series, denoted by h and p . The same distinction is made for the noise component of time series models describing the relationship between P and H , namely N and n , respectively. The error term of time series models for the relationship between P and H is denoted by ϵ . After fitting of the model to series of p and h the so obtained errors are denoted by $\hat{\epsilon}$. The residual term of regression models is denoted by ε . The difference between observed water table depths and deterministically predicted water table depths is denoted by e . The results of calibration are denoted by a hat, $\hat{\cdot}$. The results of straightforward transformations, or results obtained by ‘guessing’ rather than by calibration are denoted by a tilde, $\tilde{\cdot}$. Realisations obtained by resampling procedures are denoted by the superscript \cdot^* .

Chapter 2

Estimating fluctuation statistics of water table depths

(This chapter is based on the paper ‘Estimating fluctuation quantities from time series of water-table depths using models with a stochastic component’ by M. Knotters and P.E.V. van Walsum, which was published in the *Journal of Hydrology* **197** (1997): 25-46, copyright ©1997, with permission from Elsevier Science.)

Abstract

A method is developed to estimate fluctuation statistics of water table depths that are independent of the precipitation surplus during the monitoring period, whose length is generally limited to 4-10 years. For this purpose, one-dimensional models are calibrated with the precipitation surplus as input variable. These models include the SWATRE soil moisture accounting model, supplemented with a stochastic model for the noise series, and transfer function-noise (TFN) models. The models are used to simulate realisations of time series of water table depths with lengths of 30 years, from which the mean highest and mean lowest water tables (*MHW* and *MLW*, respectively) are calculated. These estimates can be used in water management for making strategic decisions, because they reflect conditions of the prevailing climate (i.e. average weather conditions over, say, 30 years) and not just the meteorological conditions during the groundwater monitoring period, which is usually of limited length. The results show that *MHW*s and *MLW*s which are estimated from an 8-year series may deviate more than 20 cm from those estimated from 30-year series. The results of the SWATRE models and the TFN models differ only slightly, despite having clearly different theoretical starting points. The minimum length of series needed for calibration is of practical value; series of 4 years which are not affected by human intervention were generally found to be sufficiently long to model the dynamic systems in this study adequately. Both SWATRE models and TFN models could be improved in order to obtain a constant error variance: in SWATRE models hysteresis of the soil water characteristics could be incorporated, whereas in TFN models a

non-constant variance of water table depths could be taken into account.

2.1 Introduction

Water managers are interested in the spatial and temporal variations in the water table depth. In the Netherlands, these have been measured at many locations for many years. The dynamic variation or fluctuation of the water table depth is characterised by various statistics. For instance, seasonal fluctuations in the water table depth are described by the mean highest and mean lowest water tables (*MHW* and *MLW*, respectively), calculated from time series of semi-monthly data, with lengths of at least 8 years. For each hydrologic year, the means of the three highest and the three lowest water table depths are calculated, *HG3* and *LG3*, respectively. Next, the *HG3*s and *LG3*s are combined into classes of water table depths. These are mapped concurrently with the soil survey, using correlations with profile and landscape characteristics (Van Heesen, 1970). The minimum length of 8 years is chosen arbitrarily, since only time series of this length were available at the time that mapping of the *MHW* and *MLW* started (P. van der Sluijs, personal communication). Water table classes can be translated into classes of duration, using empirical relationships (Van der Sluijs and De Guijter, 1985).

There is a growing need to improve the method described above. In this chapter one of the deficiencies experienced in practice is addressed: estimates of *MHW* and *MLW* may be affected by the precipitation surplus during the period of measurement of at least 8 years. Strategic decisions by water managers, however, require information that reflects the conditions of the prevailing climate, i.e. the average weather over a period of, say, 30 years. Therefore, it is desirable to improve the estimation of *MHW* and *MLW* accordingly.

The aim of the study described in this chapter is to obtain estimates of *MHW* and *MLW* which represent the prevailing hydrologic and climatic conditions, using time series models for water table depths with the precipitation surplus as an input variable. Two types of models are applied, with different theoretical starting points. Firstly, a physical model for vertical unsaturated flow is applied, called SWATRE (Belmans *et al.*, 1983), supplemented by an ARMA model for the series of differences between measured and simulated water table depths (ARMA, Auto-Regressive Moving Average; Box and Jenkins, 1976). Secondly, transfer function-noise models (TFN) are applied (Box and Jenkins, 1976). Both types of models are calibrated on eight time series of water table depths. The models are used to simulate a large number of time series realisations of water table depths, with lengths of 30 years. From these realisations the *MHW* and *MLW* are estimated. The stochastic component of the models is crucial in this application, because *MHW* and *MLW* are measures for extremes. The results are validated using data from an independent validation set.

This chapter is composed as follows. First, in section 2.2 the methodological principles are set out (2.2.1), the modelling procedures for SWATRE+ARMA and TFN models are explained (2.2.2 and 2.2.3), the simulation of 30-year time series of water table depths is explained (2.2.4) and the procedure to estimate *MHW*s and *MLW*s from these series is described (2.2.5). The data sets are described in section 2.3. The procedure of validating estimated *MHW*s and *MLW*s is given in section 2.4. Next,

in section 2.5 the results are presented and discussed for two representative observation wells, including an analysis of the influence of variations in precipitation surplus on *MHWs* and *MLWs* estimated from 8-year series. Concluding remarks follow in section 2.6.

Since the study described in sections 2.1 to 2.6 was published in 1997, new data became available and furthermore, new insights were obtained into some topics mentioned in this chapter. Therefore, in appendix A2.1 the minimum length of series needed for calibration of time series models will be reconsidered. In appendix A2.2 the variation of *MHWs* and *MLWs* estimated from 30-year series will be compared with the variation of 8-year estimates.

2.2 Modelling, extrapolation and stochastic simulation

2.2.1 Methodological principles

The basic idea of the methodologies used in this chapter is that models using precipitation surplus as an input variable, calibrated on time series of water table depths with limited lengths (4-10 years), enable us to simulate series of extensive length (say 30 years). From these extensive series *MHW* and *MLW* values can be estimated that are not influenced by the particular weather circumstances during the period of water table measurements. The models contain a dynamic component, describing the dynamic relationship between the input and output either physically or empirically. Additionally, the models contain a noise component, which describes the part of water table fluctuation that cannot be explained with the used physical concepts or empirically from the input series.

Since the water table depth, H , is observed with discrete time intervals, we consider discrete sequences in time, where time is denoted by the subscript t , $t = 0, 1, 2, \dots, m$ and m is the number of observations. The precipitation surplus, P_t , can be seen as the most important driving force behind the seasonal fluctuations in H_t .

Four stages can be distinguished in the weather-independent estimation procedure, the details of which are given in the following subsections:

1. The procedure starts with modelling the relationship between precipitation surplus and water table depth. The causal relationship between P_t and H_t can be described by a SWATRE model using nonlinear physical relationships and information that can be obtained at a field site, see subsection 2.2.2. Alternatively, the causal relationship between P_t and H_t can be described by a transfer function-noise model (TFN model). A TFN model describes a linear, empirical relationship between two processes, in this case P_t and H_t , with $t = 0, \pm 1, \pm 2, \dots$, see subsection 2.2.3;
2. Next, the time series of water table depths is extrapolated to a length of 30 years, by using observed data on precipitation surplus, p_t , and the relationship between p_t and h_t that was found in the first stage. It is assumed that the average weather conditions during the last 30 years represent the prevailing climate (e.g. Rudloff, 1981, p. 16-17; Schönwiese *et al.*, 1993, p. 11);

3. From the previous step a series of deterministically predicted water table depths result, generated by a model that was calibrated on observed series of p_t and h_t . It can not be expected, however, that the variation of h_t can be ‘fully’ explained from the p_t series. Since we are interested in statistics of extremes, e.g. *MHW* and *MLW*, the unexplained part or noise component, has to be taken into account in the estimation procedure. Therefore, the extrapolation stage is followed by a stage of stochastic simulation, in which a large number of realisations of the noise component is generated and next added to the deterministic series. In this way, realisations of the water table depth are obtained, see subsection 2.2.4;
4. Finally, estimates of fluctuation statistics representing the prevailing hydrologic and climatic conditions can be calculated from the stochastically simulated 30-year series of water table depths, see subsection 2.2.5.

The H_t series will show a seasonal fluctuation which is explained by the precipitation surplus. Apart from this seasonal fluctuation, the H_t series is assumed to follow a stationary process during the calibration period. This requires at least that no human interventions in the hydrologic system have taken place. Therefore, to begin with we calibrated models on time series with a length of 4 years, because time series of equidistantly observed water table depths of more than 4 years which are not influenced by human interventions are not so commonly available in the Netherlands. Besides, it is only for the last 4 years that input data on vegetation and surface water levels, needed for SWATRE models, are easily obtainable in the Netherlands. However, the calibration period should be long enough to cover the full dynamic response in order to identify appropriate models and to estimate parameters accurately. Therefore, we pay attention to the length of the calibration period in both the SWATRE modelling and the TFN modelling.

2.2.2 Modelling time series of water table depths using SWATRE

Modelling principles

The SWATRE model (Belmans *et al.*, 1983) describes the vertical hydrologic interactions in a column, providing an integration of processes in the atmosphere, vegetation, unsaturated soil and saturated subsoil. Simulation of unsaturated flow in the presence of a water table is the prime focus of the model: the other processes are treated in a more schematic manner.

Potential evapotranspiration is modelled through multiplying the Makkink reference crop evapotranspiration, e_M , by a crop factor provided by Feddes (1987). The reduction of potential to actual evapotranspiration is simulated by means of a root water uptake function that depends on the soil water potential.

The simulation of unsaturated flow is based on the Richards’ equation, which is solved with a finite-difference scheme. The flux boundary condition at the bottom of the soil profile is obtained by superposition of:

2.2 Modelling, extrapolation and stochastic simulation

- the local flux induced by the presence of trenches, drains and ditches;
- the regional groundwater flux.

The local flux is made head-dependent: the head difference between the simulated water table and the assumed drainage base is divided by the drainage resistance, yielding the drainage (or infiltration) flux. The drainage fluxes of the different orders of surface water systems (trenches, drains, ditches) are simply added up. The regional flux is obtained using a simple sine function to approximate the seasonal variation of the regional groundwater flow system. In general, this flow system reaches the peak of its intensity at the end of spring, when the regional system has received the winter precipitation surplus. The interannual variations in the regional flow, which in reality also exist, are not modelled.

Parameters

Information about crops, soil profile, drainage levels, irrigation and infiltration was collected in the field. Soil profile descriptions were translated into soil physical horizons using soil physical standard curves for the main soil horizons in the Netherlands obtained from Wösten *et al.* (1987). For technical details, we refer to Stolte *et al.* (1992, 1994). A number of parameters had to be estimated by means of calibration on the measured water table depths:

- drainage resistances;
- regional component of groundwater flow and its variation in time (mean, amplitude, time at which maximum was reached);
- effective saturated water content of the soil profile.

The saturated water content of the soil profile is the most important parameter determining the amount of water that can be stored and released during the rising and falling of the water table. Consequently, it is a key parameter in describing the fluctuation of the water table depth. The method employed to obtain its value through using soil physical standard curves has the drawback that it is not based on measurements of samples obtained locally.

Calibration was performed by minimising the root mean squared error (*RMSE*) between the measured and simulated water tables. Minimisation was achieved by means of the Levenberg-Marquardt algorithm (Marquardt, 1963), as implemented in the ZXSSQ subroutine of IMSL (IMSL, 1982).

Analysis of noise

Let $H_{S,t}$ be the part of water table depth H_t that can be explained by the SWATRE model. The unexplained part, $H_t - H_{S,t}$, forms the noise series $N_{S,t}$. The noise will show autocorrelation structure in time, because the dynamic behaviour of H will not be completely explained by the SWATRE model. The noise series $N_{S,t}$ was modelled in order to make the noise useful in simulating time series of water table depths with extensive length. ARMA models for stationary processes were appropriate for our

model study (Hipel and McLeod, 1994, p. 107) because the noise series could be assumed to form part of a stationary process. The ARMA model has the following general form:

$$(N_{S,t} - \mu) = \sum_{k=1}^p \phi_k (N_{S,t-k} - \mu) + \epsilon_{S,t} - \sum_{l=1}^q \theta_l \epsilon_{S,t-l}, \quad (2.1)$$

where:

$N_{S,t}$ is the noise series with discrete time index t ;

μ is a constant, the mean level of the series $N_{S,t}$;

$\epsilon_{S,t}$ is the error series with discrete time index t , which is considered to follow a white noise process;

ϕ_1, \dots, ϕ_p are autoregressive parameters of model Eq. (2.1) up to order p ;

$\theta_1, \dots, \theta_q$ are moving average parameters of model Eq. (2.1) up to order q .

Following the approach described by Box and Jenkins (1976) we distinguished three stages in model building: identification, calibration (fitting) and verification (diagnostic checking). The order of the ARMA model was identified by visual interpretation of the plotted autocorrelation function (ACF).

Calibration by the exact maximum likelihood method was carried out using the Genstat statistical package (Genstat 5 Committee, 1993). The model verification included diagnostic checks on the error term $\epsilon_{S,t}$ described by Box and Jenkins (1976) and Ljung and Box (1978) and a normality check described by Jarque and Bera (1980). In addition to these checks we checked for the presence of cross-correlation between $n_{S,t}$ and $\hat{h}_{S,t}$. As was indicated in subsection 2.2.1, a cross-correlation may lead to underestimation or overestimation of the fluctuation if the ARMA model is used in simulating time series of water table depths with extensive length.

2.2.3 Modelling time series of water table depths using TFN models

For a detailed description of the basic principles of TFN modelling we refer to Box and Jenkins (1976), and for application in hydrologic and environmental sciences to Hipel and McLeod (1994). Precipitation surplus and water table depth are considered to be the input and output of a linear system, which can be described by a transfer function model. The part of the system which cannot be explained by the linear relationship between input and output is called the noise component and is modelled additionally. The general form of the TFN models used in this study is given by the following equations:

$$H_{F,t} = \sum_{i=1}^r \delta_i H_{F,t-i} + \omega_0 P_{t-b} - \sum_{j=1}^s \omega_j P_{t-j-b}, \quad (2.2)$$

$$(N_{F,t} - \mu) = \sum_{k=1}^p \phi_k (N_{F,t-k} - \mu) + \epsilon_{F,t} - \sum_{l=1}^q \theta_l \epsilon_{F,t-l}, \quad (2.3)$$

$$H_t = H_{F,t} + N_{F,t}, \quad (2.4)$$

2.2 Modelling, extrapolation and stochastic simulation

where:

$H_{F,t}$ is the dynamic component of the water table depth at time t which is explained by the precipitation surplus P_t ;

$N_{F,t}$ is the noise component at time t , containing all other causes than P_t ;

μ is a constant, the mean of $N_{F,t}$;

$\epsilon_{F,t}$ is the error term with discrete time index t , which is considered to be a white noise process;

$\delta_1, \dots, \delta_r$ are autoregressive parameters of the transfer function model up to order r ;

$\omega_0, \dots, \omega_s$ are moving average parameters of the transfer function model up to order s ;

ϕ_1, \dots, ϕ_p are autoregressive parameters of the noise model up to order p ;

$\theta_1, \dots, \theta_q$ are moving average parameters of the noise model up to order q ;

b represents the delay between input and output (in this study $b = 0$).

We identified and calibrated TFN models following the procedure described by Box and Jenkins (1976). The identification of the transfer component was based on visual interpretation of the residual cross-correlation function (RCCF). The RCCF describes the cross-correlations between the residuals which remain after filtering both the h_t series and the p_t series with an appropriate univariate time series model for the p_t series ('prewhitening', Box and Jenkins, 1976, pp. 379-380). In this study, the following seasonal ARIMA model was fitted to the p_t series:

$$(p_t - p_{t-24} - c) = \hat{\phi}_1(p_{t-1} - p_{t-25} - c) + \hat{a}_t - \hat{\Theta}_{24}\hat{a}_{t-24}, \quad (2.5)$$

where c is a constant, 24 is the seasonal lag and $\hat{\Theta}_{24}$ is a seasonal moving average parameter. The series \hat{a}_t is a white noise sequence with zero mean and finite and constant variance. The RCCF is proportional to the impulse-response function and gives insight into the length of the dynamic response.

The model verification included the checks that we also applied to the error term $\epsilon_{S,t}$ of the SWATRE models. In addition to these checks, we checked the significance of estimated model parameters and for the presence of correlation between the parameters of the transfer function model and parameters of the noise model.

2.2.4 Simulating time series of water table depths with extensive length

Let l be the number of semi-monthly time steps in the period from 1961 to 1990, and let m be the number of semi-monthly time steps in the calibration period. We generated realisations of the process $\epsilon_{S,t}$ or $\epsilon_{F,t}$, $t = 1, 2, \dots, l$, through random sampling with replacement from the series $\hat{\epsilon}_{S,t}$ or $\hat{\epsilon}_{F,t}$, $t = 1, 2, \dots, m$. This approach requires no assumptions to be made about the distribution of the error terms $\epsilon_{S,t}$ and $\epsilon_{F,t}$. We avoided making assumptions about the distribution of $\epsilon_{S,t}$ and $\epsilon_{F,t}$ because estimates of MHW and MLW concern the extreme values of H_t and may therefore be sensitive to false assumptions about the distribution of the error term. Also, because the diagnostic checks indicated that the distribution of the error terms $\epsilon_{S,t}$ and $\epsilon_{F,t}$ departed from the normal distribution, we did not use the normality assumption in the simulation procedure. The number of N realisations of $\epsilon_{S,t}$ and $\epsilon_{F,t}$, $t = 1, 2, \dots, l$, were

generated by repeating the sampling procedure N times. The N realisations of $\epsilon_{S,t}$ or $\epsilon_{F,t}$ were transformed into N realisations of the noise process $N_{S,t}$ or $N_{F,t}$ by means of the noise model Eq. (2.3). Resampling of estimated error terms was described by McLeod and Hipel (1978), Kreiss (1990) and Paparoditis (1990).

N realisations of $H_t, t = 1, 2, \dots, l$ were obtained by adding N realisations of $N_{S,t}$ or $N_{F,t}$ to the deterministic series $\hat{h}_{S,t}$ or $\hat{h}_{F,t}$, depending on whether the SWATRE model or a TFN model was used, with realisations of $N_{S,t}$ or $N_{F,t}$ being generated from the error series $\hat{\epsilon}_{S,t}$ or $\hat{\epsilon}_{F,t}$ of the SWATRE model or the TFN model, respectively. Before adding realisations of the noise component to the dynamic component, one must check if there is a cross-correlation between $\hat{h}_{S,t}$ and $n_{S,t}$ or between $\hat{h}_{F,t}$ and $n_{F,t}$. Cross-correlations may give rise to an underestimation or an overestimation of the fluctuation (MHW and MLW), because the realisations of the noise component are generated independently from the dynamic component. For instance, in the case of positive cross-correlations, MHW and MLW values calculated from generated realisations of H_t will underestimate the real fluctuation.

2.2.5 Estimating MHW and MLW from simulated extensive time series of water table depths

Firstly, we generated a large number (100) of realisations of H_t for a period T of length l ; in the present study T was the period of 30 years for which the meteorological normals were calculated, i.e. the period from 1961 to 1990. A number of 100 realisations was found to be sufficiently large to obtain stable estimates of MHW and MLW (standard errors between 0.08 cm and 0.27 cm for the eight observation wells). MHW and MLW were calculated for each realisation $h_t^{*i}, t = 1, 2, \dots, l, i = 1, 2, \dots, 100$, following the procedure described by Van Heesen (1970). Realisations obtained through resampling are denoted by an asterisk (*). For each hydrologic year the three highest and the three lowest water table depths from a series of semi-monthly data were averaged to provide values of h^{*i}_h and h^{*i}_l , respectively. These series were averaged to provide values of $m_T(h^{*i}_h)$ and $m_T(h^{*i}_l)$, respectively. Then, we estimated the expectations of $m_T(H_h)$ and $m_T(H_l)$, say $\mu[m_T(H_h)]$ and $\mu[m_T(H_l)]$, by averaging the 100 values of $m_T(h^{*i}_h)$ and $m_T(h^{*i}_l)$. These estimates, say $\hat{\mu}[m_T(H_h)]$ and $\hat{\mu}[m_T(H_l)]$, are estimates of MHW and MLW conditional to the meteorological circumstances during the period from 1961 to 1990.

2.3 Time series of water table depths and precipitation surplus

We selected time series of water table depths from eight observation wells, which were located near five stations of the Royal Dutch Meteorological Institute (KNMI) so that time series of the precipitation surplus were available for a period of at least 30 years (see Figure 2.1). The geohydrologic and soil physical conditions of the wells were different and were by and large representative of large parts of the country. The water table depth was observed at approximately equidistant time steps of half a month, over a period of at least 10 years without interruptions and without changes

2.3 Time series of water table depths and precipitation surplus



Figure 2.1: Locations of meteorological stations and observation wells.

in the hydrologic regime.

P (mm day^{-1}) was calculated from daily precipitation and Makkink reference-crop evapotranspiration (De Bruin, 1987):

$$p = p' - e_p, \quad (2.6)$$

where p' = precipitation (mm day^{-1}) and e_p = potential Makkink reference-crop evapotranspiration (mm day^{-1}). We used data of daily precipitation and daily evapotranspiration. Because TFN models were used, the series of daily precipitation surplus had to be converted into the frequency of the time series of water table depths. For this purpose the daily data were averaged over the period between two semi-monthly measurements of the water table depth. Transforming the input series in such a way was found to be successful in previous studies (Van Geer and Defize, 1987). The P data were available from 1958-01-01 to 1991-12-31 for the “De Bilt” KNMI station and from 1959-01-01 to 1991-12-31 for the remaining stations (year-month-day). Data on the water table depth, from 1982-01-01 to 1991-12-31, were selected from the Groundwater Archive of the TNO Institute of Applied Geoscience (Van Bracht, 1989).

Estimating fluctuation statistics of water table depths

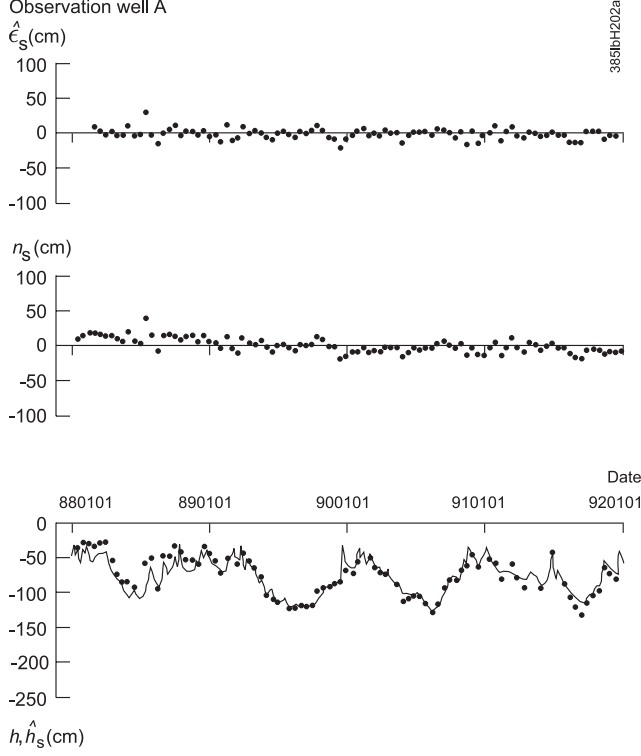


Figure 2.2: Plots of water table depths predicted by means of SWATRE (\hat{h}_S , line), observed water table depths (h , dots), noise series (n_S) and error series ($\hat{\epsilon}_S$).

2.4 Validation of estimates

Firstly we calibrated SWATRE models and TFN models on the time series of h from 1988-01-01 to 1991-12-31. The simulation results were validated using h data from 1982-04-01 to 1987-03-31, called “val.”. We tested the H_0 hypothesis that $m_{\text{val.}}(h_h)$ and $m_{\text{val.}}(h_l)$ were part of the distributions of $m_{\text{val.}}(h_h^*)$ and $m_{\text{val.}}(h_l^*)$, respectively. For this purpose we calculated 100 realisations of H_t , with $t = 1, 2, \dots, \ell$, where ℓ is the length of the validation period. We estimated 95% confidence limits for $m_{\text{val.}}(\hat{h}_h)$ and $m_{\text{val.}}(\hat{h}_l)$ from the cumulative frequency distribution of the generated 100 realisations. If $m_{\text{val.}}(h_h)$ and $m_{\text{val.}}(h_l)$, calculated from the observations, were outside these limits we concluded that the estimates differed significantly from the true values of MHW and MLW . However, significant differences between observed and estimated MHW and MLW values may or may not be relevant from a practical point of view. Therefore, we assessed the relevance by comparing differences between $m_{\text{val.}}(h_{\dots})$ and $\hat{\mu}[m_{\text{val.}}(H_{\dots})]$ with a previously chosen least relevant difference. For applications in agriculture, estimation errors of MHW and MLW larger than 10 cm are generally found to be relevant.

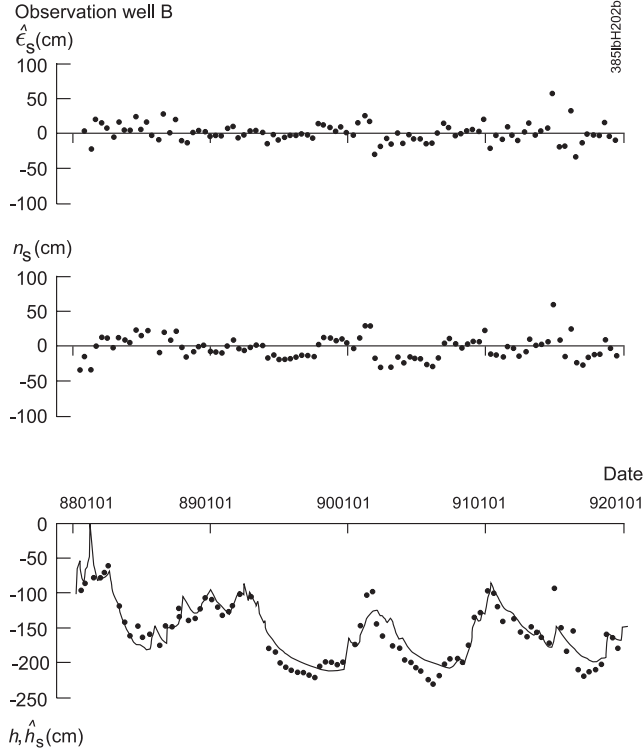


Figure 2.2: *Continued.*

2.5 Results and discussion

2.5.1 Modelling results

For economy of presentation, we will focus on the results of only two observation wells originating from hydrologically different situations. Observation well A (code 12BL0015, Figure 2.1) is situated in a drained brook valley with loamy, fine sandy soils. Water table depths fluctuate within 120 cm below the soil surface. Observation well B (code 12EP0120, Figure 2.1) is situated at the slope of a ridge of fine sand, covering glacial till. The water table depths fluctuate between roughly 100 cm and 200 cm below the soil surface.

The parameter estimates of the models for observation wells A and B are listed in Table 2.1. The parameter estimates of the SWATRE models can be physically explained only to a certain extent, because model uncertainties may be reflected in the estimates. For instance, poor representativeness of the soil physical standard curves used may be compensated in the parameter values. Figure 2.2 shows trace plots with results of the SWATRE modelling. The model fits clearly better to the data from observation well A than to those from observation well B. The poor quality of the fit may be caused by weak representativeness of the soil physical standard curves used in the modelling. Another cause may be the interannual variations in regional

Estimating fluctuation statistics of water table depths

Table 2.1: Parameter estimates of models calibrated on time series from 1988-01-01 to 1991-12-31

Model	Observation well	
	A	B
SWATRE	$\hat{\gamma}_d = 164.67$ days $\hat{\gamma}_t = 5$ days (fixed) $\hat{q}_b = 0.41$ mm day ⁻¹	$\hat{\gamma}_d = 1304.81$ days $\hat{\gamma}_t = 525.43$ days $\hat{q}_b = 0.27$ mm day ⁻¹ $\hat{A}_b = 0.22$ mm day ⁻¹ $\hat{d}_b = 224$ days
Additional noise model	$\hat{F} = 1.52$ $\hat{c} = 0.94$ cm(3.76) $\hat{\phi}_1 = 0.380(0.083)$ $\hat{\phi}_3 = 0.438(0.083)$ $\hat{\sigma}_\epsilon^2 = 54$ cm ²	$\hat{F} = 0.65$ $\hat{c} = -3.88$ cm(2.88) $\hat{\phi}_1 = 0.502(0.092)$ $\hat{\sigma}_\epsilon^2 = 198.3$ cm ²
TFN model	$\hat{\delta}_1 = 0.846(0.02)$ $\hat{\omega}_0 = 6.80(0.47)$ $\hat{\omega}_1 = 1.96(0.56)$ $\hat{c} = -94.9(2.8)$ $\hat{\phi}_1 = 0.454(0.10)$ $\hat{\sigma}_\epsilon^2 = 65$ cm ²	$\hat{\delta}_1 = 0.914(0.01)$ $\hat{\omega}_0 = 8.64(0.71)$ $\hat{\omega}_1 = 2.07(0.76)$ $\hat{c} = -208.5(7.0)$ $\hat{\phi}_1 = 0.463(0.10)$ $\hat{\sigma}_\epsilon^2 = 135.2$ cm ²

In brackets: standard errors.

$\hat{\gamma}_d$ = drainage resistance of ditches; $\hat{\gamma}_t$ = drainage resistance of trenches; \hat{q}_b = regional component of groundwater flow; \hat{A}_b = amplitude of regional component of groundwater flow; \hat{d}_b = time at which maximum regional groundwater flow is reached; \hat{F} = multiplication factor for saturated moisture content (all layers).

groundwater flux at the bottom of the soil profile which are not included in the model. This limitation will not influence the fit to the data observed in Well A, because effects of the temporal variation of the regional flux are dampened by the drainage systems. Next, we analysed the noise component, n_S . Figure 2.3 gives a summary of diagnostic checks performed on the errors, $\hat{\epsilon}_S$. The diagnostic checks indicated that no significant autocorrelation is left in the error series of both observation wells A and B. The diagnostic checks indicated absence of significant cross-correlation between the dynamic component and the noise component of observation well A, which makes the error series useful in simulating extensive H series. The error series of observation well B, however, is significantly cross-correlated with the dynamic component. In the validation stage we will investigate if these cross-correlations lead to relevant estimation errors of MHW and MLW . The checks indicated that the distributions of $\hat{\epsilon}_{S,t}$ depart from the normal distribution.

Figure 2.4 gives RCCFs of the errors which are left after prewhitening the 10-year input and output series with the filter in Eq. (2.5). The RCCFs tend to zero within the first 40 lags, which means that a calibration period of 4 years (96 semi-monthly time steps) amply covers the length of cross-correlation. Figure 2.5 shows trace plots

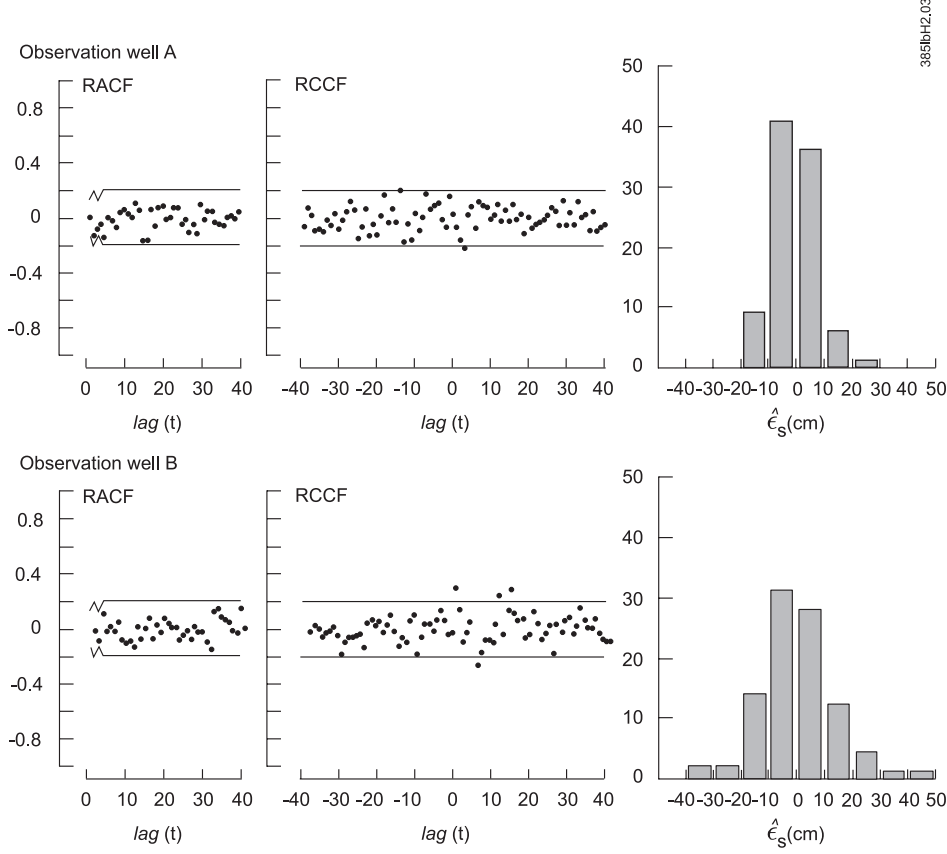


Figure 2.3: Summary of diagnostic checks, applied to the errors of the predictions by SWATRE. Residual autocorrelation function (RACF) for the error series; residual cross-correlation function (RCCF) of $\hat{h}_{S,t}$ and the errors. Before estimating the RCCF, the autocorrelated series $\hat{h}_{S,t}$ was transformed into an uncorrelated series by an appropriate time series model; histogram of the errors $\hat{\epsilon}_S$.

with results of the TFN modelling for Wells A and B. The TFN models fit the data reasonably well. Figures 2.2 and 2.5 clearly show that especially for Well B the TFN model yields a better fit than the SWATRE model. We identified TFN models of orders $b, r, s = 0, 1, 0$ or $0, 1, 1$. ARMA models for the noise of order $p, q = 1, 0$ were found to be appropriate for all series analysed. Figure 2.6 gives a summary of diagnostic checks for observation wells A and B. The checks indicated significant cross-correlations between the dynamic component and noise component for observation well B. These could not simply be removed by fitting alternative TFN models to the series. In the validation stage we will judge if this inadequacy of the models used leads to relevant estimation errors of MHW and MLW . The checks indicated that the distribution of $\hat{\epsilon}_{F,t}$ departs from the normal distribution for observation well B.

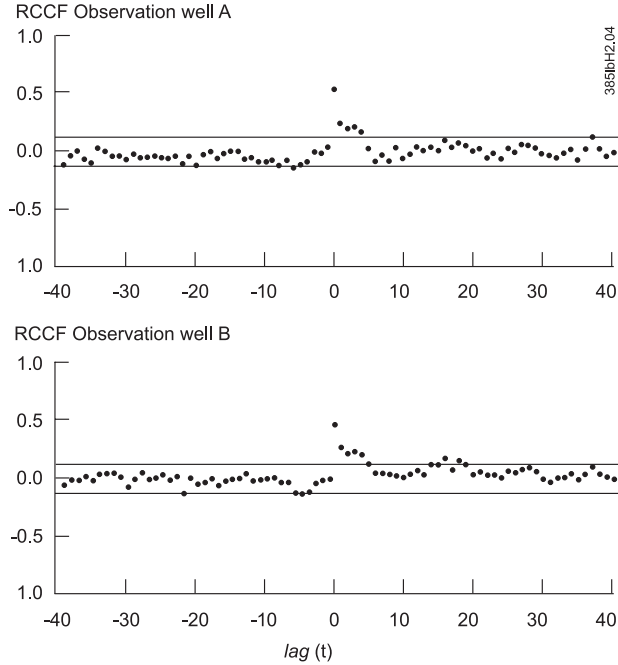


Figure 2.4: RCCF for the precipitation surplus and the water table depth. Before estimating the RCCF, both series were filtered with a seasonal ARIMA model fitted to the series of the precipitation surplus.

2.5.2 Estimated *MHW* and *MLW* values

Table 2.2 lists *MHW* and *MLW* values, estimated by using the SWATRE model and an ARMA model for the noise series and by using TFN models. In general the estimates did not differ more than 10 cm. Large differences were found between the *MHW* estimates for observation wells 32CL0064 and 57FL0019. These may be the result of non-linear relationships between input and output in the top of the fluctuation zone. These non-linearities were modelled in SWATRE but ignored in TFN modelling. In general, the results of the SWATRE models and the TFN models are quite similar, in spite of different theoretical starting-points in modelling the time series.

2.5.3 Results of validation

Table 2.3 lists results of the validation of estimated *MHW* and *MLW* values. When SWATRE models with an additional ARMA model for the noise series were used, we judged that the model estimations differed significantly from the *MHW* and *MLW* values calculated from the observed time series for four of the eight observation wells (*Pr* outside the range from 0.025 to 0.975). When TFN models were used, we concluded that the model estimates differed significantly from the *MHW* and *MLW* values calculated from the observed time series for six of the eight observation wells.

2.5 Results and discussion

Table 2.2: Estimates of the mean highest and mean lowest water tables (*MHW* and *MLW*, cm below soil surface) for the period *T*, 1961-1990

Observation well	<i>MHW</i>		<i>MLW</i>	
	[1]	[2]	[1]	[2]
12BL0015(A)	42(2)	39(1)	109(2)	106(1)
12EL0003	23(1)	26(1)	80(2)	78(1)
12EP0120(B)	103(2)	108(2)	197(2)	200(1)
28HL0045	74(2)	77(2)	173(2)	176(2)
32CL0034	98(1)	101(1)	146(1)	148(1)
32CL0064	62(1)	80(2)	138(1)	137(2)
32CL0078	101(2)	103(1)	165(2)	158(1)
57FL0019	80(2)	103(3)	208(2)	208(3)

In brackets: standard deviations of 100 realisations of *MHW* and *MLW* (cm).

Models used: [1] *SWATRE* and an *ARMA* model for the noise series; [2] *TFN* models. Calibration period: 1988-01-01 to 1991-12-31. The calibration period of Well 32CL0064 was restricted to the period 1988-01-01 to 1990-12-31 because of an intervention which took place in 1991.

Table 2.3: Validation of estimated mean highest and mean lowest water tables by calculating *Pr* values

Observation well	$Pr(m_{\text{val.}}(\hat{h}_h) > m_{\text{val.}}(h_h))$		$Pr(m_{\text{val.}}(\hat{h}_l) > m_{\text{val.}}(h_l))$	
	[1]	[2]	[1]	[2]
12BL0015(A)	0.95	0.66	0.63	0.45
12EL0003	0.93	0.91	1.00	1.00
12EP0120(B)	0.96	1.00	0.94	1.00
28HL0045	0.64	0.76	0.19	0.61
32CL0034	0.92	0.99	0.96	0.98
32CL0064	0.82	1.00	1.00	1.00
32CL0078	0.99	1.00	1.00	1.00
57FL0019	0.08	1.00	1.00	1.00

Models used: [1] *SWATRE* and an *ARMA* model for the noise series; [2] *TFN* models. Calibration period: 1988-01-01 to 1991-12-31. The calibration period of Well 32CL0064 was restricted to the period 1988-01-01 to 1990-12-31 because of an intervention which took place in 1991.

Table 2.4 lists the differences between estimated and true values of *MHW* and *MLW* for both estimation methods. If differences between true and estimated values are equal to or greater than 10 cm they are considered to be relevant. The combination of *SWATRE* and a noise model performs slightly better than the *TFN* model. Most differences were positive, i.e. *MHW* and *MLW* were estimated at a too deep level. A reasonable explanation may be that the calibration period from 1988-01-01 to 1991-12-31 did not contain all the information required to obtain an unbiased estimate of the water table depths in the preceding years. Calibration on longer periods, however, did in general not result in significantly different models (Knotters and Van

Estimating fluctuation statistics of water table depths

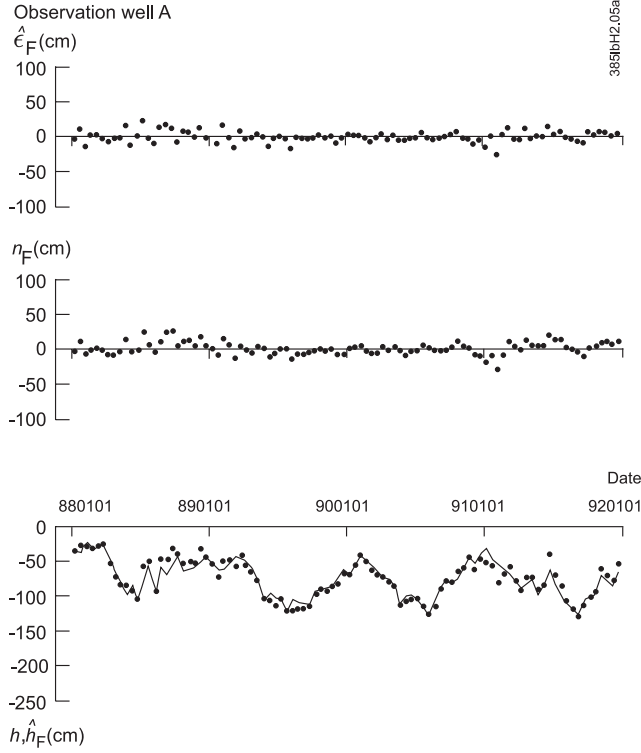


Figure 2.5: Plots of water table depths predicted by TFN models (\hat{h}_F , line), observed water table depths (h , dots), noise series (n_F) and error series ($\hat{\epsilon}_F$).

Walsum, 1994). Another explanation for the positive errors in *MHW* and *MLW* may be that the error variance is not constant. Plots of errors against the dynamic component $\hat{h}_{S,t}$ or $\hat{h}_{F,t}$ indicated a negative dependency between the error variance and the dynamic component. This can be explained physically from the storage capacity of the unsaturated zone which is reduced when the water table rises. For that reason shallow water tables vary more with the precipitation surplus than deep water tables. The TFN models in this study do not take this non-constant variance of water table depths into account, so that the error variance will be underestimated for shallow water tables and overestimated for deep water tables. Given the fact that the *MHW* is derived from the shallowest levels and the *MLW* from the deepest levels, both will be estimated at too deep levels. In contrast to the TFN models, the SWATRE models take changes in storage capacity in the unsaturated zone into account. However, the standard soil physical curves used in the SWATRE models do not describe hysteresis effects. Therefore, the variance of shallow water tables is underestimated to some extent, which also may give rise to positive errors both in *MHW* and *MLW* estimates.

The diagnostic checks indicated cross-correlations between the error term and the dynamic component for observation well B, both for the SWATRE model and the

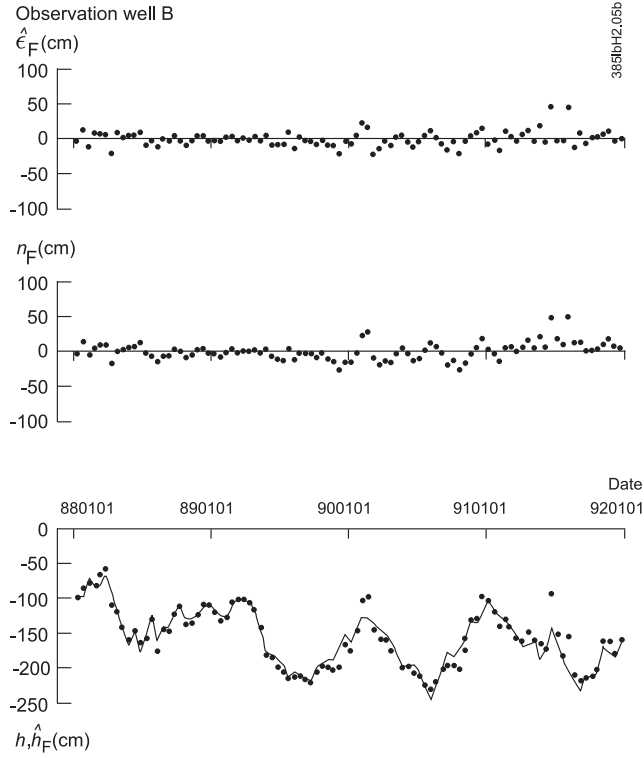


Figure 2.5: *Continued.*

TFN model. In the case of a SWATRE model these cross-correlations did not result in relevant estimation errors, as Table 2.4 shows (+9 cm and +6 cm for *MHW* and *MLW*, respectively). In the case of a TFN model relevant estimation errors were found (+20 cm and +17 cm for *MHW* and *MLW*, respectively). From the positive cross-correlations in Figure 2.6, however, one would have expected an underestimation of the fluctuation, i.e. negative errors in the *MHW* and positive errors in the *MLW*. Obviously, the effect of cross-correlation on estimates is dominated by the effect of a non-constant variance.

Summarising, the validation results make clear that both the TFN models and the SWATRE models can be improved. A first, small improvement can be made by calibrating the models on longer series. Therefore, further analysis is based on models calibrated on 8-year series. Furthermore, both the SWATRE and TFN models can be improved with respect to a non-constant variance of water table depths.

2.5.4 Influence of weather on estimated *MHW* and *MLW* values

The influence of the weather on the estimated *MHW* and *MLW* values was analysed by first calibrating SWATRE models with additional noise models and TFN models

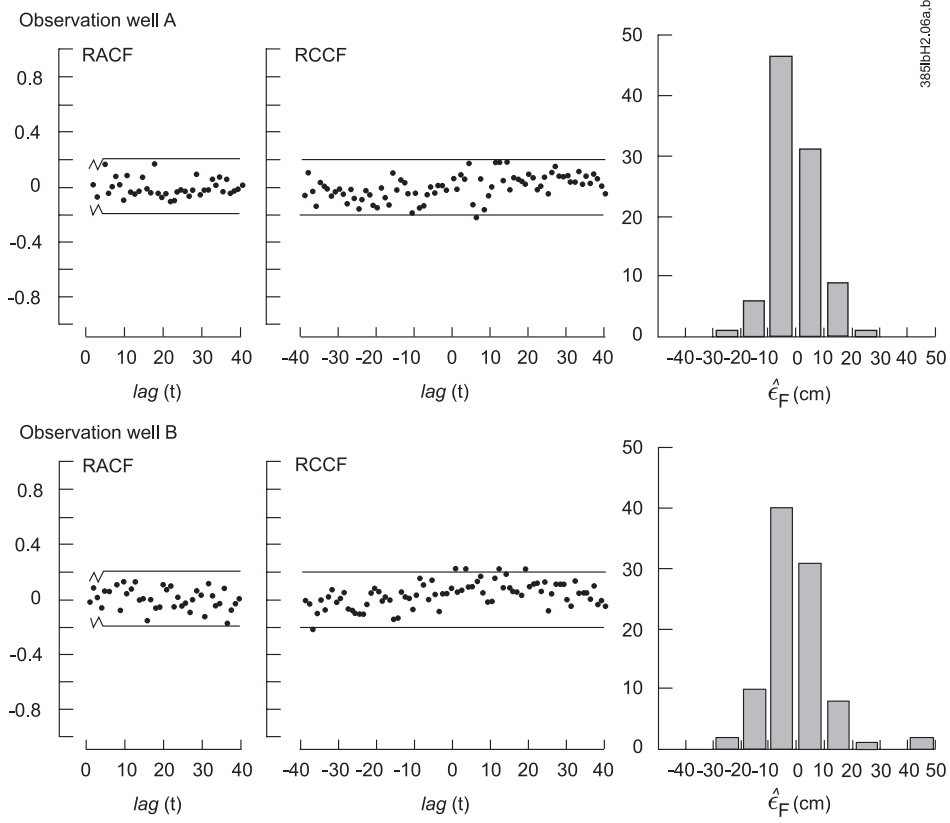


Figure 2.6: Summary of diagnostic checks applied to the residuals of the predictions by TFN models. See Figure 2.3. Before estimating the RCCF, the autocorrelated series of precipitation surplus was transformed into an uncorrelated series by an appropriate time series model.

for Wells A and B, using 8-year series. Table 2.5 shows the results of the calibrations. Next, we simulated time series of water table depths with a length of 30 years, and calculated the *MHW* and *MLW* values for periods of both 8 and 30 years, following the procedure described in subsection 2.2.4. The results were plotted against time in Figure 2.7.

As compared to Figure 2.7c, Figure 2.7a shows small differences between *MHW* values, calculated from the 8-year and the 30-year series. This is the effect of a trench close to Well A, which skims off the water table. SWATRE enabled us to model this non-linear behaviour, unlike the TFN models. Figures 2.7b and 2.7d show less obvious differences between SWATRE and TFN models. Figure 2.7b shows some damping of the *MLW* values calculated from 8-year series, because the lower boundary conditions of SWATRE are the same from year to year. Furthermore, the capillary rise towards the roots stops when the water table falls to a large depth.

The differences between *MHW* and *MLW* values calculated from the 8-year and

2.5 Results and discussion

Table 2.4: Validation of estimated mean highest and mean lowest water tables by calculating differences between estimated and true values (cm)

Observation well	$\hat{\mu}[m_{\text{val.}}(H_h)] - m_{\text{val.}}(h_h)$		$\hat{\mu}[m_{\text{val.}}(H_l)] - m_{\text{val.}}(h_l)$	
	[1]	[2]	[1]	[2]
12BL0015(A)	+7	+1	+1	-1
12EL0003	+5	+5	+14	+12
12EP0120(B)	+9	+20	+6	+17
28HL0045	+2	+4	-3	+2
32CL0034	+3	+7	+3	+5
32CL0064	+3	+20	+14	+18
32CL0078	+10	+14	+13	+9
57FL0019	-6	+27	+18	+19

Models used: [1] SWATRE and an ARMA model for the noise series; [2] TFN models. Calibration period: see Table 2.2. Validation period: from 1982-04-01 to 1987-03-31.

Table 2.5: Parameter estimates of models calibrated on time series from 1984-01-01 to 1991-12-31

Model	Observation well	
	A	B
SWATRE	$\hat{\gamma}_d = 228.34$ days $\hat{\gamma}_t = 5$ days (fixed) $\hat{q}_b = 0.38$ mm day ⁻¹ $\hat{F} = 1.45$	$\hat{\gamma}_d = 1183.35$ days $\hat{\gamma}_t = 338.56$ days $\hat{q}_b = 0.40$ mm day ⁻¹ $\hat{F} = 0.68$
Additional noise model	$\hat{c} = 0.54$ cm(1.69) $\hat{\phi}_1 = 0.3527(0.0667)$ $\hat{\phi}_3 = 0.2538(0.0672)$ $\hat{\sigma}_\epsilon^2 = 87$ cm ²	$\hat{c} = 1.59$ cm(2.42) $\hat{\phi}_1 = 0.5339(0.0647)$ $\hat{\sigma}_\epsilon^2 = 244$ cm ²
TFN model	$\hat{\delta}_1 = 0.8762(0.0217)$ $\hat{\omega}_0 = 6.878(0.384)$ $\hat{\omega}_1 = 2.752(0.469)$ $\hat{c} = -94.14(3.64)$ $\hat{\phi}_1 = 0.5110(0.0654)$ $\hat{\sigma}_\epsilon^2 = 94.5$ cm ²	$\hat{\delta}_1 = 0.9223(0.0119)$ $\hat{\omega}_0 = 8.138(0.571)$ $\hat{\omega}_1 = 1.982(0.619)$ $\hat{c} = -207.2(8.41)$ $\hat{\phi}_1 = 0.5262(0.0660)$ $\hat{\sigma}_\epsilon^2 = 192.1$ cm ²

Definition of symbols: see Table 2.1. In brackets: standard errors.

the 30-year series indicate the influence of variations in the weather on estimates from the 8-year series. For observation well A these differences could be about 10 cm for some periods, whereas those for observation well B could be as much as 20 cm.

Estimating fluctuation statistics of water table depths

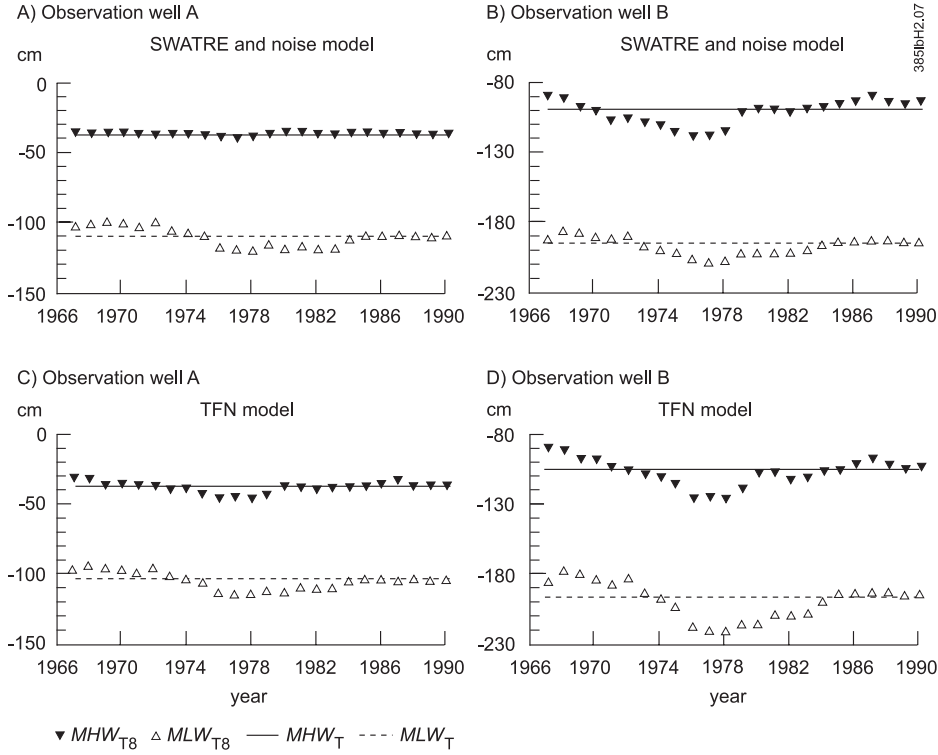


Figure 2.7: MHW and MLW values for two observation wells, calculated from 8-year series ($T8$) and a 30-year series (T , from 1961-01-01 to 1990-12-31). Series simulated by SWATRE and a noise model (Figs. A and B), and by TFN models (Figs. C and D).

2.6 Conclusions

This study shows that MHW s and MLW s estimated directly from the 8-year time series are affected by the precipitation surplus during these 8 years. Estimates based on simulated 30-year series are much more representative of the climatic conditions. The difference between estimates based on 8 years and estimates based on 30 years may be about 20 cm, which is relevant to applications in agricultural water management in the Netherlands. The method described in this paper provides estimates which represent the climatic conditions more adequately, because it enables us to generate 30-year time series of water table depths, from which MHW and MLW values can be calculated. However, the use of models to estimate fluctuation statistics will be only successful if they adequately describe the relation between input series and output series. This implies that the calibration period must be sufficiently long to identify appropriate models and to estimate parameters accurately. Analyses of two observation wells in the Netherlands indicated that a 4-year calibration period contains all information needed to provide a satisfactory description of the relation

Appendix A2.1 Minimum length of calibration period

between precipitation surplus and water table depth, for semi-monthly time steps and for shallow water table depths. In appendix A2.1 the minimum length of time series needed for calibration will be reconsidered. In appendix A2.2 some further results are presented on the variation of the 30-year estimates and the 8-year estimates of the *MHW* and the *MLW*.

In the TFN models used in this study the error variance is assumed to be constant, which is not consistent with reality, as the validation results indicate. Improvement of the models is necessary on this point. In chapter 3 the TARSO model will be presented as a nonlinear alternative for the TFN model. Although SWATRE models take account of nonlinear relationships between precipitation surplus and water table depth, the validation results were only slightly better in comparison with TFN models. This is probably due to hysteresis effects, which are not included into the model. SWATRE models could be improved at this point.

Appendix A2.1 Minimum length of calibration period

A2.1.1 Introduction

The research reported in this chapter was published in 1997. In the mean time new data became available and furthermore, new insights were obtained into some topics mentioned in this chapter. In this section the minimum length of series needed for calibration of time series models will be reconsidered.

Time series of water table depths are observed at numerous locations in the Netherlands, by several public services and for various reasons. One reason may be monitoring, to observe changes in the hydrologic system at the time they occur. Another reason may be to estimate fluctuation characteristics such as *MHW* and *MLW* which represent the actual hydrologic conditions. These characteristics are used to support strategic decisions in water management. For economic reasons the observation periods are limited to the minimum length needed to estimate the fluctuation characteristics accurately. As discussed in this chapter, time series models describing the dynamic relationship between precipitation surplus and water table depth are used in estimating fluctuation characteristics. The minimum length of the calibration period for these models is analysed in this section, using results of Knotters and Bierkens (1998) for 51 observation wells situated in open sandy soils in the Pleistocene part of the Netherlands. The results and conclusions of this subsection are therefore restricted to the soil types and hydrologic situations represented by the data set.

A2.1.2 Materials and methods

An important characteristic of the dynamic relationship between precipitation surplus and water table depth is the response time, that is, the time to which a change

Estimating fluctuation statistics of water table depths

(impulse) of precipitation surplus still influences the water table depth (response). The dynamic relationship between precipitation surplus and water table depth, given in Eq. (2.2) can be written as

$$\begin{aligned} H_{F,t} &= \nu_0 P_t + \nu_1 P_{t-1} + \nu_2 P_{t-2} + \dots \\ &= \nu(B) P_t, \end{aligned} \quad (\text{A2.1})$$

with B being the backward shift operator ($B^i P = P_{t-i}$ and $\nu(B) = \nu_0 - \nu_1 B - \nu_2 B^2 - \dots$). The weights $\nu_0, \nu_1, \nu_2, \dots$ form the impulse response function $\nu(B)$:

$$\nu(B) = \frac{\omega(B)}{\delta(B)} = \frac{\omega_0 - \omega_1 B - \omega_2 B^2 - \dots - \omega_s B^s}{1 - \delta_1 B - \delta_2 B^2 - \dots - \delta_r B^r}. \quad (\text{A2.2})$$

On the basis of the RCCFs for two observation wells (Figure 2.4) it was found in subsection 2.5.1 that a calibration period of 4 years (96 semi-monthly time steps) amply covers the length of dynamic response. Actually, Figure 2.4 indicates that the response time for both wells is not longer than half a year. In this subsection the response times are given for another 51 observation wells. The response times are calculated from the theoretical impulse response function of fitted TFN models, in contrast to subsection 2.5.1 where the response times were derived from the RCCF. It was checked that the calibration periods for all 51 TFN models covered the response times abundantly.

The TFN model that was fitted to the 51 time series of water table depths is the ARX(1,0) model (ARX: autoregressive exogenous variable), which is a special form of the TFN model. The ARX(1,0) model can easily be explained into physical terms, as is pointed out in chapter 4. The ARX(1,0) model is given by:

$$H_t - \mu = a_1(H_{t-1} - \mu) + b_0 P_t + \epsilon_t, \quad (\text{A2.3})$$

where ϵ_t is assumed to follow a white noise process with finite and constant variance σ_ϵ^2 . The model in Eq. (A2.3) was fitted to the 51 time series by using the Kalman filter algorithm as described by Bierkens *et al.* (1999). The response time r was calculated with

$$r = \frac{-3\Delta t}{\ln a_1}, \quad (\text{A2.4})$$

(Bierkens *et al.*, 1999). The response time r can also be used to determine the maximum allowable interval length between the elements of a time series. Of course, using r to determine both the minimum series length and the maximum interval length would lead to series with only two observations. Therefore, apart from the response time also the minimum number of observations needed for accurate estimation of model parameters determines both the minimum series length and the maximum interval length.

A2.1.3 Results

The response times for the 51 locations are given in Table A2.1. The groundwater table classes (GWTs) are explained in Table A2.2. An explanation of the soil pedological terms is given by De Bakker and Schelling (1989). The first group of GWTs

Appendix A2.1 Minimum length of calibration period

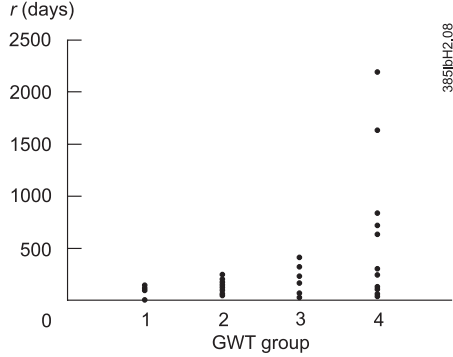


Figure A2.1: Response time vs. group of actual GWT for 51 locations in the Pleistocene part of the Netherlands. Group 1: Ia, IIa and IIIa. Group 2: IIIb, IV, Vbo and VIo. Group 3: VIIo. Group 4: VIId and VIIId.

represent wet soils with a *MHW* within 25 cm below the ground surface. These soils are often found in areas with upward seepage and limited drainage devices. The second group includes four GWTs with a *MHW* between 25 cm and 80 cm, and with a *MLW* within 180 cm below the ground surface. These GWTs are often found in artificially drained soil, or in areas where groundwater is withdrawn. The third group concerns GWT VIIo, with a *MHW* deeper than 80 cm below the ground surface and with a *MLW* within 180 cm. These are dry soils, either because of their high topographic position or because of intensive artificial drainage and groundwater withdrawal. The fourth group represents the very dry situations: the *MHW* is deeper than 80 cm below the ground surface and the *MLW* is deeper than 180 cm below the ground surface. These soils are very dry because of their topographic position, intensive drainage or excessive groundwater withdrawal. Figure A2.1 shows a graph of the response times for each group.

A2.1.4 Discussion

At first sight, in Table A2.1 no clear relationship can be found between response times and groups of actual GWT; within each group a large deviation of response times can be observed. Nevertheless, Figure A2.1 indicates that both the mean response times and the variation of response times increase with increasing depth of GWT: the mean response times are for group 1 $\bar{r} = 82$ days (standard deviation: 49 days), for group 2 $\bar{r} = 128$ days (58), for group 3 $\bar{r} = 163$ days (132), and for group 4 $\bar{r} = 597$ days (679).

Table A2.1 and Figure A2.1 show that short response times can go together with deep water tables. Possibly these time series of water table depths are observed in wells which are situated at relatively high topographic positions, for instance at a farmyard or in a road shoulder. The temporal variation of the water table depths observed in these wells reflects the groundwater dynamics in the relatively low and wet surrounding areas, however. This is possibly true for Wells 12EL0026, 21HL0018, 29CL0021, and 32GL0021 in group 2, for Well 27HP0064 in group 3 and for Wells

Estimating fluctuation statistics of water table depths

Table A2.1: Response times r in days for 51 locations in the Pleistocene part of the Netherlands.

Well code	soil type	soil texture	GWT: actual (soil map)	r
1. GWTs Ia, IIa and IIIa				
16FP7053	broekeerdgrond	very loamy, fine sand	IIIa(III)	97
21HL0019	bruine beekeerdgrond	slightly loamy, medium fine sand	IIa(IV)	12
27DP7603	vlierveengrond	peat overlaying a coarse sandy subsoil	Ia(III)	121
28DP7037	broekeerdgrond	slightly loamy, fine sand	IIa(IV)	15
32HL0105	zwarte enkeerdgrond	very loamy, fine sand	IIIa(III)	94
50AP7608	broekeerdgrond	very loamy, fine sand	Ia(I)	134
50AP7615	beekvaaggrond	slightly loamy, medium fine sand	Ia(III)	100
2. GWTs IIIb, IV, Vbo, VIo				
12EL0026	veldpodzolgrond	slightly loamy, fine sand	VIo(IIIb)	100
16GL0007	veldpodzolgrond	slightly loamy, fine sand	VIo(V)	94
21HL0018	laarpodzolgrond	slightly loamy, medium fine sand	VIo(IV)	54
27DL0031	bruine enkeerdgrond	very loamy, fine sand	VIo(IIIb)	145
28FL0125	zwarte enkeerdgrond	slightly loamy, fine sand	VIo(VII)	182
28FP7015	veldpodzolgrond	slightly loamy, fine sand	IIIb(VI)	212
28GL0008	zwarte beekeerdgrond	very loamy, fine sand	VIo(VIIb)	143
28HL0079	zwarte enkeerdgrond	very loamy, fine sand	VIo(IIIb)	205
29CL0021	laarpodzolgrond	slightly loamy, medium fine sand	VIo(Vb)	83
32GL0021	gooreerdgrond	slightly loamy, fine sand	VIo(III)	47
34BP0192	gooreerdgrond	very loamy, fine sand	VIo(VII)	124
34CP7003	veldpodzolgrond	medium fine sand	VIo(VII)	106
34CP7016	veldpodzolgrond	slightly loamy, fine sand	VIo(V)	146
34DP0155	bruine enkeerdgrond	very loamy, fine sand	VIo(Vb)	69
34GL0007	veldpodzolgrond	slightly loamy, medium fine sand	VIo(Vb)	84
34GL0012	veldpodzolgrond	slightly loamy, fine sand	VIo(Vb)	189
41BP7014	veldpodzolgrond	slightly loamy, fine sand	VIo(IIIb)	45
41BP7018	veldpodzolgrond	slightly loamy, medium fine sand	VIo(III)	200
45CL0024	zwarte beekeerdgrond	very loamy, fine sand	VIo(V)	164
50AP7618	veldpodzolgrond	very loamy, fine sand	Vbo(V)	88
51AL0003	zwarte enkeerdgrond	slightly loamy, medium fine sand	VIo(VI)	150
52DL0010	bruine beekeerdgrond	very loamy, fine sand	VIo(VI)	65
57AP7802	veldpodzolgrond	very loamy, fine sand	IV(V)	238

16EL0035, 21FL0033, and 28BL0051 in group 4. In other cases a possible explanation is the vicinity of a large ditch, which drains off the groundwater quickly, thus resulting in small response times. This explanation may apply to Wells 41BP7014, 50AP7618 and 52DL0010 in group 2, and for Wells 12FL0033 and 28BL0052 in group 3. An explanation for relatively small response times on the basis of topographic position can not be given for Well 16GL0007 (group 2), 34GL0007 (group 2), 51FL0004 (group 3) and 28DL0044 (group 4).

Appendix A2.1 Minimum length of calibration period

Table A2.1: Continued.

Well code	soil type	soil texture	GWT: actual (soil map)	<i>r</i>
3. GWT VIIo				
12FL0033	veldpodzolgrond	slightly loamy, fine sand	VIIo(V)	57
22CL0044	bruine enkeerdgrond	slightly loamy, fine sand	VIIo(VII)	229
27HP0064	zwarte beekeerdgrond	very loamy, fine sand	VIIo(V)	26
28BL0052	veldpodzolgrond	very loamy, fine sand	VIIo(VI)	53
28DL0026	veldpodzolgrond	slightly loamy, fine sand	VIIo(VI)	162
28FP7017	veldpodzolgrond	slightly loamy, fine sand	VIIo(VI)	157
33BL0003	zwarte enkeerdgrond	slightly loamy, medium fine sand	VIIo(IV)	319
51FL0004	zwarte enkeerdgrond	slightly loamy, medium fine sand	VIIo(VI)	60
51GP0087	veldpodzolgrond	slightly loamy, medium fine sand	VIIo(VI)	403
4. GWTs VIIId and VIIId				
16EL0035	zwarte enkeerdgrond	slightly loamy, fine sand	VIIId(Vb)	126
21FL0016	gooreerdgrond	loam poor, medium fine sand	VIIId(VI)	250
21FL0033	zwarte enkeerdgrond	slightly loamy, fine sand	VIIId(VI)	100
28BL0051	meerveengrond	peat overlaying a very loamy, fine sandy subsoil	VIIId(IV)	45
28DL0044	laarpodzolgrond	slightly loamy, medium fine sand	VIIId(VII)	63
28GL0075	zwarte enkeerdgrond	very loamy, fine sand	VIIId(V)	303
33FL0049	veldpodzolgrond	slightly loamy, fine sand	VIIId(VII)	638
34BL0013	veldpodzolgrond	slightly loamy, medium fine sand	VIIId(IV)	842
41AL0059	gooreerdgrond	very loamy, fine sand	VIIId(VI)	245
44HP7804	veldpodzolgrond	slightly loamy, medium fine sand	VIIId(V)	723
52CL0044	zwarte beekeerdgrond	slightly loamy, medium fine sand	VIIId(VII)	1640
57EL0006	vlakvaaggrond	very loamy, fine sand	VIIId(VII)	2183

Note: For an explanation of classes of water table depth: see Table A2.2. An explanation of the soil textures is given in Table A2.3. Soil classes are explained in Table A2.4.

The minimum length of a time series needed to characterise the groundwater dynamics is not only determined by the response time. During the observation period the water table depth must have varied only as a result of a varying precipitation surplus. Other causes of fluctuation, such as groundwater discharge and human interventions in the hydrologic regime, must be absent unless they are incorporated into the time series model. Furthermore, when modelling a time series of water table depths it is important to be aware that the relationship between precipitation surplus and water table depth may vary in time. This relationship could for instance depend on the water table depth itself, or could be different for rising and falling water tables. If a linear, time invariant, model is calibrated, the time series must contain deep and shallow, rising and falling water tables, to find an ‘average’ linear model for the entire fluctuation zone. If a nonlinear time series model is calibrated, for instance a TARSO model (chapter 3), the time series must represent the entire range of water table depths sufficiently to identify the model properly and to calibrate

Estimating fluctuation statistics of water table depths

Table A2.2: Explanation of the water table classes (GWTs).

Water table class (GWT)	mean highest water table (<i>MHW</i>)	mean lowest water table (<i>MLW</i>)
I	-	< 50
Ia	< 25	< 50
IIa	< 25	50 - 80
III	< 40	80 - 120
IIIa	< 25	80 - 120
IIIb	25 - 40	80 - 120
IV	40 - 80	80 - 120
V	< 40	> 120
Vb	25 - 40	> 120
Vbo	25 - 40	120 - 180
VI	40 - 80	> 120
VIo	40 - 80	120 - 180
VII	80 - 140	> 120
VIIo	80 - 140	120 - 180
VIII	80 - 140	> 180
VIII	> 140	> 180

Values in cm below the ground surface.

the parameter values. The same is true for calibration of the parameters of physical mechanistic models such as SWATRE (see subsection 2.2.2). Thus, a time series of water table depths must at least cover one year. It may happen that in a year the water table hardly fluctuates, due to constant meteorological circumstances during that particular year. Then, the monitoring must be continued until the entire range of water table depths has been measured. In Table A2.1 response times larger than one year are found for groups 3 and 4, which represent the dry soils. In the case of the slowest reacting groundwater system (a response time of 2183 days, that is six years), the water table depth has to be monitored during at least six years in order to determine the temporal variation.

Apart from the actual GWT, Table A2.1 gives the GWT of the polygon at the 1 : 50,000 national soil map in which the observation well is situated. In several cases the actual GWT is 'wetter' or 'drier' than the GWT of the soil map. This can be explained as follows:

- Since the soil survey took place the hydrologic system was changed by human intervention;
- The well is situated in an impure part of the map polygon;
- The GWTs of the soil map are based on *MHW* and *MLW* values which were calculated from time series of water table depths which were observed during the period of six to eight years length prior to the soil survey. However, the actual GWTs in Table A2.1 are based on *MHW* and *MLW* values which are calculated according to the method described in this chapter, thus reflecting the average weather conditions of a 30 years period.

Appendix A2.1 Minimum length of calibration period

Table A2.3: *Explanation of the soil texture terms.*

term	% < 50 μ m	median of the fraction 50-2000 μ m
loam poor	< 10	
slightly loamy	10-17.5	
very loamy	17.5-32.5	
fine sand		105-150
medium fine sand		150-210
coarse sand		210-300

Table A2.4: *Explanation of the soil taxonomy.*

Dutch classification system De Bakker and Schelling (1989)	World classification system FAO (1988)
beekvaaggrond	Gleyic Arenosols
broekeerdgrond	Umbric Gleysols
bruine beekerdgrond	Gleyic Arenosols
bruine enkeerdgrond	Fimic Anthrosols
gooreerdgrond	Umbric Gleysols
laarpodzolgrond	Gleyic Podzols
meerveengrond	Terric Histosols
veldpodzolgrond	Gleyic Podzols
vlakvaaggrond	Gleyic Arenosols
vlierveengrond	Fibric Histosols
zwarte enkeerdgrond	Fimic Anthrosols

Without further research it is impossible to assess to what extent the causes mentioned above contribute to the differences between actual GWTs and GWTs of the soil map for the locations in Table A2.1. Indeed, at 30 of the 51 locations the actual GWT is dryer than the GWT at the soil map. It is true that since the soil survey took place the drainage was often intensified and the drainage level was lowered. However, it is also true that observation wells are often situated at relatively high topographic points, such as farmyards and road shoulders. Therefore, the water table depths measured in these observation wells reflect the fluctuation of the dry, impure parts of the map polygons. Furthermore, the effect of the difference between the average precipitation surplus of a monitoring period with six to eight years length and a period of thirty years length, as illustrated by Figure 2.7, should be taken into account.

A2.1.5 Conclusion

The minimum length of observation periods needed for calibration of a time series model to establish the relationship between precipitation surplus and water table depths depends on:

1. the response time between precipitation surplus and water table depth;

2. the variation of water table depths during the monitoring period, which should reflect the entire fluctuation zone and both rising and falling water tables;
3. the absence of human intervention into the hydrologic regime which may influence the relationship between precipitation surplus and water table depth;
4. the number of observations needed to estimate the time series model parameters accurately.

The interval length must be shorter than the response time to identify a proper time series model and to estimate the model parameters accurately. It can be concluded that the semi-monthly frequency, which is usual in groundwater monitoring in the Netherlands, may not be sufficiently high for the fastest reacting groundwater systems.

Groundwater regimes which are influenced by, for instance, groundwater withdrawals and open water levels in rivers, were not considered. Regimes influenced by slowly reacting systems, nearby ice-pushed ridges and dunes, were not considered either. Furthermore, the study is restricted to soil profiles without stagnating layers, such as weakly permeable clay. Therefore, the results presented in this subsection only give a first indication of the response times which occur in a large part of the open, sandy, soils in the Netherlands, where the local precipitation surplus is the only explanatory variable of the fluctuation of water table depth.

Appendix A2.2 The effect of extrapolating time series on *MHW* and *MLW* estimates

In Figure 2.7 *MHW* and *MLW* values calculated from an 8-year series were compared with *MHW* and *MLW* values from a 30-year series (from 1961-01-01 to 1990-12-31). Figure 2.7 shows that estimates based on 8-year series may deviate up to about 20 cm from estimates based on an 30-years series. Clearly, the 8-year estimates represent the meteorological conditions during the 8-year monitoring period which deviate from the long term climatic conditions. Therefore, it is concluded that estimates based on the 30-year series are more useful to assess options for long term water policy, assuming that the average climatic conditions over the last 30 years approximate the climatic conditions over the coming 30 years. In other words, 30-year estimates are assumed to be more ‘conservative’ than 8-year estimates and therefore more useful in long term water policy. Of course, considering a *MHW* or *MLW* value as the average of *HG3* or *LG3* values for a number of successive years (section 2.1), it is obvious that 30-year estimates vary less in time than 8-year estimates. It could be interesting, however, to get more insight into the variation of 30-year estimates of *MHW* and *MLW*, because this variation could give us an idea about our uncertainty of the future long term climatic conditions which we approximate now with the average climatic conditions over the last 30 years. In Figure 2.7 only one 30-year period was considered. New data on precipitation and evapotranspiration became available since

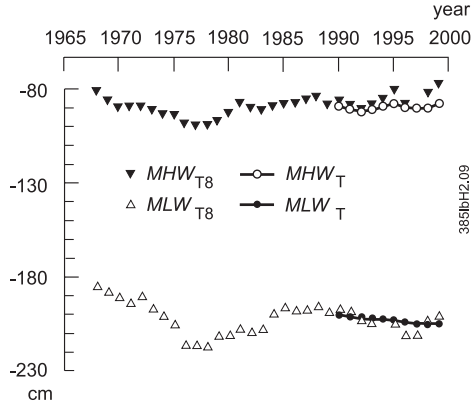


Figure A2.2: MHW and MLW estimates based on 8-year ($T8$) and 30-year periods (T) for Well B. Water table depths were simulated with a TFN model that was calibrated on data from the period 1984-01-01 to 1991-12-31.

the time that Figure 2.7 was produced. These data can be used to calculate new 30-year estimates. This was done for Well B in Figure 2.7, by using a TFN model.

The TFN model was calibrated on data for the 8-year period from 1984-01-01 to 1991-12-31, using the method described by Bierkens *et al.* (1999). This method can deal with sparsely or irregularly observed time series of water table depths, because parameters are estimated for the 1-day sampling interval of the input series on precipitation surplus. The TFN model, expressed as a state equation, is embedded in a Kalman filter algorithm. Next, a maximum likelihood criterion is minimised to obtain parameter estimates of the TFN model for the 1-day time interval of the input series. As contrasted to the method described in subsection 2.2.3, the semi-monthly series of water table depths is not assumed to be equidistant. Furthermore, there is no need to convert the series of daily precipitation surplus to a semi-monthly frequency (see subsection 2.3). A possible disadvantage is that the method described by Bierkens *et al.* (1999) is restricted to autoregressive processes of order one. In chapter 4 it will be shown that an order one autoregressive process for the relationship between precipitation surplus and water table depth can be understood from a physical point of view. The calibrated TFN model is specified as follows:

$$h_{F,t} = 0.9901h_{F,t-\Delta t} + 0.5216p_t, \quad (\text{A2.5})$$

$$(n_{F,t} + 183.8610) = 0.9787(n_{F,t-\Delta t} + 183.8610) + \hat{\epsilon}_{F,t}, \quad (\text{A2.6})$$

$$h_t = h_{F,t} + n_{F,t}, \quad (\text{A2.7})$$

with h in cm, p the average precipitation surplus between t and $t - \Delta t$ in mm day⁻¹, $\Delta t = 1$ day and $\hat{\sigma}_a^2 = 18.6104$ cm². Using the model given in Eqs. (A2.5) to (A2.7) and daily data on precipitation surplus from 1958-01-01 unto 1999-12-31 for the meteorological station Eelde (see Figure 2.1), 100 realisations of H_t series were simulated for the following periods:

- 32 successive periods of eight years length, the first one from 1961-01-01 to 1968-12-31, and the last one from 1992-01-01 to 1999-12-31;

Estimating fluctuation statistics of water table depths

- 10 successive periods of thirty years length, the first one from 1961-01-01 to 1990-12-31, and the last one from 1970-01-01 to 1999-12-31.

MHW values and MLW values were calculated for each period, using the procedure described in subsection 2.2.5. The results are shown in Figure A2.2. The 30-year estimates are clearly less variable than the 8-year estimates. The 30-year estimates indicate that the seasonal fluctuation slightly increases, because the MHW slightly tends to shallower depths, whereas the 30-year estimates of MLW slightly tends to larger depths. Because the simulated water table depths are linear transformations of the precipitation surplus (Eq. (A2.5)) this tendency is also present in the precipitation surplus series. It may be caused either by possible changes in measurement methods (for instance the introduction of measurement devices with higher resolution) or reflect a climatic change.

Note that the results in Figure A2.2 slightly differ from the results for Well B in Figure 2.7. Different TFN models were used, yielding different MHW and MLW estimates. These differences reflect model uncertainty.

Chapter 3

TARSO modelling of water table depths

(This chapter is based on the paper ‘TARSO modelling of water table depths’ by M. Knotters and J.G. de Gooijer, which was published in *Water Resources Research* **35**(3) (1999): 695-705, copyright ©1999 by the American Geophysical Union.)

Abstract

Threshold autoregressive self-exciting open-loop (TARSO) models are fitted to six time series of water table depths with precipitation surplus as input variable. Basically these models are nonlinear in structure because they incorporate several regimes which are separated by so-called thresholds. For each well a (subset-)TARSO (SSTARSO) model is selected using the Bayes Information Criterion (BIC). (SS)TARSO models are used to simulate realisations of water table depths with lengths of thirty years, from which characteristics such as durations of exceedance are computed. The simulation performance of the fitted (SS)TARSO models is compared with results obtained from transfer function noise (TFN) models, dynamic regression (DR) models, and with a physical descriptive model extended with additional ARMA processes for the noise (SWATRE+ARMA). As compared to the linear TFN and DR models the (SS)TARSO models perform better because they incorporate several regimes. These regimes are the result of different soil layers or drainage levels. Furthermore, it is interesting that (SS)TARSO models show a good relative performance as compared to the SWATRE+ARMA models. A possible reason may be that inputs of SWATRE are inaccurate.

3.1 Introduction

In chapter 2 linear transfer function noise (TFN) models were fitted to eight time series of water table depths and precipitation surplus covering a period of eight years of semi-monthly data. Using these models realisations of time series of water table depths of 30 years were simulated from which relevant statistics were calculated. The results in chapter 2 indicate that the descriptive physical model for vertical unsaturated

flow called SWATRE (Belmans *et al.*, 1983) performs only slightly better than TFN models. However, the approach described in chapter 2 is not entirely satisfactory, since TFN models describe linear relationships between precipitation surplus and water table depth. From a physical point of view nonlinear rather than linear relationships may be expected between precipitation surplus and water table depth. This may be due to different soil layers or drainage levels. SWATRE conceptually describes these nonlinear relationships. However, it needs input data on soil physical relationships which are generally uncertain and, moreover, rather expensive to collect.

The objective of this chapter is to devise time series models which are capable of representing the complex nonlinear mechanisms underlying the relationships in six relatively short time series of water table depths and precipitation surplus. Next, using these models, series of water table depths with lengths of 30 years are simulated. To this end the class of threshold autoregressive self-exciting open-loop (TARSO) models of Tong (1990) will be applied. These models identify piecewise linear functions over disjoint subregions, which are discontinuous at their boundaries of the domain of the output variable. Hipel and McLeod (1994, p. 83, 758) mentioned the usefulness of these models in modelling nonlinearities in natural time series.

Chapter 3 is organised as follows. Section 3.2 provides a brief discussion of the time series analysed in this study. In section 3.3 we introduce our main tool the TARSO model, along with some explanation regarding the selection of the “best” model. Also we explain briefly the modelling procedures for SWATRE, TFN and dynamic regression models (DR) which will be used as benchmarks for model comparison. Section 3.4 gives an outline of the procedure for selecting TARSO models. Section 3.5 describes the simulation procedure and the criteria used to evaluate the goodness-of-fit of the simulated water table depths. Section 3.6 provides details of the TARSO models fitted to six semi-monthly time series of precipitation surplus and water table depths covering a period of 5 years. Next in section 3.6 the performance of the TARSO models is evaluated by validation of simulation results and comparison is made with similar results obtained from the SWATRE, TFN and DR models fitted to these series. Section 3.7 ends with some concluding remarks.

3.2 Data

We used time series of semi-monthly observed water table depths from six observation wells, all having a length of ten years (1982-1991), i.e. 240 observations. The wells are located near three stations of the Royal Dutch Meteorological Institute (KNMI). At these stations the daily precipitation surplus has been observed for at least 30 years (1959-1991). The observed series of water table depths are representative for different geohydrologic situations in large parts of the Netherlands. The land use is pasture at all six well locations. We subtracted the daily Makkink reference-crop evapotranspiration (Winter *et al.*, 1995) from the daily precipitation in order to obtain daily precipitation surplus data. Next, we converted the daily precipitation surplus data into the semi-monthly frequency of the water table depth series, by averaging them over the periods between the successive observations on the water table depth. This conversion was found to be successful in previous studies by Van Geer and Defize (1987). As compared with chapter 2 two series have been excluded from the original

set of eight series. One series was dropped because the groundwater level hardly fluctuated during the whole period. The reason for dropping the second series is that unknown sources strongly influenced the groundwater level behaviour besides the precipitation surplus. Soil profile descriptions and drainage levels were collected at all six well sites.

3.3 Models

3.3.1 TARSO and Subset-TARSO Model

Let $\{H_t\}$ denote a time series of water table depths (output) and $\{P_t\}$ a time series of precipitation surplus (input). Now a discrete self-exciting TARSO process $\{H_t, P_t\}$ with order $(\ell; (m_1, m'_1), \dots, (m_\ell, m'_\ell))$ and delay parameter d ($d > 0$) is defined by Tong (1990) to be a solution of the equations

$$H_t = a_0^{(j)} + \sum_{i=1}^{m_j} a_i^{(j)} H_{t-i} + \sum_{i=0}^{m'_j} b_i^{(j)} P_{t-i} + \epsilon_t^{(j)}, \quad \text{if } r_{j-1} \leq H_{t-d} < r_j, \quad (3.1)$$

where $-\infty = r_0 < r_1 < \dots < r_\ell = \infty$, $a_i^{(j)}$ and $b_i^{(j)}$ ($j = 1, \dots, \ell$) are parameters, and $\{\epsilon_t^{(j)}\}$ ($j = 1, \dots, \ell$) are heterogeneous white noise sequences with zero mean and finite variances $\sigma_{\epsilon^{(j)}}^2$ and each being independent of $\{P_t\}$. The thresholds are the levels $r_1, \dots, r_{\ell-1}$. Thus, the real line is partitioned into ℓ intervals, and H_t satisfies one of ℓ dynamic regression models depending on the interval in which H_{t-d} falls.

Note that in the above formulation, the value of the output variable H_{t-d} ($d > 0$) indicates which subsystem is activated. This is slightly different from the original formulation of the TARSO model, where the regime switching is conditional on the value of the delayed observable input variable. If $b_i^{(j)} = 0$ for every $i = 1, \dots, m'_j$ and $j = 0, 1, \dots, \ell$ Eq. (3.1) is called a self-exciting threshold autoregressive (SETAR) process. SETAR processes with appropriately chosen parameters can exhibit important features of real data sets which are not exhibited by Gaussian linear models, e.g. time-irreversibility, non-Gaussian marginal distributions and occasional bursts of outlying observations. For recent work in the theory and applications of SETAR processes, the book of Tong (1990) is an excellent reference. However, in contrast with SETAR processes, relatively little attention has been given to the use of nonlinear multiple time series models.

It has been noted that when a SETAR process is fitted to real data some of the intermediate AR parameters can become small when they are compared with their respective standard errors. In such a case it may be worthwhile to fit subset-SETAR models rather than full models to the data. Indeed, De Bruin and De Gooijer (1998) noted that the out-of-sample forecast ability of SETAR models can be significantly improved by deleting insignificant parameters. Similar results may occur when TARSO models are fitted to real data. To allow for subset-TARSO processes we introduce the following additional notation. Let $a_{p_1^{(j)}}^{(j)}, a_{p_2^{(j)}}^{(j)}, \dots, a_{p_h^{(j)}}^{(j)}$ denote the non-zero parameters $a_i^{(j)}$ in Eq. (3.1) such that $p_1^{(j)}, p_2^{(j)}, \dots, p_h^{(j)}$ are subsets of integers $1, 2, \dots, m_j$ with $1 \leq p_1^{(j)} < p_2^{(j)} < \dots < p_h^{(j)} \leq m_j$. Similarly, let $b_{q_0^{(j)}}^{(j)}, b_{q_1^{(j)}}^{(j)}, \dots, b_{q_k^{(j)}}^{(j)}$ denote the

non-zero parameters $b_i^{(j)}$ in Eq. (3.1) such that $q_0^{(j)}, q_1^{(j)}, \dots, q_k^{(j)}$ are subsets of integers $1, 2, \dots, m'_j$ with $1 \leq q_0^{(j)} < q_1^{(j)} < \dots < q_k^{(j)} \leq m'_j$. Then Eq. (3.1) can be rewritten as

$$H_t = a_0^{(j)} + \sum_{i=1}^{h^{(j)}} a_{p_i^{(j)}}^{(j)} H_{t-i} + \sum_{i=0}^{k^{(j)}} b_{q_i^{(j)}}^{(j)} P_{t-i} + \epsilon_t^{(j)}, \quad \text{if } r_{j-1} \leq H_{t-d} < r_j. \quad (3.2)$$

In this case $\{H_t, P_t\}$ is said to follow a subset-TARSO (SSTARSO) process with lags $p_1^{(j)}, p_2^{(j)}, \dots, p_h^{(j)}, q_0^{(j)}, q_1^{(j)}, \dots, q_k^{(j)}$ in the j th regime. In section 3.4 we describe a procedure for selecting a (SS)TARSO model. For a description of a method to estimate (SS)TARSO models we refer to Tong (1990).

It is interesting to note that for $\ell = 1$ and $d = 0$ Eq. (3.1) reduces to

$$H_t = a_0 + \sum_{i=1}^{m_1} a_i H_{t-i} + \sum_{i=1}^{m'_1} b_i P_{t-i} + \epsilon_t \quad (3.3)$$

where for ease of notation the superscript (1) has been dropped. It is a dynamic (single-input single-output) regression (DR) model where ϵ_t is not autocorrelated. This model is also referred to as ARX (autoregressive-exogenous variables) model.

3.3.2 TFN Model

A class of linear time series models which has been commonly employed in relating one variable to another is the transfer function model with added noise (TFN) of Box and Jenkins (1976, p. 362). It consists of two parts: a transfer function part which describes a functional relationship between the input and output variables, and a noise part which is modelled in its own right using univariate linear autoregressive (integrated) moving average (AR(I)MA) processes. More specifically, the transfer function is given by the linear difference equation

$$H_{F,t} = \sum_{i=1}^r \delta_i H_{F,t-i} + \sum_{j=0}^s \omega_j P_{t-j-b} \quad (3.4)$$

where $H_{F,t}$ denotes that part of the water table depth H_t which is explained by the precipitation surplus P_t . Here b is a delay factor, which is an integer greater than or equal to zero.

In practice one could not expect that H_t can be ‘fully’ explained by movements in P_t , there will always be outside disturbances. If it is assumed that the disturbance, or noise $N_{F,t}$, is independent of the level of P_t and is additive with respect to the influence of P_t , then Eq. (3.4) can be written as

$$H_t = H_{F,t} + N_{F,t}. \quad (3.5)$$

Now $N_{F,t}$ may be related to its own past $N_{F,t-1}$, $N_{F,t-2}$, and so on. This behaviour can be described by an AR(I)MA process. Assuming that $N_{F,t}$ is stationary, this process can be written as

$$(N_{F,t} - \mu) = \sum_{i=1}^p \phi_i (N_{F,t-i} - \mu) + \epsilon_{F,t} + \sum_{j=1}^q \theta_j \epsilon_{F,t-j} \quad (3.6)$$

where $\{\epsilon_t\}$ is a white noise process with zero mean and finite variance, and where μ is the mean of the process $\{N_{F,t}\}$. We shall refer to Eq. (3.6) as an ARMA process of order (p, q) . For methods for identifying, fitting, and checking TFN models we refer to Box and Jenkins (1976).

3.3.3 SWATRE

SWATRE is a transient one-dimensional finite-difference model describing the vertical hydrologic interactions in a column of soil (Belmans *et al.*, 1983). The model can be used to simulate the unsaturated flow in the presence of a water table. The unsaturated flow is described by the so-called Richards' equation. The numerical approximation of this equation leads to a finite-difference expression for all nodal points. For the top and bottom nodes of the soil profile the expression has a special form because these nodes are at the boundaries of the flow domain. For the bottom nodal point a flux is introduced as boundary condition in the finite-difference expression. When the SWATRE model is used for simulating the unsaturated flow a distinction is made between the local fluxes induced by the presence of trenches, drains and ditches, and regional groundwater flux. The local fluxes are pressure-head dependent: the head difference between the simulated water table and the assumed drainage base is divided by the drainage resistance, yielding the drainage or infiltration flux. The drainage fluxes of the different orders of surface water systems are simply added up. To approximate the seasonal variation in the regional groundwater flow system, the regional flux is defined as a simple sine function. At the top of the soil system, boundary conditions are defined for a maximum possible flux through the canopy, and the maximum possible flux through the soil surface. For more details on the various quantities, parameters, and series needed as inputs to the SWATRE model, we refer to Belmans *et al.* (1983) and chapter 2. In SWATRE input is needed on precipitation, evapotranspiration, drainage levels and soil physical properties, whereas in empirical time series models like the (SS)TARSO, TFN and DR model only input on precipitation surplus is needed.

Now, suppose that $\{h_{S,t}\}$ denotes the deterministic part of the water table depth which is explained by the physical model SWATRE. Then, similar to Eq. (3.5), we add a noise process $\{N_{S,t}\}$ to $\{h_{S,t}\}$, i.e. $H_t = h_{S,t} + N_{S,t}$. As with the TFN model it is assumed that $\{N_{S,t}\}$ can be represented by an ARMA(p, q) process having white noise errors. Also, we assume that $\{N_{S,t}\}$ is independent of the process $\{h_{S,t}\}$. Now by modelling the noise $\{N_{S,t}\}$ with an ARMA process it becomes possible to simulate time series of water table depths. In section 3.5 we describe the simulation method in more detail. In the following we refer to the combination of the model SWATRE and an ARMA model as to the SWATRE+ARMA model.

3.4 Selection of SSTARSO models

A critical step in (SS)TARSO modelling is choosing an appropriate model from a large set of candidate models, in a systematic and reproducible way. Automatic model selection criteria, such as Akaike's Information Criterion (AIC) and Bayes Information Criterion (BIC) can be used to find a balance between lack of fit and

model complexity; see De Gooijer *et al.* (1985) for a survey. For linear univariate and multivariate time series models it has been noted that the basic difference between BIC and AIC is that the former penalises an overparametrisation of the model more strongly, in particular for large samples, than the latter. Recently Wong and Li (1998) came to a similar conclusion when studying SETAR models. Hence, we decided to use BIC as a model selection criterion.

Consider the SSTARSO process in Eq. (3.2). Let n denote the total number of observations on the process $\{H_t, P_t\}$. Similarly, let n_j be the number of observations belonging to the j th regime ($j = 1, \dots, \ell$). Then BIC can be defined as

$$BIC = \sum_{j=1}^{\ell} \{n_j \ln \hat{\sigma}_{\epsilon(j)}^2 + (h^{(j)} + k^{(j)} + 1) \ln n_j\} \quad (3.7)$$

where $\hat{\sigma}_{\epsilon(j)}^2$ is the residual variance in the j th regime.

Assume that the delay parameter d is known. In general, this assumption may not be a reasonable one. In our case, however, it can be argued that the delay is zero. Indeed from a physical point of view this means that the regime changes at the moment $\{H_t\}$ passes a boundary between two soil layers or a drainage level. Since our goal is simulating $\{H_t\}$, we can only use information from the past about the regime in which $\{H_t\}$ falls. Therefore, we decided to fix d at unity rather than zero. No other values of d will be entertained here. Based on prior information the number of regimes is fixed for each well. Although the depths of soil physical boundaries and drainage levels are known for the six well sites, we did not use this information in the model selection procedure. In subsection 3.6.1 we compare automatically selected threshold values with observed thresholds such as soil physical boundaries and drainage levels. Now, if no prior information is used on the values of the thresholds r_j ($j = 1, \dots, \ell-1$), we propose the following procedure for selecting (SS)TARSO models using Eq. (3.7):

1. Fix the number of regimes $\ell^\#$ using prior information on the number of regimes;
2. Fix the maximum orders $(M_1, M'_1), \dots, (M_{\ell^\#}, M'_{\ell^\#})$ from which the (SS)TARSO model is selected;
3. Select an interval $[r_L, r_U]$ in which the threshold values are searched, or the combinations of threshold values if there are more than two regimes. In this study, r_L and r_U are the 10th percentile and the 90th percentile of the empirical distribution of H_t respectively;
4. To guarantee that there are enough observations in each regime, search thresholds at a fixed interval (here 1 cm) between r_L and r_U such that within each j th regime n_j is at least 20. This results in a set of, say, R (combinations of) candidate threshold values $r_1, \dots, r_{\ell-1}$;
5. Select candidate subsets for the non-zero coefficients $a_i^{(j)}$ and $b_i^{(j)}$ in the input and the output series, say subsets $\{s_u\}$, where $u = 1, \dots, K$ denotes the u th of K subsets. Assign to these subsets the lags $p_1^{(j)}, p_2^{(j)}, \dots, p_h^{(j)}, q_0^{(j)}, q_1^{(j)}, \dots, q_k^{(j)}$ of the AR terms in the output and input series in the j th regime. Given $\ell^\#$ regimes, fixed threshold values, and a fixed delay there are $S = K^{\ell^\#}$ candidate SSTARSO models to represent the process $\{H_t, P_t\}$;

6. Calculate the minimum value of Eq. (3.7) over all $R \times S$ candidate models.

Note that in the above procedure the first $\max_j \{d, M_j, M'_j\}$ observations should be discarded from the two series to obtain one effective sample size for all fitted models. In the rest of the chapter we restrict the range of fitted SSTARSO models by setting $M_1 = \dots = M_{\ell^\#} = 3$, $M'_1 = \dots = M'_{\ell^\#} = 2$, and $K = 25$. For physical reasons AR terms for H_{t-2} and H_{t-3} are always combined with an AR term for H_{t-1} . For similar reasons AR terms for P_{t-1} and P_{t-2} are always put into the model specification with an AR term for P_t . Details about the choice of $\ell^\#$ will be given in subsection 3.6.1. The DR models in this study are selected by the above procedure, with $\ell^\# = 1$ and $K = 25$ candidate models. In contrast to previously proposed graphical methods to select the threshold values r (e.g. Tsay (1989); Astatkie *et al.* (1997)) the automatic selection procedure applied here is reproducible. In this sense the automatic selection procedure is equal to the procedure based on Akaike's Information Criterion (AIC) proposed by Tong (1990) for the selection of SETAR models, and to the procedure based on the modified generalised cross validation criterion described by Lewis and Stevens (1991) for modelling time series using multivariate adaptive regression splines (MARS).

3.5 Design of the simulation experiment

In the next section we compare by simulation the performance of the SSTARSO model as opposed to the performance of the TFN, DR, and SWATRE models in generating series of water table depths. Here a brief account is given of the various steps involved in the Monte Carlo experiment. First, each series of semi-monthly observed water table depths $\{h_t\}$ is divided into a validation set, covering the period 1982–1986, and a calibration set covering the period 1987–1991, i.e. each set consists of $n = 120$ observations. Next, using the various models fitted to the data in the calibration set, N artificial time series $\{h_t^{*i}\}$, $i = 1, \dots, N$ are generated. Initially, each series consists of 792 observations. However, to avoid the effect of any starting-up transients we discarded the first 72 observations. Hence, each generated series $\{h_t^{*i}\}$ has an effective sample size of $L = 720$ observations (thirty years of semi-monthly time steps).

The procedure to generate these data from SSTARSO models can be summarised as follows. Let $\{\hat{\epsilon}'_t; t = 1, \dots, n\}$ denote the set of standardised residuals obtained from fitting Eq. (3.2) to the series $\{h_t\}$ in the calibration period. Let $s = \max_{j,h,k} \{d, p_h^{(j)}, q_k^{(j)}\} + 1$ denote a starting point of the simulated time series $\{h_t^{*i}\}$. Put $h_t^{*i} = h_t$ for $(t = 1, \dots, s - 1)$, where \bar{h}_t is the mean of the series $\{h_t\}$ in the calibration period. Now, given a fitted SSTARSO model, the value \hat{h}_s is computed using computed values of $h_{s-1}^{*i}, h_{s-2}^{*i}, \dots$ and observed values of precipitation surplus p_{s-1}, p_{s-2}, \dots as inputs. Then a residual is selected through random sampling with replacement from the set $\{\hat{\epsilon}'_t; t = 1, \dots, n\}$. This residual is added to \hat{h}_s after being destandardised with respect to the regime in which h_{s-d}^{*i} falls, yielding a realisation h_s^{*i} . Next, \hat{h}_{s+1} is computed using values of $h_s^{*i}, h_{s-1}^{*i}, \dots$ and p_s, p_{s-1}, \dots as inputs. Then again a residual is selected at random with replacement from the set of available residuals. This residual is added to \hat{h}_{s+1} after being destandardised with respect to the regime in which the simulated value h_{s+1-d}^{*i} falls.

This procedure is continued until a complete set of 792 artificial realisations of $\{H_t\}$ is obtained. Discarding simulated values from time index s to 72, gives an effective set of simulated water table depths, denoted by h_t^{*i} ($t = 1, \dots, L (= 720)$). Note that by resampling from the set of fitted residuals no assumptions are being made on the distribution of the error process $\{\epsilon_t^{(j)}\}$. Also note that the same simulation procedure, apart from the conditioning on h_{t-d}^{*i} , can be applied to generate realisations h_t^{*i} for series of water table depths fitted by the DR model in Eq. (3.3).

The procedure we adopted to generate h_t^{*i} from TFN and SWATRE models is slightly different from the one given above. First, we decided that the models fitted in chapter 2 for the period 1988-1991, can be used for the slightly larger calibration period 1987-1991. The parameters of the models were re-estimated, now for the period 1987-1991. Based on these models, we generated two sets of realisations $\{\hat{h}_{F,t}\}$ and $\{\hat{h}_{S,t}\}$ ($t = 1, \dots, 792$) of respectively the deterministic processes $\{h_{F,t}\}$ and $\{h_{S,t}\}$. Now let $\{\hat{\epsilon}_{F,t}\}$ and $\{\hat{\epsilon}_{S,t}\}$ ($t = 1, \dots, n$) denote the sets of errors obtained from respectively the TFN and the SWATRE model. From these two sets we obtain the sets of realisations $\{\epsilon_{F,t}^{*i}\}$ and $\{\epsilon_{S,t}^{*i}\}$ ($t = 1, \dots, 792$) through random sampling with replacement. Next the sets $\{\epsilon_{F,t}^{*i}\}$ and $\{\epsilon_{S,t}^{*i}\}$ are transformed into sets of realisations $\{n_{F,t}^{*i}\}$ and $\{n_{S,t}^{*i}\}$ ($t = 1, \dots, 792$) of the noise processes $\{N_{F,t}\}$ and $\{N_{S,t}\}$ respectively, using the ARMA models fitted to the noise series in the calibration period. Then for each time point $t = 1, \dots, 792$ the realisations $\{n_{F,t}^{*i}\}$ and $\{n_{S,t}^{*i}\}$ are added to the series $\{\hat{h}_{F,t}\}$ and $\{\hat{h}_{S,t}\}$ respectively. Finally, discarding the first 72 observations of both series, we obtain two series of simulated realisations $\{h_t^{*i}\}$ of the process $\{H_t\}$, each having length $L = 720$.

In subsection 3.6.2 we will evaluate the accuracy of the above procedures using data on $\{h_t\}$ from the validation period. This will be mainly done with the aid of the following three statistics:

1. the mean error (ME) over N replications, i.e.

$$ME = \frac{1}{N} \sum_{i=1}^N \frac{1}{n} \sum_{t=481}^{600} (h_t - h_t^{*i})$$

where h_t^{*i} is the value of h_t^* obtained from the i th replication and where, for ease of comparison, the time-index of the validation series h_t is synchronised with the time-index of the corresponding simulated realisation h_t^* ;

2. the root mean squared error ($RMSE$) over N replications, i.e.

$$RMSE = \frac{1}{N} \sum_{i=1}^N \left\{ \frac{1}{n} \sum_{t=481}^{600} (h_t - h_t^{*i})^2 \right\}^{1/2};$$

and

3. the mean absolute error (MAE), i.e.

$$MAE = \frac{1}{N} \sum_{i=1}^N \frac{1}{n} \sum_{t=481}^{600} |h_t - h_t^{*i}|,$$

Table 3.1: Estimation results for the DR models fitted to the series $\{h_t, p_t\}$

Well	\hat{a}_0	\hat{a}_1	\hat{a}_2	\hat{a}_3	\hat{b}_0	\hat{b}_2	$\hat{\sigma}_\epsilon$
1	-26.07(2.91)	0.52(0.05)		0.19(0.04)	6.43(0.41)	1.32(0.49)	8.27
2	-36.80(3.19)	0.37(0.06)			5.78(0.52)		10.76
3	-25.83(4.52)	0.68(0.04)		0.20(0.04)	8.90(0.58)		11.64
4	-31.12(4.54)	0.59(0.06)	0.22(0.06)		7.88(0.61)		13.44
5	-28.69(4.80)	0.60(0.05)		0.16(0.05)	5.01(0.41)		9.55
6	-31.62(3.88)	0.53(0.05)	0.13(0.06)	0.14(0.04)	5.70(0.28)		6.46

Calibration period is 1987-1991 (120 observations). Standard errors are given in parentheses.

where $n = 120$ is the number of observations in the validation period, and where $L = 720$ is the effective length of the simulated realisations $h_t^{*i}, i = 1, \dots, N$. The total number N of replications will be fixed at 1000. The *ME* indicates systematic errors in the simulated water table depths. The *RMSE* and *MAE* are measures of the closeness of the simulated water table depths to the observed water table depths and reflect both systematic and random errors. The *MAE* is less sensitive for outlying observations than the *RMSE*. Note that we did not examine the residuals to verify the assumption that the error terms within each regime of Eq. (3.2) are white noise. Instead we validated this assumption through the simulation results following the procedure described above. If the residuals depart significantly from the white noise assumption this will show up in a poor simulation performance.

3.6 Results

3.6.1 Modelling results

Below, we will briefly discuss the (SS)TARSO models selected for the six series of water table depths $\{h_t\}$. Special attention will be given to physical explanations of the automatically selected threshold values and the estimated AR-coefficients for $\{p_t\}$. These physical explanations depend on the specific hydrological conditions at the well sites, which vary between the wells. All depths given in the discussion are relative to the ground surface elevation nearby the observation well. Estimation results based on the DR, TFN, and SWATRE+ARMA models are summarised in respectively Tables 3.1-3.3. Figure 3.1 shows graphs of results of (SS)TARSO modelling in the calibration period. Note that the graphs show a clear seasonal behaviour, with a seasonality of 24 semi-monthly time steps which equals one year. In the fitted models this seasonal nonstationarity is fully explained by the input variable p_t .

Well 1: The well is situated in a drained loamy, fine sandy soil. Drains are present at about -80 cm, relative to the ground surface at the well location. Moreover, at a distance of 50 cm to the well a trench with a bottom at about -50 cm is present. Therefore, three regimes ($\ell^\# = 3$) were assumed. Using the model selection procedure

TARSO modelling of water table depths

Table 3.2: Estimation results for the TFN models fitted to the series $\{h_t, p_t\}$

Well	$\hat{\delta}_1$	$\hat{\omega}_0$	$\hat{\omega}_1$	$\hat{\mu}$	$\hat{\phi}_1$	$\hat{\sigma}_\epsilon$
1	0.84(0.03)	6.48(0.44)	-1.78(0.56)	-91.20(1.93)	0.56(0.08)	8.87
2	0.40(0.08)	5.68(0.58)		-57.94(1.51)	0.32(0.09)	11.18
3	0.91(0.01)	8.51(0.62)	-2.09(0.68)	-208.08(2.19)	0.43(0.09)	11.64
4	0.85(0.02)	6.72(0.49)		-163.71(2.85)	0.53(0.08)	13.93
5	0.85(0.03)	4.91(0.43)	-1.38(0.54)	-125.53(1.24)	0.52(0.08)	9.79
6	0.90(0.01)	5.47(0.29)	-2.29(0.33)	-156.89(1.15)	0.42(0.09)	6.61

Calibration period is 1987-1991 (120 observations). Standard errors are given in parentheses.

of section 3.4, the final SSTARSO model is

$$h_t = \begin{cases} \begin{matrix} -16.10 & +0.58h_{t-1} & +0.24h_{t-3} & +6.81p_t \\ (4.17) & (0.06) & (0.05) & (0.43) \end{matrix} \\ \begin{matrix} +1.86p_{t-2} & +\hat{\epsilon}_t^{(1)}, & \text{if } h_{t-1} \leq -57 \\ (0.53) & & (-91, -57) \end{matrix} \\ \begin{matrix} -64.07 & +7.69p_t & +\hat{\epsilon}_t^{(2)}, & \text{if } -57 < h_{t-1} \leq -47 \\ (2.00) & (1.09) & & (-91, -57) \quad (-74, -44) \end{matrix} \\ \begin{matrix} -19.10 & +0.29h_{t-1} & +0.39h_{t-3} & +3.01p_t \\ (9.06) & (0.28) & (0.12) & (0.91) \end{matrix} \\ \begin{matrix} +\hat{\epsilon}_t^{(3)}, & \text{if } h_{t-1} > -47. \\ & (-74, -44) \end{matrix} \end{cases}$$

The residual standard deviations are $\hat{\sigma}_{\epsilon(1)} = 7.15$, $\hat{\sigma}_{\epsilon(2)} = 8.65$, and $\hat{\sigma}_{\epsilon(3)} = 6.13$, respectively. Thresholds are selected at -57 cm and -47 cm. The 95% confidence intervals of the estimated thresholds (in parentheses) are estimated from 10,000 bootstrap replications. The skewness of the intervals is a result of the short distance of the threshold at -47 cm to the upper limit of the range in which thresholds are searched; only 21 observations are present in regime 3. In the bootstrap replications thresholds are selected more often below -47 cm than above -47 cm, because at least 20 observations must be present in a regime. Similarly, thresholds are selected more often below -57 cm than above -57 cm.

It is interesting to note that the selected threshold values are possibly related to the drainage level of the trench at about -40 cm. The estimated AR-coefficient for p_t in regime 3 is small as compared with those in the other two regimes (3.01 vs. 6.81, 7.69). In physical terms the value 3.01 means that, starting from equilibrium conditions, a unit change of the precipitation surplus at time t , causes a change of 3.01 units in the water table depth h_t . Further note that p_t is the average daily precipitation surplus between $t-1$ and t . A physical explanation of the relatively small AR-coefficient for p_t in regime 3 may be that the fluctuation of the water table in regime 3 is damped by the drainage to the trench. This effect can be seen in Figure 3.1a, which shows a plot of observed water table depths and the interval in which 95% of the simulated water table depths fall.

Well 2: The well is situated in a peaty soil with a relatively shallow and constant surface water level. Three regimes were assumed ($\ell^\# = 3$), caused by the presence of

Table 3.3: Estimation results for the SWATRE+ARMA models fitted to the series $\{h_t, p_t\}$

Well	$\hat{\gamma}_d$	$\hat{\gamma}_t$	\hat{q}_b	\hat{A}_b	\hat{d}_b	\hat{F}	$\hat{\mu}$	$\hat{\phi}_1$	$\hat{\phi}_3$	$\hat{\theta}_1$	$\hat{\sigma}_\epsilon$
1	210.44	5	0.42			1.43	-0.19(2.32)	0.32(0.08)	0.36(0.08)		8.44
2	112.41	40.48	-2.33			1.73	-0.39(2.09)	0.23(0.09)	0.34(0.10)		9.91
3	1123.86	390.33	0.39	0.11	146	0.65	-0.92(2.67)	0.49(0.08)			14.82
4	244.83		0.21			0.68	1.87(3.63)	0.61(0.07)			15.75
5	215.46		-2.03			1.49	-0.79(1.44)	0.13(0.09)			13.59
6	239.60		1.73			1.15	-0.65(3.44)	0.94(0.05)		0.71(0.10)	8.96

Calibration period is 1987-1991 (120 observations). Standard errors are given in parentheses. Here, $\hat{\gamma}_d$ = drainage resistance of ditches (d); $\hat{\gamma}_t$ = drainage resistance of trenches (d); \hat{q}_b = regional component of groundwater flow (mm/d); \hat{A}_b = amplitude of regional component of groundwater flow (mm/d); \hat{d}_b = time at which maximum regional groundwater flow is reached (d); \hat{F} = multiplication factor for saturated moisture content (all layers) (-); $\hat{\mu}, \hat{\phi}_1, \hat{\phi}_3, \hat{\theta}_1$ and $\hat{\sigma}_\epsilon$ are parameters of the fitted ARMA models.

two drainage systems: a ditch with a water level at about -30 cm during the whole season and a trench with a bottom at about -20 cm. Following the procedure of section 3.4, the following TARSO model was fitted to the data:

$$h_t = \begin{cases} \begin{cases} -65.63 & +3.76p_t & +\hat{\epsilon}_t^{(1)}, & \text{if } h_{t-1} \leq -68 \\ (1.58) & (0.85) & & (-71, -47) \\ -35.45 & +0.32h_{t-1} & -0.26h_{t-2} & +0.30h_{t-3} \\ (14.19) & (0.25) & (0.09) & (0.09) \end{cases} \\ \begin{cases} & +7.47p_t & +\hat{\epsilon}_t^{(2)}, & \text{if } -68 < h_{t-1} \leq -47 \\ & (0.67) & & (-71, -47) \quad (-61, -35) \end{cases} \\ \begin{cases} -55.09 & +4.43p_t & +3.23p_{t-1} & +\hat{\epsilon}_t^{(3)}, & \text{if } h_{t-1} > -47 \\ (3.11) & (0.94) & (1.18) & & (-61, -35) \end{cases} \end{cases}$$

with $\hat{\sigma}_{\epsilon(1)} = 8.75$, $\hat{\sigma}_{\epsilon(2)} = 8.57$, and $\hat{\sigma}_{\epsilon(3)} = 10.94$.

Note that the selected thresholds at -68 cm and -47 cm do not seem to be related to any of the drainage levels. However, the estimated AR-coefficient for p_t in regime 1 and regime 3 is small as compared with the intermediate regime. This may be explained from the constant surface water level, which damps the fluctuation in the upper regime as a result of drainage in the wet season, and in the lower regime as a result of replenishment through surface water in the dry season. Also from Figure 3.1b it is clear that the fluctuation of the water table is damped. Finally, the relative large value of $\hat{\sigma}_{\epsilon(3)}$ seems to indicate that for the third regime the relation between H_t and P_t is less adequate than in the first two regimes.

Well 3: The well is situated at the slope of a fine-sandy ridge, overlaying glacial till. At a distance of 250 m a brook is present with a water level at about -240 cm. The second drainage system is a ditch with a bottom at about -130 cm, at 50 m distance to the well. The ditch drains in periods with shallow water tables only. Distinct soil physical boundaries are not present within the fluctuation zone of the water table. Again three regimes are assumed, caused by the two drainage systems. The final

TARSO modelling of water table depths

SSTARSO model fitted to the data is

$$h_t = \begin{cases} \begin{matrix} -55.21 & +0.73h_{t-1} & +5.95p_t & +\hat{\epsilon}_t^{(1)} \\ (26.53) & (0.13) & (0.70) & \end{matrix} & \text{if } h_{t-1} \leq -182 \\ & (-197, -137) \\ \begin{matrix} -165.39 & +8.50p_t & +\hat{\epsilon}_t^{(2)}, \\ (2.97) & (1.33) & \end{matrix} & \text{if } -182 < h_{t-1} \leq -150 \\ & (-197, -137) \quad (-172, -116) \\ \begin{matrix} -43.92 & +0.55h_{t-1} & +0.17h_{t-3} & +9.70p_t \\ (7.70) & (0.07) & (0.05) & (0.69) \end{matrix} & \\ & +\hat{\epsilon}_t^{(3)}, \quad \text{if } h_{t-1} > -150 \\ & (-172, -116) \end{cases}$$

with $\hat{\sigma}_{\epsilon^{(1)}} = 6.91$, $\hat{\sigma}_{\epsilon^{(2)}} = 14.82$, and $\hat{\sigma}_{\epsilon^{(3)}} = 9.50$.

It is clear that no direct relationship exists between the drainage level of the brook and the selected threshold values. The threshold at -150 cm may be caused by temporal drainage to the ditch. The value of the AR-coefficient for p_t increases from the lower (regime 1) to the upper regime (regime 3), which indicates that at this location shallow water tables vary more with the precipitation surplus than deep water tables. The wide interval in which 95% of the simulated water table depths fall (see Figure 3.1c) seems to indicate that the model could be further improved.

Well 4: The well is situated in a ridge of fine cover sands, with a distance of 130 m to a brook. The water level in the brook is at about -140 cm. A loamy layer is present between -150 cm and -170 cm. Three regimes are assumed, caused by drainage to the brook and stagnation of groundwater flow on the loamy layer. The SSTARSO model fitted to the data is

$$h_t = \begin{cases} \begin{matrix} -21.21 & +0.87h_{t-1} & +7.71p_t & +\hat{\epsilon}_t^{(1)}, \\ (14.98) & (0.09) & (0.70) & \end{matrix} & \text{if } h_{t-1} \leq -137 \\ & (-170, -114) \\ \begin{matrix} -46.65 & +0.33h_{t-1} & +0.29h_{t-2} & +7.83p_t \\ (12.96) & (0.12) & (0.05) & (0.59) \end{matrix} & \\ & -2.03p_{t-2} + \hat{\epsilon}_t^{(2)}, \quad \text{if } -137 < h_{t-1} \leq -91 \\ & (0.80) \quad (-170, -114) \quad (-117, -87) \\ \begin{matrix} -96.82 & +8.99p_t & +\hat{\epsilon}_t^{(3)}, \\ (4.73) & (2.13) & \end{matrix} & \text{if } h_{t-1} > -91 \\ & (-117, -87) \end{cases}$$

with $\hat{\sigma}_{\epsilon^{(1)}} = 11.36$, $\hat{\sigma}_{\epsilon^{(2)}} = 8.75$, and $\hat{\sigma}_{\epsilon^{(3)}} = 16.16$.

The threshold at -137 cm may be related to the loamy layer. A relation between the selected threshold at -91 cm and drainage levels or soil physical boundaries is not obvious. The relative large standard deviation in regime 3 ($\hat{\sigma}_{\epsilon^{(3)}} = 16.16$) is possibly the result of an ‘outlier’ in the calibration period at time $t = 604$, see Figure 3.1d. Dropping this value reduces the residual standard deviation to $\hat{\sigma}_{\epsilon^{(3)}} = 10.93$.

Well 5: This well is situated in fluvatile, heavy clay, covering fine sand at about -115 cm depth. A brook with a water level of about -80 cm is present at a distance of 15 m from the well. Three regimes were assumed in the model selection. The resulting

SSTARSO model is

$$h_t = \begin{cases} \begin{matrix} -136.84 & +3.91p_t & +\hat{\epsilon}_t^{(1)}, \\ (2.45) & (1.00) & \end{matrix} & \text{if } h_{t-1} \leq -133 \\ & (-134, -114) \\ \begin{matrix} -31.44 & +0.73h_{t-1} & +4.03p_t & +\hat{\epsilon}_t^{(2)}, \\ (11.70) & (0.10) & (0.34) & \end{matrix} & \text{if } -133 < h_{t-1} \leq -106 \\ & (-134, -114) \quad (-120, -86) \\ \begin{matrix} -54.86 & +0.32h_{t-1} & +0.17h_{t-3} & +7.60p_t \\ (9.27) & (0.10) & (0.06) & (0.89) \end{matrix} & \\ & +\hat{\epsilon}_t^{(3)}, \quad \text{if } h_{t-1} > -106 \\ & (-120, -86) \end{cases}$$

with $\hat{\sigma}_{\epsilon(1)} = 11.32$, $\hat{\sigma}_{\epsilon(2)} = 5.45$, and $\hat{\sigma}_{\epsilon(3)} = 10.04$.

The selected threshold at -133 cm cannot be explained by drainage levels or soil physical boundaries. The observed water table depths in Figure 3.1e show that the fluctuation in regime 1 (below -133 cm) is damped at a depth of about -140 cm, possibly as a result of regional groundwater flow. The model specification in regime 1 does not fully seem to capture this phenomenon, as may be noticed from the relative large value of $\hat{\sigma}_{\epsilon(1)}$. The threshold at -106 cm may be related to the distinct soil physical boundary between the heavy clay and the sandy subsoil. As compared with the other regimes, the AR-coefficient for p_t in regime 3 is large. This can be explained from the small storage capacity of clay, as compared with sand. Furthermore, the storage capacity decreases when the water table rises. The relative low value of the AR-coefficient for p_t in the first regime can be explained from a ditch, which dominates the fluctuation of the water table in the lower regime. This may also be the reason for the absence of AR-coefficients for h . The relative large value of $\hat{\sigma}_{\epsilon(3)}$ may be explained from hysteresis of the soil water characteristic in clay.

Well 6: The well is situated in fine cover sands, at the lower end of a slope to an ice-pushed ridge. Two regimes were assumed, as a result of temporal drainage to a ditch with a bottom at -160 cm, at a distance of 40 m to the well. The final SSTARSO model selected is given by

$$h_t = \begin{cases} \begin{matrix} -23.70 & +0.68h_{t-1} & +0.17h_{t-3} & +5.38p_t \\ (4.70) & (0.04) & (0.03) & (0.24) \end{matrix} & \\ & +\hat{\epsilon}_t^{(1)}, \quad \text{if } h_{t-1} \leq -121 \\ & (-158, -114) \\ \begin{matrix} -62.95 & +0.02h_{t-1} & +0.49h_{t-2} & +8.60p_t \\ (10.70) & (0.12) & (0.08) & (0.75) \end{matrix} & \\ & +\hat{\epsilon}_t^{(2)}, \quad \text{if } h_{t-1} > -121 \\ & (-158, -114) \end{cases}$$

with $\hat{\sigma}_{\epsilon(1)} = 5.13$ and $\hat{\sigma}_{\epsilon(2)} = 6.42$.

A threshold is selected at -121 cm (see Figure 3.1f), which is possibly related with temporal drainage to the ditch. As the AR-coefficients for p_t indicate, the water table depth in the upper regime varies more with the precipitation surplus than in the lower regime, which can be explained from the storage capacity of the soil that reduces when the water table rises.

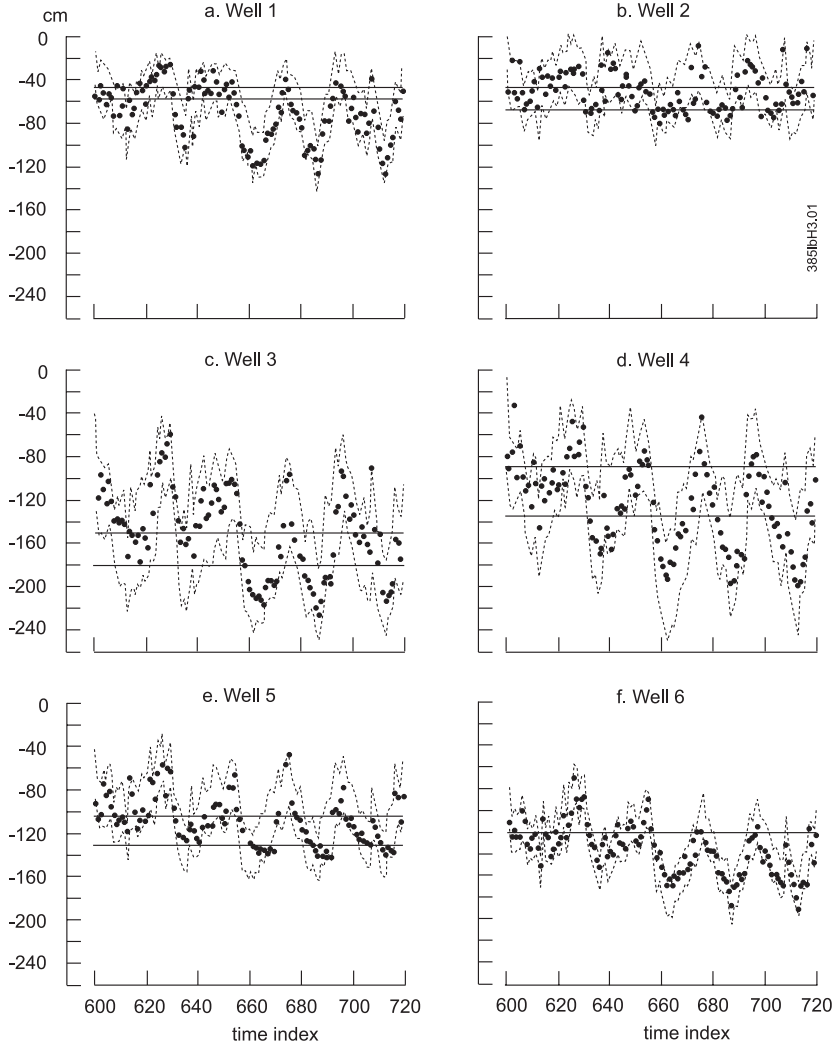


Figure 3.1: Results of (SS)TARSO modelling in the calibration period. Observed water table depths (dots), intervals in which 95% of the simulated water table depths fall (dashed lines), and selected thresholds (solid lines).

3.6.2 Results of the simulation experiment

Table 3.4 summarises the validation results of simulated water table depths. The *ME*, *RMSE* and *MAE* indicate that (SS)TARSO models perform better than the alternative models in simulations for Wells 1, 2, and 5. An improvement compared to linear models is obvious, because even if strong threshold nonlinearities are absent the (SS)TARSO models account to some extent for the non-constant relationship between precipitation surplus and water table depth caused by the non-constant storage

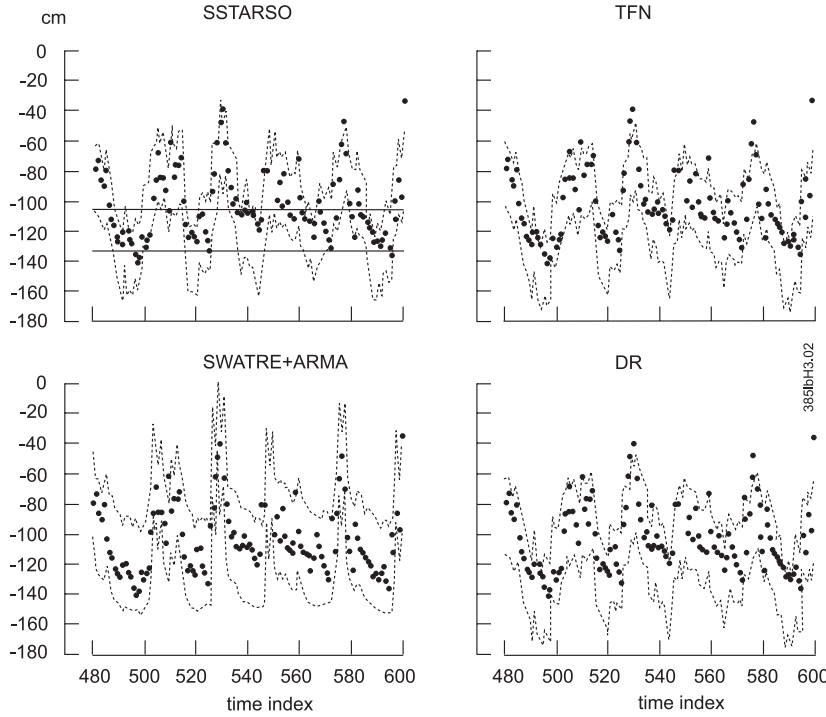


Figure 3.2: Results of model simulations in the validation period for Well 5. Observed water table depths (dots), intervals in which 95% of the simulated water table depths fall (dashed lines), and selected thresholds (solid lines).

capacity of the unsaturated zone, in contrast to linear models. More interesting the (SS)TARSO models show equal or better performance in terms of $RMSE$ and MAE than the nonlinear SWATRE+ARMA models for Wells 1, 2, 4, 5 and 6, even though in SWATRE soil physical boundaries, drainage levels and time-varying storage capacity are modelled. A reason of the relatively weak performance compared to (SS)TARSO models may be that inputs such as the standard soil physical curves do not represent the conditions at the well location.

Note from Table 3.4 that the improvement indicated by ME , $RMSE$ and MAE is largest for Well 5. As mentioned in the above subsection, the selected SSTARSO model for Well 5 contains a threshold nearby a distinct soil physical boundary between heavy clay and a sandy subsoil. The model specifications for the regimes diverge clearly. The obvious threshold nonlinearity explains why the SSTARSO model performs better than the linear TFN model and the linear DR model. A possible reason for the relatively weak performance of the SWATRE+ARMA model for Well 5 may be that the soil physical standard curves of heavy clay poorly represent the conditions at the well location.

The DR model has a poor relative performance in terms of ME , $RMSE$ and MAE as compared with the other three models. In section 3.3 we mentioned that the DR

Table 3.4: *Validation results*

Well	(SS)TARSO				TFN			
	p	ME	$RMSE$	MAE	p	ME	$RMSE$	MAE
1	8	-0.3	15.3	12.3	13	1.7	16.3	13.2
2	8	6.4	17.3	13.6	8	6.9	17.9	14.3
3	11	17.8	32.0	25.8	35	19.2	28.7	24.0
4	3	0.5	24.0	19.2	4	5.2	23.5	18.5
5	7	5.8	16.9	13.4	11	9.8	18.7	14.8
6	24	12.8	17.8	15.0	31	12.7	16.9	14.1

model can be considered as a special case of the TFN model. The main difference between the DR model (Eq. (3.3)) and the TFN model (Eq. (3.5)) is that in the TFN model the noise $N_{F,t}$ is described by an additive AR(I)MA process (Eq. (3.6)), whereas in the DR model the noise $N_{F,t}$ is assumed to follow the autoregressive structure of H_t . This will be pointed out in section 4.2, Eqs. (4.2) and (4.3). The validation results indicate that the simulation results can be improved if the noise is modelled by an additive AR(I)MA process. Therefore, it may be worthwhile to extend (SS)TARSO models with additive AR(I)MA models for the noise in each regime.

The percentages of observations outside the interval in which 95% of the simulated water table depths fall (p in Table 3.4) are generally larger than 5. In particular TFN models seem to underestimate this interval. As an example, the plots in Figure 3.2 show observed h_t and 95% intervals of simulated realisations of H_t in the validation period for Well 5. As the plots in Figure 3.2 show, the observations exceed the upper limit of the 95% interval more often than the lower limit. Furthermore, the values of ME are positive in most cases, which means that the simulated H_t is generally deep as compared with the observations. Clearly, all models simulate values of H_t which are too deep as compared with the observations in the validation period, in particular for Wells 3 and 6. We will briefly discuss some reasons for this systematic error in the simulations. A reason may be that the calibration period does not contain all information needed for simulation results free of systematic errors in the preceding validation period. However, the analyses described in subsection 2.5.1 and appendix A2.1 indicated that a four-year calibration period generally covers the dynamic response between local precipitation surplus and water table depth. Calibration of models on eight-year periods and next validation of simulation results on two-year periods reduced the systematic error only to some extent, however. Another reason may be that the water table depth shows a nonstationary behaviour during the period 1982-1991. This may be caused by the regional component of groundwater flow which influence on the water table depth may vary across years, while it is assumed to be constant in the SSTARSO, TFN and DR model, and to vary within a year only in the SWATRE+ARMA model. Temporal changes in the regional component of groundwater flow may be caused by intensified drainage in rural areas, increased groundwater withdrawal, or persistent long term meteorological fluctuations. As a result the water

3.7 Some concluding remarks

Table 3.4: *Continued.*

Well	SWATRE+ARMA				DR			
	p	ME	$RMSE$	MAE	p	ME	$RMSE$	MAE
1	9	4.3	16.4	13.0	14	1.7	16.6	13.5
2	9	7.9	17.4	13.7	6	7.2	18.1	14.4
3	12	13.2	28.6	23.0	7	19.6	33.2	27.4
4	3	3.4	25.7	20.2	3	5.7	27.3	21.7
5	6	8.6	21.4	16.7	6	9.4	20.0	15.9
6	16	12.6	20.1	16.2	19	13.0	18.1	15.1

Validation period is 1982-1986 (120 observations). p in %. ME , $RMSE$ and MAE in cm.

tables in the calibration period (1987-1991) are systematically deeper than those in the validation period (1982-1986), which cannot be explained from the input series. This is a possible reason for simulated water table depths that are systematically deeper than the observed water table depths during the validation period.

Figure 3.3 shows curves of simulated and observed durations of exceedance of water table depths in the validation period. The durations are in days per year, assuming that the cumulative frequency distributions of daily observed water table depths and semi-monthly observed water table depths are equal. The simulated durations of exceedance are calculated from 1000 simulated realisations of the water table depth in the validation period. Since it is not our purpose to compare curves between wells, we varied the scaling on the Y-axes in order to obtain a clear view on the curves of each well separately. In particular for Wells 1, 4 and 5 a good approximation of the observed durations is achieved with SSTARSO models. The duration curves computed with TFN models and DR models differ only slightly. Also from Figure 3.3 a positive systematic error is obvious in most simulations. Table 3.5 gives $RMSE$ s as a measure of the difference between simulated durations and observed durations of exceedance. Before calculating the $RMSE$, the lowest water table depth was removed from both the simulated and the observed duration curve, in order to diminish the effect of outlying values. The $RMSE$ values indicate that for Wells 1, 2 and 5 the duration curves simulated with SSTARSO models approximate the observed duration curves better than those simulated with the alternative models. As in the validation of simulated water table depths, the improvement is largest for Well 5.

3.7 Some concluding remarks

The (SS)TARSO models introduced in this chapter are adequate alternatives for the DR and the TFN models as well as the SWATRE+ARMA model in simulating water table depths and estimating durations of exceedance. The (SS)TARSO models are easy to specify and they can be interpreted physically with respect to the hydrological conditions at the well location. Physical interpretation is important if these models are generalised to a space-time context. This will be a natural step in fur-

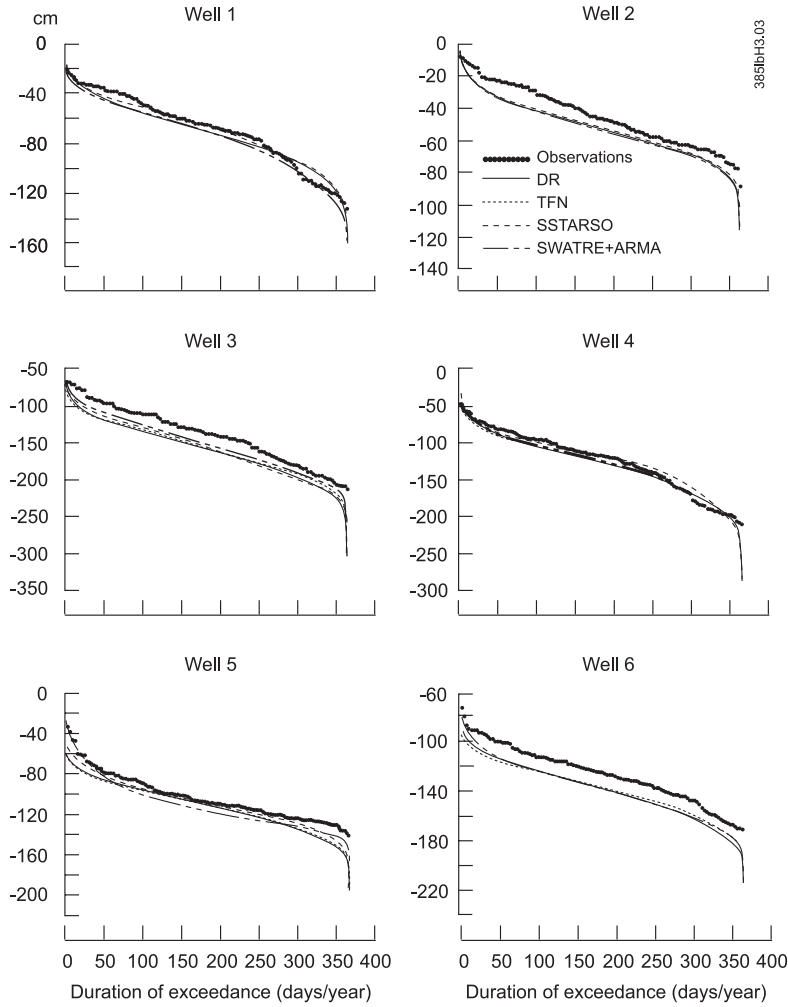


Figure 3.3: Observed and simulated durations of exceedance in the validation period.

ther research, because estimates of fluctuation characteristics are also requested for locations at which time series of water table depths are not available. Similar to the TFN model the (SS)TARSO model can be generalised to more than one input series. For instance, volumes of abstracted groundwater and surface water levels can be incorporated besides the precipitation surplus.

The comparative simulation experiment showed a systematic error in the results of all models. A possible reason may be that the regional component of groundwater flow changes in time, for instance as a result of intensified drainage in rural areas or increased groundwater withdrawal. Further research is needed to the source of this systematic error. An improvement of the (SS)TARSO models may be achieved by incorporating additive AR(I)MA processes for the noise components. The validation

3.7 Some concluding remarks

Table 3.5: Validation of duration curves by RMSEs

Well	Models			
	(SS)TARSO	TFN	SWATRE+ARMA	DR
1	4.5	6.0	5.6	5.9
2	7.3	7.8	8.7	8.1
3	18.8	19.9	13.6	20.7
4	6.9	7.4	6.2	8.0
5	7.6	11.5	9.0	11.5
6	13.4	13.1	13.0	13.5

Values are in cm. Validation period is 1982-1986 (120 observations).

results indicate that TFN models with additive AR(I)MA noise processes perform better than DR models without these components. Therefore, it is anticipated that (SS)TARSO models can be improved in this respect.

Chapter 4

Physical basis of time series models for water table depths

(This chapter is based on the paper ‘Physical basis of time series models for water table depths’ by M. Knotters and M.F.P. Bierkens, which was published in *Water Resources Research* **36**(1) (2000): 181-188, copyright ©2000 by the American Geophysical Union.)

Abstract

The relationship between precipitation surplus and water table depth can be described by empirical time series models such as transfer function-noise models (TFN), autoregressive exogenous variable models (ARX), and threshold autoregressive self-exciting open loop models (TARSO). In this paper these models are interpreted in terms of the water balance of a soil column. A physically based ARX model is used to predict the effect of an intervention on the water table dynamics at two locations. It is shown that the physically based ARX model predicts the effect of interventions reasonably well.

4.1 Introduction

In lowland areas such as the Netherlands, structural changes in the water table fluctuation will often have impact on agricultural land use and ecology. To support decision making in these areas, water managers need reliable predictions of the effects of interventions in the hydrological regime on the water table fluctuations. Preferably, these effects are expressed in terms of risks or probabilities, which implies the need of stochastic methods.

The water table depth can be related to precipitation surplus using empirical time series models such as transfer function-noise models (TFN) (chapter 2; Gehrels *et al.*, 1994; Van Geer and Zuur, 1997), transfer function models as described by Tankersley *et al.* (1993) and Tankersley and Graham (1994), ARX models or dynamic regression models (DR) (chapter 3) and threshold autoregressive self-exciting open-loop models (TARSO) (chapter 3). TFN and ARX models describe linear relationships between

input and output. The TARSO model accounts for threshold nonlinearities in the relationship between precipitation surplus and water table depth. In chapters 2 and 3, and in Bierkens and Walvoort (1998) the performance was evaluated of empirical time series models and physically based models such as the physical descriptive model SWATRE (Belmans *et al.*, 1983), a simple analytical model (Bierkens and Walvoort, 1998) and a stochastic differential equation (Bierkens, 1998). These comparative studies concluded that empirical time series models are able to simulate water table depths as well as more physically based models.

An advantage of empirical time series models is that only data on the input and output variable are required for calibration, whereas physical mechanistic models need additional information, in particular on the soil physical properties of the unsaturated zone. A disadvantage of empirical time series models as compared to physical mechanistic models is that they cannot be used to predict the effect of changes in the hydrological system on water table dynamics, because of the lack of a physical basis. Therefore, there is a need for relationships between physical quantities and the parameters of empirical time series models. If known, these relationships can be used to predict the effect of interventions in the hydrological regime on water table dynamics. Several authors have given a physical basis to time series models in hydrology. For instance, Salas and Smith (1981) demonstrated that annual streamflow time series can be represented by autoregressive moving average (ARMA) processes. Parlange *et al.* (1992) formulated a first-order autoregressive (AR(1)) model for soil water content on the basis of the hydrological budget and soil water transport equation.

The aim of this study is to analyse how parameters of empirical time series models for water table depth are related to physical quantities. In an application it will be demonstrated that physically based time series models can be used in predicting the effect of an intervention in the hydrological regime on water table dynamics.

This chapter is composed as follows. In section 4.2 the time series models describing the relationship between the potential precipitation surplus and the water table depth are given. In section 4.3 a water balance for a soil column is derived. It will be shown that the water balance can be expressed in terms of an ARX model. An extension is made to situations with threshold nonlinearities in the relationship between precipitation surplus and water table depth (TARSO model). In section 4.4 the application of a physically based ARX model to predicting the effects of interventions in the hydrological regime is illustrated by two examples from practice. Some concluding remarks will end chapter 4.

4.2 Time series models

Let $\{H_t\}$ denote a discrete time series of water table depths (output). Furthermore, let $\{P_t\}$ denote a time series of the average daily potential precipitation surplus between $t-1$ and t (input), which is calculated from the difference between daily precipitation and daily Makkink reference-crop evapotranspiration (Winter *et al.*, 1995). Now the TFN model is given by

$$H_t = H_{F,t} + N_t, \quad (4.1)$$

with

$$H_{F,t} = \sum_{i=1}^r \delta_i H_{F,t-i} + \sum_{j=0}^s \omega_j P_{t-j-b}$$

and

$$N_t - \mu = \sum_{i=1}^p \phi_i (N_{t-i} - \mu) + \epsilon_t + \sum_{j=1}^q \theta_j \epsilon_{t-j}.$$

$H_{F,t}$ denotes the part of the water table depth H_t that is explained by the precipitation surplus P_t . Here b is a delay factor, which is an integer greater or equal to zero. In this study b is assumed to be zero.

The unexplained part, or noise N_t , is independent of the level of P_t and is additive with respect to the influence of P_t . Furthermore $\{N_t\}$ is assumed to be stationary. The term $\{\epsilon_t\}$ is a discrete white noise process with zero mean and finite variance, and where μ is the mean of the process $\{N_t\}$.

If $\delta_i, i = 1, \dots, r$ equals $\phi_i, i = 1, \dots, p$ the TFN model can be written as an autoregressive moving average exogenous variable (ARMAX) model:

$$H_t - \mu = \sum_{i=1}^p \phi_i H_{F,t-i} + \sum_{j=0}^s \omega_j P_{t-j-b} + \sum_{i=1}^p \phi_i (N_{t-i} - \mu) + \epsilon_t + \sum_{k=1}^q \theta_k \epsilon_{t-k}, \quad (4.2)$$

which can be written as follows (Hipel and McLeod, 1994):

$$H_t - \mu = \sum_{i=1}^p \phi_i (H_{t-i} - \mu) + \sum_{j=0}^s \omega_j P_{t-j} + \epsilon_t + \sum_{k=1}^q \theta_k \epsilon_{t-k}.$$

If the moving average terms are taken to be zero, i.e. $\theta_k = 0, k = 1, \dots, q$, the autoregressive exogenous variable model (ARX(m, m')) results, which is given by

$$H_t - \mu = \sum_{i=1}^m a_i (H_{t-i} - \mu) + \sum_{j=0}^{m'} b_j P_{t-j} + \epsilon_t, \quad (4.3)$$

where ϵ_t is a discrete white noise process with zero mean and finite variance. In contrast to the TFN model in Eq. (4.1), the transfer component $H_{F,t}$ and the noise component N_t of the ARMAX model are interrelated, as appears from the common autoregressive operators $\phi_i, i = 1, \dots, p$ in Eq. (4.2). The TFN model is therefore more general and more flexible when fitted to observed series than the ARMAX model or the ARX model. However, in the next section it will be shown that the ARX(1,0) model can easily be related to the water balance terms of a soil column.

The above models describe linear relationships between input and output. However, the relationship between precipitation surplus and water table depth may contain several forms of nonlinearities. One form of nonlinearity is caused by the presence of thresholds which divide the relationship between precipitation surplus and water table depth into several regimes. These thresholds are, for instance, soil physical boundaries

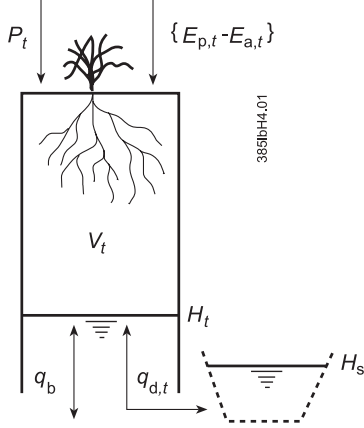


Figure 4.1: Soil column with water balance terms.

Input terms: P_t = potential precipitation surplus, i.e. precipitation–potential evapotranspiration [Lt^{-1}]; $E_{p,t} - E_{a,t}$ = difference between potential and actual evapotranspiration [Lt^{-1}]; q_b = regional groundwater flux [Lt^{-1}]. *Output terms:* V_t = volumetric water content in the unsaturated zone [L]; $q_{d,t}$ = drainage flux [Lt^{-1}]. H_s = the drainage level [L]; H_t = water table depth [L].

or drainage levels. Threshold nonlinearities in the relationship between precipitation excess and water table depth can be modelled by threshold autoregressive self-exciting open loop models (TARSO) (chapter 3; Tong, 1990). A discrete TARSO process $\{H_t, P_t\}$ with order $(\ell; (m_1, m'_1), \dots, (m_\ell, m'_\ell))$ and delay parameter d ($d > 0$) can be defined as a solution of the equations

$$H_t - \mu^{(j)} = \sum_{i=1}^{m_j} a_i^{(j)} (H_{t-i} - \mu^{(j)}) + \sum_{i=0}^{m'_j} b_i^{(j)} P_{t-i} + \epsilon_t^{(j)}, \quad \text{if } r_{j-1} \leq H_{t-d} < r_j, \quad (4.4)$$

where $-\infty = r_0 < r_1 < \dots < r_\ell = \infty$, $a_i^{(j)}$ and $b_i^{(j)}$ ($j = 1, \dots, \ell$) are parameters, and $\{\epsilon_t^{(j)}\}$ ($j = 1, \dots, \ell$) are heterogeneous white noise sequences with zero mean and finite variances $\sigma_{\epsilon^{(j)}}^2$ and each being independent of $\{P_t\}$. The thresholds are the levels $r_1, \dots, r_{\ell-1}$. The real line is partitioned into ℓ intervals, and H_t satisfies one of ℓ ARX models depending on the interval in which H_{t-d} falls. Note that Eq. (4.4) is an alternative notation of the definition of a TARSO model given by Tong (1990) with respect to the constant term.

4.3 A water balance for the phreatic groundwater zone

Figure 4.1 shows a soil profile with a water table and incoming and outgoing fluxes. At the top of the soil column two incoming fluxes are assumed: P_t , [Lt^{-1}] and

4.3 A water balance for the phreatic groundwater zone

$\{E_{p,t} - E_{a,t}\}$, $[\text{Lt}^{-1}]$. The term P_t reflects the potential precipitation surplus, which is defined by

$$P_t = P'_t - E_{p,t},$$

where P'_t $[\text{Lt}^{-1}]$ is the precipitation between $t - 1$ and t , and $E_{p,t}$ $[\text{Lt}^{-1}]$ is the potential Makkink reference-crop evapotranspiration between $t - 1$ and t . The generally unknown difference between potential and actual evapotranspiration ($E_{a,t}$ $[\text{Lt}^{-1}]$) is expressed in Figure 4.1 by the term $\{E_{p,t} - E_{a,t}\}$ $[\text{Lt}^{-1}]$. At the bottom of the soil column the flux q_b $[\text{Lt}^{-1}]$ expresses the regional component of groundwater flow, i.e. upward and downward seepage. The term $q_{d,t}$ $[\text{Lt}^{-1}]$ indicates the flux to and from surface waters such as trenches and ditches, with drainage level H_s $[\text{L}]$. The term V_t $[\text{L}]$ expresses the moisture content in the unsaturated zone, which shows a generally unknown variation in time.

Until now a discrete time scale was used, i.e. equidistant time steps indicated by subscripts t . In the following, a water balance will be constructed in terms of a linear time series model. For this purpose the continuous time scale will be used, indicated by (t) . Several terms are assumed to be independent of $H(t)$, so that a water balance with a linear structure can be constructed. First, the drainage level H_s is assumed to be independent of $H(t)$, so that the drainage flux $q_d(t)$ can be defined by

$$q_d(t) = \frac{H_s - H(t)}{\gamma},$$

where γ is the drainage resistance which is also assumed to be independent of $H(t)$. Next, the effective porosity φ $[-]$ of the porous medium in which the water table depth $H(t)$ fluctuates is assumed to be independent of $H(t)$. Finally, the regional component of groundwater flow (q_b) is assumed to be independent of $H(t)$. A water balance for the phreatic groundwater zone can now be given by

$$\varphi \frac{dH}{dt} = \frac{H_s - H(t)}{\gamma} + P(t) + q_b + \{E_p(t) - E_a(t)\} - \frac{dV}{dt}, \quad (4.5)$$

which results in the following ordinary differential equation for $H(t)$:

$$\frac{dH}{dt} = \frac{-H(t)}{\varphi\gamma} + \frac{1}{\varphi} \left\{ P(t) + q_b + \frac{H_s}{\gamma} + E_p(t) - E_a(t) - \frac{dV}{dt} \right\}. \quad (4.6)$$

The second term at the right hand side of Eq. (4.6) is called $U(t)$ in the following. If we assume that dV/dt does not depend on $H(t)$ and given that at time $t - \Delta t$ the water table depth is equal to $H(t - \Delta t)$ the solution to Eq. (4.6) is given by

$$H(t) = H(t - \Delta t)e^{-\Delta t/(\varphi\gamma)} + \int_{t-\Delta t}^t e^{-(t-\tau)/(\varphi\gamma)} U(\tau) d\tau.$$

Now, assuming that $U(t)$ is constant between $t - \Delta t$ and t , we obtain

$$H(t) = H(t - \Delta t)e^{-\Delta t/(\varphi\gamma)} + U(t)\varphi\gamma \left\{ 1 - e^{-\Delta t/(\varphi\gamma)} \right\}.$$

This results in

$$\begin{aligned}
 H(t) = & \left\{ e^{-\Delta t/(\varphi\gamma)} \right\} H(t - \Delta t) \\
 & + \gamma \left\{ 1 - e^{-\Delta t/(\varphi\gamma)} \right\} P(t) \\
 & + \{ \gamma q_b + H_s \} \left\{ 1 - e^{-\Delta t/(\varphi\gamma)} \right\} \\
 & + \gamma \left\{ [E_p(t) - E_a(t)] - \frac{\Delta V}{\Delta t} \right\} \left\{ 1 - e^{-\Delta t/(\varphi\gamma)} \right\}. \quad (4.7)
 \end{aligned}$$

The last term at the right hand side of Eq. (4.7) contains terms which are generally unknown. If this term is a white noise process then the structure of Eq. (4.7) is identical to the ARX model which is given in Eq. (4.3) with order (1,0). The first two terms at the right hand side of Eq. (4.7) describe the dynamic relationship between the input $P(t)$ and the output $H(t)$. The third term at the right hand side contains only terms that are assumed to be constant in time. The last term at the right hand side of Eq. (4.7) can be considered as the error process $\epsilon(t)$. For given Δt Eq. (4.7) can now be written as

$$H_t - \mu = a_1(H_{t-\Delta t} - \mu) + b_0 P_t + \epsilon_t, \quad (4.8)$$

with

$$\begin{aligned}
 a_1 &= e^{-\Delta t/(\varphi\gamma)}, \\
 b_0 &= \gamma(1 - a_1), \\
 \mu &= \gamma q_b + H_s. \quad (4.9)
 \end{aligned}$$

This model is referred to as the physically based ARX(1,0) model. From Eq. (4.9) it follows that $0 < a_1 < 1$ and $0 < b_0 < \infty$, as they should from a physical point of view. The physical quantities γ , φ , and q_b can now be calculated from time series parameters in the following way:

$$\begin{aligned}
 \gamma &= \frac{b_0}{1 - a_1}, \\
 \varphi &= \frac{-\Delta t}{\gamma \ln a_1}, \\
 q_b &= \frac{\mu - H_s}{\gamma}. \quad (4.10)
 \end{aligned}$$

The analysis shows that from a physical point of view the ARX(1,0) model is the most appropriate time series model among other linear alternatives. Lankester and Maas (1996) found a similar relationship between precipitation surplus and water table depth on the basis of the first two moments of the impulse response function. The parameters in Eqs. (4.9) and (4.10) can be fitted on irregularly spaced time series by using the Kalman filter application described by Bierkens *et al.* (1999).

The physically based ARX(1,0) model can easily be extended to situations with more than one drainage level. Assuming two drainage levels to be present, H_1 and

4.3 A water balance for the phreatic groundwater zone

H_2 , a water balance can be written as

$$\varphi \frac{dH}{dt} = \frac{H_1 - H(t)}{\gamma_1} + \frac{H_2 - H(t)}{\gamma_2} + P(t) + q_b + \{E_p(t) - E_a(t)\} - \frac{dV}{dt}$$

with γ_1 and γ_2 being the drainage resistances corresponding to the drainage levels H_1 and H_2 , respectively. After some derivations we obtain

$$\begin{aligned} H(t) = & \beta H(t - \Delta t) + \frac{1 - \beta}{\alpha} P(t) + \frac{1 - \beta}{\alpha} \left\{ q_b + \frac{H_1}{\gamma_1} + \frac{H_2}{\gamma_2} \right\} \\ & + \frac{1 - \beta}{\alpha} \left\{ E_p(t) - E_a(t) - \frac{\Delta V}{\Delta t} \right\}, \end{aligned} \quad (4.11)$$

with

$$\alpha = \frac{1}{\gamma_1} + \frac{1}{\gamma_2},$$

and

$$\beta = e^{-\frac{\Delta t}{\varphi} \alpha}.$$

Eq. (4.11) is analogous to the following ARX(1,0) model in continuous time:

$$H(t) - \mu' = a'_1 \{H(t - \Delta t) - \mu'\} + b'_0 P(t) + \epsilon'(t),$$

with

$$\begin{aligned} a'_1 &= \beta, \\ b'_0 &= \frac{1 - \beta}{\alpha}, \\ \mu' &= \frac{\left\{ q_b + \frac{H_1}{\gamma_1} + \frac{H_2}{\gamma_2} \right\}}{\alpha}. \end{aligned} \quad (4.12)$$

Note that the ARX(1,0) model applied in this study is chosen on the basis of physical analysis rather than selected from a large set of candidate models by a procedure of identification, estimation and diagnostic checking as described by Box and Jenkins (1976) or an automatic selection procedure using Akaike's Information Criterion (AIC) or Bayes Information Criterion (BIC) (see De Gooijer *et al.* (1985) for a survey). As a result, the ARX(1,0) model is likely to fit less well to the data than a TFN model selected from a large set of candidate models. However, as shown, the ARX(1,0) model can easily be expressed in physical quantities, which makes predicting effects of interventions in hydrological regimes possible. This will be illustrated in the next section.

The schematic soil profile in Figure 4.1 can easily be extended to a situation with more than one regime, as is shown in Figure 4.2. Threshold nonlinearities in the relationship between precipitation surplus and water table depth may be caused by soil physical boundaries, or the presence of drainage levels. Further, the regional flux

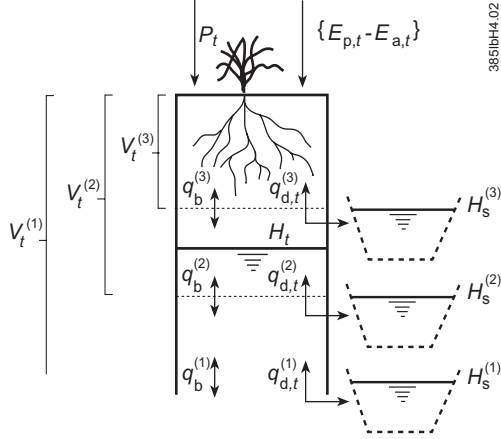


Figure 4.2: Soil column with water balance terms for three regimes. See Figure 4.1. The superscript (j) , $j = 1, \dots, 3$ indicates the regime.

may vary with the regimes. Now (4.7) can be extended to a situation with $j = 1, \dots, \ell$ regimes:

$$\begin{aligned}
 H(t) = & \left\{ e^{-\Delta t / (\varphi^{(j)} \gamma^{(j)})} \right\} H(t - \Delta t) \\
 & + \gamma^{(j)} \left\{ 1 - e^{-\Delta t / (\varphi^{(j)} \gamma^{(j)})} \right\} P(t) \\
 & + \left\{ \gamma^{(j)} q_b^{(j)} + H_s^{(j)} \right\} \left\{ 1 - e^{-\Delta t / (\varphi^{(j)} \gamma^{(j)})} \right\} \\
 & + \gamma^{(j)} \left\{ [E_p(t) - E_a(t)] - \frac{\Delta V^{(j)}}{\Delta t} \right\} \left\{ 1 - e^{-\Delta t / (\varphi^{(j)} \gamma^{(j)})} \right\} \quad (4.13)
 \end{aligned}$$

For a given Δt the structure of Eq. (4.13) equals the TARSO model given in (4.4) with order $(\ell; (1_1, 0_1), \dots, (1_\ell, 0_\ell))$. Its parameters can be written in terms of physical quantities in the following way:

$$\begin{aligned}
 a_1^{(j)} &= e^{-\Delta t / (\varphi^{(j)} \gamma^{(j)})}, \\
 b_0^{(j)} &= \gamma^{(j)} \left\{ 1 - a_1^{(j)} \right\}, \\
 \mu^{(j)} &= \gamma^{(j)} q_b^{(j)} + H_s^{(j)}.
 \end{aligned}$$

4.4 Application: predicting the effect of a hydrological intervention

In this section the physically based ARX(1,0) model is applied to predict the effect of two interventions. The first intervention took place in 1985 in the neighbourhood of an observation well in the eastern part of the Netherlands (Well 1). The data on water table depth from the period after 1985 were used to calibrate the ARX(1,0) model. The data which were collected before 1985 were used to validate the accuracy

4.4 Application: predicting the effect of a hydrological intervention

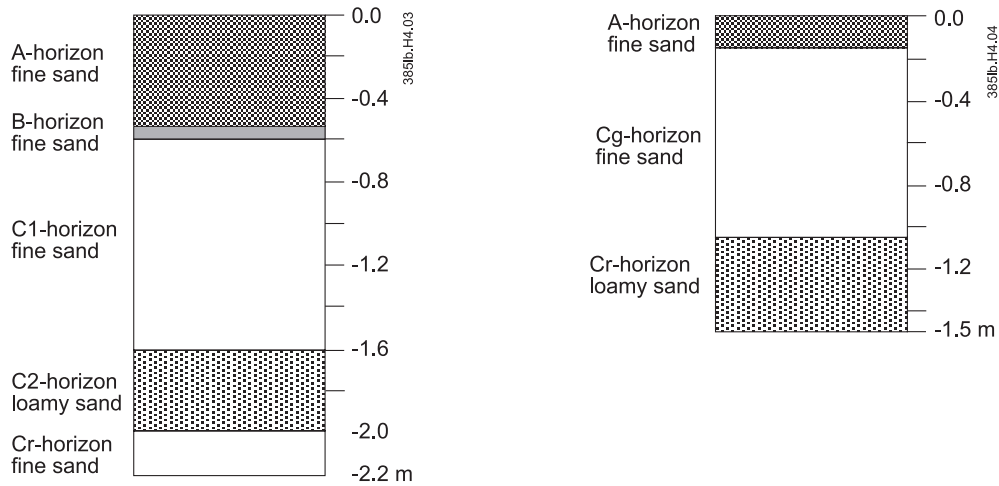


Figure 4.3: Schematic soil profiles for the sites of Well 1 (left) and Well 2 (right)

of the predicted effect of the intervention. Note that in most practical applications the data will be used the other way round, that is, calibration on data which are collected before an intervention takes place and prediction of water table depth which will occur after the intervention. However, in this application the series before the intervention was not sufficiently long enough to calibrate an ARX(1,0) model.

The second intervention took place during 1995 and 1996 in the neighbourhood of another observation well in the eastern part of the Netherlands (Well 2). The data before 1995 were used to calibrate the ARX(1,0) model, whereas the data which were observed after 1996 were used to evaluate the accuracy of the predicted effect of the intervention on the water table depth.

4.4.1 The well sites

Well 1 is situated in a slightly undulating landscape of eolian sands. The dominant land use is pasture. Observations on water table depth were available from April 28th 1982. South of the well site a water course flows. North of the well site a ditch is present. The distance between the water course and the ditch is on average 325 m. The well site is halfway between the water course and the ditch. A soil profile is given in Figure 4.3. Daily precipitation is observed at a station a few kilometers from the well site. Data on daily Makkink reference-crop evapotranspiration (Winter *et al.*, 1995) are known for the meteorological station Twente Airport at 42 km distance from the well site.

Well 2 is also situated in a slightly undulating landscape with eolian sands, mainly used as pasture. Observations on the water table depth are available from January 13th 1989. A soil profile description is given in Figure 4.3. Daily precipitation is observed 10 kilometers from the well site, daily Makkink reference crop evapotranspiration is known for the meteorological station Twente Airport at 44 kilometers distance from the well site.

Physical basis of time series models for water table depths

Table 4.1: Data on the hydrological conditions before and after the interventions took place.

Well	parameter	value before intervention	value after intervention
1	H_s	-2.2	-1.82
	H_b	-2.44	-2.60
	u	2.93	5.38
	L	325	325
2	H_s	-1.6	-1.6
	H_b	-1.5	-1.5
	u_{trenches}	0.8	0.8
	u_{ditches}	1.1	1.1
	L_{trenches}	134	108
	L_{ditches}	628	547

Values in m. All levels are relative to the ground surface at the well site. H_s = average drainage level; H_b = average bottom level of drainage devices; u = average wetted perimeter of drainage devices; L = average distance between drainage devices.

4.4.2 The interventions

Data on the hydrological conditions before and after the interventions were provided by the District Water Board “Groot Salland”.

Well 1: In 1985 the profiles of the water course and the ditch were enlarged and a pumping-engine was installed in order to maintain a higher water level in summer. As a result both the drainage resistance γ (days) and the drainage level H_s changed. The data given in Table 4.1 were used to calculate the drainage resistances before and after the intervention. For this purpose the drainage formula of Ernst (1956) was applied, which is given by

$$\begin{aligned}
 \gamma &= \gamma_v + \gamma_h + \gamma_r + \gamma_e \\
 &= \frac{D_v}{k_v} + \frac{L^2}{8 \sum_{i=1}^n k_{h,i} D_{,i}} + \frac{L}{\pi k_r} \ln \frac{\alpha D_r}{u} + \frac{L c_e}{u}
 \end{aligned} \tag{4.14}$$

where γ_v is the resistance to vertical flow (days), γ_h is the resistance to horizontal flow, γ_r is the resistance to radial flow, and γ_e is the resistance to the flow into the drainage devices, see Figure 4.4. $D_{h,i}, i = 1, \dots, n$ are the thicknesses of n layers with horizontal flow. D_v and D_r are the thicknesses of the layers with vertical and radial flow, respectively (m). $k_{h,i}, i = 1, \dots, n$ are the horizontal conductivities in the n layers with horizontal flow (m d^{-1}). k_v and k_r are the vertical and radial hydraulic conductivity in the layers with vertical and radial flow, respectively (m d^{-1}), with $k_r = \sqrt{k_h k_v}$. L is the distance between the drainage devices (m), u is the wetted perimeter (m), α is a factor accounting for the geometry of radial flow (-), and c_e is the entrance resistance (days). Three profile types can be distinguished: i) homogeneous profiles; ii) profiles with a boundary between two layers above the bottom of the drainage devices (H_b in Table 4.1), and iii) profiles with a boundary between two layers

4.4 Application: predicting the effect of a hydrological intervention

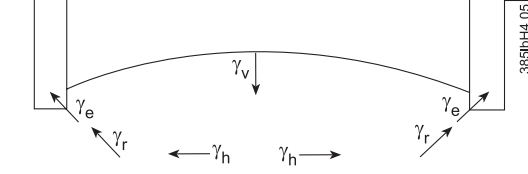


Figure 4.4: Scheme of the vertical (γ_v), horizontal (γ_h), radial (γ_r) and entrance component (γ_e) of the drainage resistance.

Table 4.2: Hydraulic conductivities used in the calculations of the drainage resistances.

Well	depth	k_h	k_v
1	0 - -1.6	15	15
	-1.6 - -2	3	1
	-2 - -2.2	15	15
2	0 - -1.15	15	15
	-1.15 - -1.5	3	1

Depths in m relative to the ground surface at the well site. k_h and k_v are horizontal and vertical saturated hydraulic conductivity, respectively, in $m\ d^{-1}$.

below the bottom of the drainage devices. Hydraulic conductivities were derived from the soil profile descriptions and a table given by Bierkens (1996), see Table 4.2. The entrance resistance c_e was assumed to be 1 day, both before and after the intervention.

Using Eq. (4.14) and the information in Tables 4.1 and 4.2 the drainage resistances before and after the intervention were estimated at 145.0 d and 90.2 d, respectively. The mean drainage levels before and after the intervention, relative to the ground surface at the well site, were estimated at -2.20 m and -1.82 m, respectively.

Well 2: In the neighbourhood of the well site trenches and ditches are present. During 1995 and 1996 the drainage in the surrounding area was intensified: the length of active trenches increased by 475 m and a trench with a length of 212 m was enlarged to the size of a ditch. Details on the hydrological situation before and after the intervention took place are given in Table 4.1. The drainage resistance before and after the intervention was calculated using Eq. (4.14) for both the trenches and the ditches. Average distances between trenches and ditches were derived from a digital topographical map, for 9 neighbourhoods with radii varying from 100 m to 900 m. The hydraulic conductivities used in Eq. (4.14) are given in Table 4.2. For a detailed description of the procedure we refer to Knotters and Bierkens (1998). The total drainage resistance before and after the intervention was calculated as follows:

$$\gamma_{tot.} = \frac{1}{\sum_{i=1}^2 \frac{1}{\min_R(\gamma_{i,R})}} \quad (4.15)$$

with $R = 100, 200, \dots, 900$ being the radius of the area for which the drainage resistance is calculated and $i = 1, 2$ indicating trenches and ditches, respectively. Following

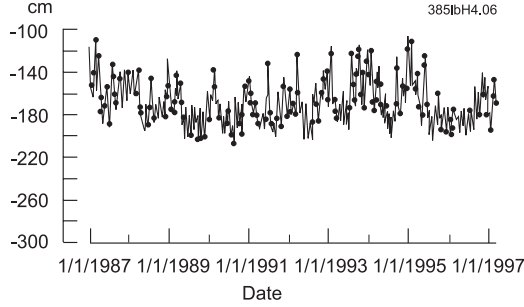


Figure 4.5: Calibrated and observed series of water table depths for Well 1. Calibration period from January 1st 1987 to March 13th 1997. ●: observed water table depth. Solid line: updated predictions.

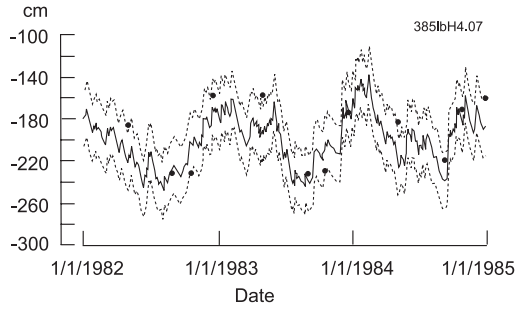


Figure 4.6: Predicted water table depths (solid line), 95% prediction interval (dashed lines), observed water table depths (●) for observation Well 1. Prediction period from January 1st 1982 to December 18th 1984.

Eqs. (4.14) and (4.15) the drainage resistances before and after the intervention were estimated at 228.3 d and 183.6 d, respectively.

4.4.3 The effect of the interventions

Well 1: Using water table depths observed between January 1st 1987 and March 13th 1997 and daily observations on precipitation and evapotranspiration, the physically based ARX(1,0) model given in Eq. (4.8) was calibrated for time intervals of one day by using the Kalman filter algorithm applied by Bierkens *et al.* (1999). With a drainage level $H_s = -182$ cm relative to the ground surface of the well site, the calibration resulted in a drainage resistance $\hat{\gamma}$ of 98.85 d, an effective porosity $\hat{\phi}$ of 0.1621 and a regional groundwater flux \hat{q}_b of 0.5775 mm d^{-1} . Following Eq. (4.9) the estimated ARX(1,0) model for $\Delta t = 1$ day is

$$h_t + 176.3 = 0.9395(h_{t-\Delta t} + 176.3) + 5.982p_t + \hat{\epsilon}_t,$$

with $\hat{\sigma}_{\epsilon} = 3.71$ cm. The results of the calibration are depicted in Figure 4.5.

Next the calibrated drainage resistance was adjusted proportionally to the drainage

4.4 Application: predicting the effect of a hydrological intervention

resistances before and after the intervention which were calculated by Eq. (4.14) (145.0 d and 90.2 d, respectively). This resulted in a drainage resistance of 158.8 d for the period before the intervention. Using the adjusted drainage resistance and assuming a drainage level of -220 cm relative to the ground surface at the well site, the water table depth was predicted for the period from January 1st 1982 to December 18th 1984, i.e. before the intervention took place. Note that the values of effective porosity φ and regional groundwater flux q_b were not adjusted. Now, using Eq. (4.9) the ARX(1,0) model adjusted for the effect of the intervention on γ and H_s is given by

$$h_t + 210.8 = 0.9619(h_{t-\Delta t} + 210.8) + 6.052p_t + \hat{\epsilon}_t.$$

The predicted time series of water table depths is depicted in Figure 4.6. The 95% prediction interval is also shown, which is estimated by

$$\left[h_{F,t} - 1.96\sqrt{\frac{\sigma_\epsilon^2}{1-a_1^2}}, h_{F,t} + 1.96\sqrt{\frac{\sigma_\epsilon^2}{1-a_1^2}} \right], \quad (4.16)$$

where $h_{F,t}$ is the (deterministically) predicted water table depth (see Eq. (4.1)), a_1 is taken from the adjusted ARX(1,0) model ($a_1 = 0.9619$) and $\sigma_\epsilon = 3.71$ cm.

Well 2: The ARX(1,0) model given in Eq. (4.8) was calibrated on water table depths observed between January 1st 1989 and December 28th 1994 and daily observations on precipitation and evapotranspiration, using the Kalman filter algorithm applied by Bierkens *et al.* (1999). The drainage level H_s both before and after the intervention is 160 cm below the ground surface at the well site, which is slightly deeper than the average bottom depths nearby the well site. The calibration resulted in a drainage resistance $\hat{\gamma}$ of 203.5 d, an effective porosity $\hat{\varphi}$ of 0.1405 and a regional groundwater flux \hat{q}_b of 1.410 mm d⁻¹. The ARX(1,0) model which describes the water table dynamics before the intervention took place is then given by

$$h_t + 131.3 = 0.9656(h_{t-\Delta t} + 131.3) + 6.997p_t + \hat{\epsilon}_t,$$

($\Delta t = 1$ day), with $\hat{\sigma}_\epsilon = 3.99$ cm. The results of the calibration are depicted in Figure 4.7.

The calibrated drainage resistance of 203.5 d was adjusted proportionally to the drainage resistances before and after the intervention which were calculated by Eqs. (4.14) and (4.15). This resulted in a drainage resistance of 163.7 d for the period after the intervention took place. Next the water table depth in the period from January 1st 1997 to November 28th 1998, i.e. after the intervention took place, was predicted by using the adjusted drainage resistance of 163.7 d in the ARX(1,0) model. The other parameters were left unchanged, i.e., $\hat{\varphi} = 0.1405$, $\hat{q}_b = 1.410$ mm d⁻¹, $H_s = -160$ cm. From Eq. (4.9) follows that the ARX(1,0) model, adjusted for the effect of the intervention on γ is given by

$$h_t + 136.9 = 0.9574(h_{t-\Delta t} + 136.9) + 6.967p_t + \hat{\epsilon}_t.$$

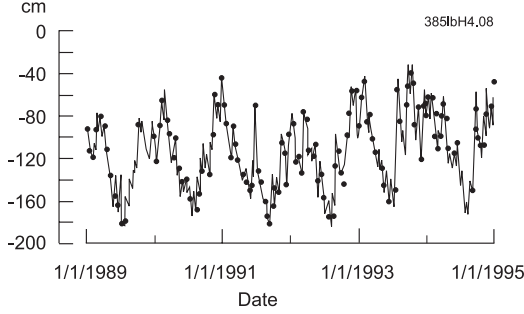


Figure 4.7: Calibrated and observed series of water table depths for Well 2. Calibration period from January 1st 1989 to December 28th 1994. •: observed water table depth. Solid line: updated predictions.

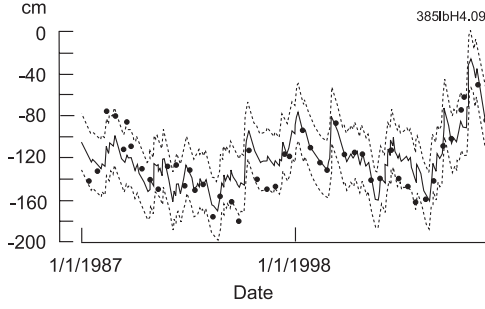


Figure 4.8: Predicted water table depths (solid line), 95% prediction interval (dashed lines), observed water table depths (•) for observation Well 2. Prediction period from January 1st 1997 to November 28th 1998.

The predicted time series of water table depths is shown in Figure 4.8. The 95% prediction interval is calculated using Eq. (4.16).

The prediction performance is validated with the following criteria:

- the mean prediction error as a measure of systematic error:

$$ME = \frac{1}{n} \sum_{i=1}^n e_i$$

where n is the number of observations and e_i is the difference between the observed and predicted water table depth.

- the standard deviation of the prediction errors as a measure of precision:

$$SDE = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (e_i - ME)^2}$$

4.4 Application: predicting the effect of a hydrological intervention

Table 4.3: Validation results of the predictions with physically based ARX(1,0) models.

Well, model	<i>ME</i>	<i>SDE</i>	<i>RMSE</i>	<i>MAE</i>
1, 1	10.58	21.76	23.37	19.93
1, 2	-21.78	25.54	32.75	24.93
2, 1	-4.28	14.30	14.77	11.33
2, 2	-13.68	13.85	19.35	15.98

Values in cm. *ME* = mean error, *SDE* = standard deviation of errors, *RMSE* = root mean squared error, *MAE* = mean absolute error.

Well 1, model 1 (adjusted for the intervention): $H_s = -2.2$ m, $\gamma = 158.8$ d, $\varphi = 0.1621$, $q_b = 0.5775$ mm d⁻¹.

Well 1, model 2 (benchmark, not adjusted for the intervention) : $H_s = -1.82$ m, $\gamma = 98.85$ d, $\varphi = 0.1621$, $q_b = 0.5775$ mm d⁻¹.

Well 2, model 1 (adjusted for the intervention): $H_s = -1.6$ m, $\gamma = 163.7$ d, $\varphi = 0.1405$, $q_b = 1.410$ mm d⁻¹.

Well 2, model 2 (benchmark, not adjusted for the intervention) : $H_s = -1.6$ m, $\gamma = 203.5$ d, $\varphi = 0.1405$, $q_b = 1.410$ mm d⁻¹.

- the root mean squared error as a measure of accuracy:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n e_i^2};$$

- the mean absolute error as measure of accuracy which is less sensitive for outlying values than the *RMSE*:

$$MAE = \frac{1}{n} \sum_{i=1}^n |e_i|.$$

Table 4.3 lists the results for the above described validation criteria. Predictions with the models that were not adjusted for the interventions are used as a benchmark. From both Table 4.3 and Figures 4.6 and 4.8 it is clear that the adjusted ARX(1,0) model is predicting the effect of the interventions reasonably well. As compared with the benchmark model that does not account for the intervention, the prediction performance clearly improved if the parameters of the ARX(1,0) model were adjusted for the intervention. The *ME* of 10.58 cm which was found for Well 1 can be explained from the fact that only the drainage resistance and the drainage level were adjusted for the intervention, whereas it may be expected that the upward regional groundwater flux q_b decreases due to higher surface water levels after the intervention. This may be mended by adjusting q_b on the basis of expert knowledge or the results of a stationary model for groundwater flow.

The results for Well 2 in Table 4.3 show that the predictions are improved in terms of *ME*, *RMSE* and *MAE* if the ARX(1,0) model is adjusted for the intervention in the drainage system. In the case of Well 2 the decreased drainage resistance mainly influences the mean water table depth: the *SDE* increases slightly if the adjusted model is applied in the predictions, whereas the *ME* reduces with 9.4 cm.

4.5 Some concluding remarks

In this chapter, empirical time series models describing the relationship between precipitation surplus and water table depth are interpreted physically. It is shown that the ARX(1,0) model can be expressed in water balance terms. Furthermore, it is shown that the TARSO model, which accounts for threshold nonlinearities in the relationship between precipitation surplus and water table depth, can be expressed in water balance terms also. For each regime, separated by thresholds, a physically based ARX(1,0) model can be derived.

The ARX(1,0) model applied in this study is chosen on the basis of physical analysis. A model identification procedure as described by Box and Jenkins (1976) or an automatic selection procedure with for instance Akaike's Information Criterion or Bayes Information Criterion (De Gooijer *et al.*, 1985) may result in models that fit better to the data than the ARX(1,0) model. However, because the physically based ARX(1,0) model can easily be expressed in physical quantities they can be attractive tools for predicting the effect of interventions in the hydrological regime on the water table dynamics. In contrast to mechanistic groundwater models only limited physical information is required. The predictions can further be improved by accounting for the effect of interventions on the regional groundwater flux. In situations with a strong nonlinear relationship between precipitation surplus and water table depth, the physically based TARSO model can be applied.

Chapter 5

Predicting water table depths in space and time using a regionalised time series model

(This chapter is based on the paper ‘Predicting water table depths in space and time using a regionalised time series model’ by M. Knotters and M.F.P. Bierkens, *Geoderma* (in press), copyright ©2001, with permission from Elsevier Science.)

Abstract

A regionalised autoregressive exogenous variable (RARX) model is presented for the relationship between precipitation surplus and water table depth. The parameters of the RARX model are ‘guessed’ at unvisited locations using auxiliary information such as soil profile descriptions, a topographic map and a digital elevation model (DEM). In the Direct Method the guessed parameters are used to predict time series of water table depths at unvisited locations; observed water table depths are not used in the prediction procedure. In the Indirect Method, observed water table depths are used to correct the predictions resulting from the Direct Method for systematic prediction errors. The prediction performance is evaluated by cross-validation. The validation results show small random errors (standard deviation on average 10 cm), but large systematic errors (absolute mean error on average 18 cm). The root mean squared error of the predicted time series is on average 22 cm. Taking the uncertainty of both the future weather conditions and the RARX-model predictions into account, a map reflecting the risk that a critical depth will be exceeded at a critical day in a future year is constructed. Furthermore, maps showing the components of uncertainty in predicted water table depths are given.

5.1 Introduction

In the Netherlands, the phreatic water table is generally found at shallow depths, say, between 0 - 2 m below the ground surface. As a consequence, the water table

depth is often critical to agriculture and nature reserves. Therefore, water managers want to be informed about the variation in both space and time of water table depth. Information in terms of risks may be useful in taking strategic decisions in land use planning and water management. Stochastic methods are able to provide this risk information.

The water table fluctuates in time mainly as a result of the precipitation surplus. Empirical time series models such as linear transfer-function noise (TFN) models (Box and Jenkins, 1976; chapter 2), and autoregressive exogenous variable (ARX) models (Hipel and McLeod, 1994; chapter 4) have proved to be useful tools in describing the dynamic relationship between precipitation surplus and water table depth. The parameters of these models can be estimated for well sites, where sufficiently long time series of precipitation surplus and water table depth are available. Since precipitation surplus data are usually available over long periods, the models can be used to simulate long time series of water table depths, say 30 years, from which all climate representative characteristics of the dynamic behaviour can be calculated.

It may be desirable, however, to not only be informed about the dynamic behaviour of the water table at the well sites, but also at other locations. To this end, the water table depth needs to be predicted in both time and space. Statistical spatio-temporal prediction methods can roughly be divided into three approaches: i) methods starting from geostatistical methodology (see Kyriakidis and Journel (1999) for a review), ii) methods based on multivariate time series modelling (see for instance, Pfeifer and Deutsch, 1980), and iii) methods based on time series models with regionalised parameters (for instance Van Geer and Zuur, 1997). Combinations of these three approaches are also possible. In some way, all these methods parametrise the spatio-temporal random function $X(\mathbf{u}, t)$, where X is the target variable, \mathbf{u} is a vector representing the spatial co-ordinates x, y , and t is the time index. Generally, these parametrisations demand many observations on the target variable X in both space and time. However, in most regional surveys in the Netherlands, only few time series of water table depths are available, and the opportunities for collecting additional groundwater data in space and time are limited. Therefore, there is a need to use widely available auxiliary information in parametrising the random function $X(\mathbf{u}, t)$. Since simple time series models describing the relationship between precipitation surplus and water table depth can be expressed in physical quantities (chapter 4), this study focuses on regionalised time series models using relationships between widely available physical information and parameters of time series models. These relationships may enable spatial prediction of time series models. This way, input series on precipitation surplus can be transformed into output series of water table depths at all points in space. From these spatially predicted or simulated series, all needed characteristics can be calculated.

The aim of this study is to predict time series of water table depths spatially using physically based regionalised time series models and databases, such as digital elevation models (DEM), digital topographic maps and soil maps, and to evaluate the accuracy of such predictions. Furthermore, it will be shown that information in terms of risks can be provided by regionalised time series models.

The chapter is composed as follows. In section 5.2, the study area and the data set are described. In section 5.3, the regionalised ARX model (RARX) is introduced and

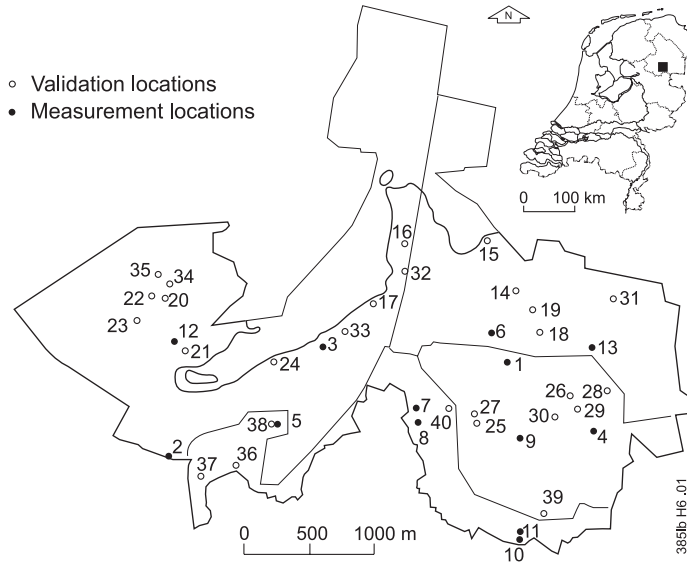


Figure 5.1: Study area and its location in the Netherlands. Closed circles: 13 locations where time series of water table depths have been observed in permanently installed wells. Open circles: 27 locations where water table depths have been observed incidentally in temporarily installed observation wells.

its physical basis is given. In section 5.4, we describe how the parameters of the RARX model are guessed from auxiliary information such as DEMs, soil profile descriptions and digital topographic maps. The set up of the prediction experiment is described in section 5.5. The results are discussed in section 5.6. In section 5.7 the sources of uncertainty are discussed. In section 5.8, two maps based on the RARX model are presented: i) a map of the risks that the water table is present within a depth of 50 cm at April 1st in any future year, given the actual hydrological conditions, and ii) a map with expected water table depths at April 1st in any future year. Furthermore, two maps with components of uncertainty about the future water table depth at April 1st are given. The chapter ends with a summary and discussion in section 5.9.

5.2 Data

The study area of 1,375 hectares is situated in the northern part of the Netherlands (Figure 5.1). A brook valley with relatively wet and peaty soils forms the southern part of the area. Here, the water table depth fluctuates roughly between 0 and 50 cm in winter and 80 and 120 cm in summer. The northern part is built up of slightly undulating cover sands, which locally cover boulder clays. In this part, the water table depth fluctuates roughly between 50 and 120 cm in winter and 120 and > 200 cm in summer. In 1997, a detailed soil map was made as part of a rural development project. At 1,185 locations soil profile descriptions were made, to a depth of 180 cm by augering (Van Dodewaard, 1997). The soil descriptions include field observations

Predicting water table depths in space and time

Table 5.1: Information on the observed time series of water table depths.

Well number	n	starting date (d-m-y)	ending date (d-m-y)	mean (cm)	st. dev. (cm)	min. (cm)	max. (cm)
1	57	30- 7-1995	16- 3-1998	-151	10.6	-172	-115
2	41	9- 5-1995	16- 3-1998	-140	18.7	-164	-86
3	40	9- 5-1995	16- 3-1998	-192	15.2	-217	-160
4	58	30- 7-1995	2- 3-1998	-138	11.3	-165	-105
5	33	1-10-1996	16- 3-1998	-117	14.8	-141	-80
6	88	29- 1-1991	28- 4-1998	-80	13.1	-108	-29
7	61	12- 2-1991	28- 5-1998	-14	10.9	-37	12
8	58	27- 3-1991	28- 5-1998	-13	9.8	-44	7
9	97	14- 1-1991	28- 5-1998	-93	13.2	-116	-29
10	90	14- 1-1991	28- 5-1998	-63	14.5	-86	-15
11	77	29- 1-1991	28- 5-1998	-43	13.4	-64	-2
12	28	2-11-1995	15- 3-1997	-118	25.9	-150	-55
13	28	2-11-1995	15- 3-1997	-115	12.8	-135	-80
14	25	16- 6-1997	12- 6-1998	-61	10.6	-79	-37
15	25	16- 6-1997	12- 6-1998	-225	22.5	-254	-180
16	25	16- 6-1997	12- 6-1998	-250	31.2	-250	-152
17	25	16- 6-1997	12- 6-1998	-184	20.4	-211	-145
18	25	16- 6-1997	12- 6-1998	-81	9.4	-95	-56
19	25	16- 6-1997	12- 6-1998	-46	11.2	-63	-17
20	25	16- 6-1997	12- 6-1998	-93	27.5	-135	-53

of texture, organic matter content, soil horizons, and gley marks such as rust mottles and the depth to the grey coloured, permanently saturated, reduced zone (G horizon).

A DEM of the area was made by laser scanning. The elevation ranges globally from 4.5 m above Dutch Ordnance Datum in the brook valley to 13.5 m above Dutch Ordnance Datum in the northern part. The accuracy of the elevation data, expressed as root mean squared error, is approximately 22 cm at a point scale.

Up-to-date digital topographic maps (scale 1 : 10,000) are available for all parts of the Netherlands. These maps distinguish several classes of drainage devices, such as trenches, and ditches of various widths. Therefore, these maps may be useful in regionalising RARX parameters that are related to the drainage characteristics. For each of the classes the lengths of drainage devices within previously defined areas can be calculated, and hence the average distance between drainage devices which belong to the same class. Data on the wetted perimeters and bottom depths of drainage devices are provided by the local water authorities.

Figure 5.1 shows the locations of 40 observation wells present in the area. In 13 permanently installed wells the water table depth is observed semi-monthly. For these wells, time series are available with a length varying from 28 to 97 observations. In 27 temporarily installed wells, the water table depth has been observed daily during 5 separate weeks covering, in 1997 and 1998, both a wet and a dry season.

Table 5.1: Continued.

Well number	n	starting date (d-m-y)	ending date (d-m-y)	mean (cm)	st. dev. (cm)	min. (cm)	max. (cm)
21	25	16- 6-1997	12- 6-1998	-77	14.9	-103	-53
22	25	16- 6-1997	12- 6-1998	-78	28.9	-118	-33
23	25	16- 6-1997	12- 6-1998	-88	26.6	-125	-45
24	25	16- 6-1997	12- 6-1998	-190	23.3	-221	-146
25	25	16- 6-1997	12- 6-1998	-90	6.1	-99	-75
26	24	16- 6-1997	12- 6-1998	-110	5.3	-117	-98
27	24	16- 6-1997	12- 6-1998	-67	9.1	-83	-46
28	25	16- 6-1997	12- 6-1998	-158	10.2	-170	-134
29	25	16- 6-1997	12- 6-1998	-93	5.9	-101	-78
30	25	16- 6-1997	12- 6-1998	-90	5.8	-98	-75
31	20	16- 6-1997	13- 3-1998	-70	13.9	-88	-40
32	25	16- 6-1997	12- 6-1998	-168	27.0	-168	-120
33	25	16- 6-1997	12- 6-1998	-128	17.8	-128	-93
34	25	16- 6-1997	12- 6-1998	-87	28.5	-87	-46
35	25	16- 6-1997	12- 6-1998	-44	29.8	-90	-4
36	25	16- 6-1997	12- 6-1998	-105	17.9	-129	-71
37	24	16- 6-1997	12- 6-1998	-86	18.6	-109	-51
38	25	16- 6-1997	12- 6-1998	-104	17.8	-125	-73
39	25	16- 6-1997	12- 6-1998	-112	13.4	-131	-81
40	25	16- 6-1997	12- 6-1998	-51	6.7	-63	-36

Table 5.1 gives details of the observed time series of water table depths. The 27 temporarily installed wells are located following a stratified random sampling design, with catchments as strata. The boulder clay area in the northern part was excluded from the network of observation wells, because in these soils the water table depth could not be accurately measured (see Figure 5.1). The remaining area in which the wells are located amounts 976 hectares.

The groundwater data set is roughly representative for local studies in areas of comparable size in the Netherlands.

Daily data on precipitation are available from the station Dedemsvaart nearby the study area. Daily data on the potential Makkink crop reference evapotranspiration (Winter *et al.*, 1995) are available from the station Eelde at about 55 km distance of the study area. The precipitation surplus is calculated by

$$p_t = p'_t - e_{p,t}, \quad (5.1)$$

where p'_t is the daily precipitation [Lt^{-1}] and $e_{p,t}$ is the daily potential Makkink crop reference evapotranspiration [Lt^{-1}].

5.3 A regionalised ARX model (RARX)

We apply the first order autoregressive exogenous variable model (ARX) for the relationship between precipitation surplus and water table depth, because this model can easily be related to physical quantities, as shown the previous chapter. These relationships to physical quantities will be used in regionalising the time series model. The ARX model used in this study is defined by

$$H_t - \mu = a_1\{H_{t-\Delta t} - \mu\} + b_0P_t + \epsilon_t,$$

where t indicates the time step, Δt indicates the time interval ($\Delta t = 1$ day in this study), H_t [L] is the water table depth at time t , P_t is the mean potential precipitation surplus between $t - \Delta t$ and t [L t^{-1}], a_1 [-], b_0 [t] and μ [L] are parameters and ϵ_t [L] is an error term which is assumed to be a white noise sequence with zero mean and finite and constant variance σ_ϵ^2 . This model is referred to as the ARX(1,0) model, where 1 is the time lag in the autoregressive term of the output variable H and 0 is the time lag in the autoregressive term of the input variable P .

The above defined ARX(1,0) model describes the dynamic relationship between precipitation surplus and water table depth at one point in space. However, it may be expected that this relationship varies in space, depending on hydrologic conditions and soil physical conditions. This means that the parameters a_1 , b_0 and μ are space dependent. Furthermore, it is likely that the error term ϵ_t and the error variance σ_ϵ^2 vary in space. Now, a regionalised form of the ARX(1,0) model is given by

$$H_t(\mathbf{u}) - \mu(\mathbf{u}) = a_1(\mathbf{u})\{H_{t-\Delta t}(\mathbf{u}) - \mu(\mathbf{u})\} + b_0(\mathbf{u})P_t + \epsilon_t(\mathbf{u}), \quad (5.2)$$

with error variance $\sigma_\epsilon^2(\mathbf{u})$, and with $\mathbf{u} = (x, y)$ indicating the spatial co-ordinates x and y . We refer to Eq. (5.2) as to the regionalised autoregressive exogenous variable ($\mathbf{u}; 1, 0$) model (abbr. RARX($\mathbf{u}; 1, 0$)). Note that P_t is assumed to be global, that is, space invariant. This assumption is quite reasonable in relatively small study areas, as is often the case in the Netherlands. For larger areas a spatially varying $P_t(\mathbf{u})$ can, for instance, be obtained through spatial interpolation using radar data.

The RARX($\mathbf{u}; 1, 0$) model given in Eq. (5.2) is a special form of the more general extension of the transfer-function noise model (TFN) to the space domain given by Van Geer and Zuur (1997). Van Geer and Zuur (1997) consider first order autoregressive processes for both the transfer component and the noise component for reasons of simplicity, that is, to easily incorporate them into a Kalman Filter algorithm. We consider in Eq. (5.2) a first order autoregressive model for reasons of physical comprehensibility, because physical relationships may be helpful in regionalisation of the RARX parameters. As explained in the previous chapter, the ARX model can be considered as a special form of the transfer function-noise model: the transfer component and the noise component have equal autoregressive parameters and therefore, both components are interrelated in the ARX model. The more general TFN model allows for a noise component which is independent of the transfer component.

Based on the water balance of a soil column given in chapter 4, relationships between ARX(1,0) model parameters and physical quantities were derived, assuming that these physical quantities are independent of $H_t(\mathbf{u})$. For the RARX($\mathbf{u}; 1, 0$) model these

5.4 Guessing RARX parameters from physical information

relationships are:

$$a_1(\mathbf{u}) = e^{-\Delta t / (\varphi(\mathbf{u})\gamma(\mathbf{u}))}, \quad (5.3)$$

where $\varphi(\mathbf{u})$ is the effective porosity [-], and $\gamma(\mathbf{u})$ is the drainage resistance [t],

$$b_0(\mathbf{u}) = \gamma(\mathbf{u}) \{1 - a_1(\mathbf{u})\}, \quad (5.4)$$

and

$$\mu(\mathbf{u}) = \gamma(\mathbf{u})q_b(\mathbf{u}) + H_s(\mathbf{u}), \quad (5.5)$$

where $q_b(\mathbf{u})$ is the flux to the shallow groundwater from the deeper groundwater systems [Lt^{-1}], and $H_s(\mathbf{u})$ is the drainage level (e.g. the surface water level in ditches and trenches, relative to the ground surface at location \mathbf{u}) [L].

The error term can be expressed in physical quantities by

$$\epsilon_t(\mathbf{u}) = \gamma(\mathbf{u}) \left\{ [E_{p,t} - E_{a,t}(\mathbf{u})] - \frac{\Delta V(\mathbf{u})}{\Delta t} \right\} \left\{ 1 - e^{-\Delta t / (\varphi(\mathbf{u})\gamma(\mathbf{u}))} \right\} + \epsilon_t^\bullet(\mathbf{u}), \quad (5.6)$$

where $[E_{p,t} - E_{a,t}(\mathbf{u})]$ is the difference between potential and actual evapotranspiration, and $\frac{\Delta V(\mathbf{u})}{\Delta t}$ is the change of the water content in the unsaturated zone during Δt . Note that the error term contains those water balance terms that are usually not very well known: both the actual evapotranspiration and the water content in the unsaturated zone are generally neither measurable nor accurately predictable against reasonable costs. The term $\epsilon_t^\bullet(\mathbf{u})$ contains the remaining unknown influences.

5.4 Guessing RARX parameters from physical information

If only a limited number of observed time series of water table depths is available, then it makes sense to use auxiliary information extensively when predicting water table depths in space and time. The physical relationships that are given in the previous section can be used to ‘guess’ the RARX($\mathbf{u};1,0$) parameters from physical information. The term ‘guess’ instead of ‘estimate’ or ‘predict’ is used to emphasise that physical information is straightforwardly transformed into RARX($\mathbf{u};1,0$) parameter values, instead of executing a calibration or optimisation. Guessed parameters are denoted as $\tilde{\cdot}$ in this thesis. In this section, it will be explained how the RARX($\mathbf{u};1,0$) parameters $a_1(\mathbf{u})$, $b_0(\mathbf{u})$, and $\mu(\mathbf{u})$ can be guessed from physical information in data sources such as DEMs, soil profile descriptions, information on surface water levels, and digital topographic maps. In Eqs. (5.3) and (5.4) it can be seen that $a_1(\mathbf{u})$ and $b_0(\mathbf{u})$ both depend on the drainage resistance $\gamma(\mathbf{u})$ and the effective porosity $\varphi(\mathbf{u})$. The methods by which these two physical quantities can be guessed are described in the next two subsections. In the third subsection it is explained how $\mu(\mathbf{u})$ can be guessed from physical information. Using the procedures given in the next three subsections guessed RARX parameter values can be obtained for the 1,185 locations with soil profile descriptions.

5.4.1 Guessing drainage resistance

The drainage resistance cannot be observed directly in the field, but can be calculated from information on drainage devices and saturated conductivities of the subsoil. In this study the formula of Ernst (1956) is used, which is given by

$$\begin{aligned}\gamma &= \gamma_v + \gamma_h + \gamma_r + \gamma_e \\ &= \frac{D_v}{k_v} + \frac{L^2}{8 \sum_{i=1}^n k_{h,i} D_{h,i}} + \frac{L}{\pi k_r} \ln \frac{\alpha D_r}{u} + \frac{L c_e}{u}\end{aligned}\quad (5.7)$$

where γ_v is the resistance to vertical flow, γ_h is the resistance to horizontal flow, γ_r is the resistance to radial flow, and γ_e is the resistance to the flow into the drainage devices. $D_{h,i}$, $i = 1, \dots, n$ are the thicknesses of n layers with horizontal flow. D_v and D_r are the thicknesses of the layers with vertical and radial flow, respectively (m). $k_{h,i}$, $i = 1, \dots, n$ are the horizontal conductivities in the n layers with horizontal flow (m d^{-1}). k_v and k_r are the vertical and radial hydraulic conductivity in the layers with vertical and radial flow, respectively (m d^{-1}), with $k_r = \sqrt{k_h k_v}$. L is the distance between the drainage devices (m), u is the wetted perimeter (m), α is a factor accounting for the geometry of radial flow (-), and c_e is the entrance resistance (d). Three profile types can be distinguished in determining the value of α : i) homogeneous profiles; ii) profiles with a boundary between two layers above the bottom of the drainage devices and iii) profiles with a boundary between two layers below the bottom of the drainage devices.

Three types of drainage device were distinguished:

1. trenches and periodically dry ditches;
2. ditches with a width up to 3 m;
3. ditches with a width larger than 3 m.

The average distance between drainage devices was derived from digital topographic maps (1:10,000) for areas with a specified radius. Bottom depths and wetted perimeters of the ditches were provided by the local water authorities. Bottom depths relative to the ground surface were calculated using the DEM. Hydraulic conductivities were derived from the soil profile descriptions and a table given by Bierkens (1996). Effective horizontal and vertical conductivities for multilayered profile types are calculated in the following way:

$$k_{h,\text{eff.}} = \frac{1}{n} \sum_{i=1}^n k_{h,i},$$

$$k_{v,\text{eff.}} = \frac{1}{\sum_{i=1}^n \frac{1}{k_{v,i}}},$$

where n is the number of layers. The entrance resistance c_e was assumed to be 1 day.

5.4 Guessing RARX parameters from physical information

In the presence of 3 drainage types the total drainage resistance can be calculated as follows:

$$\tilde{\gamma} = \frac{1}{\sum_{i=1}^3 \frac{1}{\min_R(\gamma_{i,R})}} \quad (5.8)$$

with $R = 100, 200, \dots, 1000$ m being the radii of the area for which the drainage resistances are calculated. This way, for each type of drainage device (1, 2, or 3), the smallest drainage resistance is selected from average drainage resistances of neighbourhoods of varying size.

5.4.2 Guessing effective porosity

The effective porosity is the ratio between a change in the water balance during Δt and the resulting change in water table depth. The effective porosity not only depends on the porosity of the medium in which the water table fluctuates, but also on the storage of water in puddles and preferential flow or ‘bypass flow’. Prior information about the porosity is available from the soil profile descriptions and the soil physical standard curves for the main soil horizons in the Netherlands (Wösten *et al.*, 1987). From this information various quantities which may be related to the effective porosity can be calculated, for instance the storage coefficient if the pressure head profile is at equilibrium and the water table depth is equal to the depth to the permanently saturated, reduced zone. This depth is observed by augering during the soil survey.

The storage coefficient μ_s [-] was calculated using the program CAPSEV (Wesseling, 1991), as follows:

$$\mu_s = \frac{\sum_{i=1}^n \{\theta_{\text{sat},i} - \theta(h_p(z_i)) + \theta_{\text{residual},i}\} \Delta z_i}{z_g} \quad (5.9)$$

where n is the number of discretisation intervals (0.01 m) between the ground surface and the top of the permanently saturated zone, z_g is the water table depth (m) which is set equal to the depth to the permanently saturated zone here, z_i is the depth of the i th discretisation interval of 0.01 m, $\theta_{\text{sat},i}$ is the saturated moisture content in the i th discretisation interval (m^3m^{-3}), $\theta(h_p(z_i))$ is the moisture content corresponding to the pressure head h_p at depth z_i (m^3m^{-3}), and $\theta_{\text{residual},i}$ is the residual moisture content in the i th discretisation interval (m^3m^{-3}).

The permanently saturated zone may not be present within the augering depth of 1.80 m. Omitting these values or replacing them by the censor depth (1.80 m) would lead to a bias in the guessed effective porosities. Therefore, these censored observations were replaced by their maximised likelihood value (2.04 m) in the calculations (Knotters *et al.*, 1995).

5.4.3 Guessing $\mu(\mathbf{u})$

As indicated by Eq. (5.5) the RARX parameter $\mu(\mathbf{u})$ depends on the drainage level $H_s(\mathbf{u})$, the drainage resistance $\gamma(\mathbf{u})$ and the regional groundwater flux $q_b(\mathbf{u})$. The value of $H_s(\mathbf{u})$ can easily be calculated as the difference between the surface elevations from the DEM and the surface water levels provided by the local water authorities.

The value of $\gamma(\mathbf{u})$ can be guessed in the way described in subsection 4.1. However, the value of $q_b(\mathbf{u})$ can not be easily and accurately guessed from available data (Knotters and Bierkens, 1998). Therefore we guessed the value of $\mu(\mathbf{u})$ directly from the hydromorphic soil profile characteristics that were observed in the field. These characteristics roughly indicate the fluctuation zone of the water table. The value of $\mu(\mathbf{u})$ was guessed from the average of the depth to the permanently saturated zone and the depth to the top of the hydromorphic characteristics, such as rust mottles.

5.5 Spatial prediction of time series and validation

5.5.1 A direct method

A first approach in predicting time series of water table depths spatially may be to interpolate the guessed RARX parameters ‘directly’ from the 1,185 augering locations to a location \mathbf{u} and next to transform a p_t series into a $\tilde{h}_t(\mathbf{u})$ series. The performance of this method is evaluated as follows:

1. Guess the values of $a_1(\mathbf{u})$, $b_0(\mathbf{u})$, and $\mu(\mathbf{u})$ at the 1,185 augering locations, using the methods described in section 5.4 and Eqs. (5.3) and (5.4);
2. Model the spatial structures of the guessed parameter values;
3. Interpolate the guessed parameter values to the 27 validation locations which are located following a stratified random sampling design (Figure 5.1);
4. Transform the p_t series (Eq. (5.1)) into a series of water table depths at the 27 validation locations by using the interpolated, guessed RARX parameters and Eq. (5.2);
5. Calculate the differences between observed and predicted water table depths and validation criteria:

$$e_j(\mathbf{u}_i) = h_j(\mathbf{u}_i) - \tilde{h}_j(\mathbf{u}_i), j = 1, \dots, n(\mathbf{u}_i), \quad (5.10)$$

where $n(\mathbf{u}_i)$ is the number of observed water table depths at the i th validation location. The systematic error in a predicted time series is calculated by

$$ME(\mathbf{u}_i) = \frac{1}{n(\mathbf{u}_i)} \sum_{j=1}^{n(\mathbf{u}_i)} e_j(\mathbf{u}_i). \quad (5.11)$$

The standard deviation of error in a predicted time series is calculated by

$$SDE(\mathbf{u}_i) = \sqrt{\frac{1}{n(\mathbf{u}_i) - 1} \sum_{j=1}^{n(\mathbf{u}_i)} (e_j(\mathbf{u}_i) - ME(\mathbf{u}_i))^2}. \quad (5.12)$$

$SDE(\mathbf{u}_i)$ can be interpreted as a measure of the error in predicting the temporal variation of the water table depth at location \mathbf{u}_i .

The root mean squared error of a predicted time series is calculated by

$$RMSE(\mathbf{u}_i) = \sqrt{\frac{1}{n(\mathbf{u}_i)} \sum_{j=1}^{n(\mathbf{u}_i)} e_j(\mathbf{u}_i)^2}. \quad (5.13)$$

The mean absolute error of a predicted time series is calculated by

$$MAE(\mathbf{u}_i) = \frac{1}{n(\mathbf{u}_i)} \sum_{j=1}^{n(\mathbf{u}_i)} |e_j(\mathbf{u}_i)|. \quad (5.14)$$

Both $RMSE(\mathbf{u}_i)$ and $MAE(\mathbf{u}_i)$ are measures of the overall closeness of the predictions to the observations. $MAE(\mathbf{u}_i)$ is less sensitive to outlying observations than $RMSE(\mathbf{u}_i)$. Note that first the RARX parameters are interpolated and next the water table depths are calculated, that is, ‘interpolate first, calculate later’ or IC. The other way round (CI) would imply that for every time step t the spatial structure of $\tilde{H}_t(\mathbf{u})$ has to be modelled (see for instance Bierkens, 2001).

5.5.2 An indirect method

In the ‘direct’ method the observed water table depths are not used for prediction. It may be attractive in practice, however, to improve the predictions by using observed water table depths. The observations can be used to correct the predictions for systematic errors. Furthermore, the observations can be used to predict the standard deviation of the error process, $\sigma_\epsilon(\mathbf{u})$, which is used to construct a risk map (section 5.8). The ‘indirect’ prediction method is evaluated by a cross-validation procedure:

1. See the above steps 1 and 2;
2. Interpolate the guessed RARX parameters to the 40 locations where water table depths are observed (Figure 5.1);
3. Transform the p_t series into 40 series of water table depths by using the interpolated, guessed, RARX parameters using Eq. (5.2);
4. Calculate errors, mean errors $ME(\mathbf{u}_k)$ and standard deviations of error $SDE(\mathbf{u}_k)$, $k = 1, \dots, 40$ indicating the observation well, by replacing i by k in Eqs. (5.10) to (5.12);
5. Estimate the standard deviation of the error term $\epsilon(\mathbf{u})$ of the RARX model (Eq. (5.2)) for the k locations in the following way:

$$\tilde{\sigma}_\epsilon(\mathbf{u}_k) = \sqrt{(1 - \tilde{a}_1^2(\mathbf{u}_k))SDE^2(\mathbf{u}_k)}, \quad (5.15)$$

with $k = 1, \dots, 40$. Eq. (5.15) is true for first order autoregressive processes with finite variance, which implies that $0 < |\tilde{a}_1(\mathbf{u}_k)| < 1$ (see for instance Chatfield, 1989, p. 36). $\tilde{\sigma}_\epsilon(\mathbf{u}_k)$ will be used in constructing a risk map (section 5.8);

6. Model the spatial structure of $ME(\mathbf{u})$ and $\tilde{\sigma}_\epsilon(\mathbf{u})$ using $ME(\mathbf{u}_k)$ and $\tilde{\sigma}_\epsilon(\mathbf{u}_k)$, $k = 1, \dots, 40$;
7. Leave one of the 27 validation locations out, say $\mathbf{u}^\#$, and interpolate the mean error and $\tilde{\sigma}_\epsilon$ from the remaining 39 locations to this location, say $\widetilde{ME}(\mathbf{u}^\#)$ and $\tilde{\sigma}_\epsilon(\mathbf{u}^\#)$;
8. Correct the predicted water table depths at location $\mathbf{u}^\#$ for $\widetilde{ME}(\mathbf{u}^\#)$;
9. Repeat steps 7 and 8 for all 27 validation locations;
10. Calculate the differences between observed and predicted water table depths and the validation criteria, see Eqs. (5.10) to (5.14).

In this cross-validation experiment the models for the spatial structures of $\widetilde{ME}(\mathbf{u})$ and $\tilde{\sigma}_\epsilon(\mathbf{u})$ were estimated once on the basis of all 40 well locations, instead of 27 times for 39 locations for each cross-validation run apart. It may be expected, however, that a bias resulting from this is limited.

5.6 Results

5.6.1 Regionalised RARX parameters

Guessed values of $\tilde{a}_1(\mathbf{u})$ were calculated from guessed drainage resistance and effective porosity using Eq. (5.3), for the 1,185 augering locations. The spatial structure of $\tilde{a}_1(\mathbf{u})$ was modelled by a nested spherical isotropic variogram:

$$\gamma(h) = \begin{cases} 0.001223 + 0.003190 \left(1.5 \frac{h}{365.6} - 0.5 \left(\frac{h}{365.6} \right)^3 \right) \\ \quad + 0.01662 \left(1.5 \frac{h}{2363} - 0.5 \left(\frac{h}{2363} \right)^3 \right), & \text{if } h \leq 365.6, \\ 0.001223 + 0.003190 + 0.001662 \left(1.5 \frac{h}{2363} - 0.5 \left(\frac{h}{2363} \right)^3 \right), & \text{if } 365.6 < h \leq 2363, \\ 0.001223 + 0.003190 + 0.001662, & \text{if } h > 2363, \end{cases} \quad (5.16)$$

where h is the lag distance. The sample variogram and the fitted model (Eq. (5.16)) are shown in Figure 5.2a. The model was fitted by the weighted least squares method, using the numbers of pairs of observations in the distance classes as weights. Figure 5.2b shows a map of $\tilde{a}_1(\mathbf{u})$, obtained by ordinary kriging using Eq. (5.16) (Deutsch and Journel, 1992). As Figure 5.2b indicates, the value of $\tilde{a}_1(\mathbf{u})$ increases in the northern direction. The small values in the southern part are caused by the relatively small drainage resistances and a relatively small effective porosity. This is representative for fast reacting groundwater systems with shallow water tables and intensified drainage, as can be found in the southern part of the study area. In the northern part, the drainage resistances are generally larger and the effective porosities are generally

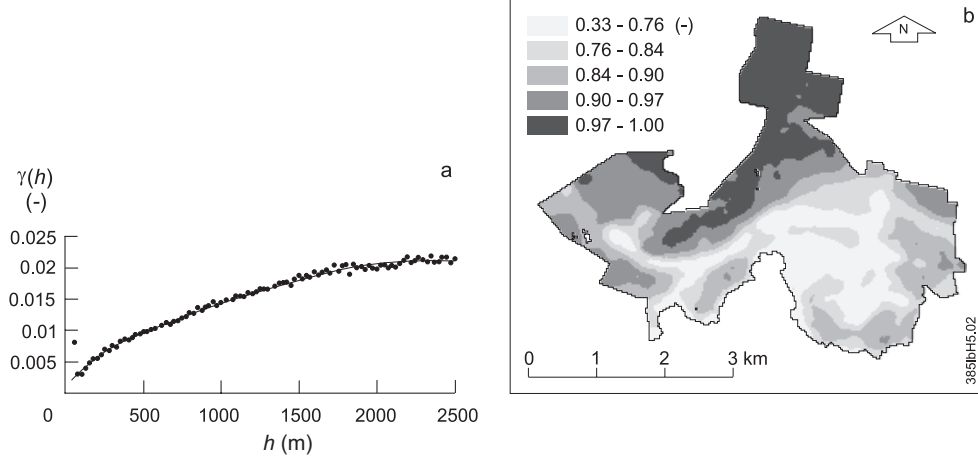


Figure 5.2: Spatial structure of $\tilde{a}_1(\mathbf{u})$, **a** Sample variogram (dots) and fitted variogram model (line), **b** Map of $\tilde{a}_1(\mathbf{u})$.

larger too, because of thicker unsaturated zones. Therefore, in this part of the study area larger values of $\tilde{a}_1(\mathbf{u})$ are found.

Guessed values of $\tilde{b}_0(\mathbf{u})$ were calculated from guessed drainage resistance and guessed effective porosity using Eq. (5.4). The spatial structure of $\tilde{b}_0(\mathbf{u})$ was modelled by the following spherical isotropic variogram:

$$\gamma(h) = \begin{cases} 3.363 + 4.801 \left(1.5 \frac{h}{1834} - 0.5 \left(\frac{h}{1834} \right)^3 \right), & \text{if } h \leq 1834, \\ 3.363 + 4.801, & \text{if } h > 1834. \end{cases} \quad (5.17)$$

Figure 5.3a shows the sample variogram of $\tilde{b}_0(\mathbf{u})$ and the fitted model given in Eq. (5.17). Figure 5.3b shows a map of $\tilde{b}_0(\mathbf{u})$, obtained by ordinary kriging using Eq. (5.17). Figure 5.3b indicates large values of $\tilde{b}_0(\mathbf{u})$ in the southern part. This can be explained from the relatively small effective porosities; starting from equilibrium, a change in precipitation surplus will result in a larger immediate change of water table depth in the southern part than in the northern part. In the northern part, the effective porosity is generally larger because of thicker unsaturated zones. Of course, the relatively small drainage resistances in the southern part will reduce the effect of a small effective porosity on the values of $\tilde{b}_0(\mathbf{u})$, but only to some extent, as Figure 5.3b indicates.

Guessed values of $\tilde{\mu}(\mathbf{u})$ were obtained from the 1,185 soil profile descriptions (subsection 4.3). The spatial structure of $\tilde{\mu}(\mathbf{u})$ was modelled by an exponential isotropic variogram:

$$\gamma(h) = 197.1 + 407.2 \left\{ 1 - \exp \left(-\frac{3h}{1244} \right) \right\} \quad \text{if } h \leq 500. \quad (5.18)$$

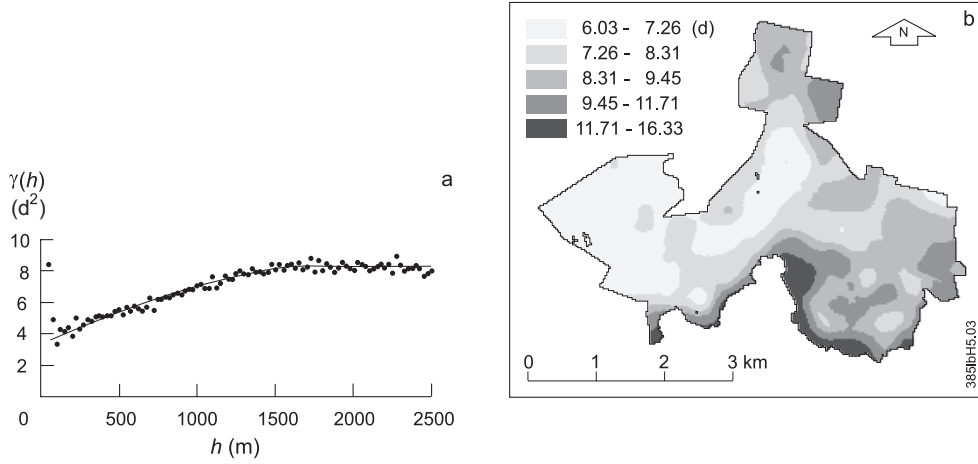


Figure 5.3: Spatial structure of $\tilde{b}_0(\mathbf{u})$, **a** Sample variogram (dots) and fitted variogram model (line), **b** Map of $\tilde{b}_0(\mathbf{u})$.

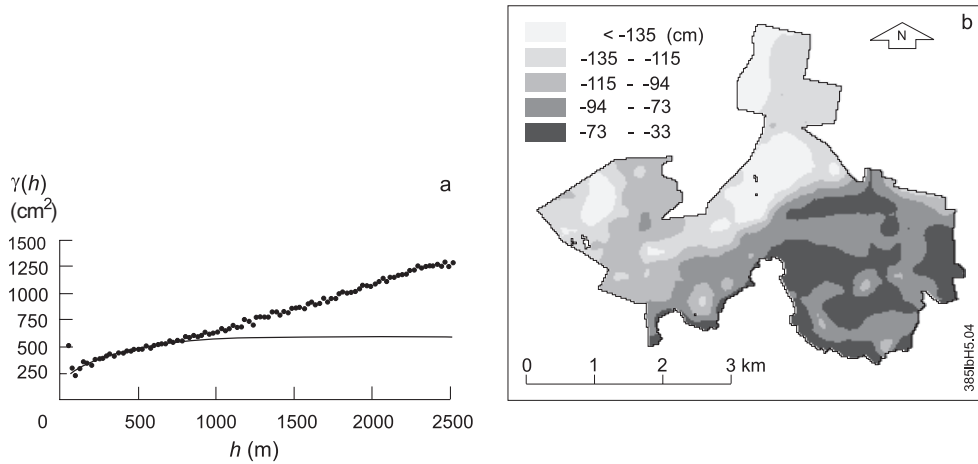


Figure 5.4: Spatial structure of $\tilde{\mu}(\mathbf{u})$, **a** Sample variogram (dots) and fitted variogram model (line), **b** Map of $\tilde{\mu}(\mathbf{u})$.

Figure 5.4a shows the sample variogram of $\tilde{\mu}(\mathbf{u})$ and the fitted model given in Eq. (5.18). For distances larger than 500 m a trend seems to be present, as the sample variogram indicates. Therefore the spatial structure is described locally up to a distance of 500 m, and ordinary kriging with a search neighbourhood is used. Figure 5.4b shows a map of $\tilde{\mu}(\mathbf{u})$, obtained by local ordinary kriging and Eq. (5.18). Figure 5.4b reflects the relatively shallow water tables in the southern part of the study area and the relatively deep water tables in the northern part, from which the presence trend in $\mu(\mathbf{u})$ can be explained (see section 5.2).

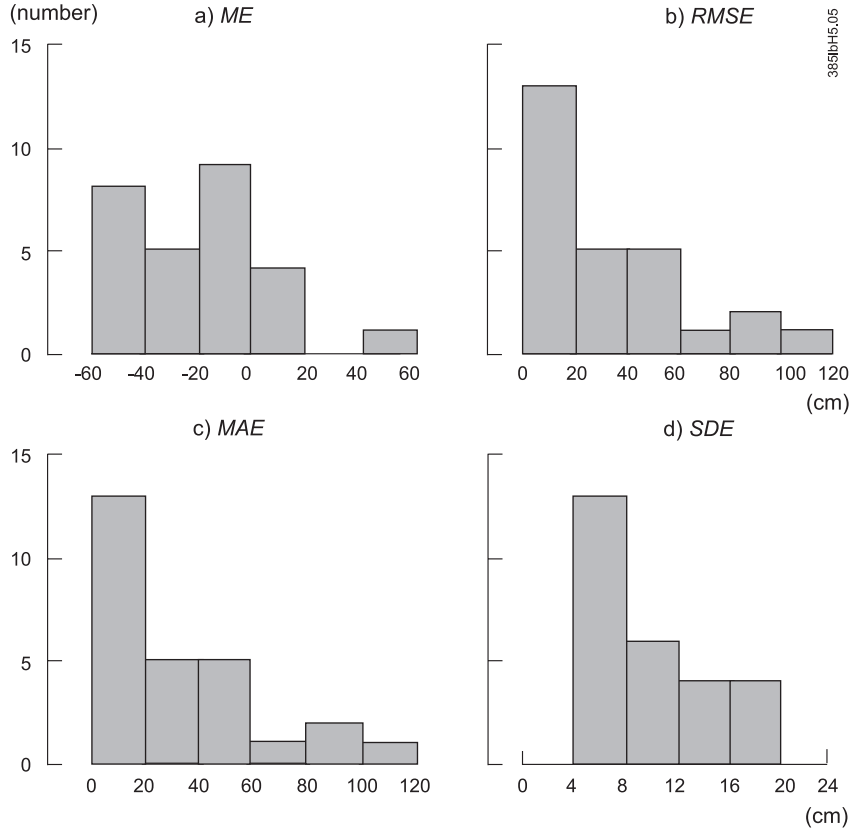


Figure 5.5: Validation results of the Direct Method (values in cm), **a** mean error (ME), **b** root mean squared error (RMSE), **c** mean absolute error (MAE), **d** standard deviation of error (SDE).

5.6.2 Validation results for the direct method

Figures 5.5a to 5.5d show the validation results of the direct approach in predicting time series of water table depths. On average, the *RMSE* is equal to 34 cm (Figure 5.5b). Several *RMSE* values are large, due to large systematic errors (*ME*) in the predicted water table depths. The *SDE* values (Figure 5.5d) are small as compared to the *ME* and *RMSE* values. Since the *SDE* can be interpreted as a measure of the error in predicting the temporal variation of water table depth, percentages of variance accounted for were calculated, using the standard deviations of water table depths given in Table 5.1. The percentage of variance accounted for is 64 % on average. Note that the predictions are deterministic, that is, the observed p_t series is straightforwardly transformed into $\tilde{h}(\mathbf{u}_i)$ series using Eq. (5.2) and the guessed, interpolated RARX($\mathbf{u};1,0$) parameters.

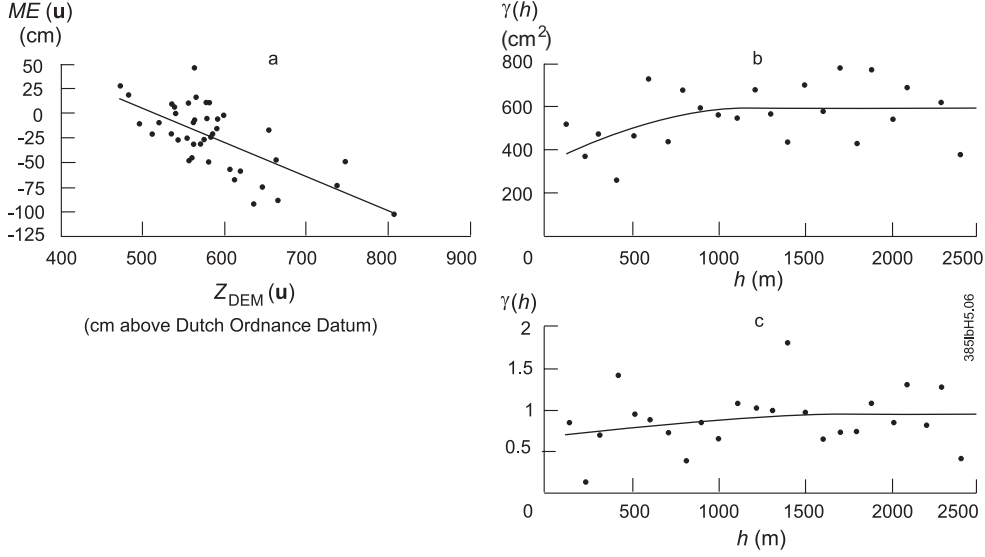


Figure 5.6: Spatial structure of $ME(u)$ and $\tilde{\sigma}_\epsilon(u)$, **a** $ME(u)$ against $z_{DEM}(u)$, **b** Sample variogram (dots) and fitted variogram model (line) of regression residuals of the regression $ME(u)$ against $z_{DEM}(u)$, **c** Sample variogram (dots) and fitted variogram model (line) of normalised $\tilde{\sigma}_\epsilon(u)$.

5.6.3 Validation results for the indirect method

The ME values observed at the 40 well locations appeared to be related to the ground surface elevation. The following relation was found between systematic errors and digital elevation data (z_{DEM}) (Figure 5.6a):

$$ME(u) = \begin{matrix} 180.1 \\ [34.5] \end{matrix} - \begin{matrix} 0.3487z_{DEM}(u) \\ [0.0586] \end{matrix} + \varepsilon(u) \quad \hat{\sigma}_\varepsilon = 25.03 \text{ cm}, \quad (5.19)$$

(standard errors in brackets) with ME and z_{DEM} in cm and $R_{adj}^2 = 46.9\%$. The spatial structure of the residual term $\varepsilon(u)$ in Eq. (5.19) was modelled by

$$\gamma(h) = \begin{cases} 343.1 + 245.0 \left(1.5 \frac{h}{1089.7} - 0.5 \left(\frac{h}{1089.7} \right)^3 \right), & \text{if } h \leq 1089.7, \\ 343.1 + 245.0, & \text{if } h > 1089.7, \end{cases} \quad (5.20)$$

see Figure 5.6b. The ME values were interpolated to the validation locations by means of External Drift Kriging (Deutsch and Journel, 1992), using the variogram model given in Eq. (5.20) and with z_{DEM} as an external drift.

The spatial structure of $\tilde{\sigma}_\epsilon(u)$ was modelled by the following spherical variogram

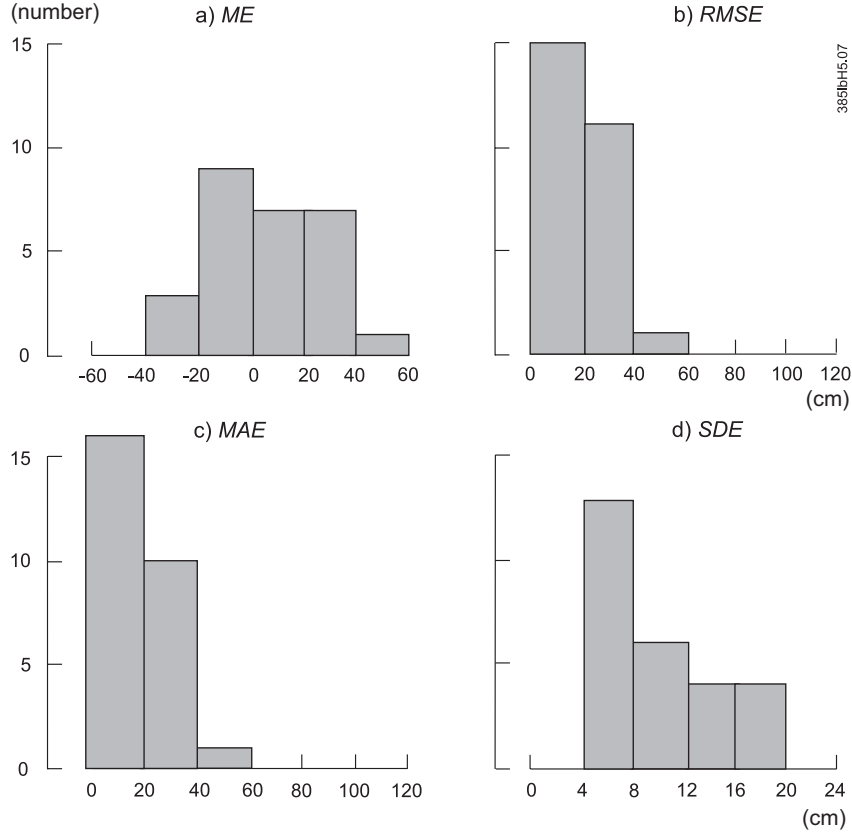


Figure 5.7: Validation results of the Indirect Method (values in cm), **a** mean error (*ME*), **b** root mean squared error (*RMSE*), **c** mean absolute error (*MAE*), **d** standard deviation of error (*SDE*).

model after transformation to a standard normal distribution:

$$\gamma(h) = \begin{cases} 0.6616 + 0.2240 \left(1.5 \frac{h}{1455} - 0.5 \left(\frac{h}{1455} \right)^3 \right), & \text{if } h \leq 1455, \\ 0.6616 + 0.2240, & \text{if } h > 1455, \end{cases} \quad (5.21)$$

see Figure 5.6c. Eq. (5.21) was used in constructing a risk map (section 5.8).

Figures 5.7a to 5.7d show the results of the indirect method in predicting time series of water table depths. As compared to Figure 5.5a a reduction of systematic errors can be seen in Figure 5.7a. Keeping in mind the accuracy of the DEM ($RMSE = 22$ cm), the predictions seem reasonably accurate with an average $RMSE$ value of 22 cm (Figure 5.7b). The precision of the predictions (Figure 5.7d) is the same as in the direct method (Figure 5.5d), because only the systematic error was reduced.

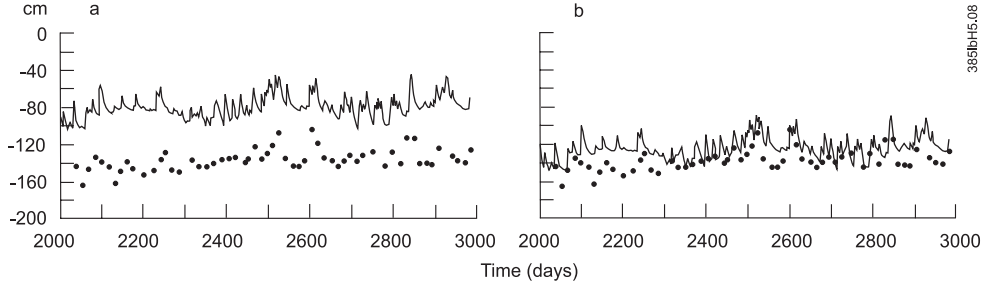


Figure 5.8: Prediction results for Well 4. Horizontal axes: day number (day 1 = 1 January 1990); Vertical axes: depth in cm relative to the ground surface. **a** Direct Method, **b** Indirect Method.

Figures 5.8a and 5.8b show time series plots of observed and predicted water table depths, obtained by the Direct Method and the Indirect Method, respectively, for Well 4 (see Table 5.1). The reduction of the systematic error obtained by the Indirect Method is obvious: in this example, $ME = -58.2$ cm in the Direct Method, whereas $ME = -14.2$ cm in the Indirect Method. The standard deviation of error, SDE , is 9.2 cm for both methods, which corresponds with a percentage of (temporal) variance accounted for of 33 %. This possibly indicates that the linear ARX(1,0) model not adequately describes the relationship between precipitation surplus and water table depth for the location of well number 4.

5.7 Sources of uncertainty

In the errors that are observed in the cross-validation experiments, the following sources can be distinguished:

1. the assumption that the p_t series represents the precipitation surplus at all locations in the study area;
2. inadequacy of the ARX(1,0) model to describe the relationship between precipitation surplus and water table depth;
3. the guessed values $\tilde{a}_1(\mathbf{u})$, $\tilde{b}_0(\mathbf{u})$, and $\tilde{\mu}(\mathbf{u})$ at the 1,185 augering locations. These guessed values may contain errors because of possible weak representativeness of the physical relationships which are used in the guessing procedures (section 5.4), and errors in the data sets;
4. the interpolation of the values of $\tilde{a}_1(\mathbf{u})$, $\tilde{b}_0(\mathbf{u})$, and $\tilde{\mu}(\mathbf{u})$, from the 1,185 augering locations to the ‘unvisited’ locations, such as the 27 validation locations;
5. the correction for systematic error, because $ME(\mathbf{u})$ was observed at a limited sample and its spatial structure was modelled using a limited number of locations (only in the Indirect Method).

5.7 Sources of uncertainty

Table 5.2: Error components in $RARX(\mathbf{u};1,0)$ model predictions

Well number	A		B		DM		IM
	ME	SDE	ME	SDE	ME	SDE	ME
1	-2.3	8.4	-32.9	8.4	-66.4	8.5	-35.1
2	0.9	12.5	-10.3	12.7	-49.3	14.7	-24.7
3	0.5	9.8	-4.4	13.3	-74.0	14.5	-36.2
4	-3.2	8.6	-38.0	11.3	-58.2	9.2	-14.2
5	-0.2	9.7	-2.1	10.1	-30.4	11.2	-9.2
6	0.0	10.4	-6.1	10.6	7.4	10.5	17.7
7	0.3	6.6	0.8	12.6	19.0	10.4	7.9
8	0.3	6.2	1.4	14.9	26.7	10.3	13.8
9	0.0	11.0	-9.2	11.1	-7.4	11.2	16.6
10	-0.3	10.4	2.1	10.5	-21.0	11.0	-23.0
11	-0.4	9.1	3.9	14.7	9.3	11.4	31.1
12	0.7	13.4	-30.0	18.5	-26.3	20.1	-24.9
13	0.2	7.4	-51.2	8.6	-47.2	8.6	-28.9

ME = mean error. SDE = standard deviation of error. *A*: $RARX(\mathbf{u};1,0)$ model parameters fitted on observed time series. *B*: $RARX(\mathbf{u};1,0)$ model parameters guessed from physical information obtained at the well sites. *DM* (Direct Method): Guessed $RARX(\mathbf{u};1,0)$ model parameters, interpolated from augering locations to unvisited locations. *IM*: Results of the Indirect Method, that is, results of *DM* corrected for the systematic error, interpolated to unvisited locations. Results for *DM* and *IM* obtained by cross-validation. Values in cm.

The effects of these sources of uncertainty on the predictions were analysed by cross-validation for the 13 locations, where long time series of water table depths are observed. Table 5.2 gives the results of this analysis. The ME and SDE values arising from the fitted $RARX(\mathbf{u};1,0)$ model indicate the error resulting from sources 1 and 2. If $RARX(\mathbf{u};1,0)$ model parameters are guessed using physical information from the well site, then the ME and SDE values increase in comparison with the values arising from the fitted models, as a result of source 3. At unvisited locations, where soil profile descriptions are not available, the guessed $RARX(\mathbf{u};1,0)$ model parameters are predicted by interpolation. The ME and SDE values arising from these interpolated, guessed, $RARX(\mathbf{u};1,0)$ model parameters are larger than those arising from the fitted model, but are not necessarily larger than those arising from the guessed values, as Table 5.2 indicates. This can be explained from the weak representativeness of local physical information and from the limited sample size used in observing the errors. From the results in Table 5.2, it is clear that the systematic error is predominant. The last column of Table 5.2 shows the reduction of the systematic error that is obtained by interpolating observed systematic errors to unvisited locations in order to correct the preliminary predictions (in the Indirect Method).

Table 5.2 shows generally small differences in SDE values. For instance, for Well 4 the SDE arising from both the Direct and the Indirect Method is 9.2 cm which

corresponds with a percentage of variance accounted for of 33 % (see Figure 5.8). The performance of the fitted model is only slightly better: a SDE of 8.6 cm, which corresponds with a percentage of variance accounted for of 41 %. This indicates that in this situation, the linear ARX(1,0) model may not adequately describe the relationship between precipitation surplus and water table depth.

5.8 Mapping risks

In section 5.5, it was described how time series of water table depths can be predicted spatially at 40 validation locations. In a similar way, time series of water table depths can be predicted at each node of the DEM-grid. Hence, for all grid nodes, the dynamic behaviour of the water table can be characterised. Stochastic methods enable us to characterise the water table dynamics in terms of risk, which may be valuable in taking strategic decisions on water management. The map presented in this section reflects the risk that the water table is within a critical depth, at a critical day in a future year. The procedure for constructing risk maps starts with the following three assumptions:

- the daily precipitation surplus over the last 30 years is representative for the climatic conditions in the future 30 years;
- the mean prediction errors ($ME(\mathbf{u})$, Eq. (5.11)) and standard deviations of errors ($SDE(\mathbf{u})$, Eq. (5.12)) are independent of the predicted water table depths;
- $ME(\mathbf{u})$ is independent of $SDE(\mathbf{u})$.

The procedure is as follows:

1. Interpolate the guessed RARX parameters $\tilde{a}_1(\mathbf{u})$, $\tilde{b}_0(\mathbf{u})$, and $\tilde{\mu}(\mathbf{u})$ from the 1,185 augering locations to the M nodes of the DEM grid, say $\tilde{a}_1(\mathbf{u}_g)$, $\tilde{b}_0(\mathbf{u}_g)$, and $\tilde{\mu}(\mathbf{u}_g)$, with $g = 1, \dots, M$ indicating the g th grid location ($M = 21,547$);
2. Simulate N realisations of the mean errors $ME(\mathbf{u})$ conditional to the observed mean errors at the 40 well locations (Eq. (5.11)), say $ME^{(i)}(\mathbf{u}_g)$, with $i = 1, \dots, N$ indicating the i th realisation ($N = 100$ in this study), using the variogram given in Eq. (5.20) and External Drift Kriging with z_{DEM} as an external drift in combination with sequential simulation (Deutsch and Journel, 1992);
3. Transform the 30-year input series P_t into $N \times M$ series $\tilde{h}_t^{(i)}(\mathbf{u}_g)$ using

$$\begin{aligned} \tilde{h}_t^{(i)}(\mathbf{u}_g) - (\tilde{\mu}(\mathbf{u}_g) + ME^{(i)}(\mathbf{u}_g)) = \\ \tilde{a}_1(\mathbf{u}_g) \left\{ \tilde{h}_{t-1}^{(i)}(\mathbf{u}_g) - (\tilde{\mu}(\mathbf{u}_g) + ME^{(i)}(\mathbf{u}_g)) \right\} + \tilde{b}_0(\mathbf{u}_g) p_t, \end{aligned}$$

with t representing the t th day in the 30-year series;

4. Select the 30 values of $\tilde{h}_t^{(i)}(\mathbf{u}_g)$ for which t corresponds with the predefined critical day in the year, for instance April 1st, say $\tilde{h}_d^{(i)}(\mathbf{u}_g)$, $d = 1, \dots, 30$;

5. Simulate N realisations of the standard deviation of the error process $\tilde{\sigma}_\epsilon(\mathbf{u})$ conditional to the values estimated at the 40 well locations (Eqs. (5.12) and (5.15)), say $\sigma_\epsilon^{(i)}(\mathbf{u}_g)$, $i = 1, \dots, N$, using the model of spatial structure given in Eq. (5.21) and sequential Gaussian simulation (Deutsch and Journel, 1992);
6. Standardise the observed prediction errors at the 40 well sites (Eq. (5.10)) in the following way:

$$e'_k(\mathbf{u}_o) = \frac{e_k(\mathbf{u}_o) - ME(\mathbf{u}_o)}{SDE(\mathbf{u}_o)},$$

with $o = 1, \dots, 40$ indicating the o th observation well and $k = 1, \dots, n_o$ indicating the k th error from the o th series. This results in a series of standardised errors $e'_j, j = 1, \dots, K, K = \sum_{o=1}^{40} n_o$;

7. Destandardise the series $e'_j, j = 1, \dots, K$ for each grid location \mathbf{u}_g and each realisation $\tilde{\sigma}_\epsilon^{(i)}(\mathbf{u}_g)$ as follows:

$$e_j^{(i)}(\mathbf{u}_g) = e'_j \sqrt{\frac{\tilde{\sigma}_\epsilon^{2,(i)}(\mathbf{u}_g)}{1 - \tilde{a}_1^2(\mathbf{u}_g)}}; \quad (5.22)$$

8. Add $e_j^{(i)}(\mathbf{u}_g), j = 1, \dots, K$ to each $\tilde{h}_d^{(i)}(\mathbf{u}_g)$. In this way, the distribution of the predicted water table depth for the critical day in any future year at each location \mathbf{u}_g is approximated by $K \times N \times 30$ values, say $\tilde{h}_d^{(i,j)}(\mathbf{u}_g)$ with $i = 1, \dots, N, j = 1, \dots, K$ and $d = 1, \dots, 30$. Hence, the frequency of exceeding a predefined critical level can be estimated.

The following two components of uncertainty in predicted water table depths at a critical day in a future year can be distinguished:

1. Uncertainty about the precipitation surplus in a future year. This uncertainty can be expressed by the variance, say σ_1^2 , of the predicted water table depths at the 30 critical days:

$$\hat{\sigma}_1^2(\mathbf{u}_g) = \frac{1}{29} \sum_{d=1}^{30} \left\{ \tilde{h}_d^{(i)}(\mathbf{u}_g) - \left(\frac{1}{30} \sum_{d=1}^{30} \tilde{h}_d^{(i)}(\mathbf{u}_g) \right) \right\}^2.$$

This variance is independent of $ME(\mathbf{u}_g)$ and can be estimated from any of the N series $\tilde{h}_d^{(i)}(\mathbf{u}_g), i = 1, \dots, N$;

2. Uncertainty about the relationship between precipitation surplus and water table depth as described by the RARX($\mathbf{u}; 1, 0$) model with guessed parameters. This source of uncertainty can be divided into a systematic component ($ME(\mathbf{u})$) and a random component ($SDE(\mathbf{u})$). These components are observed at the well sites (Eqs. (5.11) and (5.12)). The uncertainty about the relationship between precipitation surplus and water table depth described by RARX($\mathbf{u}; 1, 0$) model with guessed parameters can be quantified by the following variance:

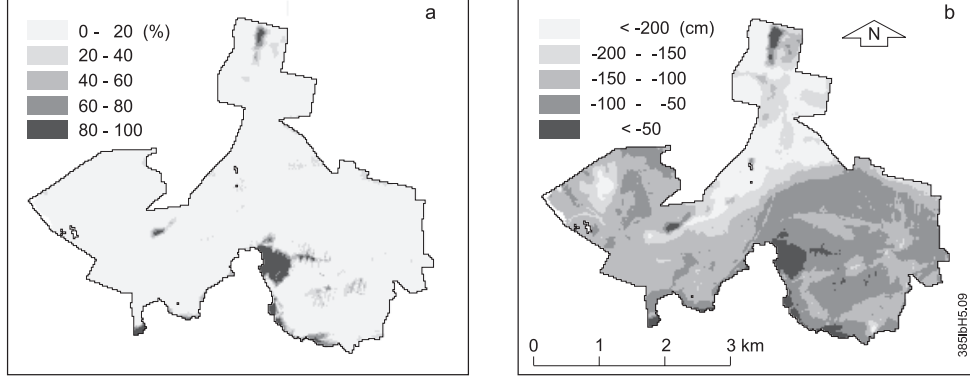


Figure 5.9: **a** Map of the risk that at April 1st in a future year the water table depth will be shallower than 50 cm; **b** Map of the expected water table depth at April 1st in a future year. Values in cm.

$$\hat{\sigma}_2^2(\mathbf{u}_g) = \frac{1}{(N-1) \times (K-1) \times 29} \sum_{k=1}^N \sum_{l=1}^K \sum_{d=1}^{30} \left\{ \tilde{h}_d^{(k,l)}(\mathbf{u}_g) - \left(\frac{\sum_{i=1}^N \sum_{j=1}^K \sum_{m=1}^{30} \tilde{h}_m^{(i,j)}(\mathbf{u}_g)}{N \times K \times 30} \right) \right\}^2 - \hat{\sigma}_1^2(\mathbf{u}_g).$$

A map showing the risk that the water table depth at April 1st in a future year will be shallower than 50 cm is shown in Figure 5.9a. April 1st is representative for the start of the growing season. At that time, shallow water tables will negatively influence the soil trafficability and soil temperature. A map of the expected water table depth at April 1st in a future year is shown in Figure 5.9b.

If the risk is close to 50%, it will be difficult to take a decision in water management or land use planning. Risks close to 50% arise from the choice of a critical level close to the median of the probability distribution of $\tilde{h}_t(\mathbf{u})$, or from high uncertainty, either about the future weather conditions or the predictions by the $\text{RARX}(\mathbf{u};1,0)$ model. These high uncertainties result in a median of the probability distribution of $\tilde{h}_t(\mathbf{u})$, close to the chosen critical level. Figures 5.10a and 5.10b show both components of uncertainty. The highest uncertainties concern the northern part of the study area, where the temporal variation of water table depth is large and the mean water table is found at large depths. Furthermore, the parameter $a_1(\mathbf{u})$ has large values in the northern part, see Figure 5.2b. Consequently, the noise component has large values, as follows from Eq. (5.22), and thus uncertainty component 2 is large. Besides this, the large values of $a_1(\mathbf{u})$ result in large variation between years, and thus the uncertainty component 1 is large also.

If the main source of uncertainty is in the relationship between precipitation surplus and water table depth as described by the $\text{RARX}(\mathbf{u};1,0)$ model, additional effort can

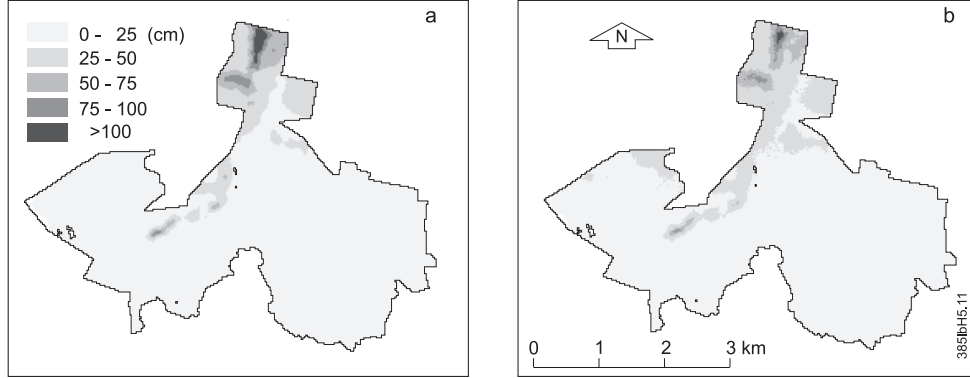


Figure 5.10: Maps of uncertainty components in predicted water table depth at April 1st in a future year, expressed in standard deviations (cm). **a** uncertainty arising from unknown weather conditions, **b** uncertainty arising from the $RARX(\mathbf{u};1,0)$ model.

be made to reduce this uncertainty. For instance, a possible reduction of uncertainty can be achieved by intensifying the groundwater monitoring network with respect to the patterns shown by Figures 5.9a and 5.10b. Additional observations should be made in the areas with risks close to 50% as well as high uncertainty. Note that in this example, there is little need to improve the predictions of water table depth, since the area of risks close to 50% and large uncertainty is very limited.

5.9 Summary and discussion

A regionalised and physically interpretable autoregressive exogenous variable model was presented to predict and to simulate time series of water table depths spatially, in an area where only limited observations of water table depth were available. The validation results indicate that the predictions are reasonably accurate (The areal mean $RMSE = 22$ cm), as compared to the accuracy of the digital elevation data ($RMSE = 22$ cm). The $RARX(\mathbf{u};1,0)$ model was applied in mapping the risk that a critical level will be exceeded at the start of the growing season, taking the uncertainty about the future weather conditions and the uncertainty arising from the $RARX(\mathbf{u};1,0)$ model into account. The mapped risks may form the input of a cost function, which should inform water managers whether a decision on water management is cost-effective or not. The closer the risks to 50%, the more difficult it will be to take a decision. The risk map only shows small areas having risks nearby 50%, which implies that in this example, the method can be helpful to water managers in taking strategic decisions. The risks were calculated under the following assumptions: i) the precipitation surplus over the last 30 years is representative for the precipitation surplus over the coming 30 years, ii) mean errors ($ME(\mathbf{u})$) and standard deviations of errors ($SDE(\mathbf{u})$) of predicted water table depths are independent of the predicted water table depths, and iii) $ME(\mathbf{u})$ is independent of $SDE(\mathbf{u})$. The accuracy of the

estimated risks depends on the quality of these three assumptions. A validation of estimated risks would be an interesting topic for further research.

In this study, the $\text{RARX}(\mathbf{u};1,0)$ model has been applied to calculate water table dynamics under the actual hydrologic conditions. Using the physical relationships given in Eqs. (5.3) to (5.5), the RARX model can be used to evaluate the effects of interventions in the hydrologic regime on the water table dynamics (chapter 4). The method can be extended to situations where strong threshold nonlinearities are present (chapter 3), using the physical basis of threshold autoregressive models (TARSO) given in chapter 4. The amount of information needed for parametrising the $\text{RARX}(\mathbf{u};1,0)$ model is small when compared to physically based mechanistic models for groundwater flow. However, soil profile information needs to be widely available. The proposed method is restricted to areas with relatively shallow water tables and precipitation surplus being the main explanatory variable of water table fluctuation. Extensions to situations with more explanatory variables, for instance, surface water levels and groundwater withdrawals, have not yet been tried.

The procedure of ‘guessing’ $\text{RARX}(\mathbf{u};1,0)$ parameters and interpolating them to unvisited locations can be improved at several points. In this study, the $\text{RARX}(\mathbf{u};1,0)$ parameters were first calculated and next interpolated to unvisited locations, i.e. CI (calculate first, interpolate later). The other way round, i.e. IC (interpolate first, calculate later) may have several advantages in this case, however. If the physical information from which the drainage resistance, the effective porosity, and hence $\tilde{a}_1(\mathbf{u})$, and $\tilde{b}_0(\mathbf{u})$ are calculated is available at the unvisited locations, then the calculated parameter values will ‘automatically’ satisfy the stationarity conditions, that is $0 < |\tilde{a}_1(\mathbf{u})| < 1$ and $\tilde{b}_0(\mathbf{u}) > 0$. Furthermore, possible relationships between $\tilde{a}_1(\mathbf{u})$, $\tilde{b}_0(\mathbf{u})$, and $\tilde{\mu}(\mathbf{u})$ play no role in IC. IC needs more calculation than CI, however. In particular, the interpolation of the soil profile descriptions and the calculation of average distances between drainage devices would be very labour intensive. Therefore, we decided to use CI in this study. We neglected possible relationships between $\tilde{a}_1(\mathbf{u})$, $\tilde{b}_0(\mathbf{u})$, and $\tilde{\mu}(\mathbf{u})$, because i) scatter plots indicated that these relationships are not strong; ii) possible interpolation errors due to neglecting possible relations between the RARX parameters will be small, because of the high density of points (1,185 in an area of 1,375 hectares) from which is interpolated to unvisited locations, and iii) the Indirect Method corrects for the systematic part of possible errors. Checks indicated that all interpolated values of $\tilde{a}_1(\mathbf{u})$ and $\tilde{b}_0(\mathbf{u})$ satisfied the stationarity conditions, which is expected since we interpolated from a dense grid. In case of convex interpolators this is also ensured.

Chapter 6

Accuracy of spatio-temporal RARX model predictions of water table depths

(This chapter is based on the paper ‘Accuracy of spatio-temporal RARX model predictions of water table depths’ by M. Knotters and M.F.P. Bierkens, accepted for publication in *Stochastic Environmental Research and Risk Assessment*.)

Abstract

Time series of water table depths (H_t) are predicted in space using a regionalised autoregressive exogenous variable (RARX) model with precipitation surplus (P_t) as input variable. Because of their physical basis, RARX model parameters can be guessed from auxiliary information such as a digital elevation model (DEM), digital topographic maps and digitally stored soil profile descriptions. Three different approaches to regionalising RARX parameters are used. In the ‘direct’ method (DM), P_t is transformed into H_t using the guessed RARX parameters. In the ‘indirect’ method (IM), the predictions from DM are corrected for observed systematic errors. In the Kalman filter approach, the parameters of regionalisation functions for the RARX model parameters are optimised conditional to observations on H_t . These regionalisation functions describe the dependence on spatial co-ordinates of the RARX parameters. External drift kriging and simple kriging with varying means are applied as regionalisation functions, using guessed RARX model parameters or DEM data as secondary variables. Predictions of H_t at given days, as well as estimations of expected water table depths are made for a study area of 1,375 hectares. The performance of the three approaches is tested by cross-validation using observed values of H_t in 27 wells which are positioned following a stratified random sampling design. IM performs better with respect to systematic errors than the alternative methods in estimating expected water table depths. The Kalman filter methods perform better than both DM and IM in predicting the temporal variation of H_t , as is indicated by lower random errors. Particularly the Kalman filter method that uses DEM data as an external drift outperforms the alternative methods with respect to the prediction

of the temporal variation of the water table depth.

6.1 Introduction

Agricultural and ecological water management in the Netherlands requires accurate information on water table fluctuation. Preferably, the uncertainty about water table fluctuation is quantified, in order to enable risk assessment. The information may concern either *prediction* of the water table depth on a given day, or *estimation* of statistics. These statistics are for instance *MHW* and *MLW* (chapter 2), duration curves (section 3.6.2, Figure 3.3), percentiles or percentages (section 5.8, Figure 5.9a), or expected water table depths (section 5.8, Figure 5.9b). Predictions of water table depths on a given day are more and more required in the current water management practice in the Netherlands. By using predictions of actual water table depths, water managers can anticipate actual trends. For instance, in the wet spring of 1998 up-to-date predictions of water table depths, for instance at March 13, might have been helpful to water managers in aiming for optimal conditions in the start of the growing season. In contrast, statistics of the water table fluctuation are needed to assess options for long term water policy. For instance, to decide on investments to lower the water table depth in the start of the growing season, an estimate of the expected water table depth at March 13 in any future year, given the actual hydrological regime, may be relevant information. Agricultural or ecological water management in the Netherlands is generally restricted to areas of limited size, in which mostly often only a few time series of water table depth are available. On average, one suitable observation well is present per 750 to 1,250 ha (Finke, 2000). Therefore, auxiliary information may be needed to predict the water table depth accurately in these areas. In chapter 5 a regionalised autoregressive exogenous variable model (RARX) for the relationship between precipitation surplus and water table depth was introduced. In essence, the RARX model is a linear time series model which parameters are made space dependent or ‘regionalised’. In other words, the value of a RARX parameter depends on its location. For locations at which time series of water table depths were observed, the RARX model parameters can be calibrated. At other locations the RARX model parameters can be guessed from auxiliary physical information which is generally widely available, as explained in chapter 5. This auxiliary information includes a Digital Elevation Map (DEM), digitally stored soil profile descriptions, soil physical standard curves and digital information on the positions and sizes of ditches.

Using the guessed RARX parameter fields it is possible to predict time series of water table depth spatially. This can be done straightforwardly, without use of observed water table depths. However, applying this ‘direct’ method can result in large systematic prediction errors (subsection 5.6.2). The direct method can be improved by interpolating the observed systematic prediction errors to unvisited locations in order to correct the preliminary results. Furthermore, if repeated observations in time are available, the noise term, containing the part of water table fluctuation that cannot be explained from the precipitation surplus, can be quantified. In section 5.8 it was shown that this ‘indirect’ method can be applied in estimating the risk of shallow water tables at the start of the growing season.

Bierkens *et al.* (2001) incorporated the RARX model into a space-time Kalman

filter algorithm. This application of the space-time Kalman filter is a mathematical framework which enables the prediction of water table depths in space and time conditionally to observed water table depths, optionally using auxiliary information. For every time step at which the water table depth was observed an update of the predictions is made. This makes the Kalman filter attractive in practising daily water management. As the ‘indirect’ method, the Kalman filter approach can also be used for estimation of statistics of the water table fluctuation.

The aim of this study is to evaluate the accuracy of prediction methods in either predicting water table depths, or estimating water table depths which are expected in the prevailing hydrological regime. To this end a cross-validation experiment is carried out. The accuracy of estimates of expected water table depths ‘in any future year’ is evaluated as follows: expected water table depths are estimated for days in the monitoring period at which observations were taken, given the meteorological conditions during the monitoring period. Next, the estimates are compared with the observations.

The set up of this chapter is as follows. In section 6.2 the study area and the data set are introduced. In section 6.3 the ‘direct’ and ‘indirect’ spatio-temporal prediction method, and three applications of the space-time Kalman filter algorithm are explained. An outline of the validation procedure is given in section 6.4. The results are presented and discussed in section 6.5. Conclusions follow in section 6.6. In appendix A6.1 the accuracy of quantified uncertainty is analysed.

6.2 Study area; data set

The study area of 1,375 hectares is situated in the north-eastern part of the Netherlands. The set of observed water table depths H_t (Figure 5.1) can be divided into two parts:

1. a preferential sample of 13 time series observed in permanently installed observation wells (measurement locations);
2. a stratified random sample (approximately proportionally allocated to the stratum areas) of 27 short time series (of two years length) observed in temporarily installed observation wells (validation locations), representing a part of the study area of 976 hectares. The four strata are catchments with controlled surface water levels, see Figure 5.1.

The northern ‘appendix’ of the study area (Figure 5.1) is not a part of the area of 976 hectares from which the stratified random sample has been taken. In the northern part boulder clay is present at shallow depth. Because of the temporary occurrence of perched water tables at impermeable layers, the water table depth can not be measured accurately in observation wells placed in the boulder clay. Therefore, the northern part of the study area was excluded from the validation. Predicted water table depths are mapped for the northern part, however, in order to illustrate the differences in the predictions obtained by the evaluated methods.

The ground surface elevation was determined at all 40 groundwater observation points. Figure 6.1 shows the sampling scheme in time. Note that the interval lengths are not constant.

Accuracy of spatio-temporal RARX model predictions

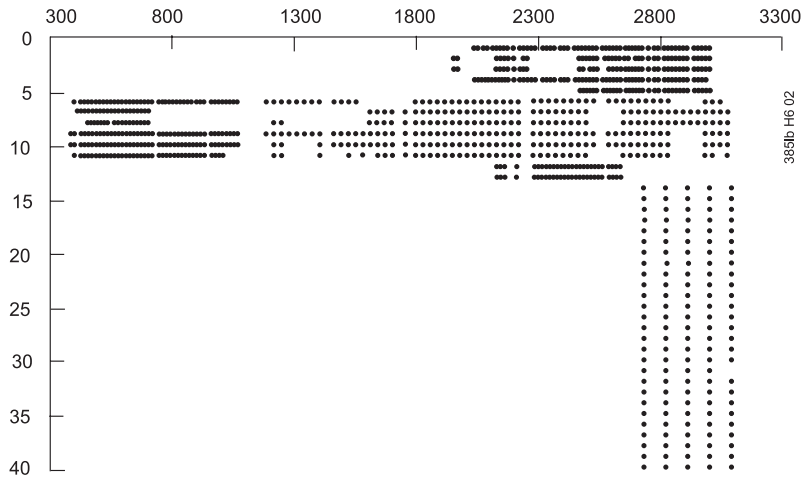


Figure 6.1: Sampling scheme. Vertical axis: well number. Horizontal axis: day number (day 1 = January 1st, 1990)

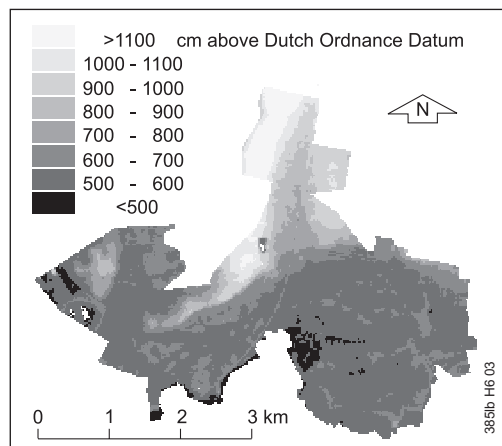


Figure 6.2: Digital elevation model of the study area. Values in cm above Dutch Ordnance Datum.

The soil of the study area is built up of relatively wet and peaty sediments in a brook valley in the southern part, and relatively dry cover sands, locally covering boulder clays in the northern part. A digital elevation model (DEM, Figure 6.2) reflects the geomorphologic structure of the landscape: a relatively flat and low brook valley in the south and relatively high, slightly undulating cover sand ridges in the north.

Daily data on precipitation are available from the station Dedemsvaart nearby the study area. Daily data on the potential Makkink crop reference evapotranspiration (Winter *et al.*, 1995) are available from the station Eelde at approx. 55 km distance

from the study area. The precipitation surplus is calculated as follows:

$$p_t = p'_t - e_{p,t},$$

where p'_t is the daily precipitation [Lt^{-1}] and $e_{p,t}$ is the daily potential Makkink crop reference evapotranspiration [Lt^{-1}].

At 1,185 locations in the study area soil profile descriptions were made by experienced soil surveyors, as a part of a regional survey for land development. A DEM, made by laser scanning, is available with a resolution of 25×25 m, see Figure 6.2. On a point scale, the DEM deviates on average 22 cm from the elevation determined by surveying. Information on ditches, such as average distances and perimeters, are derived from digital topographic maps. Information on surface water levels is provided by the local water authorities.

6.3 Spatio-temporal prediction methods

6.3.1 The RARX model

The RARX model is explained in detail in chapter 5. Here the model will be recapitulated. The RARX($\mathbf{u};1,0$) model is given by

$$H_t(\mathbf{u}) - \mu(\mathbf{u}) = a_1(\mathbf{u})\{H_{t-\Delta t}(\mathbf{u}) - \mu(\mathbf{u})\} + b_0(\mathbf{u})P_t + \epsilon_t(\mathbf{u}), \quad (6.1)$$

with $\mathbf{u}^T = (x, y)$ indicating the spatial co-ordinates x and y . Δt is the time step. Note that the precipitation surplus P_t is assumed to be global, that is, space invariant. This assumption is reasonable for relatively small areas. The error process, $\epsilon_t(\mathbf{u})$, is assumed to form a discrete sequence of mutually independent disturbances with zero expectation and finite and constant variance $\sigma_\epsilon^2(\mathbf{u})$. In space, the error process is assumed to be continuous and structured.

According to chapters 4 and 5, the physical basis of the RARX($\mathbf{u};1,0$) model parameters is derived from the water balance of a soil column. The RARX($\mathbf{u};1,0$) model parameters can be expressed in physical quantities as follows:

$$a_1(\mathbf{u}) = e^{-\Delta t / (\varphi(\mathbf{u})\gamma(\mathbf{u}))}, \quad (6.2)$$

where $\varphi(\mathbf{u})$ is the effective porosity [-], and $\gamma(\mathbf{u})$ is the drainage resistance [t],

$$b_0(\mathbf{u}) = \gamma(\mathbf{u})\{1 - a_1(\mathbf{u})\}, \quad (6.3)$$

and

$$\mu(\mathbf{u}) = \gamma(\mathbf{u})q_b(\mathbf{u}) + H_s(\mathbf{u}), \quad (6.4)$$

where $q_b(\mathbf{u})$ is the flux to the shallow groundwater from the deeper groundwater systems [Lt^{-1}], and $H_s(\mathbf{u})$ is the drainage level (e.g. surface water level) [L]. $\varphi(\mathbf{u})$, $\gamma(\mathbf{u})$, $q_b(\mathbf{u})$, and $H_s(\mathbf{u})$ are assumed to be time invariant. Furthermore, it is assumed that the water table fluctuation depends on precipitation surplus only. Other influences such as groundwater discharge are assumed to be absent or forming part of the time invariant flux $q_b(\mathbf{u})$.

The physical relationship for $\sigma_\epsilon^2(\mathbf{u})$ is:

$$\sigma_\epsilon^2(\mathbf{u}) = \text{var} \left\{ b_0(\mathbf{u}) \left[(E_{p,t} - E_{a,t}(\mathbf{u})) - \frac{\Delta V(\mathbf{u})}{\Delta t} \right] + \epsilon_t^\bullet(\mathbf{u}) \right\}, \quad (6.5)$$

where $E_{a,t}(\mathbf{u})$ is the actual evapotranspiration [LT^{-1}] and $V(\mathbf{u})$ is the moisture volume in the unsaturated zone. Both quantities can not be observed against reasonable costs and therefore generally are unknown. The term $\epsilon_t^\bullet(\mathbf{u})$ contains the remaining unknown influences. Thus, to estimate $\sigma_\epsilon^2(\mathbf{u})$ measured water table depths are needed.

The RARX($\mathbf{u};1,0$) model in Eq. (6.1) is the basis of the three prediction methods that are evaluated in this study: the ‘direct’ method, the ‘indirect’ method and the space-time Kalman filter.

6.3.2 The ‘direct’ method (DM) and the ‘indirect’ method (IM)

In DM, widely available auxiliary information is transformed directly into RARX parameters. In section 5.4 it was described how $\gamma(\mathbf{u})$ and $\varphi(\mathbf{u})$ are guessed from auxiliary information on ground surface elevation, drainage devices and soil profile descriptions. Using the relationships in Eqs. (6.2) and (6.3), the guessed physical parameters $\tilde{\gamma}(\mathbf{u})$ and $\tilde{\varphi}(\mathbf{u})$ are transformed into guessed RARX parameters $\tilde{a}_1(\mathbf{u})$ and $\tilde{b}_0(\mathbf{u})$. The parameter $\mu(\mathbf{u})$ is guessed directly from hydromorphic soil characteristics which were determined at the 1,185 augering locations. Eq. (6.4) was not used in guessing $\mu(\mathbf{u})$, because it was not possible to guess the regional groundwater flux $q_b(\mathbf{u})$ accurately from auxiliary information.

The RARX parameters can be guessed for locations where soil profile descriptions were made. Next, the guessed parameter values are interpolated to unvisited locations (e.g. a grid). Using these interpolated, physically based RARX parameters $\tilde{a}_1(\mathbf{u})$, $\tilde{b}_0(\mathbf{u})$, $\tilde{\mu}(\mathbf{u})$ and the model given in Eq. (6.1), time series on precipitation surplus are transformed into time series on water table depths. Note that DM is a purely deterministic method; observed water table depths are not used. As shown in subsection 5.6.2 and Figure 5.5, DM may result in large systematic prediction errors. Furthermore, a prediction of the standard deviation of the error process, $\sigma_\epsilon(\mathbf{u})$, is not provided, which makes the direct method unsuitable for application in risk assessment.

As in DM, IM starts with transformation of physical information into RARX parameters and subsequently, after spatial interpolation, transformation of time series on precipitation surplus into time series on water table depths. Next, the predicted water table depths are compared with observed water table depths. The observed systematic errors are spatially interpolated to correct the preliminary, direct, predictions. In this study, the systematic errors were interpolated by external drift kriging (Deutsch and Journel, 1992; Goovaerts, 1997), using elevation data from the DEM as an external drift.

If time series of water table depths are observed, the standard deviation of the error process, $\sigma_\epsilon(\mathbf{u})$, can be estimated. The estimated $\sigma_\epsilon(\mathbf{u})$ can be interpolated spatially to unvisited locations, and thus it can be used in constructing maps reflecting the risk that a critical water table depth is exceeded (chapter 5).

6.3.3 The space-time Kalman filter (KF)

In DM, observed water table depths are not used, whereas in IM observed water table depths are only used to correct afterwards for systematic errors, and to predict the variance of the error term, $\sigma_\epsilon(\mathbf{u})$. However, spatio-temporal predictions of water table depths could be improved if they are optimised with respect to observed water table depths. The space-time Kalman filter algorithm enables one to predict water table depths in space and time, conditional to the observed water table depths. Furthermore, the space-time Kalman filter can be used in conditional simulation and in network optimisation. Bierkens *et al.* (2001) described in detail a space-time Kalman filter algorithm for water table depths. Here, this algorithm will briefly be recapitulated.

The basis of the algorithm is the RARX($\mathbf{u};1,0$) model given in Eq. (6.1). In the space-time Kalman Filter applied in this study, the RARX($\mathbf{u};1,0$) model is rewritten in a state-space formulation:

$$\begin{aligned} \begin{pmatrix} H_t(\mathbf{u}_1) \\ H_t(\mathbf{u}_2) \\ \vdots \\ H_t(\mathbf{u}_N) \end{pmatrix} &= \begin{pmatrix} \hat{a}(\mathbf{u}_1) & 0 & \cdots & 0 \\ 0 & \hat{a}(\mathbf{u}_2) & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \hat{a}(\mathbf{u}_N) \end{pmatrix} \begin{pmatrix} H_{t-1}(\mathbf{u}_1) \\ H_{t-1}(\mathbf{u}_2) \\ \vdots \\ H_{t-1}(\mathbf{u}_N) \end{pmatrix} \\ &+ \begin{pmatrix} 1 - \hat{a}_1(\mathbf{u}_1) & 0 & \cdots & 0 & \hat{b}_0(\mathbf{u}_1) \\ 0 & 1 - \hat{a}_1(\mathbf{u}_2) & \ddots & \vdots & \hat{b}_0(\mathbf{u}_2) \\ \vdots & \ddots & \ddots & 0 & \vdots \\ 0 & \cdots & 0 & 1 - \hat{a}_1(\mathbf{u}_N) & \hat{b}_0(\mathbf{u}_N) \end{pmatrix} \begin{pmatrix} \hat{\mu}(\mathbf{u}_1) \\ \hat{\mu}(\mathbf{u}_2) \\ \vdots \\ \hat{\mu}(\mathbf{u}_N) \\ p_t \end{pmatrix} \\ &+ \begin{pmatrix} \epsilon_t(\mathbf{u}_1) \\ \epsilon_t(\mathbf{u}_2) \\ \vdots \\ \epsilon_t(\mathbf{u}_N) \end{pmatrix}, \end{aligned} \quad (6.6)$$

with error standard deviations $(\hat{\sigma}_\epsilon(\mathbf{u}_1), \hat{\sigma}_\epsilon(\mathbf{u}_2), \dots, \hat{\sigma}_\epsilon(\mathbf{u}_N))$. Here, N indicates the number of locations for which the water table depth is to be predicted for discrete time steps, say $t = 0, 1, \dots, K$. The hats in Eq. (6.6) indicate that the RARX parameters $a_1(\mathbf{u}_i)$, $b_0(\mathbf{u}_i)$, $\mu(\mathbf{u}_i)$, and $\sigma_\epsilon(\mathbf{u}_i)$, $i = 1, \dots, N$ are estimated. The RARX parameters can be calibrated for the 13 locations where time series of water table depth are available which are sufficiently long to calibrate an ARX(1,0) model (see chapter 2, appendix A2.1). For all other locations, the parameters have to be estimated by so-called regionalisation functions. These functions describe the dependency of the RARX parameters on the spatial co-ordinates \mathbf{u} . The RARX parameters could only be calibrated for 13 locations in a study area of 1,375 hectares. Therefore, it should be possible to include auxiliary variables into the regionalisation functions. In this study, the guessed RARX parameter values, which are derived from physical auxiliary information by using Eqs. (6.2) to (6.4) and the procedures described in section 5.4, are used as auxiliary variables. Alternatively, ground surface elevation data from the DEM (Figure 6.2) are used as auxiliary variable. An estimator of a RARX($\mathbf{u};1,0$)

parameter, for instance $\hat{a}_1(\mathbf{u})$, is given by

$$\hat{a}_1(\mathbf{u}) = r_a[\mathbf{u}; a_1(\mathbf{u}_i), i = 1, \dots, n; s_j(\mathbf{u}), j = 1, \dots, m; \boldsymbol{\theta}_a], \quad (6.7)$$

(Bierkens *et al.*, 2001), where $a_1(\mathbf{u}_i)$ are parameter values calibrated on time series observed at n locations $\mathbf{u}_i, i = 1, \dots, n$, and $s_j(\mathbf{u}), j = 1, \dots, m$ are the values of m auxiliary variables at location \mathbf{u} . In this study, $m = 1$, i.e., only one auxiliary variable at a time is considered. $\boldsymbol{\theta}_a$ are the parameters of the regionalisation function. For instance, if some type of kriging is used as a regionalisation function, $\boldsymbol{\theta}_a$ are the variogram parameters.

The dependence of the RARX parameter values on the regionalisation functions is explicitly denoted in the following matrix-vector notation of Eq. (6.6):

$$\mathbf{h}_t = \mathbf{A}(\boldsymbol{\theta}_a)\mathbf{h}_{t-1} + \mathbf{B}(\boldsymbol{\theta}_a, \boldsymbol{\theta}_b)\mathbf{x}_t(\boldsymbol{\theta}_\mu) + \mathbf{e}_t(\boldsymbol{\theta}_\sigma, I), \quad (6.8)$$

where $\boldsymbol{\theta}_a, \boldsymbol{\theta}_b, \boldsymbol{\theta}_\mu$ and $\boldsymbol{\theta}_\sigma$ are parameter vectors of the regionalisation functions, see Eq. (6.7), and I is the integral scale. Eq. (6.8) is referred to as the state equation, representing the ‘state’ of the water table depth at time t . \mathbf{h}_t is the $N \times 1$ ‘state’ vector, \mathbf{A} is the $N \times N$ system matrix, \mathbf{B} is the $N \times (N+1)$ input matrix, \mathbf{x}_t is the $(N+1) \times 1$ input vector, and \mathbf{e}_t is the $N \times 1$ vector with additive noise or ‘system noise’.

In section 6.3.1 it was mentioned that the error term $\epsilon_t(\mathbf{u})$ is assumed to form a white noise sequence in time, but assumed to be continuous and structured in space. It is further assumed that the spatial structure of $\epsilon_t(\mathbf{u})$ can be described with an isotropic exponential spatial covariance. The properties of the system noise vector \mathbf{e}_t can now be summarised as:

$$\begin{aligned} E[\mathbf{e}_t] &= \mathbf{0} \\ E[\mathbf{e}_t \mathbf{e}_s^T] &= \begin{cases} \mathbf{Q}(\boldsymbol{\theta}_\sigma, I) & \text{if } t = s \\ \mathbf{0} & \text{if } t \neq s, \end{cases} \end{aligned} \quad (6.9)$$

where \mathbf{Q} is a $N \times N$ covariance matrix, which elements $[Q_{ij}]$ are given by

$$[Q_{ij}] = \hat{\sigma}(\mathbf{u}_i; \boldsymbol{\theta}_\sigma) \hat{\sigma}(\mathbf{u}_j; \boldsymbol{\theta}_\sigma) \exp\left(\frac{-|\mathbf{u}_j - \mathbf{u}_i|}{I}\right). \quad (6.10)$$

Here, I denotes the integral scale.

The ‘state’ of the water table depth (Eq. (6.8)) is linked to observed water table depths by the measurement equation

$$\mathbf{y}_t = \mathbf{C}_t \mathbf{h}_t + \mathbf{v}_t. \quad (6.11)$$

The measurement vector \mathbf{y}_t contains the water table depths that are observed at time t . Because the number of observations may vary in time, \mathbf{y}_t is a $M_t \times 1$ vector with M_t being the number of observations taken at time t . The observations are linked to the state by the $M_t \times N$ measurement matrix \mathbf{C} as follows: if the i th element of \mathbf{y}_t is an observation of the j th state variable in \mathbf{h}_t , then the element (i, j) in \mathbf{C} is set to one, and the remaining elements of the i th row in \mathbf{C} are set to zero. The noise vector \mathbf{v}_t represents the measurement errors. These are assumed to have zero expectation.

6.3 Spatio-temporal prediction methods

Furthermore, the measurement errors are assumed to be independent in time, and to be independent of the system noise. In summary, the properties of \mathbf{v}_t are:

$$\begin{aligned} E[\mathbf{v}_t] &= \mathbf{0} \\ E[\mathbf{v}_t \mathbf{v}_s^T] &= \begin{cases} \mathbf{R}_t & \text{if } t = s \\ \mathbf{0} & \text{if } t \neq s \end{cases} \\ E[\mathbf{v}_t \mathbf{e}_s^T] &= \mathbf{0} \quad \forall t, s \end{aligned} \quad (6.12)$$

Here, $\mathbf{0}$ is a vector or a matrix, filled with zeros. Usually it is assumed that the measurement errors are spatially independent, so that \mathbf{R}_t is a diagonal matrix with variances of the observation errors on the diagonal. Given the state equation, Eq. (6.8), the measurement equation, Eq. (6.11), and the properties of the error processes, Eqs. (6.9), (6.10) and (6.12), the Kalman filter can be used to predict the water table depth optimally. A detailed description of the Kalman filter can be found, for instance, in Van Geer *et al.* (1991).

Optimal predictions of water table depths can be obtained by an optimisation procedure in which model parameters are estimated, using observed water table depths. It is important to note that the RARX parameters are estimated via optimisation of the parameters of their regionalisation functions. Following Bierkens *et al.* (2001), the regionalisation parameters θ_a , θ_b , θ_μ , θ_σ , and the integral scale I are optimised. The RARX parameter values in the space-time Kalman filter algorithm depend on these regionalisation parameters: the system matrix \mathbf{A} , the input matrix \mathbf{B} and the input vector \mathbf{x} depend on θ_a , θ_b and θ_μ (Eq. (6.8)), whereas the covariance matrix \mathbf{Q} (Eq. (6.9)) depends on θ_σ and I . The parameters $\Theta^T = (\theta_a, \theta_b, \theta_\mu, \theta_\sigma, I)$ can be estimated by running the Kalman filter and minimising the prediction errors. Assuming that both the system noise \mathbf{e}_t and the observation errors \mathbf{v}_t are Gaussian distributed, Θ can be estimated by minimising a maximum likelihood criterion given by Schweppe (1973):

$$J(L; \Theta) = LM_{\mathbf{y}} \ln(2\pi) + \sum_{l=1}^L \ln(|\mathbf{Z}_l(\Theta)|) + \sum_{l=1}^L \mathbf{n}_l^T(\Theta) \mathbf{Z}_l^{-1}(\Theta) \mathbf{n}_l(\Theta). \quad (6.13)$$

Here, L is the number of times at which the water table depth was observed. Since the water table depth is not measured at every time step at all locations, $M_{\mathbf{y}}$ is the *average* number of observations per observation time. \mathbf{n}_l is a vector of differences between the observations \mathbf{y}_l and the so-called time update $\hat{\mathbf{h}}_l$ at time step l . The time update $\hat{\mathbf{h}}_l$ is the prediction using all observations up to and including the previous time step, $l-1$. Thus, \mathbf{n}_l can be considered as a vector of one-step-ahead prediction errors. \mathbf{Z}_l is the covariance matrix of \mathbf{n}_l , that is, $\mathbf{Z}_l = E[\mathbf{n}_l \mathbf{n}_l^T]$.

Two types of regionalisation functions that account for an auxiliary variable were applied: kriging with an external drift, and simple kriging with varying means (Goovaerts, 1997; Deutsch and Journel, 1992). In kriging with an external drift the regionalisation function for the RARX parameter $a_1(\mathbf{u})$, $r_a(\mathbf{u}; \dots)$ in Eq. (6.7), has the following form:

$$r_a(\mathbf{u}; \dots) = \beta[s(\mathbf{u})] + \sum_{i=1}^n \lambda_i(\theta_a) \{a_1(\mathbf{u}_i) - \beta[s(\mathbf{u}_i)]\}, \quad (6.14)$$

where $\beta[\cdot]$ is a drift function depending on the auxiliary variable $s(\mathbf{u})$, $\lambda(\boldsymbol{\theta}_a)$ are kriging weights depending on the parameter vector $\boldsymbol{\theta}_a$. Here $\boldsymbol{\theta}_a$ is the vector of parameters of the variogram of the residuals. The drift function has a linear form:

$$\beta[s(\mathbf{u})] = b'_0 + b'_1 s(\mathbf{u}), \quad (6.15)$$

where b'_0 and b'_1 are estimated implicitly when solving the kriging system (cf. Goovaerts, 1997).

In this study, the parameters $\tilde{a}_1(\mathbf{u})$, $\tilde{b}_0(\mathbf{u})$, and $\tilde{\mu}(\mathbf{u})$ which were guessed on the basis of physical information (subsection 6.3.2) are used as an external drift. As explained in subsection 6.3.1, $\sigma_\epsilon(\mathbf{u})$ cannot easily be guessed from physical auxiliary information against reasonable costs. However, it can be expected that a relationship exists between $\sigma_\epsilon(\mathbf{u})$ and the average water table depth. This can be pointed out as follows. The effective porosity will increase with increasing water table depth, because the fraction of pores filled with air will increase. Consequently, the value of the term $b_0(\mathbf{u})$ in Eq. (6.5) will decrease, as follows from Eqs. (6.2) and (6.3). Thus, it can be expected that $\sigma_\epsilon(\mathbf{u})$ is related to the average water table depth. Therefore, the values of $\tilde{\mu}(\mathbf{u})$ which were guessed from hydromorphic soil characteristics which roughly reflect the average water table depth, can be useful as auxiliary information in regionalising $\sigma_\epsilon(\mathbf{u})$. The parameter $\sigma_\epsilon(\mathbf{u})$ was regionalised using external drift kriging with $\tilde{\mu}(\mathbf{u})$ as an external drift.

In simple kriging with varying means the regionalisation function is in its structure equal to the function in Eq. (6.14). However, the methods have different definitions of the trend component $\beta[\cdot]$. In external drift kriging, the trend component is a simple linear regression function, see Eq. (6.15). The parameters b'_0 and b'_1 in Eq. (6.15) are implicitly estimated through the kriging system within each neighbourhood. In contrast, in simple kriging with varying means the trend can have any form. This trend, or ‘varying mean’, can for instance be the result of arbitrary *ad-hoc* procedures, whereas in external drift kriging inference relies on the generalised least squares method. In this study, the ‘varying mean’ is formed by guessed RARX parameter values, which can be considered as the outcomes of *ad-hoc* procedures. Thus, for the RARX parameter $a_1(\mathbf{u})$, the trend component $\beta[\tilde{a}_1(\mathbf{u})]$ is defined as

$$\beta[\tilde{a}_1(\mathbf{u})] = \tilde{a}_1(\mathbf{u}),$$

whereas in external drift kriging the trend component is defined as

$$\beta[\tilde{a}_1(\mathbf{u})] = b'_0 + b'_1 \tilde{a}_1(\mathbf{u}),$$

where b'_0 and b'_1 are estimated through the kriging system.

6.3.4 Summary of prediction methods

The methods described in subsections 6.3.2 and 6.3.3 require different minimum sizes of groundwater data sets. DM does not need observed water table depths at all. IM needs observed water table depths at a sufficiently large number of points in space, in order to accurately estimate the model of spatial structure of the systematic prediction error. Preferably, time series on water table depths are observed at these points so

6.3 Spatio-temporal prediction methods

that the standard deviation of the error process $\sigma_\epsilon(\mathbf{u})$ can be estimated and next spatially predicted at unvisited locations. If in KF a kriging interpolation method is used as a regionalisation function, calibrated ARX(1,0) parameters are needed at the well locations. This implies that sufficiently long time series on water table depth must be available. Additional incidental observations are needed to calibrate the regionalisation parameters Θ (Bierkens *et al.*, 2001).

Both in DM and IM auxiliary information on drainage resistance, effective porosity and hydromorphic soil characteristics is needed, as is explained in subsection 6.3.2. In KF the RARX model parameters that are guessed following the same procedure as in DM can be used as auxiliary variables. Alternatively, in situations where the physical information needed for guessing RARX model parameters is absent, other auxiliary information such as a DEM can be used in KF. The prediction methods evaluated in this study are now summarised as follows:

- DM: the direct method. Physically based, guessed RARX parameters are used in the predictions, observed water table depths are not used;
- IM: the indirect method. Physically based, guessed RARX parameters and all observed water table depths are used;
- KF1: space-time Kalman filter, with external drift kriging as a regionalisation function for $a_1(\mathbf{u})$, $b_0(\mathbf{u})$, $\mu(\mathbf{u})$. Physically based, guessed RARX parameters are used as external drifts. $\sigma_\epsilon(\mathbf{u})$ is regionalised by external drift kriging with $\tilde{\mu}(\mathbf{u})$ as an external drift. All observed water table depths are used;
- KF2: space-time Kalman filter, with simple kriging with varying means as a regionalisation function for $a_1(\mathbf{u})$, $b_0(\mathbf{u})$, and $\mu(\mathbf{u})$. Physically based, guessed RARX parameters are used as varying means. $\sigma_\epsilon(\mathbf{u})$ is regionalised by external drift kriging with $\tilde{\mu}(\mathbf{u})$ as an external drift. All observed water table depths are used;
- KF3: space-time Kalman filter, with external drift kriging as a regionalisation function for $a_1(\mathbf{u})$, $b_0(\mathbf{u})$, $\mu(\mathbf{u})$, and $\sigma_\epsilon(\mathbf{u})$. The ground surface elevation from the DEM is used as auxiliary variable. All observed water table depths are used.

Two types of information are distinguished: predicted water table depths, and estimates of statistics for the water table fluctuation, for instance expected water table depths. In DM and IM, one and the same procedure is followed for both predicting water table depths and estimating expected water table depths: in both cases the input series is transformed into an output series, using guessed RARX parameter values. In IM, observed water table depths are only used to correct for systematic errors, both in prediction and in estimation. However, if KF methods are used for predicting water table depths, for each time step at which observations on water table depths are available an update of the spatio-temporal predictions is made. Alternatively, if KF is used for estimation of expected water table depths, first the RARX(\mathbf{u} ;1,0) parameters are estimated using the available observations on water table depths and the auxiliary information used in the regionalisation function. Next, time series on precipitation surplus are transformed into time series on water table depth by using the estimated RARX(\mathbf{u} ;1,0) parameters, without updating.

6.4 Set up of the validation of spatio-temporal predictions

Time series of water table depth were predicted (or estimated) at the 27 validation locations by means of a cross-validation procedure. In IM and the KF methods all observed water table depths were used in the prediction (see Figures 5.1 and 6.1), except those observed at the validation location. This procedure was repeated 27 times, each time leaving one validation location out for which prediction errors were calculated. The interval length in Eq. (6.1) is one day ($\Delta t = 1$ day) in the prediction procedures, since daily precipitation surplus is the input of the RARX model.

For each test location the following validation measures were calculated:

- the mean error,

$$ME_{hi} = \frac{1}{n_{hi}} \sum_{j=1}^{n_{hi}} e_{j,hi}, \quad (6.16)$$

where hi indicates the i th test location in the h th stratum, e is the difference between observed and predicted water table depth resulting from the cross-validation, n_{hi} is the length of the observed time series at location hi . The absolute value, $|ME|_{hi}$, is a measure for the closeness of the predicted to the observed mean water table depth;

- the standard deviation of error,

$$SDE_{hi} = \sqrt{\frac{1}{n_{hi} - 1} \sum_{j=1}^{n_{hi}} [e_{j,hi} - ME_{hi}]^2}, \quad (6.17)$$

which is a measure for the closeness of predicted to observed temporal fluctuation;

- and the root mean squared error,

$$RMSE_{hi} = \sqrt{\frac{1}{n_{hi}} \sum_{j=1}^{n_{hi}} e_{j,hi}^2}, \quad (6.18)$$

which is a measure for the overall closeness of predicted to observed water table depths.

The three measures in Eqs. (6.16) to (6.18) describe the closeness of predicted time series of water table depth to observed time series at location \mathbf{u}_{hi} . Areal means of these measures can be estimated by

$$m_A(y) = \sum_{h=1}^L W_h \left[\frac{\sum_{i=1}^{n_h} y_{hi}}{n_h} \right], \quad (6.19)$$

(Cochran, 1977), where L denotes the number of strata, and W_h denotes the stratum weight. The stratum weight is proportion of the stratum area to the total area. y can be replaced by ME , $|ME|$, SDE or $RMSE$.

Table 6.1: Validation results for the prediction of water table depths.

	DM	IM	KF1	KF2	KF3
$m_A(ME)$	-23.6	4.5	8.3	-10.0	14.4
$m_A(ME)$	30.7	17.5	24.1	28.1	24.2
$m_A(SDE)$	9.5	9.5	6.9	7.0	5.7
$m_A(RMSE)$	33.5	21.5	26.1	29.6	25.9

Values are in cm.

6.5 Results and Discussion

6.5.1 Predicting water table depths

The areal means of ME , $|ME|$, SDE , and $RMSE$ in predicting water table depths are listed in Table 6.1. The SDE can be interpreted as the closeness of the predicted to the observed temporal fluctuation. It is indicated that KF methods perform well with respect to SDE . In particular KF3 performs well with respect to SDE , despite the use of only DEM data as auxiliary variable, whereas in the alternative methods all available physical information was used.

The systematic errors (ME) are relatively large for all methods except IM. As explained in subsection 6.3.2, correction for systematic errors is part of the prediction procedure in IM. The negative ME values found for DM and KF2 can be explained to some extent from possible bias in the guessed RARX parameter $\tilde{\mu}(\mathbf{u})$. The parameter $\tilde{\mu}(\mathbf{u})$ has been guessed from hydromorphic soil characteristics observed in the field by augering at 1,185 locations as part of a soil survey. $\tilde{\mu}(\mathbf{u})$ is calculated as the average of the depth to the top of the permanently reduced zone and the depth to the top of the hydromorphic characteristics, such as rust mottles. This average value should approximate the mean water table depth. However, hydromorphic characteristics are often ‘fossil’, that is, the hydromorphic characteristics do not represent the actual water table, but reflect the water table in the past which generally fluctuated at shallower depths. Besides, the mean water table depth differs from the RARX parameter $\mu(\mathbf{u})$ by a factor $b_0(\mathbf{u})\bar{p}/(1 - a_1(\mathbf{u}))$, where \bar{p} is the mean precipitation surplus. The contribution of this factor to the negative systematic errors is generally limited, however. Bias in $\tilde{\mu}(\mathbf{u})$ will be reflected in the predictions of KF2, because here $\tilde{\mu}(\mathbf{u})$ is used as an auxiliary variable in simple kriging with a varying mean (see subsection 6.3.3). In KF1, however, $\tilde{\mu}(\mathbf{u})$ is used as an auxiliary variable in kriging with an external drift. In this method the values of the auxiliary variable are linearly transformed by a regression model (Eq. (6.15)), which possibly eliminates bias in $\tilde{\mu}(\mathbf{u})$.

Figure 6.3 shows maps of the predicted actual water table depth at March 13, 1998, as an example of predicting actual water table depths. The relatively detailed patterns resulting from IM and KF3 can be explained by the DEM data from a 25×25 m grid which were used as an external drift variable in both methods. It is interesting to compare the patterns obtained by DM with the patterns obtained by IM, KF1, and KF2, because differences are the effect of using observed water table depths in the predictions. Most resemblance is found for DM and KF2, particularly in the

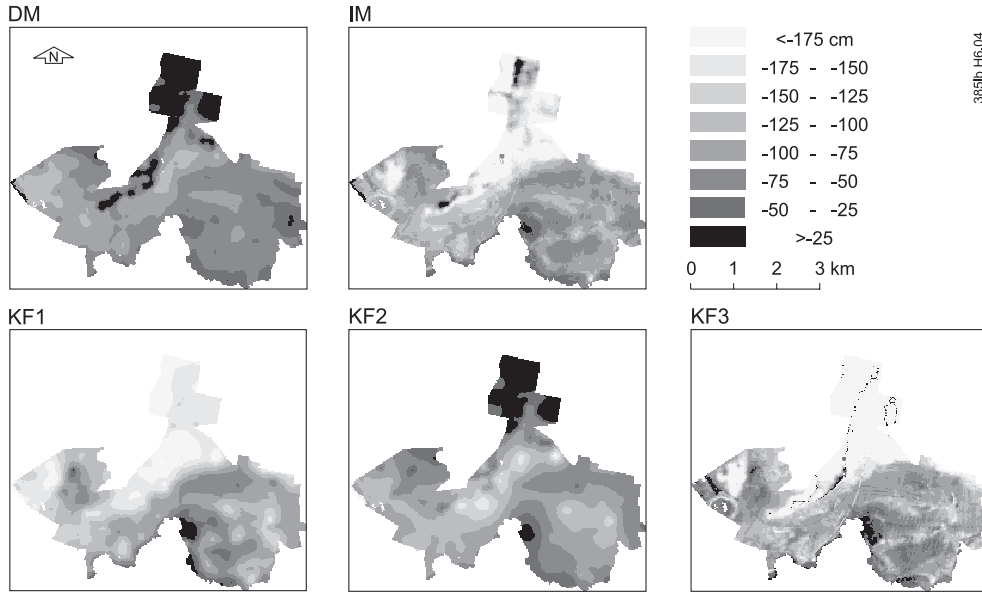


Figure 6.3: Predicted water table depth at March 13, 1998. Values in cm below ground surface. DM: direct method, IM: indirect method, KF1: space-time Kalman Filter with guessed RARX parameters as an external drift, KF2: space-time Kalman Filter with guessed RARX parameters as varying means, KF3: space-time Kalman Filter with DEM data as an external drift.

northern part of the study area. In section 6.2 it is explained that in this northern ‘appendix’ the water table depth was not observed. The relatively small differences between DM and KF2 indicate that the observed water table depths contribute less to the predictions in KF2 than in IM and KF1. Clearly, the stochastic components of the regionalised RARX parameters, for example $\{a_1(\mathbf{u}_i) - \beta[s(\mathbf{u}_i)]\}$ in Eq. (6.14), are small in KF2. In KF1 kriging with an external drift is used as a regionalisation function. Here, the drift $\beta[s(\mathbf{u}_i)]$ is a linear transformation of the guessed RARX parameter values, see Eqs. (6.14) and (6.15). The patterns in Figure 6.3 indicate that in KF1, with kriging with an external drift as a regionalisation function, the observed water table depths contribute more to the predictions than if kriging with varying means is used as a regionalisation function in KF2.

6.5.2 Estimating expected water table depths

Figure 6.4 shows maps of the expected water table depth at March 13 in any future year, given the prevailing hydrological regime. The future meteorological conditions were approximated by the daily precipitation surplus series from 1969 to 1998. As compared to the actual water table depths at March 13, 1998 (Figure 6.3), the water table at March 13 in any future year (Figure 6.4) is predicted at larger depths. Indeed, this result reflects the relatively wet meteorological conditions in the Netherlands in

6.5 Results and Discussion

Table 6.2: Validation results for the estimation of expected water table depths.

	DM	IM	KF1	KF2	KF3
$m_A(ME)$	-23.6	4.5	16.3	-8.8	25.0
$m_A(ME)$	30.7	17.5	30.2	29.1	32.3
$m_A(SDE)$	9.5	9.5	9.5	9.3	9.9
$m_A(RMSE)$	33.5	21.5	32.6	31.8	35.0

Values are in cm.

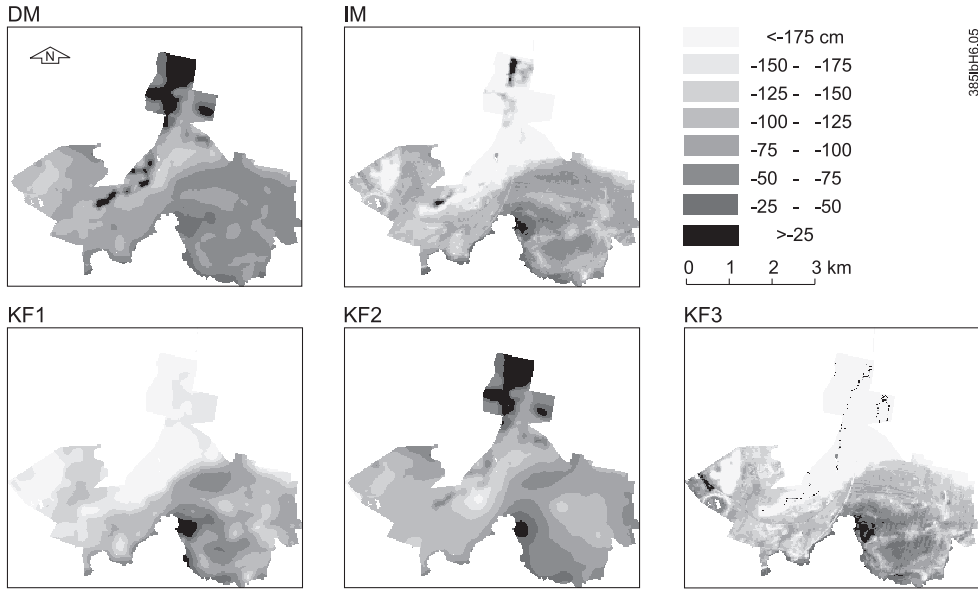


Figure 6.4: Estimates of the water table depth, expected for March 13 in any future year, given the prevailing hydrological regime. DM: direct method, IM: indirect method, KF1: space-time Kalman Filter with guessed RARX parameters as an external drift, KF2: space-time Kalman Filter with guessed RARX parameters as varying means, KF3: space-time Kalman Filter with DEM data as an external drift.

the spring of 1998. Again, resemblance is found for the maps obtained by DM and KF2.

In Table 6.2 the areal means of ME , $|ME|$, SDE , and $RMSE$ for estimates of expected water table depths are given. Again, negative ME values are found for DM and KF2; the possible causes are discussed in subsection 6.5.1. As compared to prediction of actual water table depths (Table 6.1), the performance of KF methods decreases in estimating expected water table depths. This can be explained from the fact that in this case no updates are made for each time step at which observations become available. Instead, the time series on precipitation surplus is transformed straightforwardly into a time series of estimates of expected water table depths, by using the preliminary calibrated $RARX(u;1,0)$ parameters.

The results in Table 6.2 indicate that the methods perform equally well with respect to SDE . The results on $|ME|$ in Table 6.2 indicate that IM approximates the mean water table depth better than KF1 and KF3. In particular, KF3 results in large systematic errors. Note that in KF3 only DEM data were used as auxiliary variable, whereas in the alternative methods all available physical information was incorporated.

6.6 Conclusions

The performance of five spatio-temporal prediction methods, based on the $RARX(u;1,0)$ model for water table depths, was tested in a case study. The performance was evaluated by means of a cross-validation procedure, since the number of observed time series of water table depths was too small to separate an independent validation set. A disadvantage of cross-validation is that the distances from the observation points to the prediction points are overrated, because an incomplete set of observation points is used in the interpolation. Thus, the prediction errors may be overestimated. For this reason the cross-validation results are mainly used to evaluate the relative performance of the five methods.

Distinction was made between prediction of water table depths and estimation of expected water table depths. Predictions of water table depths are needed to support short term decisions in daily water management. It is concluded that the temporal *fluctuation* of the water table is predicted more precisely by methods which include the space-time Kalman filter algorithm than by the ‘direct’ and ‘indirect’ method which are based on guessed RARX parameter fields and geostatistically interpolated corrections for systematic error. This may be expected, because in the space-time Kalman filter methods an update of the predictions is made for every time step at which the water table depth has been observed. However, the Kalman filter methods fail in predicting the mean water table depth, whereas the indirect method predicts the mean level accurately.

Estimates of expected water table depths are needed to support decision making in long term water policy. The case study indicates that the indirect method (IM) is an accurate alternative for the Kalman filter methods in predicting water table depths which are expected in a given hydrological regime. In particular in the Kalman filter method that makes use of DEM data as an auxiliary variable in external drift kriging (KF3) relatively large systematic errors can occur.

Maps of predicted water table depths and estimates of expected water table depths indicate that observations contribute more to the predictions if kriging with an external drift is used as a regionalisation function for the RARX parameters in the space-time Kalman filter algorithm (KF1) than if simple kriging with varying means (KF2) is used.

Appendix A6.1 Accuracy of quantified uncertainty about water table depths

In the previous sections we evaluated the accuracy of predicted water table depths. Apart from predictions, stochastic methods also provide information on the uncertainty about these predictions. Quantified uncertainty can be helpful in risk assessment. It is therefore interesting to evaluate the accuracy of the information on uncertainty. This will be done in this section by means of cross-validation.

First, let us reconsider the RARX($\mathbf{u};1,0$) model, now expressed as a TFN model, analogous to Eqs. (4.1) and (4.2):

$$H_t(\mathbf{u}) = H_{F,t}(\mathbf{u}) + N_t(\mathbf{u}), \quad (\text{A6.1a})$$

with

$$(H_{F,t}(\mathbf{u}) - \mu(\mathbf{u})) = a_1(\mathbf{u})(H_{F,t-\Delta t}(\mathbf{u}) - \mu(\mathbf{u})) + b_0(\mathbf{u})P_t, \quad (\text{A6.1b})$$

and

$$N_t(\mathbf{u}) = a_1(\mathbf{u})N_{t-\Delta t}(\mathbf{u}) + \epsilon_t(\mathbf{u}), \quad (\text{A6.1c})$$

where

$H_t(\mathbf{u})$ is the ‘true’ water table depth at time t and location \mathbf{u} ;

$H_{F,t}(\mathbf{u})$ is the dynamic component, that is, that part of the water table depth that can be explained from the precipitation surplus P ;

N_t is the noise component, containing that part of the ‘true’ water table depth that can not be explained from the precipitation surplus P ;

$\epsilon_t(\mathbf{u})$ is the error term, which is assumed to form a white noise sequence with zero mean and finite and constant variance $\sigma_\epsilon^2(\mathbf{u})$;

$a_1(\mathbf{u})$, $b_0(\mathbf{u})$, $\mu(\mathbf{u})$, and $\sigma_\epsilon^2(\mathbf{u})$ are the RARX($\mathbf{u};1,0$) parameters;

Δt is the time interval, which equals one day in this study.

Note that the parameter $\mu(\mathbf{u})$ now forms part of the dynamic component, which is somewhat different from Eqs. (4.1) and (4.2). Expressing the RARX($\mathbf{u};1,0$) model as a TFN model in Eqs. (A6.1a) to (A6.1c) has the advantage that the noise component $N_t(\mathbf{u})$ is clearly distinguished from the error term $\epsilon_t(\mathbf{u})$. Furthermore, the noise component is clearly defined as the difference between the ‘true’ water table depth and the dynamic component:

$$N_t(\mathbf{u}) = H_t(\mathbf{u}) - H_{F,t}(\mathbf{u}). \quad (\text{A6.2})$$

In both the present and the previous chapter, methods are discussed which assign values to the RARX($\mathbf{u};1,0$) parameters either by ‘guessing’ or by calibration. If $a_1(\mathbf{u})$, $b_0(\mathbf{u})$, and $\mu(\mathbf{u})$ are assigned, a time series of water table depths at location \mathbf{u} can be predicted simply by transforming the input series P_t into a series of the dynamic component $H_{F,t}(\mathbf{u})$, by using Eq. (A6.1b). These predictions, say $\tilde{h}_{F,t}(\mathbf{u})$, are referred to as ‘deterministic’, since they result from straightforward transformation of the input series. Consider k locations where time series of water table depths were observed, having lengths n_1, \dots, n_k . In the present case study k equals 40. The differences

between the observed water table depths at independent validation locations and the deterministically predicted water table depths result from our uncertainty about the true water table depth:

$$e_j(\mathbf{u}_k) = h_j(\mathbf{u}_k) - \tilde{h}_{F,j}(\mathbf{u}_k), \quad (\text{A6.3})$$

with k indicating the k th location where a time series of water table depths was observed and $j = 1, \dots, n_k$. The observed series $e_j(\mathbf{u}_k)$ can be considered as a realisation of the noise component $N_t(\mathbf{u}_k)$. For locations where time series of water table depths were observed, the standard deviation of noise, $\sigma_N(\mathbf{u}_k)$, can be estimated by the standard deviation of $e_j(\mathbf{u}_k)$ ($SDE(\mathbf{u}_k)$, see Eqs. (5.11) and (5.12), with i replaced by k). It is important to note that $SDE(\mathbf{u}_k)$ only reflects uncertainty about the temporal variation of the water table depth at location \mathbf{u}_k . The uncertainty about the mean water table depth at location \mathbf{u}_k is reflected by the mean error $ME(\mathbf{u}_k)$ (Eq. (5.11)). Both the standard deviation of noise, $\sigma_N(\mathbf{u})$, and the mean error, $ME(\mathbf{u})$, are predicted at locations where no observations of water table depths are available. The accuracy of quantified uncertainty about the temporal variation of water table depths (predictions of $\sigma_N(\mathbf{u})$) will be discussed in subsection A6.1.1. The accuracy of information on uncertainty about the mean water table depth (predictions of $ME(\mathbf{u})$) will be discussed in subsection A6.1.2.

A6.1.1 Accuracy of quantified uncertainty about the temporal variation

The noise component, $N_t(\mathbf{u})$ in Eq. (A6.1c), represents the uncertainty about the dynamic relationship between precipitation surplus and water table depth. The standard deviation of the noise, $\sigma_N(\mathbf{u})$, is a measure of uncertainty: a large value of $\sigma_N(\mathbf{u})$ implies high uncertainty about the dynamic relationship between precipitation surplus and water table depth.

In the indirect method (IM, see subsections 5.5.2 and 6.3.2) $\sigma_N(\mathbf{u})$ can be guessed at unvisited locations as follows:

1. Calculate $SDE(\mathbf{u}_k)$, $i = 1, \dots, 40$ for the 40 locations where time series of water table depths were observed, see Eqs. (5.11) and (5.12);
2. Model the spatial structure of $SDE(\mathbf{u})$, and interpolate $SDE(\mathbf{u}_k)$ to unvisited locations, say $\tilde{\sigma}_N(\mathbf{u})$.

An alternative procedure may be to interpolate $\tilde{\sigma}_\epsilon(\mathbf{u}_k)$ (Eq. (5.15)) and $\tilde{a}_1(\mathbf{u}_k)$ to unvisited locations, by using the variograms given in Eqs. (5.21) and (5.16), respectively. Then estimate the standard deviation of the noise process by

$$\tilde{\sigma}'_N(\mathbf{u}) = \sqrt{\frac{\tilde{\sigma}_\epsilon^2(\mathbf{u})}{1 - \tilde{a}_1^2(\mathbf{u})}}. \quad (\text{A6.4})$$

In the space-time Kalman Filter methods (KF1 to KF3, see subsections 6.3.3 and 6.3.4) $\sigma_N(\mathbf{u})$ can be calculated from the calibrated RARX($\mathbf{u};1,0$) parameters $\hat{\sigma}_\epsilon(\mathbf{u})$

Appendix A6.1 Accuracy of quantified uncertainty

Table A6.1: Summary of differences between $SDE(\mathbf{u})$ and $\sigma_N(\mathbf{u})$ predicted by IM, KF1, KF2 or KF3. Results for wells 1 to 13.

Method	mean difference (cm)	mean absolute difference (cm)
IM, $\tilde{\sigma}_N(\mathbf{u})$	3.1	3.1
IM, $\tilde{\sigma}'_N(\mathbf{u})$	3.9	3.9
KF1, $\hat{\sigma}_N(\mathbf{u})$	0.6	2.5
KF2, $\hat{\sigma}_N(\mathbf{u})$	3.1	4.3
KF3, $\hat{\sigma}_N(\mathbf{u})$	1.5	3.0

and $\hat{a}_1(\mathbf{u})$:

$$\hat{\sigma}_N(\mathbf{u}) = \sqrt{\frac{\hat{\sigma}_\epsilon(\mathbf{u})}{1 - \hat{a}_1^2(\mathbf{u})}}. \quad (\text{A6.5})$$

The $\text{RARX}(\mathbf{u};1,0)$ parameters $\sigma_\epsilon(\mathbf{u})$ and $a_1(\mathbf{u})$ are only ‘known’ at the 13 locations where time series are available which are sufficiently long to accurately calibrate an $\text{ARX}(1,0)$ model. In subsection 6.3.3 it was explained that $\hat{\sigma}_\epsilon(\mathbf{u})$ and $\hat{a}_1(\mathbf{u})$ are obtained at unvisited locations by using kriging methods as a regionalisation function.

Standard deviations of noise, either $\tilde{\sigma}_N(\mathbf{u})$ or $\tilde{\sigma}'_N(\mathbf{u})$ for IM, and $\hat{\sigma}_N(\mathbf{u})$ for KF methods, are calculated by means of cross-validation for the 13 well locations where relatively long time series of water table depths are available. The results are compared with the ‘observed’ $SDE(\mathbf{u})$ values in Table A6.1. Assuming that the values of $SDE(\mathbf{u})$ observed at the 13 well locations approximate the true value of $\sigma_N(\mathbf{u})$ accurately, a positive difference indicates an underestimation of $\sigma_N(\mathbf{u})$. The mean differences in the second column of Table A6.1 indicate that all methods underestimate $\sigma_N(\mathbf{u})$ on average. Comparing the mean differences in the second column of Table A6.1 with the mean absolute differences in the third column, it is indicated that IM underestimates $\sigma_N(\mathbf{u})$ for all 13 locations. A possible reason is that the observed errors do not represent the entire deviation of the model predictions from the true water table depth (see Figures 5.1 and 6.1). In particular the errors observed at the 27 validation locations may underestimate the temporal variation, because the observations are clustered in time; the water table depth was observed daily during five separate weeks. Although covering both a wet and a dry season in 1997 and 1998, these observations may not represent the full temporal variation. Nevertheless, the errors observed at all 40 locations were used in the spatial modelling of $\sigma_N(\mathbf{u})$. Underestimation of $\sigma_N(\mathbf{u})$ is also found for the KF methods: KF2 tends to underestimate $\sigma_N(\mathbf{u})$ more than KF1 and KF3. KF1 approximates $\sigma_N(\mathbf{u})$ more accurately than the alternative methods. The underestimation of $\sigma_N(\mathbf{u})$ in the KF methods can possibly be explained from the uncertainty of the regionalised $\text{RARX}(\mathbf{u};1,0)$ parameters that was not taken into account in $\hat{\sigma}_N(\mathbf{u})$.

A common way of expressing the uncertainty about the true value of the water table depth is calculating a prediction interval, in which the true value is expected to fall with a certain probability, often 95%. Consequently, it is expected that a 95% prediction interval covers 95% of the true values. Assuming a normal distribution of

Accuracy of spatio-temporal RARX model predictions

Table A6.2: Percentage of observed water table depths covered by 95% prediction intervals.

Well number	IM, $\tilde{\sigma}_N(\mathbf{u})$	IM, $\tilde{\sigma}'_N(\mathbf{u})$	KF1	KF2	KF3
1	86.0	87.7	94.7	98.2	87.7
2	85.4	87.8	90.2	70.7	80.5
3	90.0	90.0	95.0	97.5	95.0
4	87.9	91.4	96.6	98.3	93.1
5	93.9	84.8	100.0	75.8	97.0
6	76.1	67.0	72.7	73.9	70.5
7	91.8	73.8	100.0	44.3	100.0
8	94.8	75.9	100.0	69.0	100.0
9	83.5	80.4	80.4	79.4	89.7
10	92.2	78.9	85.6	76.7	90.0
11	92.2	85.7	90.9	94.8	93.5
12	71.4	67.9	85.7	50.0	75.0
13	89.3	89.3	89.3	100.0	92.9

the noise $N_t(\mathbf{u})$, a 95% prediction interval can be calculated by

$$[h_{F,t}(\mathbf{u}) - 1.96\sigma_N(\mathbf{u}), h_{F,t}(\mathbf{u}) + 1.96\sigma_N(\mathbf{u})]. \quad (\text{A6.6})$$

The percentages of values covered by the 95% prediction intervals were calculated by means of cross-validation for the 13 well locations where relatively long time series of water table depths are available. The results for IM, KF1, KF2, and KF3 are given in Table A6.2. The intervals resulting from IM are clearly too ‘narrow’. In IM prediction intervals are calculated by using a sample of errors $e_j(\mathbf{u}_k)$, which may not fully represent the entire deviation of the model predictions from the true values. Apart from this, the distribution of $e_j(\mathbf{u}_k)$ may deviate from the normal distribution. Furthermore, the spatial structure $\sigma_N(\mathbf{u})$ is approximated by a model. For these reasons the prediction intervals in IM may not cover the previously declared percentage of the true values. This problem may be partly tackled by simulation methods and by resampling techniques, just like the procedure which was followed in constructing a risk map (see section 5.8), because in that case no assumption about the distribution of $N_t(\mathbf{u})$ needs to be made. The underestimation would be less serious, however, if water table depths were randomly sampled in time at the 27 validation locations, perhaps with some stratification to cover both the wet and the dry seasons.

Until now we focused on the standard deviation of the noise, $\sigma_N(\mathbf{u})$, at 13 locations where extensive time series of water table depths were observed. Now the accuracy of the standard deviation of the error series ($\epsilon_t(\mathbf{u})$ in Eq. (6.1)), will be evaluated for KF methods. If KF methods are applied in predicting actual water table depths, an update of the spatio-temporal predictions is made for each time step t at which water table depths are observed, as explained in subsection 6.3.4. Starting from the update at time step t , a prediction is made for the next time step, $t + \Delta t$. This prediction is referred to as the one-step-ahead prediction. The difference between the observed water table depth at time step $t + \Delta t$ and the one-step-ahead prediction is called the

Appendix A6.1 Accuracy of quantified uncertainty

one-step-prediction error $\epsilon_{t+\Delta t}(\mathbf{u})$:

$$\epsilon_{t+\Delta t}(\mathbf{u}) = \{H_{t+\Delta t}(\mathbf{u}) - \mu(\mathbf{u})\} - \{a_1(\mathbf{u})\{H_t(\mathbf{u}) - \mu(\mathbf{u})\} + b_0(\mathbf{u})P_{t+\Delta t}\}. \quad (\text{A6.7})$$

In this study Δt equals one day. If water table depths are observed at a large number of successive days, then $\sigma_\epsilon(\mathbf{u})$ can be estimated from observed one-step-ahead prediction errors. This is the case for the 27 validation points, as described in sections 5.2 and 6.2. At these locations the water table depth was observed daily during five separate 5-day periods in 1997 and 1998, representing both wet and dry meteorological conditions, see Table 5.1 and Figure 6.1. The number of observations for these series is 25, 24 or 20, so that 20, 19 or 16 one-step-ahead prediction errors could be calculated, respectively. Next $\sigma_\epsilon(\mathbf{u}_i)$ was estimated by

$$\hat{\sigma}'_\epsilon(\mathbf{u}_i) = \sqrt{\frac{1}{n_i - 1} \sum_{j=1}^{n_i} (\epsilon_j - \bar{\epsilon}_i)^2}, \quad (\text{A6.8})$$

where $i = 1, \dots, 27$ indicates the validation location, $j = 1, \dots, n_i$ indicates the j th one-step-ahead prediction error at the i th validation location, and $\bar{\epsilon}_i$ is the mean of the one-step-ahead prediction errors at validation location i . The thus obtained estimates $\hat{\sigma}'_\epsilon(\mathbf{u}_i)$ are compared with estimates $\hat{\sigma}_\epsilon(\mathbf{u}_i)$ which result from calibration of the Kalman filter algorithm in Figure A6.1. It is clear that both estimates are not related to each other. When compared to $\hat{\sigma}'_\epsilon(\mathbf{u}_i)$, which is based on observed one-step-ahead prediction errors, the KF methods underestimate both the mean level of $\sigma_\epsilon(\mathbf{u})$ and its spatial variation. This can be explained from kriging of the ‘known’ values of $\sigma_\epsilon(\mathbf{u})$ at the 13 locations of permanently installed observation wells to any other location in the study area, see Figure 5.1. The spatial mean of the $\hat{\sigma}_\epsilon(\mathbf{u}_i)$ values resulting from the KF methods roughly corresponds with the mean of the values of $\sigma_\epsilon(\mathbf{u})$ observed at these 13 locations (4.2 cm), which can be considered as the effect of smoothing. In other words, the values of $\hat{\sigma}_\epsilon(\mathbf{u})$ are ‘dominated’ by the values that were observed at the 13 locations of permanently installed observation wells, which do not represent the water table dynamics in the entire area. The values of $\hat{\sigma}_\epsilon(\mathbf{u})$, estimated by the Kalman filter, show less spatial variation than the values of $\hat{\sigma}'_\epsilon(\mathbf{u})$, estimated from the one-step-ahead prediction errors, see Figure A6.1. The underestimated spatial variation of $\sigma_\epsilon(\mathbf{u})$ can also be explained from the smoothing effect of kriging. This is particularly true for KF3, where relatively small effective ranges were calibrated for the variograms of $\sigma_\epsilon(\mathbf{u})$. As a result, at short distances from the 13 permanently installed observation wells the values of $\hat{\sigma}_\epsilon(\mathbf{u})$ equal the mean of the values of $\sigma_\epsilon(\mathbf{u})$ that were observed at the 13 well locations. Obviously, the effect of updating on the basis of observed water table depths from the 27 locations of temporarily installed wells, is very limited on spatial predictions of $\sigma_\epsilon(\mathbf{u})$. Therefore, the values of $\hat{\sigma}_\epsilon(\mathbf{u})$ mainly reflect the values of $\sigma_\epsilon(\mathbf{u})$ that were observed at the locations of the 13 permanently installed observation wells. It should be noted, however, that due to the cross-validation procedure followed here the distance to these 27 locations is overrated. At smaller distances to the 27 temporarily well locations the effect of updating may be larger, resulting in less smoothed values of $\sigma_\epsilon(\mathbf{u})$.

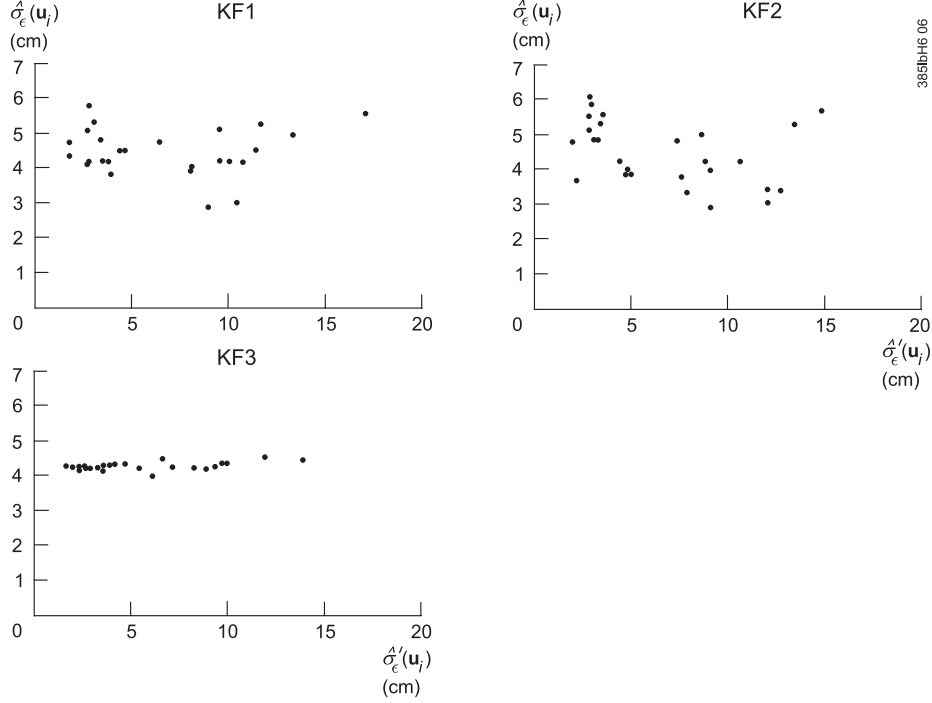


Figure A6.1: $\sigma_\epsilon(\mathbf{u})$ resulting from calibration of a Kalman filter algorithm ($\hat{\sigma}_\epsilon(\mathbf{u})$) vs. $\sigma_\epsilon(\mathbf{u})$ estimated from observed one-step-ahead prediction errors ($\hat{\sigma}'_\epsilon(\mathbf{u})$).

A6.1.2 Accuracy of quantified uncertainty about the mean depth

The error in the prediction of the mean water table depth at location \mathbf{u} is quantified by the mean error, $ME(\mathbf{u})$. For a location k where a time series of water table depths with length n_k was observed the mean error is calculated by:

$$ME(\mathbf{u}_k) = \frac{1}{n_k} \sum_{j=1}^{n_k} h_j(\mathbf{u}_k) - \frac{1}{n_k} \sum_{j=1}^{n_k} h_{F,j}(\mathbf{u}_k). \quad (\text{A6.9})$$

It is shown in Tables 6.1 and 6.2 that large systematic errors can be present in spatially predicted time series of water table depths. In section 5.8 a procedure is given to simulate possible values of $ME(\mathbf{u})$ at unvisited locations, as part of a procedure to construct a risk map. However, the uncertainty about $ME(\mathbf{u}_k)$ which arises from the limited size of the sample taken at location k is neglected in this procedure. The sample size is limited in two ways. Firstly, the samples cover periods of limited length, and may therefore not represent the full fluctuation zone of the water table. Secondly, the time steps at which the water table depth was observed may be selected with unequal and unknown probability, as is the case in this study (see Figure 6.1). Because of the latter, it is not easy in this case study to quantify the uncertainty about $ME(\mathbf{u}_k)$. If the water table depth were observed continuously in time, for

Appendix A6.1 Accuracy of quantified uncertainty

instance by using automatic devices such as divers, then the uncertainty related to sample size could be reduced to zero, provided that the water table showed maximum fluctuation during the monitoring period. If it is not possible to record the water table depth continuously, then it may be worthwhile to collect data following some random sampling design, because these designs enable us to make statistical inference with respect to our uncertainty about $ME(\mathbf{u}_k)$ (Cochran, 1977). At a location \mathbf{u}_k we consider the water table depth during a predefined period in time as the population from which a random sample is taken. If the sample has size n , the observed water table depths $\{h_1(\mathbf{u}_k), h_2(\mathbf{u}_k), \dots, h_n(\mathbf{u}_k)\}$ can be compared with deterministically predicted water table depths $\{h_{F,1}(\mathbf{u}_k), h_{F,2}(\mathbf{u}_k), \dots, h_{F,n}(\mathbf{u}_k)\}$, resulting in a random sample $\{e_1(\mathbf{u}_k), e_2(\mathbf{u}_k), \dots, e_n(\mathbf{u}_k)\}$ following Eq. (A6.3). This sample can thus be considered to be taken from a population of differences between observed and predicted water table depths. For a simple random sampling design the accuracy of $ME(\mathbf{u}_k)$ can now be quantified by its standard error:

$$se(ME(\mathbf{u}_k)) = \sqrt{\frac{SDE^2(\mathbf{u}_k)}{n}}, \quad (\text{A6.10})$$

where $SDE(\mathbf{u}_k)$ is calculated by Eq. (5.12), replacing i by k .

Chapter 7

General discussion and conclusions

7.1 Introduction

In this chapter the results and conclusions of the studies reported in chapters 2 to 6 are discussed in the perspective of the objectives as formulated in chapter 1. Apart from these main objectives, several problems encountered during the work on this thesis are discussed. In section 7.2, the estimation of fluctuation characteristics of the water table depth representing the prevailing climatic and hydrologic conditions is dealt with. In section 7.3 attention is paid to time series models for the relationship between precipitation surplus and water table depth. Special attention will be given to the extent of knowledge about the physical system incorporated into the models. Section 7.4 deals with the different ways to find an appropriate time series model. Section 7.5 discusses the physical basis of the ARX(1,0) model. The prediction performance of the regionalised time series model applied in this thesis (the RARX(\mathbf{u} ;1,0) model) is discussed in section 7.5. Section 7.7 treats the accuracy of quantified uncertainty about the water table depth. Finally, in section 7.8 some thoughts about future research needs are given.

7.2 Estimation of fluctuation statistics

In *chapter 2* a method was presented to estimate fluctuation statistics of the water table depth representing the prevailing climatic and hydrologic conditions, independent of the precipitation surplus during the monitoring period. The length of the monitoring period is generally restricted to 4 to 10 years. In essence the procedure is:

1. model the relationship between precipitation surplus and water table depth, by using models with a stochastic component;
2. transform an extensive time series of precipitation surplus into a series of water table depths by using the model resulting from step 1. In this study series of

General discussion and conclusions

30 years are used, assuming that the average weather conditions during those 30 years represent the prevailing climate;

3. generate a large number of realisations by using the stochastic model component;
4. calculate fluctuation statistics from the realisations of 30-year series of water table depths.

Fluctuation statistics such as the mean highest and mean lowest water table (*MHW* and *MLW*, respectively) used to be estimated from observed time series covering at least 8 years (Van Heesen, 1970). However, in practice these lengths are generally restricted to 4 to 10 years. The results in chapter 2 show that estimates based on an 8-year series may differ more than 20 cm from estimates based on a 30-year series (see Figures 2.7 and A2.2). As the results in Figure A2.2 indicate, the 30-year estimates show less variation than the 8-year estimates. It can be concluded that the methodology presented in chapter 2 meets the first objective formulated in chapter 1: **to develop stochastic methods for estimating fluctuation characteristics of the water table depth representing the prevailing climatic and hydro-logic conditions**. Apart from *MHW* and *MLW*, this methodology can be used to estimate various fluctuation characteristics, for instance duration curves (chapter 3). The structures of the models described in chapters 2 and 3 imply that water table depths can only be simulated with semi-monthly time intervals. Hence, the fluctuation characteristics represent the semi-monthly frequency only. Recently, this restriction of time series models was eliminated by Bierkens *et al.* (1999). They embedded the TFN model in a Kalman filter algorithm, enabling the prediction or simulation of time series of water table depths at the daily frequency of the input series on precipitation surplus.

7.3 Time series models for water table depths

Two models were applied in chapter 2 to describe the relationship between precipitation surplus and water table depth: i) the physical descriptive model SWATRE (Belmans *et al.*, 1983) supplemented with a stochastic model for the noise series (SWATRE+ARMA), and ii) transfer-function models with added noise (TFN, Box and Jenkins, 1976). TFN models are referred to as *black box models* (Bohlin, 1991), because they are lacking a physical basis. They only describe empirical relationships between input and output series. Physical mechanistic models such as SWATRE, are called *white box models*, because they are based on specific knowledge of the physical system. The performance of both models in estimating fluctuation statistics was evaluated in a validation experiment. Although SWATRE is conceptually based on specific knowledge of the physical system, the validation results were only slightly better than those obtained by the TFN models. This is probably due to errors in the input data on soil physical relationships. Moreover, hysteresis effects were not included into the model and, for the sake of mathematical simplicity, physical processes are schematised. For these reasons water table depths simulated by using the SWATRE model may differ from observed water table depths. If these differences are minimised by calibrating some of the model parameters, their values may lose

their physical meaning, since they counterbalance errors in the physical data and the model structure. Thus, once calibrated, the SWATRE model can be seen as a *grey box model*, as it is based on knowledge of the physical structure as well as on observed time series.

The validation results in chapter 2 indicated that both the SWATRE+ARMA model and the TFN model can be improved. More specifically, TFN models describe linear relationships between precipitation surplus and water table depth, which is not consistent with reality. A nonlinear rather than a linear relationship may be expected between precipitation surplus and water table depth, due to the presence of different soil layers or drainage levels and a time-varying storage capacity of the unsaturated zone.

In *chapter 3* an attempt was made to improve the estimation of fluctuation statistics, by applying a time series model that accounts for threshold nonlinearities in the relationship between precipitation surplus and water table depth. The performance of this so-called threshold autoregressive self-exciting open-loop (TARSO) model (Tong, 1990) was evaluated in a validation experiment. The SWATRE model, the TFN model and the dynamic regression model (DR), which is a special form of the TFN model, were used as benchmarks for model comparison. It is concluded that TARSO models are adequate alternatives for the DR and the TFN models as well as the SWATRE+ARMA model to simulate water table depths and to estimate durations of exceedance. The differences in the validation results for the four evaluated models are small, however. The TARSO models incorporate several regimes which are separated by thresholds. As explained in subsection 3.3.1, the active regime depends on the water table depth at the previous time step. Within each regime the relationship between precipitation surplus and water table depth is described by a linear dynamic regression model. An automatic procedure was followed in determining the threshold values and the autoregressive orders of the dynamic regression model for each of the regimes, using an automatic selection criterion (BIC, see section 3.4). The role of prior knowledge on the physical system in the model selection procedure will be elaborated upon in the following section. In fact, the selected subset-TARSO models in chapter 3 can be considered as grey box models, because prior physical knowledge was used in the model selection procedure. Moreover, the selected models could be interpreted physically.

TARSO models can be classified in between the physical descriptive model SWATRE and linear time series models such as TFN and DR models. SWATRE explicitly describes threshold nonlinearities such as drainage levels and soil physical boundaries, in addition to other nonlinear relationships. TARSO models only take threshold nonlinearities into account, whereas TFN models and DR models have a linear structure. However, the validation results in chapters 2 and 3 indicated only small differences in performance between the various models, despite the clearly different model concepts.

7.4 Finding an appropriate model

In chapters 2 to 4 appropriate time series models for the relationship between precipitation surplus and water table depth were found in three different ways:

General discussion and conclusions

1. a procedure of identification, estimation (calibration) and diagnostic checking (verification) as given by Box and Jenkins (1976) (chapter 2);
2. automatic model selection using an automatic selection criterion (Bayes Information Criterion, chapter 3);
3. a physical analysis (chapter 4).

The first two methods are based on observations on the input and output variables. Knowledge about the physical system is not used. The disadvantage of the first method is that its results depend on experts' insights and experience. Therefore the results may not be reproducible, in contrast to the second method which allows models to be selected in an objective and reproducible way. Both methods or parts of them can be used to explore data, in order to gain more insight into a physical system. The selected model structures may confirm existing theory about the underlying physics, but may also give rise to new insights into physical processes. This is extensively discussed in P.C. Young's work on the data-based mechanistic modelling (DBM) approach, see for instance Young and Beven (1994). The DBM approach starts with the selection of a model structure by applying objective statistical inference to time series data. Next, only those model equations are accepted which are found to be physically relevant. The procedure followed in chapter 3 is similar to the DBM approach. First, in chapter 3 appropriate subset-TARSO models were selected from a large set of candidate models, by using an automatic model selection criterion (BIC, section 3.4). Next, the selected models were interpreted physically (subsection 3.6.1).

An important step in automatic model selection procedures is the compilation of the set of candidate models. The number of candidate models should not be too large. The larger the set, the higher the risk of finding a model which fits well to the data, but which describes the underlying process inadequately. This risk can be lowered by restricting the set of candidate models to models which are relevant from a physical point of view. It should be noted, however, that if the set of candidate models is restricted too drastically on the basis of existing physical insights, little opportunity is left to gain new physical insights from the selected models. Obviously, in the compilation of the set of candidate models some subjective choices cannot be avoided. Apart from that, automatic selection procedures can be considered as objective and reproducible. Prior knowledge on the physical system was used in the selection procedure for subset-TARSO models in chapter 3. Firstly, the delay parameter, d in Eq. (3.2), was fixed at unity, based on prior physical knowledge as explained in section 3.4. Secondly, the number of regimes was fixed on the basis of prior physical information as well. It was assumed that thresholds in the relationship between precipitation surplus and water table depth are related to drainage levels and soil physical boundaries. These levels were observed in the field; the number of regimes was derived from the number of observed drainage levels and determined soil physical boundaries. The candidate threshold levels were restricted only with respect to a minimum of observations within the regimes, as explained in section 3.4. The selected threshold values were compared with drainage levels and levels of soil physical boundaries in subsection 3.6.1, in order to evaluate the theory that threshold nonlinearities are caused by these drainage levels and soil physical boundaries. Indeed, several threshold values

were close to drainage levels and soil physical boundaries. If drainage levels or soil physical boundaries were absent, or not so pronounced, the selected threshold values could only be explained from the non-constant relationship between precipitation surplus and water table depth caused by the non-constant storage capacity of the unsaturated zone, as pointed out in subsection 3.6.2.

Diagnostic checks are performed in chapter 2 to verify lack of fit, see subsections 2.2.2 and 2.2.3 and Figures 2.3 and 2.6. The procedure of identification, estimation and diagnostic checking is repeated until a suitable model is found which explains all data and satisfies all model assumptions. This procedure may be sound from a scientific point of view. However, from an engineering point of view it may make sense to keep the purpose of the model in mind during the modelling process. For instance, diagnostic checks applied to the residuals may indicate the presence of correlations which are unacceptable from a theoretical point of view. However, if the magnitude of these residuals is not relevant given the purpose of the model, then the model may still be appropriate for practical application. Therefore, apart from diagnostic checking, the simulation results were *validated* in chapter 2. In chapter 3 the diagnostic checks were replaced by a validation procedure with respect to the purpose of the model. It was assumed that if the residuals depart from the white noise assumption, this would show up in a poor simulation performance, see section 3.5. Of course, validation procedures depend on the availability of independent validation sets which are sufficiently large and which represent the processes under investigation.

7.5 The physically based ARX(1,0) model

Compared to the automatic model selection procedure applied in chapter 3, in chapter 4 a time series model was *chosen* rather than selected. In fact, the number of candidate models was restricted to *one* on the basis of a physical analysis. It was shown that a simple first order autoregressive-exogenous variable model (ARX(1,0) model) can easily be expressed in terms of a water balance for a soil column. The physical interpretation could straightforwardly be extended to TARSO models. Interestingly, both the procedure of identification, estimation and diagnostic checking in chapter 2 and the automatic model selection procedure in chapter 3 indicated first order autoregressive processes for the relationship between precipitation surplus and water table depth. The process of physical interpretation reveals the physical meaning of the assumptions made in linear time series modelling. The linear model structure implies that drainage levels, the effective porosity of the soil, and the regional component of groundwater flow are assumed to be independent of the water table depth. However, drainage levels will be independent of the water table depth only in areas where constant surface water levels are maintained. In other areas drainage levels will show temporal fluctuations depending on the groundwater storage. Furthermore, the effective porosity of the soil will vary in time, depending on the fraction of pores filled with air, and thus also depending on the water table depth. Finally, the regional component of groundwater flow will probably show temporal fluctuations, depending on the gradient in the groundwater head and thus also depending on the water table depth. Besides this, the regional component of groundwater flow will probably show temporal fluctuations, reverberating persistent long term meteorological fluctuations.

The noise component of the ARX(1,0) model contains remaining influences which are not incorporated explicitly into the model. These include, for instance, the difference between potential and actual evapotranspiration, surface runoff, flow through macropores or cracks and preferential flow. Furthermore, the noise component includes the temporal variation of those terms which are assumed to be time invariant: the drainage level, the effective porosity and the regional component of groundwater flow. The noise component of the ARX(1,0) model is assumed to follow the same autoregressive structure as the dynamic component, as explained in section 4.2 and Eqs. (4.2) and (4.3). In contrast to the ARX model, in a TFN model the noise component does not necessarily follow the same autoregressive structure as the dynamic component. The validation results in chapter 3 indicated that TFN models perform slightly better than DR models, which are equivalent to ARX models with respect to the autoregressive structure of the noise component. Therefore, the autoregressive structure of the unknown influences forming the noise component possibly differs from the autoregressive structure of the dynamic component and thus, ARX models may not be adequate at this point.

In a case study (section 4.4) it is demonstrated that the physically interpreted ARX(1,0) model can be applied to predict the effects of human interventions on the water table dynamics. The validation results indicate that the ARX(1,0) model predicts the effects reasonably well, despite its drastically schematised representation of the physical system. Because of its physical basis the ARX(1,0) model can easily be linked to additional information. The physically based time series model forms the basis of the spatio-temporal prediction methods described in chapters 5 and 6, because in areas where only a limited number of observed time series is present it makes sense to incorporate additional information into the prediction methods.

7.6 Spatio-temporal predictions of water table depths

The second objective of the study reported in this thesis is: **to develop methods to predict water table depths in space and time in terms of probabilities, for application in regions where suitable observation wells are scarce.**

For that purpose in chapter 5 a regionalised time series model was devised. Here ‘regionalised’ means that the parameters of the time series model are made dependent of location. The physically based ARX(1,0) model for the relationship between precipitation surplus and water table depth, presented in chapter 4, is regionalised, because its parameters can be ‘guessed’ from widely available information on the physical system. This makes the method suitable for application in areas where only a limited number of observed time series is available. For such areas it is not possible to model the spatial structure of time series model parameters which are obtained through calibration on observed time series. By using the physical basis of the ARX model, additional information on the physical system can be transformed into time series model parameters. Additional information on the physical system is obtained from digital elevation models (DEM), detailed digital topographic maps which contain information on the locations and sizes of ditches, and digitally stored soil profile

descriptions which are combined with soil physical standard curves. In section 5.4, methods are described to obtain time series model parameters from this additional information. The thus obtained parameter values are referred to as ‘guessed’ parameter values, because they result from straightforward transformation rather than from calibration on observed time series. The regionalised ARX(1,0) model is referred to as the RARX(\mathbf{u} ;1,0) model, with \mathbf{u} being a vector of spatial co-ordinates.

Two spatio-temporal prediction methods based on the RARX(\mathbf{u} ;1,0) model are presented in chapter 5: the ‘direct’ method (DM) and the ‘indirect’ method (IM). DM is a purely deterministic method which provides predicted water table depths only. Apart from predictions IM provides quantitative information on the uncertainty about the true water table depths. IM can thus be considered as a stochastic method, which enables us to predict water table depths in terms of probabilities.

The third objective of this thesis is: **to evaluate the accuracy of the stochastic methods developed with respect to objectives 1 and 2.** Therefore in chapter 6 the accuracy of four alternative spatio-temporal prediction methods based on the RARX(\mathbf{u} ;1,0) model is assessed in a case study. Two types of information are distinguished:

1. predictions of water table depths. Using these predictions, water managers can immediately anticipate actual trends;
2. estimates of statistics of the water table fluctuation, e.g., fluctuation statistics like *MHW* and *MLW* (chapter 2), duration curves (chapter 3), probabilities of exceedance (chapter 6) and expected water table depths (chapter 6). These statistics can help water managers in assessing options for long term water policy.

Apart from DM and IM, three applications of a RARX(\mathbf{u} ;1,0) model, embedded in a spatio-temporal Kalman filter algorithm, are evaluated in chapter 6, the details of which are explained in Bierkens *et al.* (2001). Three alternative applications of the Kalman filter algorithm, referred to as KF1, KF2 and KF3, respectively, are described in subsection 6.3.3. In KF1 and KF2 physically based, guessed RARX parameters are used as auxiliary information in the interpolation of the RARX parameter values which are ‘known’ for 13 locations in the study area. At these 13 locations time series were observed which were suitable to calibrate the RARX model parameters. In KF1 external drift kriging is applied in the interpolation of all RARX parameters. In KF2 simple kriging with varying means is applied to interpolate $a_1(\mathbf{u})$, $b_0(\mathbf{u})$ and $\mu(\mathbf{u})$, whereas external drift kriging is applied to interpolate $\sigma_\epsilon(\mathbf{u})$, with guessed values $\tilde{\mu}(\mathbf{u})$ as an external drift. In KF3 external drift kriging is applied, with ground surface elevation from the DEM as an external drift for all RARX parameters. It should be noted that KF methods depend on the presence of time series which are suitable to calibrate RARX model parameters, in contrast to DM and IM.

In KF methods the spatio-temporal predictions can be updated at any moment that observed water table depths become available. In other words, KF methods enable us to predict water table depths spatially at time t , conditionally to the observed water table depths up to and including time t . This property makes the KF methods attractive for the prediction of actual water table depths. DM and IM provide predictions which, in fact, are linear transformations of the precipitation surplus only.

In DM observed water table depths are not at all taken into account. In IM observed water table depths are only used to correct for systematic errors and to quantify the uncertainty about the true water table depths.

The validation results in chapter 6 indicate that, at least in this case study, KF methods predict the temporal fluctuations of the actual water table depth more precisely than DM and IM. This is clearly the effect of the ‘updating’ in the KF methods. However, IM shows less large *systematic* errors than the alternative methods. All methods estimate the expected water table depth equally well. It is indicated that IM estimates the expected (temporal) mean depth of the water table in a given hydrologic regime more accurately than the KF methods. A possible explanation is that the 13 locations for which RARX model parameters were calibrated on observed time series, did not represent the average conditions in the study area. In general, the validation results show that the temporal variation of the water table depth is predicted reasonably precisely, but that relevant errors can occur in estimating the mean depth around which the water table fluctuates. Thus it is concluded that the spatio-temporal prediction methods need to be improved with respect to the prediction of the temporal mean.

In the methods described in chapters 5 and 6 time series modelling and geostatistical methodology are combined to parametrise the spatio-temporal function $X(\mathbf{u}, t)$. A similar approach was followed for weather data by Hutchinson (1995), who interpolated the parameters of stochastic point models for rainfall spatially by thin plate smoothing splines, accounting for dependence on elevation. Van Geer and Zuur (1997) modelled the spatial structure of parameters of TFN models for the relationship between precipitation surplus and water table depth. In contrast to the methods described in chapters 5 and 6, they did not use auxiliary information in the interpolation, because in their case time series model parameters were calibrated for a sufficiently large number of locations. In section 5.1 two alternative approaches for regionalised time series models were mentioned. Multivariate geostatistical methods form the first alternative, see for instance Rouhani and Hall (1989) for an application to piezometric heads. The second alternative is formed by methods based on multivariate time series modelling (for instance Pfeifer and Deutsch, 1980; Dalezios and Adamowski, 1995). As compared to regionalised time series models, multivariate geostatistical approaches as well as multivariate time series modelling generally demand many observations on the target variable in both time and space. Because the studies in this thesis concern regional applications in which observed water table depths are scarce, methods based on a regionalised time series model were devised for spatio-temporal prediction of water table depths.

7.7 The accuracy of quantified uncertainty

The stochastic methods devised in this thesis enable us to describe the dynamics of the water table in terms of probabilities. Therefore, the methods can be applied in risk assessment. An example is given in Figure 5.9, which shows a map reflecting the risks that the water table is present within a depth of 50 cm on April 1st in any future year, given the actual hydrologic and climatic conditions. Two components of uncertainty contribute to these risks: i) the uncertainty about the future precipitation

surplus, and ii) the uncertainty about the relationship between precipitation surplus and water table depth as described by the $\text{RARX}(\mathbf{u};1,0)$ model. The uncertainty about the future precipitation surplus is approximated by the variation of the precipitation surplus over the last 30 years. The uncertainty about the representativeness of the $\text{RARX}(\mathbf{u};1,0)$ model is quantified by using the differences between observed and predicted water table depths at 40 (cross-)validation locations. In section 5.8 the procedure for drawing a risk map is explained and three assumptions in calculating the risks are given. A fourth assumption can be added to these three: the uncertainty on the relationship between precipitation surplus and water table depth as described by the $\text{RARX}(\mathbf{u};1,0)$ model is assumed to be completely covered by the prediction errors observed at the 40 (cross-)validation locations.

If the risk of taking a wrong decision is close to 50%, it will be difficult to decide. Because Figure 5.9 shows only small areas with risks of around 50%, the map can be helpful to water managers when they have to take strategic decisions. Of course, the calculated risks may deviate from the true risks, since they are based on several assumptions. Evidently, the practical value of stochastic spatio-temporal prediction methods in decision making depends on the accuracy of the quantified uncertainty. Therefore in appendix A6.1 the accuracy of quantified uncertainty is assessed, by means of cross-validation.

In summary, the analyses in appendix A6.1 indicate the following:

1. The uncertainty about the temporal variation of the water table depth is underestimated by the spatio-temporal prediction methods applied in this study. This may be caused by the sampling design, which did not fully represent the temporal variation of the water table depth. Apart from that, in the KF methods the uncertainty of regionalised parameters $a_1(\mathbf{u})$, $b_0(\mathbf{u})$ and $\mu(\mathbf{u})$, was not taken into account in the spatial predictions of $\sigma_N(\mathbf{u})$. Furthermore, the analyses indicated that the interpolated values of $\sigma_\epsilon(\mathbf{u})$ are smoothed. In this case study the smoothing effect of kriging resulted in underestimation of both the mean level and the spatial variation of $\sigma_\epsilon(\mathbf{u})$;
2. The accuracy of the uncertainty about the mean depth, around which the water table fluctuates in time, cannot be assessed using the sampling design in the present case study. Hence, it can be concluded that the sampling design can be improved in this respect. More specifically, to provide estimates of the uncertainty about the mean error, the water table depth must be monitored continuously in time or, at least following a random sampling design.

7.8 Some topics for further research

During the work on this thesis several problems were faced which could not be investigated within the scope of the studies reported here. In this final section some ideas for future research are given. First, some opportunities to improve the methods presented in this thesis are discussed. Next, some ways to improve or to complete data are suggested. Finally, some thoughts are given on the tuning of stochastic methods to support decision making in groundwater policy.

General discussion and conclusions

The results in chapters 2 to 4 indicated that the time (series) models applied in this thesis did not explain all data. Furthermore, the diagnostic checks in chapter 2 indicated that the models did not satisfy all model assumptions. Thus, provided that the data are correct, the models can be improved from a scientific point of view. Besides, the validation results in chapters 2 and 3 indicated that the models could be improved keeping the purposes of the models in mind. However, the validation results also indicate that linear time series models, threshold nonlinear models and the physical descriptive model SWATRE perform only slightly differently. Therefore, possible improvements may be expected from other types of models than those applied in this thesis.

In chapter 2 it was pointed out that shallow water tables vary more with the precipitation surplus than deep water tables, due to the varying storage capacity of the unsaturated zone. As a result, the errors remaining from a TFN model are heteroscedastic (i.e., the errors have non-constant variance). Time series models accounting explicitly for heteroscedasticity are the GARCH models (Generalised Autoregressive Conditional Heteroscedasticity, e.g. Franses and Van Dijk, 1996). Essentially, the error variance is modelled by an ARMA model in GARCH models. Given the heteroscedastic errors resulting from TFN models, it might be interesting to apply GARCH models to the noise component of TFN models. It would be particularly interesting to give a physical basis to the GARCH model, analogous to chapter 4. This physical basis can be used by predicting the error variance component of the GARCH model spatially. The TARSO model (chapter 3) also accounts for non-constant error variance. However, the TARSO model is not smooth around the threshold levels which divide the fluctuation of the water table depth into regimes. The Multivariate Adaptive Regression Splines algorithm (MARS) for time series, introduced by Lewis and Stevens (1991), can be used to obtain continuous (smooth) nonlinear threshold autoregressive models. A MARS algorithm for time series (TSMARS) was applied to sea surface temperature series by Lewis and Ray (1997). They applied a semi-multivariate adaptive spline threshold autoregressive (SMASTAR) model, that incorporates several predictor variables (wind direction and wind speed) to model the sea surface temperature. Interestingly, this model could be interpreted by oceanographers. The development of a SMASTAR model for water table depths that can be interpreted physically would be an interesting topic for further research.

The validation results in chapters 2 and 3 indicate systematic errors in the simulated water table depths. In subsection 3.6.2 it was suggested that systematic errors appearing from the validation experiment could possibly be explained from persistent long term meteorological fluctuations, which cause long range fluctuations of the regional groundwater flux. Another possible cause of the systematic errors could be the lowering of water tables as a result of intensified drainage and groundwater withdrawal. The causes of these systematic errors should be further examined. The next question to be answered is how to deal with these phenomena in predicting actual water table depths and estimating statistics. Influences like groundwater abstraction could be incorporated into time series models as input series, next to precipitation surplus, provided that groundwater abstraction data are available. If the history of drainage works is known, they could be treated as interventions (e.g. Hipel *et al.*, 1975; Hipel and McLeod, 1994). Recently, Coulibaly *et al.* (2001) applied artificial

neural network models (ANN) to forecast deep water tables during the dry season in Burkina Faso. In this part of the Sahel region, the hydrological regime changes in time, because of the decreasing trend in rainfall and increasing water demands. The results indicate that ANN models can be applied to a hydrologic regime that changes in time. Apart from a time series of water table depths, time series on precipitation, river water level, maximum and mean temperatures were used in ANN modelling of water table depths. It might be interesting to investigate the performance of ANN models in predicting water table depths in parts of the Netherlands, where the hydrological system is changing because of increasing water demands and intensified drainage. A disadvantage of ANN models is that they cannot easily be interpreted physically, they really are ‘black box’ models. This may be a disadvantage when the models are extended to a space-time context and observed time series of water table depths are scarce.

The ‘echo’ of persistent long term meteorological fluctuations in the water table depth could be modelled by including autoregressive terms with long lags into the time series model. Another way to account for these meteorological fluctuations may be to incorporate an input series of observed groundwater heads, which reflect these long term meteorological fluctuations. However, in general it will not be possible to identify these models and to calibrate their parameters, because observed time series of sufficient length are scarce.

In chapter 6 the accuracy of three spatio-temporal prediction methods based on the $RARX(\mathbf{u};1,0)$ model was evaluated. The accuracy may increase if the methods described in chapters 5 and 6 are combined. For instance, values of $\tilde{a}_1(\mathbf{u})$, $\tilde{b}_0(\mathbf{u})$ and $\tilde{\mu}(\mathbf{u})$ are obtained by the indirect method, IM, firstly. Next, only the ranges of the variograms for $\sigma_\epsilon(\mathbf{u})$ and $\epsilon(\mathbf{u})$ are calibrated by the space-time Kalman filter.

In appendix A6.1 and section 7.7 it was pointed out that the accuracy of quantified uncertainty depends on the sampling design. It was suggested that the uncertainty can be quantified more accurately if the water table depths are observed continuously in time at the validation points, or at least following a random sampling design. In chapters 5 and 6 an existing sampling design was used with which the uncertainty could not be quantified accurately, as appeared from the results in appendix A6.1. An interesting topic for further research would be to optimise spatio-temporal sampling designs with respect to accurate estimates of uncertainty, in particular the application of classical sampling theory both in space and in time (e.g. Brus and De Gruijter, 1997; De Gruijter, 2000).

The accuracy of predictions or estimates depends not only on models or sampling designs, but also on the quality of the observations used. In this thesis, input series of potential precipitation surplus were calculated from daily data on precipitation and on potential Makkink crop evapotranspiration (De Bruin, 1987), observed at the nearest meteorological station. The effects of more advanced interpolation methods on the prediction of water table depths should be investigated. A discussion on the quality of data on precipitation and on evapotranspiration is beyond the scope of this thesis. It should be noted that the availability of accurate meteorological data is crucial in the application of the methods described in this thesis. The same holds for observations on water table depths. The measurement error in observed water table depths is neglected in this thesis. This can be justified, because the observation wells

are checked on defects like stoppages. Apart from that, the measurements, taken with a bell and a tape measure (chapter 1), can be considered as precise. However, little is known about the accuracy of measured water table depths. Therefore, it is advisable to investigate the accuracy of measured water table depths, preferably for various soil types. A specific problem is the presence of perched, temporary, water tables above impermeable layers, of which the depths cannot accurately be measured in the existing network of observation wells. For this reason the boulder clay area was excluded from the validation study in chapters 5 and 6. For applications in agriculture and ecology it is important to know the depth to these temporary water tables, because these temporary water tables are relevant to plant growth. A possible way to measure the depth to temporary water tables at a certain location is to install a number of wells with small filters at various depths, both above and below impermeable layers. The depth to perched, temporary, water tables can be observed in tubes with filters above an impermeable layer.

In chapters 1 and 6 it was mentioned that on average one suitable observation well per 750 to 1,250 hectares is present (Finke, 2000). The accuracy of predicted water table depths will increase if more observations are done. However, the accuracy of predicted water table depths will not necessarily decrease if the number of observed time series is reduced. If only those wells are removed in which erroneous time series of water table depths (i.e., negatively contributing to the predictions) are observed, the accuracy of predictions will increase. In other cases, the monitoring design has to be reconsidered in order to maintain or to improve the accuracy of predicted water table depths (e.g. Angulo *et al.*, 2000).

In observation wells and boreholes, water table depths can be measured at points in space. Water table depths can be observed continuously in space with ground-penetrating radar devices (e.g. Shih *et al.*, 1986; Anderson Jr. *et al.*, 2000). It would be interesting to investigate if water table depths in Dutch soil types can be accurately measured by ground-penetrating radar. An interesting application would be to measure the depth to perched water tables by ground-penetrating radar. Besides this, ground-penetrating radar can be used to make ‘snapshots’ of the water table at a large number of locations on a certain day.

Electronic divers can be installed in observation wells to measure water table depths continuously in time, or at least with a high sampling frequency, say hourly or at each quarter of an hour. Measurements can also be taken following some random sampling design. For these reasons, divers can be useful in quantifying the uncertainty about the water table depth, see section 7.7 and subsection A6.1.2.

Finally, some ideas are presented about a better tuning of stochastic methods for the prediction of water table depths to support decision making in groundwater policy and water management. By using stochastic methods it is possible to quantify uncertainty about the true water table depths. Of course, hydrologists are challenged to reduce this uncertainty, by improving models and by optimising monitoring strategies. However, some part of the water table depth will always be left unexplained, and furthermore, the water table depth will never be observed at all times and places. Quantified uncertainty can be useful in supporting strategic decisions in water policy. For instance, probabilities can be calculated which can be helpful in choosing between alternatives. An example is given in section 5.8, where a map is presented reflecting

the risk that the water table is within a critical depth at the start of the growing season. Note that ‘risk’ is used here as a synonym for ‘probability’. Risks can be estimated by multiplying probabilities with costs, for instance costs related to agricultural production losses. Apart from risks, probabilities can be translated into opportunities, for instance the opportunity that an ecosystem can be restored. Summarising, once translated into probabilities, risks or opportunities, quantified uncertainty can be used in water management and water policy.

As discussed in section 5.9, high uncertainty implies probabilities of approx. 50%, which makes it difficult to take decisions. Thus, probabilities or risks highlight the need for reduction of uncertainty. Next, an additional survey can be executed to provide more accurate information for decision makers. It might be more cost-effective, however, to decide on the accuracy needed for application in water management or policy before a survey or monitoring program starts. This implies that water managers, policy makers and scientists need to communicate about uncertainty, probabilities and risks in advance. For instance, a question on monitoring the effects of an intervention can possibly be translated into a testing problem with H_0 : no effect. Then, water managers or policy makers rather than scientists must decide on the least relevant difference, that is, the slightest effect that must be indicated. Furthermore, they should indicate the maximum tolerated probability of wrongly accepting H_0 , that is, wrongly concluding that an effect is absent (type II error). These decisions are based on the costs of the monitoring, and the costs which are made if it is wrongly concluded that an intervention had no effect. These costs can be considered as the ‘consumer’s risk’, which is equivalent to the probability of a type II error in hypothesis testing (Kendall and Buckland, 1982). In conclusion, to improve the fit of stochastic methods to problems in water management and groundwater policy, it is extremely important to translate practical questions carefully into statistical terms, as well as to translate statistical inference into answers which can be understood by water managers and policy makers. The utilisation of statistical knowledge in water management and water policy is an interesting topic for further research. Experts in statistics and hydrology as well as experts in management science, communication studies and psychology should co-operate in finding ways to improve the utilisation of knowledge about uncertainty. After all, being uncertain is a human condition.

Chapter 8

Samenvatting

Als je ergens op het platteland in Nederland een gat in de grond boort, bereik je na enige tijd meestal een diepte waarop de bodem is verzadigd met water. Het water dat niet is gebonden aan bodemdeeltjes en niet onder druk door een afsluitende laag staat, wordt 'freatisch', vrij, grondwater genoemd. Het freatische grondwater, dat zich vrij in de poriën van de grond bevindt, stroomt het boorgat in tot zich een bepaald niveau instelt. Op dit niveau is de drukhoogte van het grondwater gelijk aan de atmosferische druk. Als we een paar meter verderop een tweede boorgat maken, dan vinden we daarin waarschijnlijk op ongeveer dezelfde diepte freatisch grondwater als in het eerste boorgat. We gaan door met het maken van boorgaten, totdat we uiteindelijk een heel gebied hebben verkend. Na enige tijd, afhankelijk van de doorlatendheid van de grond, kunnen we in al deze gaten de diepte tot het freatische grondwater meten. Meestal gebeurt dit met behulp van een meetlint waaraan een cilindervormig klokje is bevestigd. Het klokje wordt langzaam neergelaten in het boorgat. Als het klokje de waterspiegel raakt hoor je 'plop' en kan de diepte van de grondwaterspiegel worden afgelezen van het meetlint. De waarnemingen in de boorgaten kunnen worden beschouwd als puntwaarnemingen van een continu oppervlak van freatisch grondwater, dat de grondwaterspiegel wordt genoemd. De diepte van de grondwaterspiegel ten opzichte van het maaiveld wordt aangeduid als de grondwaterstand. Als we in alle boorgaten tegelijkertijd de grondwaterstand zouden meten, dan leverde dat een interessante verzameling waarnemingen op van de grondwaterstand die in een gebied op een bepaald tijdstip is opgetreden. Hoogstwaarschijnlijk zal uit een eerste analyse van deze waarnemingen blijken dat de grondwaterstand varieert van plaats tot plaats. Vervolgens plaatsen we buizen met doorlatende filters in alle boorgaten, zodat we ook in de toekomst grondwaterstanden kunnen meten op diezelfde plekken. Keer op keer meten we de grondwaterstand in alle boorgaten, bijvoorbeeld met tussenpozen van enkele dagen. Analyseren we ook deze waarnemingen, dan zal blijken dat de grondwaterstand niet alleen van plaats tot plaats varieert, maar tevens door de tijd fluctueert, met een piek in de winter en een dal in de zomer. We kennen nu de grondwaterstand op lokaties waar grondwaterstandsbuizen staan en op tijdstippen waarop de grondwaterstand is gemeten. Voor lokaties waar geen buizen staan en voor tijdstippen waarop de grondwaterstand niet is gemeten zullen we voorspellingen moeten

Samenvatting

doen. Die voorspellingen kunnen afwijken van de grondwaterstand in werkelijkheid. We zijn dus onzeker over de echte grondwaterstand op de lokaties en tijdstippen waarvoor we voorspellingen hebben gedaan.

In het grootste deel van Nederland bevindt de grondwaterspiegel zich al binnen twee meter beneden maaiveld. Vanwege deze geringe diepte is informatie over de grondwaterstand belangrijk voor de landbouw en het natuurbeheer. Grondwaterbeheerders moeten rekening houden met meerdere, vaak tegengestelde belangen ten aanzien van de grondwaterstand. Een verandering in het grondwaterregime kan voor de één een gewenste ontwikkeling inluiden, maar voor de ander een risico betekenen. Om goede afwegingen te kunnen maken dienen grondwaterbeheerders te beschikken over nauwkeurige informatie over de grondwaterstand. Bij voorkeur is ook de betrouwbaarheid van deze informatie bekend, zodat zij kan worden gebruikt in risicoanalyses. In Nederland wordt de grondwaterstand al sinds 1948 regelmatig waargenomen in een netwerk van duizenden grondwaterstandsbuizen. Grondwaterstandsreeksen weerspiegelen de dynamiek van de grondwaterstand die hoort bij de hydrologische en meteorologische omstandigheden gedurende de waarnemingsperiode. Een waarnemingsperiode met constante hydrologische omstandigheden is meestal beperkt tot circa 2 tot 10 jaar. Om strategische beslissingen op het gebied van grondwaterbeheer te kunnen ondersteunen is echter informatie gewenst die de dynamiek weerspiegelt onder de heersende klimatologische omstandigheden, wat meestal benaderd wordt met de gemiddelde weersomstandigheden over een periode van 30 jaar. Er is dus behoefte aan methoden waarmee het mogelijk is om grondwaterstandsreeksen te extrapoleren naar reeksen van 30 jaar lang, waaruit vervolgens karakteristieken kunnen worden berekend.

Een grondwaterstandsreeks bevat alleen informatie over de dynamiek van de grondwaterspiegel op de plaats van de grondwaterstandsbuis waarin deze reeks is waargenomen. Voor toepassingen in het grondwaterbeheer is echter informatie nodig voor elke lokatie in een bepaald gebied. De dichtheid van het netwerk van grondwaterstandsbuizen is echter beperkt tot één per 750 tot 1 250 hectare. Regionale studies waarin informatie over de grondwaterstand gebiedsdekkend gewenst is hebben meestal betrekking op gebieden met een omvang van ca. 2 000 tot ca. 15 000 ha. Dit kunnen bijvoorbeeld ruilverkavelingen, (delen van) waterschappen of natuurgebieden betreffen. Het aantal grondwaterstandsbuizen in deze gebieden is te klein om de grondwaterstand nauwkeurig in ruimte en tijd te voorspellen op basis van uitsluitend waargenomen grondwaterstandsreeksen.

Tot nu toe werd de variatie van de grondwaterstand weergegeven op zogenaamde grondwatertrappenkaarten (Gt-kaarten). Deze kaarten, die tegelijk met de bodemkaart worden vervaardigd, kunnen worden beschouwd als ruimte-tijdmodellen van de grondwaterstand. Gt-kaarten worden gemaakt op basis van grondwaterstands-waarnemingen, bodemkundige profielkenmerken en landschappelijke kenmerken. Gt-kaarten kunnen verouderd zijn, doordat inmiddels de grondwaterstand is verlaagd, bijvoorbeeld omdat de grondwaterstand is aangepast aan landbouwkundige eisen, of omdat er grondwater wordt onttrokken voor drinkwater, industrie of beregening van landbouwgewassen. Een ander nadeel van Gt-kaarten is dat de indeling in grond-

watertrappen geënt is op landbouwkundige toepassing, terwijl voor toepassingen in bijvoorbeeld het natuurbeheer informatie gewenst is die niet door de Gt-kaarten wordt verschaft. Vanuit wetenschappelijk oogpunt heeft de methode waarmee Gt-kaarten vervaardigd worden het nadeel dat zij niet reproduceerbaar is. Tenslotte is de betrouwbaarheid van de Gt-kaarten onbekend, tenzij er extra inspanningen worden geleverd om deze alsnog te schatten. Daardoor zijn Gt-kaarten niet zonder meer geschikt voor gebruik in risicoanalyses. Risicoanalyse kan in het bijzonder nodig zijn als er in een gebied tegengestelde belangen bestaan ten aanzien van de grondwaterstand. Daarom is er behoefte aan stochastische methoden waarmee de dynamiek van de grondwaterstand in termen van waarschijnlijkheden kan worden beschreven.

De belangrijkste **doelen** van het onderzoek dat in dit proefschrift is beschreven zijn:

1. stochastische methoden te ontwikkelen waarmee de fluctuaties van de grondwaterstand zoals die zich voordoet onder de heersende hydrologische en klimatologische omstandigheden kan worden gekarakteriseerd;
2. methoden te ontwikkelen waarmee de grondwaterstand in ruimte en tijd kan worden voorspeld in termen van waarschijnlijkheden, in gebieden waar weinig bruikbare grondwaterstandsruizen aanwezig zijn;
3. de nauwkeurigheid te evalueren van de methoden die met het oog op de doelen 1 en 2 zijn ontwikkeld.

In *hoofdstuk 2* is een stochastische methode ontwikkeld waarmee fluctuatiekarakteristieken van de grondwaterstand kunnen worden geschat die representatief zijn voor de heersende hydrologische en klimatologische omstandigheden. De methode maakt gebruik van eendimensionale modellen die de relatie beschrijven tussen het neerslagoverschot en de grondwaterstand. Twee modellen met een verschillende theoretische basis zijn toegepast. Het eerste model is het fysisch-mechanistische model SWATRE voor de fysische beschrijving van stroming van water in de onverzadigde zone, uitgebreid met een tijdreeksmodel voor de residuen (SWATRE+ARMA). Het tweede model is het transfer-ruismodel (TFN), dat de empirische relatie tussen neerslagoverschot en grondwaterstand beschrijft. Met deze modellen worden grondwaterstandsreeksen gesimuleerd van 30 jaar lang, waaruit fluctuatiekarakteristieken kunnen worden berekend. In hoofdstuk 2 zijn dat de gemiddeld hoogste en gemiddeld laagste grondwaterstand (resp. GHG en GLG), waarop de indeling in grondwatertrappen is gebaseerd. De methode is toegepast op acht verschillende grondwaterstandsreeksen in een validatie-experiment. De resultaten van de SWATRE+ARMA-modellen verschillen maar weinig van die van de TFN-modellen, ondanks de duidelijke theoretische verschillen. Uit de resultaten blijkt dat GHG's en GLG's die zijn berekend uit reeksen van 8 jaar lang met meer dan 20 cm kunnen afwijken van GHG's en GLG's die berekend zijn uit reeksen van 30 jaar lang. Van praktisch belang is de minimale lengte die een grondwaterstandsreeks moet hebben om er een model op te kunnen kalibreren die de relatie met het neerslagoverschot adequaat beschrijft. De minimale

Samenvatting

lengte hangt af van 1) de responstijd tussen het neerslagoverschot en de grondwaterstand; 2) de variatie van de grondwaterstand tijdens de waarnemingsperiode; 3) de afwezigheid van een ingreep in de waterhuishouding en 4) het aantal waarnemingen dat nodig is om de modelparameters nauwkeurig te kunnen schatten. Een reeks moet tenminste een jaar lang zijn, mits in dat jaar een seizoensfluctuatie optreedt waarbij de top en het dal van de grondwaterstandsfluctuatie worden bereikt. Als de responstijd tussen het neerslagoverschot en de grondwaterstand langer is dan een jaar, dan is de minimale reekslengte gelijk aan de responstijd. Voor 51 reeksen afkomstig het Pleistocene deel van Nederland zijn de responstijden berekend. In zes gevallen bleek de responstijd langer dan een jaar te zijn.

In *hoofdstuk 3* is een zogenaamd TARSO-model toegepast op zes grondwaterstandsreeksen. Dit model verdeelt het grondwaterstandsverloop in verschillende regimes, die worden gescheiden door drempels. De relatie tussen het neerslagoverschot en de grondwaterstand in een bepaald regime wordt beschreven met een dynamische regressievergelijking. Passeert de grondwaterstand een drempel, dan treedt een ander regime in werking en geldt een andere dynamische regressievergelijking. Voor elk van de zes grondwaterstandsreeksen is een TARSO-model geselecteerd met behulp van een automatisch criterium voor modelselectie: het Bayes Informatie Criterium (BIC). Dit criterium maakt een afweging tussen enerzijds de mate waarin het model bij de waarnemingen past en anderzijds de complexiteit van het model. De geselecteerde TARSO-modellen zijn gebruikt om grondwaterstandsreeksen van 30 jaar lang te simuleren, waaruit karakteristieken zoals de GHG, de GLG en duurlijnen van de grondwaterstand kunnen worden berekend. De resultaten van de TARSO-modellen zijn in een validatie-experiment vergeleken met die van SWATRE+ARMA-modellen, TFN-modellen en dynamische regressiemodellen (DR). Uit het validatie-experiment blijkt dat de resultaten van de verschillende modellen weinig van elkaar verschillen. Als er duidelijke drempels in het grondwaterstandsverloop zijn te onderscheiden, bijvoorbeeld als gevolg van de aanwezigheid van drainageniveaus of lagen met verschillende bodemfysische eigenschappen, dan leveren TARSO-modellen betere resultaten op dan de lineaire TFN- en DR-modellen. De kleine verschillen tussen de validatieresultaten van TARSO-modellen en SWATRE+ARMA-modellen kunnen mogelijk worden verklaard uit onnauwkeurigheden in de invoergegevens die nodig zijn voor het fysisch-mechanistische model SWATRE.

In *hoofdstuk 4* is een eenvoudig lineair tijdreeksmodel voor de relatie tussen het neerslagoverschot en de grondwaterstand uitgedrukt in termen van de waterbalans voor een bodemkolom. De parameters van het *autoregressive exogenous variable*-model (ARX-model) kunnen worden geschreven in termen van drainageweerstand, effectieve porositeit, regionale grondwaterflux en drainageniveau. Uit twee voorbeelden blijkt dat met het fysisch geïnterpreteerde ARX-model het effect van ingrepen in de waterhuishouding op de grondwaterstand goed kan worden voorspeld.

In *hoofdstuk 5* wordt een geregionaliseerd tijdreeksmodel gepresenteerd voor de relatie tussen het neerslagoverschot en de grondwaterstand. Als basis dient het fysisch

geïnterpreteerde ARX-model dat in hoofdstuk 4 is beschreven. Als dit ARX-model wordt ‘geregionaliseerd’ tot een RARX-model dan wil dat zeggen dat de parameters van het tijdreeksmodel afhankelijk worden van de lokatie. Voor lokaties waar grondwaterstandsreeksen zijn waargenomen kunnen de waarden van de RARX-parameters worden geschat door middel van kalibratie. Op andere lokaties kunnen de waarden van de RARX-parameters worden gegist met behulp van de fysische relaties die in hoofdstuk 4 zijn gegeven en hulpinformatie zoals een digitale hoogtecijferkaart, een digitale topografische kaart met lokaties en afmetingen van sloten, digitaal opgeslagen bodemkundige profielbeschrijvingen en bodemfysische standaardcurven. Twee methoden om de grondwaterstand in ruimte en tijd te voorspellen worden in hoofdstuk 5 beschreven: de ‘directe’ methode (DM) en de ‘indirecte’ methode (IM). DM is een deterministische methode en maakt geen gebruik van waargenomen grondwaterstanden. In IM wordt als vervolg op DM gebruik gemaakt van waargenomen grondwaterstanden om te corrigeren voor systematische fouten en om de onzekerheid over de werkelijke grondwaterstand te kwantificeren. De nauwkeurigheid van de voorspellingen is onderzocht door middel van kruisvalidatie. Bij kruisvalidatie wordt telkens een lokatie waar een grondwaterstandsreeks is waargenomen buiten beschouwing gelaten. Met behulp van de overige gegevens wordt een grondwaterstandsreeks voorspeld voor deze lokatie. De voorspellingen worden vervolgens vergeleken met de waargenomen grondwaterstanden. Deze procedure wordt net zolang herhaald totdat voor alle lokaties waar grondwaterstanden zijn waargenomen ook voorspellingen zijn gedaan. De verschillen tussen de waargenomen en voorspelde grondwaterstanden zeggen iets over de nauwkeurigheid van een voorspellingsmethode. Uit de validatieresultaten blijkt dat er grote systematische fouten kunnen voorkomen in DM, die in IM aanmerkelijk kunnen worden gereduceerd. De fouten in de voorspelling van de temporele fluctuatie zijn relatief klein; gemiddeld wordt 64 % van de temporele variatie verklaard. Met behulp van IM is een kaart gemaakt die het risico weergeeft dat in een toekomstig jaar een kritische diepte aan het begin van het groeiseizoen wordt overschreden. Hierbij zijn als componenten van onzekerheid onderscheiden: i) onzekerheid over het neerslagoverschot in een toekomstig jaar, en ii) onzekerheid over de relatie tussen het neerslagoverschot en de grondwaterstand zoals die beschreven is door het RARX-model. Hoe groter de onzekerheid, hoe dichter risico’s bij 50 % zullen liggen en hoe moeilijker het is voor grondwaterbeheerders om een beslissing te nemen. Kaartjes met de componenten van onzekerheid laten de gebieden zien waar in dat geval de voorspellingen mogelijk te verbeteren zijn, bijvoorbeeld door aanvullende waarnemingen te verrichten.

In *hoofdstuk 6* wordt de nauwkeurigheid van een aantal methoden om de grondwaterstand in ruimte en tijd te voorspellen met elkaar vergeleken. Naast de directe en de indirecte methode die in hoofdstuk 5 werden besproken, zijn dit methoden die zijn gebaseerd op een zogenaamd ruimte-tijd Kalmanfilter. Met een ruimte-tijd Kalmanfilter is het mogelijk om voorspellingen te optimaliseren met behulp van de beschikbare waarnemingen van de grondwaterstand. Daarnaast kan het ruimte-tijd Kalmanfilter worden gebruikt om grondwaterstanden te simuleren, gegeven de waarnemingen. Verder kunnen met het ruimte-tijd Kalmanfilter monitoringnetwerken worden geoptimaliseerd. In tegenstelling tot DM en IM dienen bij het ruimte-tijd

Samenvatting

Kalmanfilter de RARX-parameters op een aantal lokaties bekend te zijn. In het studiegebied zijn de RARX-parameters bekend voor 13 lokaties waar tijdreeksen beschikbaar zijn die voldoende lang zijn om de parameters te kalibreren. Vervolgens worden deze ‘bekende’ RARX-parameters ruimtelijk geïnterpoleerd, waarbij gebruik wordt gemaakt van hulpinformatie. Drie varianten zijn onderscheiden: KF1, KF2 en KF3. In KF1 en KF2 wordt dezelfde hulpinformatie gebruikt als in DM en IM, namelijk waarden voor RARX-parameters die zijn gegist uit hulpinformatie zoals de digitale hoogtecijferkaart, de digitale topografische kaart met lokaties en afmetingen van sloten, digitaal opgeslagen bodemkundige profielbeschrijvingen en bodemfysische standaardcurven. Het verschil tussen KF1 en KF2 is gebaseerd op de manier waarop de RARX-parameters ruimtelijk worden geïnterpoleerd met behulp van die hulpinformatie. In KF1 wordt gebruik gemaakt van *kriging with an external drift*, waarbij gelijktijdig met de interpolatie een lineaire samenhang tussen de doelvariabele en de hulpvariabele wordt aangepast. In KF2 wordt gebruik gemaakt van *kriging with varying means*, waarbij de gegiste RARX-parameterwaarden worden beschouwd als het gemiddelde niveau en de afwijkingen ten opzichte van dit gemiddelde worden geïnterpoleerd. In KF3 worden digitale maaiveldhoogtecijfers gebruikt als *external drift* in *kriging with an external drift*.

De nauwkeurigheid van DM, IM, KF1, KF2 en KF3 is met elkaar vergeleken door middel van kruisvalidatie. Hieruit blijkt dat de KF-methoden de temporele fluctuatie van de grondwaterstand preciezer voorspellen dan DM en IM. Dit is het gevolg van het feit dat de voorspellingen worden geoptimaliseerd met behulp van beschikbare grondwaterstandswaarnemingen (*updating*). Het gemiddelde grondwaterstandsniveau blijkt echter door IM beter geschat te worden. IM blijkt een nauwkeurig alternatief voor de drie KF-methoden te zijn bij het schatten van de grondwaterstand die wordt verwacht op een bepaalde dag in enig toekomstig jaar, gegeven de heersende hydrologische en klimaatomstandigheden.

Tenslotte is in hoofdstuk 6 de nauwkeurigheid onderzocht waarmee de onzekerheid over de werkelijke grondwaterstand is gekwantificeerd. Het blijkt dat de onzekerheid over de temporele variatie van de grondwaterstand vooral in IM wordt onderschat. Dit kan worden verklaard uit het feit dat op 27 lokaties de waarnemingen in de tijd geclusterd waren, waardoor de temporele variatie van de voorspelfout wordt onderschat. De onderschatting van de onzekerheid over de temporele variatie van de grondwaterstand die bij KF-methoden werd geconstateerd, kan worden verklaard uit het feit dat de onzekerheid over de geregionaliseerde RARX-parameters niet in rekening is gebracht. De standaardafwijking van de foutencomponent van het RARX-model wordt onderschat omdat de 13 lokaties van waaruit geïnterpoleerd wordt niet representatief zijn voor het gehele studiegebied en de *kriging*-interpolatiemethoden een te ‘glad’ oppervlak voorspellen. De nauwkeurigheid van de geschatte systematische fout in de grondwaterstandsvoorspellingen kon niet goed worden bepaald vanwege het steekproefontwerp waarmee de grondwaterstand in te tijd was waargenomen. Als de grondwaterstanden waren waargenomen volgens een aselechte steekproefmethode was het wel goed mogelijk geweest om de nauwkeurigheid van de systematische fout te bepalen.

In *hoofdstuk 7* worden de resultaten van de studies uit hoofdstuk 2 tot en met 6 bediscussieerd en worden de belangrijkste conclusies samengevat. Ook worden er een aantal ideeën voor vervolgonderzoek gegeven. Om de fluctuatie van de grondwaterstand te karakteriseren voor de heersende hydrologische en klimatologische omstandigheden worden modellen gebruikt die de samenhang tussen het neerslagoverschot en de grondwaterstand beschrijven. De TFN-modellen die in hoofdstuk 2 worden toegepast worden ook wel *black box*-modellen genoemd, omdat zij slechts een empirische relatie beschrijven tussen de invoerreeks van het neerslagoverschot en uitvoerreeks van de grondwaterstand. Fysisch-mechanistische modellen zoals SWA-TRE worden daarentegen *white box*-modellen genoemd, omdat zij gebaseerd zijn op kennis van het fysische systeem. Een fysisch-mechanistisch model zal echter nooit volledig het grondwaterstandsverloop kunnen verklaren, omdat fysische processen in het model zijn geschematiseerd en omdat invoergegevens fouten bevatten. Zodra parameters van een fysisch-mechanistisch model worden gekalibreerd op waargenomen grondwaterstanden, kunnen deze parameters inboeten aan fysische betekenis omdat zij een tegenwicht moeten bieden tegen schematisaties en fouten in de invoer. Zodra een fysisch-mechanistisch model gekalibreerd is, is het dus geen *white box*-model meer maar een *grey box*-model, dat gebaseerd is op zowel fysische kennis als waargenomen tijdreeksen. De TARSO-modellen die in hoofdstuk 3 worden beschreven kunnen ook worden beschouwd als *grey box*-modellen, omdat de set kandidaatmodellen die is gebruikt in de automatische modelselectieprocedure is samengesteld op basis van fysische kennis. Bovendien kunnen de geselecteerde modellen fysisch worden uitgelegd.

In hoofdstuk 2 werden TFN-modellen gevonden op basis van een procedure van identificatie, kalibratie en verificatie. Bij de identificatie en verificatie speelt de visuele interpretatie van geschatte autocorrelatiefuncties en kruiscorrelatiefuncties een belangrijke rol. Een nadeel van deze methode is dat zij niet geheel reproduceerbaar is. Dit in tegenstelling tot de procedure op basis van een automatisch selectiecriteria (BIC) die in hoofdstuk 3 is gevolgd bij de selectie van TARSO-modellen. Toch vallen bij de samenstelling van de set kandidaatmodellen waaruit met behulp van het criterium een model wordt geselecteerd min of meer subjectieve keuzes niet te vermijden. Wordt de set kandidaatmodellen namelijk te groot, dan neemt het risico toe dat er een model wordt geselecteerd dat weliswaar goed bij de waarnemingen past, maar het onderliggende proces niet goed beschrijft. Dit risico kan worden verkleind door de set kandidaatmodellen te beperken tot modellen die op basis van bestaande fysische kennis relevant zijn. Als de set kandidaatmodellen echter te drastisch wordt beperkt op basis van bestaande inzichten in de fysische processen, dan verdwijnt de mogelijkheid om met behulp van de resultaten van de modelselectie nieuwe inzichten in het fysische systeem op te doen.

Bij verificatie of *diagnostic checking* wordt gecontroleerd of het model goed bij de waarnemingen past en of aan alle modelveronderstellingen wordt beantwoord. Dit is een correcte wetenschappelijke benadering, maar het kan nuttig zijn om ook de doelen van het model in gedachten te houden. Uit de verificatie kan bijvoorbeeld blijken dat in de residuen correlaties voorkomen die vanuit theoretisch oogpunt ontoelaatbaar zijn. Als de residuen echter gering zijn gelet op het doel van het model, dan kan het model nog steeds geschikt zijn voor toepassing in de praktijk. Door de prestatie

Samenvatting

van het model te testen op een onafhankelijke validatieset, zoals in hoofdstuk 2 en 3, kan de geschiktheid voor praktische toepassingen worden beoordeeld. In hoofdstuk 4 is een model voor de relatie tussen het neerslagoverschot en de grondwaterstand niet geselecteerd, maar *gekozen* op basis van een fysische analyse. In feite is de set kandidaatmodellen beperkt tot één. Een eenvoudig eerste-orde autoregressief model met het neerslagoverschot als invoervariabele (een ARX(1,0)-model) blijkt eenvoudig in termen van de waterbalans voor een bodemkolom te kunnen worden vertaald.

In hoofdstuk 5 en 6 is een geregionaliseerde tijdreeksmodel toegepast om de grondwaterstand in ruimte en tijd te voorspellen. Dit geregionaliseerde tijdreeksmodel is een combinatie van tijdreeksmodellering en geostatistische interpolatie. Het is ook mogelijk om met behulp van multivariate geostatistische methoden of multivariate tijdreeksmodellen de grondwaterstand in ruimte en tijd te modelleren. Het nadeel van deze twee benaderingen is echter dat er veel waarnemingen van de grondwaterstand nodig zijn om de modelparameters te schatten, in tegenstelling tot de geregionaliseerde tijdreeksmodellen die in hoofdstuk 5 en 6 zijn toegepast.

Tijdens het werk aan dit proefschrift rezen een aantal vragen die niet direkt konden worden beantwoord, maar die wel interessant zijn om in de toekomst te worden onderzocht. Deze vragen hebben betrekking op verbeteringen of alternatieven voor de methodieken die in dit proefschrift zijn toegepast, op verbeteringen of aanvullingen aan de gegevensbestanden die zijn gebruikt en op de mogelijkheden om de resultaten van stochastische methodieken beter te benutten in het grondwaterbeheer en het grondwaterbeleid.

Naast de modellen die zijn gebruikt in hoofdstuk 2 en 3 om de relatie tussen het neerslagoverschot en de grondwaterstand te beschrijven zijn er nog een aantal alternatieven die de moeite waard zijn om te onderzoeken. Een probleem is dat de grondwaterstand niet constant varieert met het neerslagoverschot: hoge grondwaterstanden variëren meer dan lage, omdat het bergingsvermogen van de bodem bij hoge grondwaterstanden kleiner is. GARCH-modellen (*Generalised Autoregressive Conditional Heteroscedasticity*) zijn tijdreeksen die met deze niet-constante variatie rekening kunnen houden. Het is misschien de moeite waard om te onderzoeken of GARCH-modellen tot betere resultaten leiden dan de modellen die tot nu toe zijn toegepast. De TARSO-modellen uit hoofdstuk 3 houden ook rekening met niet-constante variatie door het grondwaterstandsverloop in verschillende regimes te verdelen. Als de grondwaterstand een bepaalde drempelwaarde passeert dan treedt er een ander regime in werking. Met MARS-modellen (*Multivariate Adaptive Regression Splines*) is het mogelijk om deze overgang van het ene naar het andere regime geleidelijker te laten verlopen, wat wellicht meer in overeenstemming is met de werkelijkheid dan de plotselinge overschakeling op een ander regime bij de TARSO-modellen. Uit de validatieresultaten in hoofdstuk 2 en 3 bleek dat er systematische fouten voorkwamen in de gesimuleerde grondwaterstanden. Mogelijk hangen deze systematische fouten samen met lange-termijn schommelingen in het weer die vertraagd in het grondwaterstandsverloop tot uiting komen. Een andere oorzaak kan zijn dat de grondwaterstand daalt doordat er meer gedraineerd wordt of meer grondwater wordt gewonnen. Met modellen die zijn gebaseerd op zogenaamde neurale netwerken (ANN, *artificial neural*

networks) is het mogelijk gebleken om in Burkina Faso grondwaterstandsreeksen te modelleren die geleidelijk steeds dieper werden doordat de neerslag afnam en de waterwinning toenam. Wellicht kunnen ANN-modellen ook worden toegepast op reeksen in gebieden in Nederland waar de grondwaterstand geleidelijk verandert.

In hoofdstuk 6 bleek dat de nauwkeurigheid waarmee onzekerheid kan worden gekwantificeerd afhangt van het steekproefontwerp waarmee verschillen tussen gemeten en voorspelde grondwaterstanden worden waargenomen. Dit is een interessant en belangrijk onderwerp voor verder onderzoek, met name de toepassingsmogelijkheden van klassieke steekproeftheorie, zowel in ruimte als in tijd. Niet alleen de steekproef- of monitoringontwerpen en de modellen voor grondwaterstanden zijn belangrijk voor de nauwkeurigheid van schattingen en voorspellingen, maar ook de kwaliteit van de waarnemingen. Voor de toepassing van de methoden die in dit proefschrift zijn beschreven is het van belang om over neerslag- en verdampingscijfers van goede kwaliteit te beschikken. Over de nauwkeurigheid van grondwaterstandsmetingen in buizen en boorgaten is weinig bekend. Het is belangrijk om hier onderzoek naar te doen, zodat eventuele meetfouten in rekening kunnen worden gebracht bij het kwantificeren van onzekerheid. Met name het meten van de grondwaterstand in bodems waarin slecht doorlatende lagen voorkomen, zoals keileem, is een probleem. De tijdelijke grondwaterspiegels die op deze stagnerende lagen voorkomen zijn relevant voor de plantengroei, en dienen daarom gemeten te worden. Wellicht dient in dit soort bodems de grondwaterstand te worden waargenomen in niet een maar een aantal buizen met filters op verschillende dieptes, zowel boven als onder stagnerende lagen. Nieuwe meetapparatuur biedt mogelijkheden om de grondwaterstand nauwkeuriger te meten en onzekerheid beter te kwantificeren. Met name grondradar en zogenaamde elektronische *divers*, waarmee het mogelijk is om de grondwaterstand continu te meten in resp. de ruimte en de tijd, bieden nieuwe mogelijkheden.

Tenslotte is het belangrijk om de benutting van kennis over onzekerheid, die met stochastische methoden kan worden opgedaan, te verbeteren in het grondwaterbeheer en het grondwaterbeleid. Dit is een belangrijk onderwerp voor toekomstig onderzoek, waarin deskundigen op uiteenlopende gebieden zoals hydrologie, statistiek, bestuurskunde, communicatiewetenschappen en psychologie vruchtbaar zouden kunnen samenwerken. Uiteindelijk is onzekerheid een menselijke eigenschap.

References

- Anderson Jr., W.P., D.G. Evans and S.W. Snyder, 2000. The effects of Holocene barrier-island evolution on water-table elevations, Hatteras Island, North Carolina, USA. *Hydrogeology Journal* **8**: 390-404.
- Angulo, J.M., M.C. Bueso and F.J. Alonso, 2000. A study on sampling design for optimal prediction of space-time stochastic processes. *Stochastic Environmental Research and Risk Assessment* **14**(6): 412-427.
- Astatkie, T., D.G. Watts and W.E. Watt, 1997. Nested threshold autoregressive (NeTAR) models. *International Journal of Forecasting* **13**: 105-116.
- Bannink, M.H., B.J. Bles and A.F. van Holst, 1985. *A survey of soil and agrohydrology in eastern Noord-Brabant, Map 51 East* (in Dutch). Stiboka Rapport 1777, Wageningen, 97 pp.
- Belmans, C., J.G. Wesseling and R.A. Feddes, 1983. Simulation of the water balance of a cropped soil: SWATRE. *Journal of Hydrology* **63**: 271-286.
- Bierkens, M.F.P., 1996. Modeling hydraulic conductivity of a complex confining layer at various spatial scales. *Water Resources Research* **32**(8): 2369-2382.
- Bierkens, M.F.P., 1998. Modeling water table fluctuations by means of a stochastic differential equation. *Water Resources Research* **34**(10): 2485-2499.
- Bierkens, M.F.P., 2001. Spatio-temporal modelling of the soil water balance using a stochastic model and soil profile descriptions. *Geoderma* (accepted).
- Bierkens, M.F.P. and D.J.J. Walvoort, 1998. Simple stochastic models for fluctuations of the water table depth. Part 2: A combined soil-groundwater model with stochastic input (in Dutch). *Stromingen* **4**(3): 5-20.
- Bierkens, M.F.P., M. Knotters and F.C. Van Geer, 1999. Calibration of transfer function-noise models to sparsely or irregularly observed time series. *Water Resources Research* **35**(6): 1741-1750.
- Bierkens, M.F.P., M. Knotters and T. Hoogland, 2001. Space-time modeling of water table depth using a regionalized time series model and the Kalman filter. *Water Resources Research* **37**(5): 1277-1290.
- Bohlin, T., 1991. *Interactive system identification: prospects and pitfalls*. Communications and control engineering series. Springer, Berlin, 365 pp.
- Boucneau, G., M. van Meirvenne, J. Desmet and G. Hofman, 1996a. A methodology to evaluate the reliability of the Belgian soil map for predicting the actual water table characteristics. *Geoderma* **69**: 193-207.
- Boucneau, G., M. van Meirvenne, J. Desmet and G. Hofman, 1996b. Regressive modelling of soil water table fluctuations. *Mathematics and Computers in Simulation* **42**: 271-277.
- Box, G.E.P. and G.M. Jenkins, 1976. *Time series analysis: Forecasting and control*. Revised edition. Holden-Day, San Francisco, 575 pp.
- Brus, D.J., 1994. Improving design-based estimation of spatially means by soil map stratification. A case study of phosphate saturation. *Geoderma* **62**: 233-246.
- Brus, D.J. and J.J. de Gruijter, 1997. Random sampling or geostatistical modelling? Choosing between design-based and model-based sampling strategies for soil (with discussion). *Geoderma* **80**: 1-44.
- Chatfield, C., 1989. *The Analysis of Time Series. An Introduction*. Fourth Edition. Chapman and Hall, London, 241 pp.

References

- Cochran, W.G., 1977. *Sampling techniques*. John Wiley and Sons, New York, 428 pp.
- Coulibaly, P., F. Anctil, R. Aravena and B.Bobée, 2001. Artificial neural network modeling of water table depth fluctuations. *Water Resources Research* **37**(4): 885-896.
- Dalezios, N.R., K. Adamowski, 1995. Spatio-temporal precipitation modelling in rural watersheds. *Hydrological Sciences -Journal- des Sciences Hydrologiques* **40**(5): 553-568.
- De Bakker, H. and J. Schelling, 1989. *A system of soil classification for the Netherlands, the higher levels* (in Dutch). Revised edition. Pudoc, Wageningen, 209 pp.
- De Bruin, H.A.R., 1987. From Penman to Makkink. In: J.C. Hooghart (Editor), *Evaporation and weather*. CHO-TNO Mededeling 39, Den Haag, 98 pp.
- De Bruin, P.T. and J.G. De Gooijer, 1998. *A comparison of ARMA and SETAR forecasts*. Amsterdam, Institute of Actuarial Science and Econometrics, University of Amsterdam Report AE 6/98, Amsterdam, 18 pp.
- De Gooijer, J.G., B. Abraham, A. Gould and L. Robinson, 1985. Methods for determining the order of an autoregressive- moving average process: a review. *International Statistical Review* **53**(3): 301-329.
- De Gruijter, J.J., 2000. *Sampling for spatial inventory and monitoring of natural resources*. Alterra Report 070, Wageningen, 90 pp.
- De Haan, M.W.A., 1992. *The characteristics of duration curves of some ground-water dependent vegetations of the Littorelletea, Isoeto-Nanajuncetea, Orycocio-Sphagnetetea en Scheuchzerietetea* (in Dutch). KIWA Rapport SWE92.015, Nieuwegein, 85 pp.
- Deutsch, C.V. and A.G. Journel, 1992. *GSLIB. Geostatistical Software Library and User's Guide*. Oxford University Press, New York, 369 pp.
- Ernst, L.F., 1956. Calculation of the steady flow of groundwater in vertical cross-sections. *Netherlands Journal of Agricultural Science* **4**: 126-131.
- FAO, 1988. FAO-UNESCO Soil map of the world, revised legend. World Soil Resources Report 60, FAO, Rome. Reprinted as Technical Paper 20, ISRIC, Wageningen, 109 pp.
- Feddes, R.A., 1987. Crop factors in relation to Makkink reference-crop evapotranspiration. In: J.C. Hooghart (Editor), *Evaporation and weather*. CHO-TNO Mededeling 39, Den Haag, 98 pp.
- Finke, P.A., 2000. Updating the (1:50,000) Dutch groundwater table class map by statistical methods: an analysis of quality versus cost. *Geoderma* **97**: 329-350.
- Franses, P.H. and D. van Dijk, 1996. Forecasting stock market volatility using (non-linear) Garch models. *Journal of Forecasting* **15**(3): 229-235.
- Gehrels, J.C., 1999. *Groundwater level fluctuations, separation of natural from anthropogenic influences and determination of groundwater recharge in the Veluwe area, the Netherlands*. Ph.D. Thesis Free University Amsterdam, 269 pp.
- Gehrels, J.C., F.C. Van Geer, and J.J. De Vries, 1994. Decomposition of groundwater level fluctuations using transfer modeling in an area with shallow to deep unsaturated zones, *Journal of Hydrology* **157**: 105-138.
- Genstat 5 Committee, 1993. *Genstat 5 Release 3 Reference Manual*. Clarendon Press, Oxford, 796 pp.
- Goovaerts, P., 1997. *Geostatistics for natural resources evaluation*. Oxford University

-
- Press, New York, 483 pp.
- Hangos, K.M., (Editor), 1995. *International Journal of Adaptive Control and Signal Processing*, Vol. 9: Special Issue on Grey Box Modelling. John Wiley and Sons, New York.
- Hipel, K.W., W.C. Lennox, T.E. Unny and A.I. McLeod, 1975. Intervention analysis in water resources. *Water Resources Research* **11**(6): 855-861.
- Hipel, K.W. and A.I. McLeod, 1994. *Time series modelling of water resources and environmental systems*. Elsevier, Amsterdam, 1013 pp.
- Hutchinson, M.F., 1995. Stochastic space-time weather models from ground-based data. *Agricultural and Forest Meteorology* **73**: 237-264.
- IMSL, 1982. *IMSL Reference Manual*. Houston.
- Jarque, C.M. and A.K. Bera, 1980. Efficient tests for normality, homoscedasticity and serial independence of regression residuals. *Economics Letters* **6**: 255-259.
- Kendall, M.G. and W.R. Buckland, 1982. *A dictionary of statistical terms*. Longman Group Ltd., London, 213 pp.
- Knotters, M., D.J. Brus and J.H. Oude Voshaar, 1995. A comparison of kriging, co-kriging and kriging combined with regression for spatial interpolation of horizon depth with censored observations. *Geoderma* **67**: 227-246.
- Knotters, M. and M.F.P. Bierkens, 1998. *The relationship between time series models for water table depth and physical information*. DLO Winand Staring Centre Report 167, Wageningen, 112 pp.
- Knotters, M. and P.E.V. van Walsum, 1994. *Elimination of weather influences in the characterisation of the water table fluctuation* (in Dutch). SC-DLO Rapport 350, Wageningen, 122 pp.
- Kreiss, J.P., 1990. Bootstrap procedure for $AR(\infty)$ -processes. In: K.-H. Jöckel, G. Rothe and W. Sendler (Editors), *Bootstrapping and related techniques, Proceedings of an International Conference, 4-8 June 1990, Trier*. Springer Verlag, Berlin, pp. 107-113.
- Krul, W.F.J.M., 1952. A map of water table depths in The Netherlands (in Dutch). *Verslagen technische bijeenkomsten CHO-TNO*, pp. 85-93.
- Kyriakidis, P.C. and A.G. Journel, 1999. Geostatistical Space-Time Models: A Review. *Mathematical Geology* **31**(6): 651-684.
- Lankester, J. and K. Maas, 1996. A research to characterize sites on the basis of impulse responses for ecological purposes (in Dutch). *Stromingen* **2**(3): 5-17.
- Lewis, P.A.W. and J.G. Stevens, 1991. Nonlinear modeling of time series using multivariate adaptive regression splines (MARS). *Journal of the American Statistical Association* **86**(416): 864-877.
- Lewis, P.A.W. and B.K. Ray, 1997. Modeling long-range dependence, nonlinearity, and periodic phenomena in sea surface temperatures using TSMARS. *Journal of the American Statistical Association* **92**(439): 881-893.
- Ljung, G.M. and G.E.P. Box, 1978. On a measure of lack of fit in time series models. *Biometrika* **65**(2): 297-303.
- Marquardt, D.W., 1963. An algorithm for least squares estimation of nonlinear parameters. *Journal of the Society of Industrial and Applied Mathematics* **11**: 431-441.
- McLeod, A.I. and K.W. Hipel, 1978. Simulation procedures for Box-Jenkins models. *Water Resources Research* **14**(5): 969-975.

References

- Mew Jr., H.E., M.A. Medina Jr., R.C. Heath, K.H. Reckhow and T.L. Jacobs, 1997. Cost-effective monitoring strategies to estimate mean water table depth. *Ground Water* **35**(6): 1089-1096.
- Muñoz-Reinoso, J.C., 2001. Vegetation changes and groundwater abstraction in SW Doñana, Spain. *Journal of Hydrology* **242**: 197-209.
- Paparoditis, E., 1990. Bootstrapping some statistics useful in identifying ARMA models. In: K.-H. Jöckel, G. Rothe and W. Sendler (Editors), *Bootstrapping and related techniques, Proceedings of an International Conference, 4-8 June 1990, Trier*. Springer Verlag, Berlin, pp. 115-119.
- Parlange, M.B., G.G. Katul, R.H. Cuenca, M.L. Kavvas, D.R. Nielsen and M. Mata, 1992. Physical basis for a time series model of soil water content. *Water Resources Research* **28**(9): 2437-2446.
- Pfeifer, P.E. and S.J. Deutsch, 1980. A three-stage iterative procedure for space-time modeling. *Technometrics* **22**(1): 35-47.
- Price, L.E., P. Goodwill, P.C. Young and J.S. Rowan, 2000. A data-based mechanistic modelling (DBM) approach to understanding dynamic sediment transmission through Wyresdale Park Reservoir, Lancashire, UK. *Hydrological Processes* **14**(1): 63-78.
- Rouhani, S. and T.J. Hall, 1989. Space-time kriging of groundwater data. In: M. Armstrong (Editor), *Geostatistics, Vol. 2*. Kluwer Academic Publishers, Dordrecht, pp. 639-650.
- Rudloff, W., 1981. *World-Climates: with tables of climatic data and practical suggestions*. Wissenschaftliche Verlagsgesellschaft, Stuttgart, 632 pp.
- Runhaar J., 1999. *Impact of hydrological changes on nature conservation areas in the Netherlands*. Ph.D. Thesis Leiden University, 190 pp.
- Salas, J.D., and R.A. Smith, 1981. Physical basis of stochastic models of annual flows. *Water Resources Research* **17**(2): 428-430.
- Scheweppe, F.C., 1973. *Uncertain Dynamic Systems*. Prentice-Hall Inc., Englewood Cliffs, New Jersey.
- Schönwiese, C.-D., J. Rapp, T. Fuchs and M. Denhard, 1993. *Klimatrend-atlas Europa 1891-1990*. Berichte des Zentrums für Umweltforschung Nr. 20, Frankfurt, 218 pp.
- Shih, S.F., J.A. Doolittle, D.L. Myhre and G.W. Schellentrager, 1986. Using radar for groundwater investigation. *Journal of Irrigation and Drainage Engineering* **112**(2): 110-118.
- Stolte, J., G.J. Veerman and M.C.S. Wopereis, 1992. *Manual soil physical measurements, version 2.0*. DLO Winand Staring Centre Technical Document 2, Wageningen, 49 pp.
- Stolte, J., J.J. Freijer, W. Bouten, C. Dirksen, J.M. Halbertsma, J.C. van Dam, J.A. Van den Berg, G.J. Veerman and J.H.M. Wösten, 1994. Comparison of six methods to determine unsaturated soil hydraulic conductivity. *Soil Science Society of America Journal* **58**(6): 1596-1603.
- Sun, G., H. Riekerk and L.V. Kornhak, 1995. Shallow groundwater table dynamics of Cypress wetland/Pine upland systems in Florida flatwoods. *Soil and Crop Science Society of Florida Proceedings* **54**: 66-71.
- Sun, G., H. Riekerk and L.V. Kornhak, 2000. Ground-water-table rise after harvesting on Cypress-Pine flatwoods in Florida. *Wetlands* **20**(1): 101-112.

-
- Tankersley, C.D., W.D. Graham and K. Hatfield, 1993. Comparison of univariate and transfer function models of groundwater fluctuations. *Water Resources Research* **29**(10): 3517-3533.
- Tankersley, C.D. and W.D. Graham, 1994. Development of an optimal control system for maintaining minimum groundwater levels. *Water Resources Research* **30**(11): 3171-3181.
- Ten Cate, J.A.M., A.F. van Holst, H. Kleijer and J. Stolp, 1995. *Soil survey manual. Guidelines and regulations. Part B: Groundwater* (in Dutch). DLO Winand Staring Centre Technical Document 19B, Wageningen, 101 pp.
- Tong, H., 1990. *Non-Linear Time Series: A Dynamical System Approach*. Oxford University Press, 564 pp.
- Tsay, R.S., 1989. Testing and modeling threshold autoregressive processes. *Journal of the American Statistical Association* **84**(405): 231-240.
- Van Bracht, M.J., 1989. *OLGA (On Line Groundwater Archive). An information system for groundwater data in the Netherlands*. TNO Institute of Applied Geoscience Report PN 89-14-A, Delft, 10 pp.
- Van Dam, J.C., 2000. *Field-scale water flow and solute transport, SWAP model concepts, parameter estimation, and case studies*. Ph.D. Thesis Wageningen University, 167 pp.
- Van der Sluijs, P. and J.J. de Gruijter, 1985. Water table classes: a method to describe seasonal fluctuation and duration of water-table classes on Dutch soil maps. *Agricultural Water Management* **10**: 109-125.
- Van Dodewaard, E., 1997. *Composition of the soil in the rural development areas Zuidwolde-Zuid, Beneden-Egge and Zuidwolde-Noord: results of a soil survey* (in Dutch). DLO-Winand Staring Centre Rapport 157, Wageningen, 136 pp.
- Van Geer, F.C. and P.R. Defize, 1987. Detection of natural and artificial causes of groundwater fluctuations. The influence of climate change and climatic variability on the hydrologic regime and water resources. *IAHS Publ.* **168**: 597-606.
- Van Geer, F.C., C.B.M. te Stroet and Z. Yangxiao, 1991. Using Kalman filtering to improve and quantify the uncertainty of numerical groundwater simulations 1. The role of system noise and its calibration. *Water Resources Research* **27**(8): 1987-1994.
- Van Geer, F.C. and A.F. Zuur, 1997. An extension of Box-Jenkins transfer/noise models for spatial interpolation of groundwater head series. *Journal of Hydrology* **192**: 65-80.
- Van Heesen, H.C., 1970. Presentation of the seasonal fluctuation of the water table on soil maps. *Geoderma* **4**: 257-278.
- Visser, W.C., 1958. *The agrohydrology of The Netherlands* (in Dutch). COLN-TNO, Delft, 159 pp.
- Wesseling, J.G., 1991. *CAPSEV. Steady state moisture flow theory. Program description. User manual*. DLO-Winand Staring Centre Report 37, Wageningen, 51 pp.
- Winter, T.C., D.O. Rosenberry and A.M. Sturrock, 1995. Evaluation of 11 equations for determining evapotranspiration for a small lake in the north central United States. *Water Resources Research* **31**(4): 983-993.
- Wong, C.S. and W.K. Li, 1998. A note on the corrected Akaike information criterion

References

- for threshold autoregressive models. *Journal of Time Series Analysis* **19**(1): 113-124.
- Wösten, J.H.M., M.H. Bannink and J. Beuving, 1987. *Water retention and hydraulic conductivity characteristics of topsoils and subsoils in the Netherlands: the Staring series* (in Dutch). Institute for Land and Water Management Research (ICW) Report 18, Wageningen, 73 pp.
- Young, P.C. and K.J. Beven, 1994. Data-based mechanistic modelling and the rainfall-flow non-linearity. *Environmetrics* **5**: 335-363.

Curriculum vitae

Martin Knotters werd geboren op 27 maart 1963 in Schoonhoven (ZH). Op zijn vijfde wist hij al zeker dat hij later accordeonist zou worden. Wat ging er mis?

In 1980 behaalde hij het diploma Hoger Algemeen Voortgezet Onderwijs aan de reformatorische scholengemeenschap ‘De Driestar’ te Gouda. In datzelfde jaar stond hij voor de keuze naar het Voorbereidend Wetenschappelijk Onderwijs door te stromen, naar het conservatorium te gaan of cultuurtechniek te studeren aan de Hogere Bosbouw en Cultuurtechnische School (HBCS) in Velp. Hij koos voor het laatste. In 1984 studeerde hij af aan de HBCS met als vakken onder meer hydrologie, waterbeheersing en bodemkunde. Direct na zijn militaire diensttijd trad hij in dienst bij de Stichting voor Bodemkartering, dat later met enkele andere instituten opging in het DLO-Staring Centrum. Met de onderbreking van een jaar waarin hij werkzaam was bij de Landinrichtingsdienst in Utrecht maakte hij tot 1991 bodemkaarten. Aangemoedigd door collega’s en door een PHLO-cursus Onderzoekstechniek kreeg hij steeds meer interesse in de toepassing van statistiek in het bodemkundig en hydrologisch onderzoek, wat in 1995 resulteerde in zijn eerste publicatie in een internationaal wetenschappelijk tijdschrift. In die periode ontstond ook het idee om aan een promotieonderzoek te beginnen op het gebied van de voorspelling van de grondwaterstand in ruimte en tijd. Tegenwoordig is Martin als wetenschappelijk onderzoeker werkzaam bij Alterra in Wageningen, waar hij zich hoofdzakelijk met de onzekerheid over de grondwaterstand bezighoudt. Zijn passie is muziek.