# Groundwater

through

# System

Time Series Analysis

# Identification

Jos von Asmuth

# Groundwater System Identification
### through
# Time Series Analysis

motto:

*Kees gaf de ruimte,*
*Jos nam de tijd*

# Groundwater System Identification
### through
# Time Series Analysis

Proefschrift

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft;
op gezag van de Rector Magnificus Prof. ir. K.C.A.M. Luyben;
voorzitter van het College voor Promoties

in het openbaar te verdedigen op maandag 5 maart 2012 om 15.00 uur

door Jost-René VON ASMUTH

ingenieur in de biologie
geboren te Eindhoven

# Acknowledgements

As a long time has passed since the start of this research, people have come and gone, and contributed to different phases of it. There is, coming to think of it, only one person that was really there at the start, and will be so at the finish. That person, of course, is Kees Maas. Kees, your farewell speech at KWR made me realize again how much our work has been interwoven over the last years, also for you. You asked me and KWR to treat *Menyanthes* as your baby too, and to take good care of it. To me, however, *Menyanthes* is more an attempt to show the world the power and beauty of the equations and ideas, which I think are your true babies (or crown jewels, as I once called them). Through these babies, one may see the beauty of the world itself, which perhaps is a necessity for an ecologist to devote so much time to them. Thanks for all the things I could learn from you, and for your unquestioning support in better and lesser times. At this point, I must also not forget André Jansen, as it was his impassioned plea (and way of acting later on) that convinced me of taking on this task in the first place.

My thanks is also due to (the rest of) my advisory board, who guided me in my first years. These are Ab Grootjans, Evert-Jan Lammerts, Ella de Hullu, Jan Hoogendoorn, and of course Cees van den Akker, initially my promotor, and Marc Bierkens, who is my second promotor. Whereas the subject of this thesis is partly physically oriented, the other part has its basis in stochastics and statistics, and is much indebted to Marc's guidance and inspiration. Again, I want to express my apologies to the ecologists in my advisory board (including myself), for the lack of ecology in my research. I hope that in spite of this, the ecological applications and cooperation in later (and hopefully also coming) years can partly make up for this. My thanks also go to Harry Rolf of the Provincial water supply company of North-Holland, who initiated me in his 'down to earth' viewpoint on ARMA time series models, which results from many years of experience.

Next, I want to thank my colleagues at both the Delft University of Technology and KIWA or KWR Watercycle Research Institute for their cooperation and support, including Ed Veling, Frank Smits, Huub Savenije, Hanneke de Jong, Christophe Obergfell, Gijsbert Cirkel, Jan Willem Kooiman and Arthur Meuleman. This also includes Arjen Kruithof, Otto de Keizer, Danneke Bakker, Vanessa Sternitzke, Marieke Eeuwe, Geerte van der Meijden, Harm Brinkhof and Arend Terluin, who contributed as MSc students to my work. Theo Olsthoorn, thanks for the flexibility in adopting me as a Ph.D. Student, after the superannuation of Cees van den Akker, and for your enthousiasm and support on every level. Mark Bakker, thanks for the fruitful cooperation and the unfailing acuteness of your judgment (and the joy whenever Ajax gets beaten!). Peter Hesen, thanks for cooperating, having the patience and sharing the pleasure in making a businesslike success out of *Menyanthes* too. In this respect, also the support and faith of Ron van Megen, whenever it was needed, was highly valued and has been crucial. Frans Schaars, thanks for teaching me all the tips and tricks, and primarily the fun in using Matlab. Hopefully, we may again join forces one

day. Inke Leunk, thanks a.o. for the firm way in which you kept *Menyanthes* and its helpdesk going, especially during my leave. Although not being a colleague literally, I also want to thank Martin Knotters for his colleague-like cooperation through the years (and unparalleled sense of scientific humor). Thomas de Meij, I consider your assignment that allowed me to spend another seven months in Delft and finish this thesis, to be sort of a miracle. Thanks for thinking along the way you did (at this and it seems all other occasions).

Logically, these final words are for my friends and family. How naïve can one be? Many years ago, I started this research with the intention to treat it as just another job, experienced as project manager and determined to work disciplined. Why then go into the trouble of extensively thanking those that are closest, as all other thesis acknowledgements seem to do? But clichés are of course never without cause, and perhaps it is impossible to avoid a struggle in one way or another when pushing oneself to the limit, in relative isolation and over such a long period as in a Ph.D. research. Thanks to everyone who supported me and my little family in this period! I hope to find everybody still alive, when I soon finish this chapter in my life and look out the window again. Special thanks go to Laurens and Wendy, for the beautiful representation of an 'impulse response' on the cover. Thanks, Daniel, for cooperating in *Menyanthes* as well as being a brother. Miep, you are the best grand parent in the world! And I end here with deeply thanking my wife Karen, and my kids Hannah and Abel, for everything. Hopefully, in near future, I can somehow and somewhat correct the image that I am married to my research and have *Menyanthes* as a child.

Woerden, January 17th 2012                                             Jos von Asmuth

# Contents

# Symbols and notation

In this thesis, the symbols and notation of related literature is followed where possible. On the other hand, the equations in this thesis should of course be unambiguous, every symbol having a clear and unique meaning. In some cases, both objectives proved to be incompatible and some overlap is allowed. Examples are $\delta$ for both the Dirac delta function and autoregressive parameters of ARMA models, and $r$ for both radius and recharge. Where this is the case, the meaning of symbols is clear from or clarified in the text.

| Space and time | | |
|---|---|---|
| $t$ | : | time, either discrete ($t \in \mathbb{N}$, [-]) or continuous ($t \in \mathbb{R}$, [T]) |
| $Y$ | : | year [T] |
| $D$ | : | Julian day [T] |
| $\xi$ | : | frequency [T$^{-1}$] |
| $\eta$ | : | phase shift [T] |
| $\lambda$ | : | amplitude |
| $xyz$ | : | Cartesian coordinates [L] |
| $r$ | : | radius [L] |
| $L$ | : | length [L] |
| $A$ | : | area [L$^2$] |
| $C$ | : | relative position or centrality [-] |

| Operators and special functions | | |
|---|---|---|
| $\Sigma$ | : | summation operator |
| $\Pi$ | : | product operator |
| E | : | expectation operator |
| B | : | backward shift operator |
| d | : | differential operator |
| $\partial$ | : | partial differential operator |
| $\Delta$ | : | difference operator |
| $\nabla$ | : | gradient vector |
| $\Gamma$ | : | gamma function |
| $K_0$ | : | modified Bessel function of the second kind and zeroth order |
| $\delta(t)$ | : | Dirac delta function |
| $H(t)$ | : | Heaviside step function |

| Statistics | | |
|---|---|---|
| $N$ | : | number |
| $O$ | : | set of observations |
| $\Psi$ | : | set of parameters |
| $P$ | : | probability |
| $\Lambda$ | : | log likelihood |
| $J$ | : | Jacobian matrix |
| $C$ | : | covariance matrix |
| $R^2$ | : | coefficient of determination |
| $S^2$ | : | sum of weighted squared innovations |
| $\overline{x}$ | : | average of $x$ |
| $\tilde{x}$ | : | deviates of $x$ from $\overline{x}$ |
| $\hat{x}$ | : | prediction of $x$ (deterministic, or in time update) |
| $\check{x}$ | : | prediction of $x$ (in measurement update) |
| $M_n$ | : | temporal moment of order $n$ |
| $\overset{c}{M}_n$ | : | centralized temporal moment |
| $\mu$ | : | mean |
| $\sigma^2$ | : | variance |
| $\overset{1}{\gamma}$ | : | skewness |
| $\overset{2}{\gamma}$ | : | kurtosis |

**ARMA approach**

$y_t$ : explained variable

$x_t$ : explanatory variable

$n_t$ : error or residual series [L]

$a_t$ : discrete white noise process [L]

$\Theta_t$ : transfer function

$\Phi_t$ : noise response function

$\omega$ : transfer moving average parameter

$\delta$ : transfer autoregressive parameter

$\theta$ : noise moving average parameter

$\varphi$ : noise autoregressive parameter

$nr$ : number of $\omega$'s in an ARMA model

$ns$ : number of $\delta$'s in an ARMA model

$np$ : number of $\theta$'s in an ARMA model

$nq$ : number of $\varphi$'s in an ARMA model

$b$ : dead or delay time [-]

**PIRFICT approach**

$\theta(t)$ : impulse response function

$\Theta(t)$ : block response function

$\Omega(t)$ : step response function

$\Xi(t)$ : frequency response function

$\phi(t)$ : noise impulse response function

$W(t)$ : continuous white noise (Wiener) process [L]

$\nu(t)$ : innovation series [L]

$f$ : evaporation factor [-]

$d$ : local drainage level [L]

$A, a, b, n, \alpha, \beta, \gamma, \upsilon$ : parameters in various response functions

**Physical processes**

$h(t)$ : groundwater head [L]

$p(t)$ : precipitation intensity [$LT^{-1}$]

$P_t$ : precipitation amount [L]

$e(t)$ : evaporation [$LT^{-1}$]

$q(t)$ : specific discharge [$LT^{-1}$]

$r(t)$ : specific recharge [$LT^{-1}$]

$w(t)$ : groundwater withdrawal [$L^3T^{-1}$]

$s(t)$ : surface water level [L]

$b(t)$ : barometric pressure [L]

$m(t)$ : hydrologic interventions [-]

$\vartheta(t)$ : volumetric water content [-]

$R_i(t)$ : value of stress $i$

**Physical parameters**

$c$ : drainage resistance [T]

$S$ : storativity [-]

$H$ : saturated thickness [L]

$K$ : permeability [$LT^{-1}$]

$KH$ : transmissivity [$L^2T^{-1}$]

$D$ : effective diffusion-dispersion coefficient [$L^2 T^{-1}$]

$v$ : effective propagation velocity [$LT^{-1}$]

# Chapter

# 1

# General introduction

## 1.1    Initial objectives and scope

As perhaps is not uncommon in PhD research, the initial objectives differ rather strongly from the final results presented in here. In the original proposals [*Maas, 1995; Von Asmuth,* 2000], the primary objective was to improve eco-hydrologic modeling methods with respect to the way in which the relationship between groundwater dynamics and the vegetation composition in groundwater dependent ecosystems was modeled. At that time but also to date, common



*figure* **1.1**: *Overview of the method of impulse response moments, showing the three fields of modeling from which moments can be derived and be mutually exchanged [Von Asmuth and Maas, 2001].*

methods to present and characterize groundwater level fluctuations are the so-called duration lines [e.g.,*Tüxen,* 1954*; Grootjans,* 1985*; De Haan,* 1992], regime curves and overall characteristics like MxGL statistics [*Van der Sluijs and De Gruijter,* 1985]. Such methods, however, have two important drawbacks:

- Statistics based on groundwater level series over limited periods are sensitive to long term variation in the driving forces [*Knotters and Van Walsum,* 1997*; Bartholomeus et al.,* 2008].
- Graphs or lines are essentially vectors, and less convenient for data storage, (spatial) presentation, analysis and modeling purposes than scalars. Overall statistics, on the other hand, only capture certain aspects of the dynamics [*Von Asmuth and Knotters,* 2004].

Because of these drawbacks, it was hypothesized that an alternative route based on the so-called impulse response function (see section 2.3.4) could improve ecohydrologic modeling methods. In this alternative, a key notion is that spatial differences in groundwater level dynamics are mainly determined by spatially variable system properties, while temporal dynamics are mainly driven by spatially less variable meteorologic dynamics. Consequently, it was hypothesized that spatial differences in vegetation could be modeled more accurately using system properties alone, or in other words by 'filtering out' temporal, meteorologic dynamics. To be more specific, the use of time series models was proposed for inferring impulse response functions from series of groundwater level observations, and in turn to characterize these by their moments (see section 2.3.4). Moments are scalars and constants in linear, time-invariant systems, and together with the spatially less variable driving forces, they completely characterize the (deterministic part of the) dynamics at a certain location. As an additional advantage, moments can also be simulated directly and spatially using distributed groundwater models (see section 2.4.3).
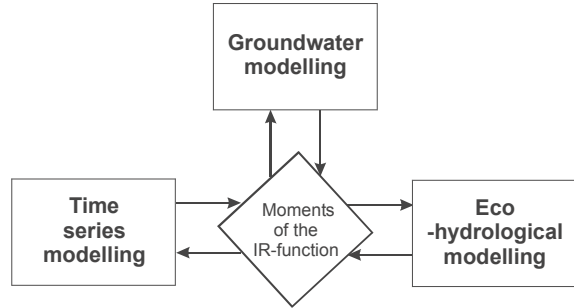
All in all, the initial scope of this research encompassed three fields of modeling, i.e. time series, groundwater and eco-hydrologic modeling. These fields of modeling together constitute the method of impulse response moments, from which moments can be derived and be mutually exchanged (figure 1.1). The three 'leaves' in this research led to naming the computer program which was to be developed after *Menyanthes trifoliata*, the scientific name of the Marsh trefoil or Bogbean. As the natural habitat of *Menyanthes trifoliata* is on the verge of ground and surface water, as it is threatened by anthropogenic influences and also is one of the wonderful flowers in such areas, it serves as the natural ambassador for the initial objectives of this research.

## 1.2    Contributions and outline of thesis

Although it was envisaged that the link between time series models and groundwater models would be a subject in its own right, at the time the existing ARMA time series models were thought to be well developed and be usable without modification. As it turns out, however, most of the contributions of this thesis are in the field of time series analysis. The main reasons for this are:

- The practical limitations of ARMA models described in section 2.2.3. These are especially problematic when having to process large numbers of time series of irregular frequency, as would have been necessary considering the initial objectives of this research.

- From a continuous time viewpoint, the transfer functions of ARMA models are block responses (see chapter 3). Inferring moments of continuous time impulse responses using ARMA models at least would have required a transformation of their results and the development of a continuous time framework.

- As envisaged in [*Maas,* 1995], making moments from time series models comparable to those generated by a groundwater model is not straightforward. This is a.o. due to a difference in scale and due to effects of the unsaturated zone. This of course requires insight and work on the groundwater modeling part, but also on the methods, accuracy and principles of inferring moments from time series.

- In principle, the possible number of parameters in an ARMA transfer function is infinite (see section 2.2.1). Consequently, ARMA transfer functions cannot be determined uniquely from a limited number of moments. This is required when aiming to create time series models from moments generated by a groundwater model, as was envisaged in the method of impulse response moments.

- Moments are not, or not always, uniquely and directly usable for eco-hydrologic modeling (see section 8.5). The methods developed in chapter 6 can serve as an alternative, but require a direct link between moments and

the actual groundwater level fluctuations, which in turn requires the use of a response function with a limited number of parameters (see previous point).

What followed from these ascertainments was a process of improving and adapting time series analysis methods to the needs set by the initial objectives of this research (but not limited to that). As such, one of the contributions of this thesis is that it brings together different viewpoints, methods and aspects related to (geohydrologic) time series and time series analysis. In an introductory fashion, these different viewpoints and methods are treated in the next sections (the systems approach in section 2.1, statistical time series analysis in section 2.2, the physical viewpoint in section 2.3, use of distribution functions in section 2.4). The views and methods presented are united in a consistent, continuous time framework, together referred to as *groundwater system identification*. By combining aspects of all methods, a new method of time series analysis was developed and named the PIRFICT method. This method has several advantages over the ARMA method. Perhaps most importantly, it allows for the application of time series analysis methods on a large scale and in a standardized, less knowledge and labor intensive fashion. Within the PIRFICT method, physically-based, analytic solutions to geohydrologic problems can be used, which enhances the physical insight in the model. Distribution functions of skew-Gaussian nature, like the scaled gamma distribution, prove to fit the behavior of a wide range of systems quite well. As for application in practice, an important step is that a user friendly computer program named *Menyanthes* was developed, which contains most of the methods presented. By using the methods and tools developed, the effects, impulse responses and moments of different driving forces can be effectively inferred from a given data set and be separated. Simply put, this thesis provides methods and tools that are better put up to the original tasks set for it.

In the next chapters, the different steps taken and peer reviewed papers published[1] are presented. In chapter 3, the PIRFICT method of time series analysis is presented. Its practical advantages are discussed, and its structure and performance are compared to those of ARMA models. In chapter 4, the noise part of the PIRFICT model is presented and its methods and criteria for parameter optimization are discussed. Furthermore, it is compared to (the pure prediction form of) the Kalman filter, to which it is related. In chapter 5, the methods behind the PIRFICT model are extended to cover the complex situations that occur in real world, where multiple factors generally influence groundwater heads simultaneously. In this chapter, also the modeling procedure and methods for interpreting and checking model results are discussed. In chapter 6, the relationship between moments and groundwater level characteristics is discussed, by regarding the dynamics and response of systems in the frequency domain. It is shown that the commonly used MxGL-statistics have some important drawbacks, as they filter out the low-frequency dynamics of a system and mix-up annual with higher frequencies. In chapter 7, the computer program *Menyanthes* is presented and a synthesis of the methods developed and results

---

[1] Some relatively minor changes were made to the text that was published. First, the symbols and notation were made uniform and are now summarized at the beginning of this thesis. Also, the terminology was harmonized.

obtained is given, while specifically focusing on the interface between data-based and physically-based methods. Finally, chapter 8 contains the general conclusions on and summary of the methods and results presented in this thesis.

**Abstract:**

Appreciating the contents of this thesis requires some knowledge of its background, i.e. the methods and theory on (geohydrologic) time series analysis and system identification. In this chapter, a basic introduction is given, and different views on the subject are presented. First, some thought is given to the systems approach in general, as to date, most (geo)hydrologist are less familiar with system identification methods than with e.g., spatially-distributed groundwater models. In short, the system viewpoint can be characterized by the fact that in essence it is top-down. The system viewpoint treats a groundwater system first as a 'whole' and not bottom-up, as an aggregate of cells, layers and/or elements, which is still the mainstream viewpoint. Having said this, also from the systems perspective, there are different ways in which a groundwater system, or any system for that matter, can be perceived, modeled and/or analyzed.

Time series analysis is a method that originates from the statistical sciences. In principle, it does not require any knowledge of the physical functioning of the system under consideration. In its basics, it can be seen as a variant of simple, linear regression, and the coefficients in the regression equation, either autoregressive or moving average parameters, do not have a physical meaning *a priori*. From a physical point of view, on the other hand, a central concept is the so-called impulse response function, as it completely characterizes the functioning of a linear time-invariant system at a certain point in space. Impulse response functions can be inferred from a data set through time series analysis, but also using 'purely' physically-based methods, either analytic or numeric (in case of linear systems). This means that impulse response functions can also be derived from the differential equation and boundary conditions that belong to a certain geohydrologic system and its schematization.

In this thesis, a 'mix' between both worlds is developed and presented. In this approach, the time series analysis problem is formulated in a continuous time domain. It allows for the use of (continuous) distribution functions that have a statistical origin, as well as physically-based analytic response functions. Distribution functions of skew-Gaussian nature, like the scaled gamma distribution, prove to fit the behavior of a wide range of systems quite well. Next to that, a further link between the physically-based world of groundwater modeling and time series analysis is established using moments of impulse response functions, as these can also be generated directly and spatially using moment-generating differential equations, implementable in any standard groundwater model.

# Chapter 2

# General background, methods and theory

## 2.1 Facets of the system approach

### 2.1.1 On the system concept

As it is a widely used term in and outside the scientific community, 'system' can have many different meanings. Originally, system stems from the Greek word σύστημα, which translates as 'a whole compounded of several parts or members' or 'composition' [*Perseus Digital Library*]. Being a central concept in some disciplines, there have been several attempts to give a formal definition of system and related terms [see e.g., *Marchal,* 1975; *Willems,* 1991; *Backlund,* 2000; *Klir,* 2001]. From his review, Backlund concludes that previous

*figure 2.1: Schematic representation of an open, physical system. The boundary separates the system from the environment, with which it exchanges matter, energy and/or information.*

definitions of system are imprecise. He argues that only via a definition based on the constituting parts and their formal interrelationships, it is possible to unambiguously define what is part of a system and what is not (e.g., what are external stimuli), or what are actually two systems in stead of one. In short, only such a definition would define what is a system, and what is not. For our purposes, however, we prefer a more general usage of the term, as does [*Ljung,* 1999] who defines a system in loose terms as '…an object in which variables of different kinds interact and produce observable signals'. In a physical or thermodynamic context, system has a technical meaning and simply refers to the portion of the physical universe chosen for the analysis. What is inside and outside a system is not fixed by some objective, formal definition, as attempted by Backlund, but is in contrast a free choice, generally made to simplify the analysis. Everything outside the system is then referred to as the environment, which is ignored in the analysis except for its effects on the system. Albeit the fact that it is a very general concept, physical systems share common characteristics. In general, systems for instance have (figure 2.1):

- a boundary (defining what is part of a system and what is environment)
- structure (defined by the internal components and their composition, determining the operation)
- behavior (involving the (linear or non-linear) response to inputs of matter, energy and/or information)

In a system identification or time series analysis context, it is common practice to denote forcing variables as 'input' and forced variables as 'output'. Variables related to systems, however, do not necessarily represent something physically entering it or coming out. The 'output' may for instance well be the evolution of a certain
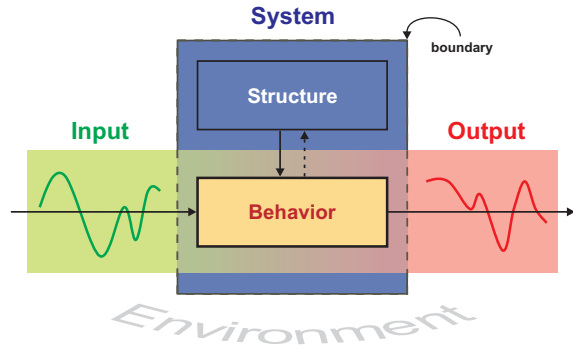
parameter or 'state variable' that describes a condition in or of the system itself. In simple mechanical systems, position coordinates and their derivatives are typical state variables. Knowing these and the external forces, it is possible to determine their future state (which is a central concept in the so-called state space representation of dynamical systems).  Other terms that are used to denote input and output are independent and dependent variables, forcing and forced, stresses, excitations and stimuli for the input, or explanatory and explained variables in a general statistical context. Furthermore, in stead of behavior, terms like operation, processing, transfer or response are used. The diversity in terminology reflects the diversity in sciences in which theory related to systems has been parallelly applied and developed (see section 2.1.4). See also [*Willems,* 1991] for a more general, formal and mathematical framework of systems theory, in which behavior is a central concept and variables are either latent or manifest (i.e. forcing or forced), analogous to the usage of these terms in other sciences. Behavior, however, is used by Willems in a more general sense, as the ensemble of input, response and output. Here we choose to use behavior to denote the way in which the systems responds or 'behaves' when it is stimulated. Depending on their relationships with the environment, systems can be classified as:

- open       (exchange of energy and matter)
- closed     (exchange of energy, not matter)
- isolated   (no exchange)

Systems can furthermore be:

- time-varying
- time-invariant

depending on whether or not their structural properties change over time. In case of linearity, this means that their response or transfer function is time-invariant also. This is not be confused with the issue of stationarity, which is a property of a signal or a stochastic process. Stationarity indicates whether different observations have the same joint probability distribution (strong or strict stationarity) and/or the same mean and covariance (weak stationarity). A non-linear system, for instance, can be time-invariant but at the same time harbor a non-stationary process. Last but not least, the operation of systems can be (approximated as being):

- linear
- non-linear

In the first case, the output is simply a sum of the effects of all individual events and inputs, whereas in the latter case, it is not. The issue of linearity has large practical implications, as linearity allows for the use of powerful linear algebra. Although there are solutions for modeling non-linear groundwater systems (e.g., the threshold or TARSO approach, which is also incorporated in the computer program *Menyanthes* described in chapter 7), this thesis exclusively deals with linear (or linearized) systems.

**HYDROLOGIC SIMULATION**

**SYSTEM INVESTIGATION**

**PHYSICAL HYDROLOGY**

**PHYSICAL MODELS**

**PHYSICALLY-BASED MATHEMATICAL METHODS**
Boundary-value problems using partial differential equations and potential theory.

**PARAMETRIC METHODS**
Statiscal treatment of systems with deterministic inputs and outputs

**STOCHASTIC METHODS**
Statiscal treatment of systems on the basis of the variables, without reference to the system on which they operate.

**MATHEMATICAL MODELS**

**ANALOG COMPUTER SOLUTIONS**

**DIGITAL COMPUTER SOLUTIONS**

**PHYSICALLY-BASED DIGITALLY - SIMULATED HYDROLOGIC RESPONSE MODEL**

**Spatially and Sequentially**

**LUMPED**

**DISTRIBUTED**

*figure* **2.2:** *Blueprint for a physically-based, digitally-simulated hydrologic response model [Blue line, Freeze and Harlan, 1969, in part after Amorocho and Hart, 1964]. 'System investigation' is holded to be the opposite of 'physical hydrology'.*

In that case, the operation of a system in response to a single stimulus can be described by a single response function.

### 2.1.2 The ultimate system versus the elementary volume scale

To date, the mainstream approach for modeling the dynamics of hydrologic systems, or any physical system for that matter, is based on an appropriate partial differential equation for the process under consideration (see also section 2.3.2). Generally, the differential equation is based on continuum mechanics and applies to a so-called Representative Elementary Volume (REV). If no direct analytic solution is available, the system is modeled by dividing space into cells, elements and/or layers, and defining their dynamically relevant properties. After setting the initial and boundary conditions, the differential equation is then solved using the finite difference or finite element method. As a final step, the match between model predictions and available observations is optimized using some automated optimization routine. For hydrologic processes, this approach or 'blueprint for a physically-based, digitally-simulated hydrologic response model (figure 2.2)' has been laid down by [*Freeze and Harlan,* 1969]. Although the accomplishments of this approach are vast and undisputed, there also has been strong and increasing criticism. In short, the main problems with the REV-scale approach boil down to two points:

- First, complete physical determination of the properties of complex systems is not only a practical, but also a fundamental impossibility. To start with, the continuum approach already neglects the variability below the REV-scale, e.g., consisting of pores and grains in a soil, or molecules on an even lower scale. Laws and properties that apply to the REV-scale are therefore in essence some average or statistical outcome of the underlying lower scale processes, as described by statistical mechanics. Volume properties that are generally considered to be physical can in fact only be derived empirically. Next to that, methods for directly determining these properties are inherently invasive. When the REV-scale approach is applied to environmental systems, complete determination of all physical properties would mean mining out the entire region being modeled, as pointed out by, e.g., [*Oreskes et al.,* 1994]. For use in a model, the REV-scale physical properties that were determined at individual locations have to be extrapolated to the scale of the cells in the model, and those in turn to layers or other elements. In this step, based on the necessarily limited data, much of the heterogeneity above the REV-scale is also lost, and the remainder is the result of correlations, extrapolations and assumptions. As a consequence, these larger scale properties are almost never effectively derived from physical data alone. Alternatively, they are estimated as model parameters through calibration or inverse modeling. Because of the ubiquitous simplifications and errors in the model, generally the parameters thereby loose their direct physical meaning.
- Second and building on the first point, inverse modeling methods using observations on state variables are in most cases neither sufficient to uniquely determine the 'physical' properties or parameters. Generally there are more degrees of freedom in a model than that there is independent information in

the observations, and/or not every model parameter contributes meaningfully and independently to the solution. In technical terms, the inverse modeling problem is said to be mathematically ill-posed [*Hadamard,* 1902]. In light of this problem, there has been discussion on the amount of complexity that can be warranted in a model [e.g., *Jakeman and Hornberger,* 1993*; Young et al.,* 1996]. As a possible solution, there were calls for a revival of simpler, larger scale and/or more strongly data based models [e.g., *Young,* 1998*; Savenije,* 2001*; Kirchner,* 2006*; Kirchner,* 2009]. As models that start bottom-up at the REV-scale almost inherently invite the incorporation of all detail above it, it has lead some to the prophecy that the Freeze and Harlan blueprint will be abandoned in future, and/or to the design of alternative blueprints on larger aggregation scales [e.g., *Reggiani et al.,* 1998*; Reggiani et al.,* 1999*; Beven,* 2002]. Next to that, there have been pleas for treating multiple model structures and parameter sets as equally acceptable or equifinal (see next section).

The problems mentioned above have lead some to the somewhat philosophical conclusion that environmental models are in essence scientific hypotheses. Like these [*Popper,* 1959], they can not be verified or validated, only invalidated [*Konikow and Bredehoeft,* 1992*; Oreskes et al.,* 1994]. Being as it may, models are in practice indeed almost never invalidated, but only adjusted to better fit newly available data, which of course has also been a dispute with regard to scientific theories in general [*Kuhn,* 1962]. As opposed to the 'physical hydrology' or REV-scale approach of figure 2.2, which in essence is bottom-up, we can also approach a system top-down and firstly as a whole. This is perhaps a direct consequence of the definition and usage of the term system, and the 'holistic' nature of systems theory is often stressed (while trying to avoid the common and negative association of holism with metaphysics [e.g., *Von Bertalanffy,* 1950]). In figure 2.2, the term 'system investigation' primarily refers to unit hydrograph and linear theory of hydrologic systems [e.g., *Sherman,* 1932*; Nash,* 1958*; Dooge*, 1973], primarily applied to surface water hydrology. Such methods were common at the time, and they are highly comparable to the more modern, more general and statistically better developed fields of time series analysis and system identification (see section 2.1.4). Whereas [*Freeze and Harlan,* 1969] characterize their approach with the term 'physically-based', unit hydrograph and other large scale methods are often denoted by terms like 'lumped', 'conceptual', 'data-based' or 'black-box'. Response functions, however, may be equally directly derived from physical laws and principles (see section 2.3), whereas 'physically-based' models may be equally strongly calibrated on available data. We therefore consider the distinction made to be primarily an issue of scales, whereas by definition 'system' may be regarded to be the ultimate scale in a scientific problem. Furthermore, the modeling approach followed can start top-down from the system-scale, or bottom up from the REV(or any other)-scale. Also, models can be distinguished that only describe the temporal behavior of systems from those that (also) describe their spatial structure. An important advantage of the top-down system viewpoint is that it naturally leads to a parsimonious approach to modeling. Time series models come with relatively few assumptions, as in principle they do not require any information on the physical structure or properties of systems. Because of this, they are easy to construct and at

the same time they generally show a higher model fit than spatially distributed models. As both the system and REV-scale approach have their own advantages and limitations (see also section 2.2.3), their integration could be a means to advance hydrologic response modeling methods. We therefore consider the distinction made by [*Freeze and Harlan,* 1969] on the highest level of 'hydrologic simulation' to be non-fundamental or gradual, and this thesis aims to contribute in bridging the gap.

### 2.1.3    Simple behavior versus complex structure (equifinality)

At many occasions and in many different forms, scientists have noticed that different physical systems can exhibit behavior that is to a large extent comparable. In fact, different physical processes like conduction of heat, flow of water and electric currents are governed by laws that are remarkably equivalent in a mathematical sense. This equivalent behavior of physically very different systems has led to an attempt to formulate a science of systems in general at a higher level of abstraction, named general systems theory [*Von Bertalanffy,* 1950*; Skyttner,* 2005]. One of the outcomes of general systems theory that is still used today is the concept of equifinality. Von Bertalanffy defines equifinality as the fact that in open thermodynamic systems, the same final state can be reached irrespective of the initial conditions. In such cases, the final state is determined only by the inflow and outflow of the system and the constants of reaction. The term equifinality was adopted for the environmental modeling context by Keith Beven [e.g., *Beven,* 1993*; Beven,* 2006]. Equifinality is used by Beven in a broader sense, as the expectation that the same acceptable model prediction might be achieved in many different ways, i.e. with many different model structures or parameter sets. If the same behavior can be produced by systems of different structure, inversely the structure of a system can logically not be determined uniquely from its behavior. In other words, the notion of equifinality implies that the underlying, complex structure and properties of hydrologic systems can not be determined from time series of observations on state and forcing variables alone (given that the state can only be partially observed). Beven's primary conclusion out of the concept of equifinality is a plea against the assumption of a single optimal model and for allowing multiple representations of a system (inputs, model structures, parameter sets and errors) to be equally acceptable or 'behavioral', in the sense of being consistent with the observations. This approach forms the basis of the Generalized Likelihood Uncertainty Estimation (GLUE) methodology of [*Beven and Binley,* 1992]. Beven, however, also notices that hydrologic systems often show relatively simple response characteristics and that '…various parametric models with all their potential for equifinality and different process interpretations, are just different attempts to simulate the same response characteristics' [*Beven,* 2006, see also *Young et al.,* 1996]. He furthermore identifies the problem of model dimensionality reduction as an important area for further research.

Parallel to the above, also in the field of solute transport modeling it has been recognized that mere agreement between observed and predicted output is not sufficient to test the validity of a model [e.g., *Beck,* 1987; *Jury and Roth,* 1990*; Sánchez-Vila and Carrera,* 2004]. Jury and Roth consider it a curious phenomenon that physically-based laws like the convection-dispersion equation became commonly

accepted merely because experiments have shown that its solution could be manipulated into matching the shape of the outflow concentrations. In fact, [*Jury,* 1982] and also [*Maas,* 1994] go one step further, and suggest the use of statistical distribution functions for modeling the transfer of solutes through soil. When using such distribution functions (discussed are the lognormal, gamma and Pearson type III distribution) the temporal behavior of a system is explicitly defined, but the exact physical or spatial structure is not. In [*Jury and Roth,* 1990] it is shown that the shapes of the lognormal and gamma distributions have a striking similarity with that of an impulse response solution to the convection-dispersion equation (see also section 2.3.5.3). As a consequence, Jury and Roth state that the main difference between statistical distribution functions and the convection-dispersion equation is that in the latter, also the relationship between the shape of the transfer function and distance is defined. However, they also point out that many experiments have shown that the convection-dispersion equation is not completely adequate at this point, as the dispersion coefficient is not constant but increases with distance [see also the review by *Gelhar et al.,* 1985]. Jury and Maas independently suggested the use of moments to characterize the properties of transfer or response functions, as is common practice for distribution functions in statistics (and was also proposed for unit hydrographs in [*Nash,* 1959]). Although these authors do not refer to the transfer function noise models discussed in section 2.1.4, their choice for describing the temporal behavior of systems with functions in which no explicit physical structure is assumed, is in fact equivalent to what is common practice in the field of statistical time series analysis (see next sections).

### 2.1.4 Time series analysis, signal processing and system identification

The theory on dynamic systems can be seen as a branch of systems theory. As dynamic systems are ubiquitous in our (natural and man-made) environment, they form the subject of many different scientific disciplines. In spite of the large differences in the type of system addressed, their physical functioning, the available data and the problem to be solved, the theory and techniques used for analyzing and modeling collected time series data are, however, often remarkably similar. If we focus on the analysis of separate time series here, and exclude spatially-distributed models, we may distinguish two main approaches in dealing with and presenting the problem, in line with the general distinction made in sections 2.2 and 2.3. The viewpoint in the first approach is primarily statistical and the



*figure* **2.3**: *Back of the textbook by the 'founding fathers' of time series analysis, Box and Jenkins.*

emphasis lies on the properties of the data itself. The second approach in itself is more diverse, but it shares the notion of and stronger emphasis on the physical system that generated the observations.

In the first approach, generally referred to as *time series analysis*, a time series is simply defined as a set of observations of a certain variable that is arranged chronologically. As such, time series are primarily regarded as data sets and methods are presented for analyzing and modeling their statistical properties. The application of time series analysis methods developed rapidly after the publication of the comprehensive text book by Box and Jenkins [1970, figure 2.3]. The so-called ARIMA (AutoRegressive Integrated Moving Average) time series models described in this book, or models having a related structure, are still popular in many fields of science like statistics, econometrics or social sciences. Such sciences

*figure 2.4: Textbook on system identification by Ljung, former vice president of IFAC (International Federation Of Automatic Control).*

have in common that the behavior of the systems that are studied cannot (or cannot easily) be described by physical laws, and/or their physical properties are largely unknown. Moreover, it is often not the functioning of the system that is of primary interest, but the state and behavior of the variable itself. In econometrics, for instance, a classical problem rests in adequately predicting the course of currencies, stock prices and the like, in order to optimize the return on investments. The relationship between the explained variable and other variables like consumer trust may be of primary importance in that respect, because of its predictive capabilities. In typical time series literature [e.g., *Box and Jenkins,* 1970*; Hipel and McLeod,* 1994], much attention is paid to exploratory data analysis, model identification, parameter estimation, the modeling of errors and uncertainties and the diagnostic checking of model results. Furthermore, various types of statistical behavior of time series are identified (e.g., periodic, seasonal, long memory, series with interventions or external stimuli), and models are presented for dealing with each one of them. The underlying system and various physical processes generating the observations, however, generally receive little attention. An exception to this are studies on the physical justification of time series models and/or the use of physical knowledge for identifying the appropriate model order in certain cases [see e.g., *Klemes,* 1978*; Parlange et al.,* 1992*; Knotters and Bierkens,* 2000]. Furthermore, as digital records contain discrete observations, usually discrete-time mathematics are used and little reference is made to sampling or frequency issues, or to the often continuous time processes occurring in real world.

Second, there are the kindred disciplines of *system identification*, *control engineering*, *signal processing* and *filtering*, which, as stated before, have a stronger physical orientation. System identification, a term that has been coined by [*Zadeh,* 1956], deals
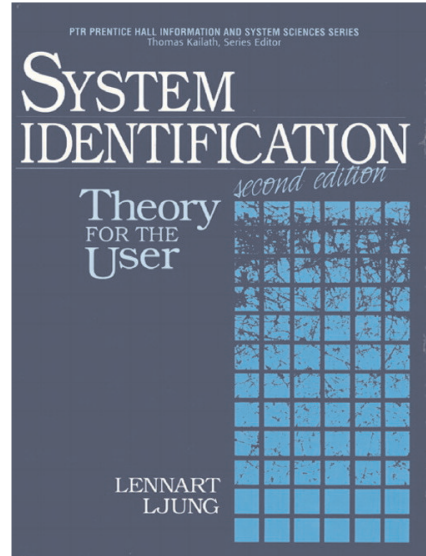
with the problem of building mathematical models of dynamic systems based on observed data. In general, system identification methods are applied when a purely physical model would be overly complex or impossible to obtain in reasonable time, due to the complex nature of many systems and processes. Foremost advocate of the system identification viewpoint is Lennart Ljung. Ljung is also main author of the perhaps leading (and one of the few) software tools for time series analysis, i.e. Matlab's system identification toolbox [*Ljung,* 1999, figure 2.4]. Control engineering or cybernetics, on the other hand, is a discipline that is very close to system identification. Control engineering not only deals with the behavior of dynamic systems, but an additional issue here is that systems also have a desired behavior or output, called the reference. An archetypical problem in control engineering is how to manipulate the inputs into a system, to



**figure 2.5:** *Textbook by Ziemer et al., representative of the signal processing viewpoint.*

obtain the desired effect on its output. In signal processing, as in time series analysis, the focus primarily lies on the signal itself. The term signal, however, implicitly acknowledges the presence of a system, and may be defined as a function of time that represents a physical variable associated to it. As such, signal processing is primarily an area of communication theory and electric engineering. In the theory on signals and systems [e.g., *Ziemer et al.,* 1998, figure 2.5], attention is for instance paid to several transformations and representations of signals (e.g., Fourier, Laplace, State-variable representation) and the difference in continuous and discrete-time systems comes naturally. Filtering, last but not least, can be seen as a branch of both system identification, signal processing and control engineering, as in all these disciplines separation of the useful signal from the corrupting noise is essential. Filtering in short can mean removing some frequencies from a signal and allowing others to pass, in order to suppress interfering signals. Here, the frequency domain representation and frequency response of a filter is a central issue. A second approach to filtering takes a statistic or stochastic viewpoint. In that approach, an optimal filter is sought whose output would come as close to the original signal as possible, by minimizing some error or performance criterion, or maximizing the likelihood. Well known examples are the Wiener and Kalman filter, of which the latter allows for the distinction between system and measurement noise and can also be used for adaptive filtering (see also section 2.2.2).

### 2.1.5 Hydrologic applications and physical insight

As discussed in the previous section, ARIMA-type time series analysis methods have been around for quite some time now. Although it took a while before the methods and principles were picked up in the hydrologic sciences, this was amply compensated by the publication of the comprehensive book by [*Hipel and McLeod,* 1994, figure 2.6]. In this book, not only several applications can be found, but also extensions to the ARIMA models of Box and Jenkins. Here, specific models are published and discussed for dealing with several aspects commonly found in hydrologic and environmental time series, like long memory systems or periodic influences. In the Netherlands, time series analysis methods were adopted relatively early and were applied on a relatively large scale, with an emphasis on groundwater level series [e.g., *Van Geer et al.,* 1988*; Rolf,* 1989*; Gehrels et al.,* 1994*; Van de Vliet et al.,* 2000]. More recently, attention is paid to the fact that the transfer function of

*figure 2.6: Comprehensive textbook by Hipel and McLeod, devoted to hydrologic and environmental time series.*

time series models, which describes the response of systems through time, should have a physical basis, at least when applied to physical phenomena like groundwater level fluctuations (this, however, was already one of the propositions in [*Van Geer,* 1987]). In the process of incorporating physical insight in the use of time series models, the following approaches can be distinguished:

- When applying ARMA models, the so-called model order or number of parameters in het model has to be defined (see section 2.2.1). As a first approach, which is called data-based mechanistic modeling, physical knowledge is used to select the correct model order using results from a set of candidates *a posteriori*. Here, selection is based on the question whether the results and transfer functions found are plausible from a physical-hydrologic perspective or not [*Young and Beven,* 1994*; Young, 1998; Price et al.,* 2000]. A simple example is precipitation, which of course should lead to a rise and not a lowering of the water level.
- A second approach is the reversal of this process. Here, the model order is selected *a priori* based on a physical-hydrologic analysis of the system. In their paper, [*Knotters and Bierkens,* 2000] selected the AR(1) model (which is equivalent to an exponential response function) for modeling groundwater level series, based on an analysis of the functioning of a simplified soil column (see also section 2.3.5.1).

- • As a third approach, in this thesis the statistically oriented ARMA model structure is abandoned and replaced by a model structure in which continuous time response functions are used (see chapter 3). By doing so, more complex, analytic solutions to geohydrologic problems can be used as response function. The physical functioning of a system, in response to a specific excitation, can in principle thus be imposed on the model. In the computer program *Menyanthes* (see chapter 7), Hantush' well formula is for instance used for modeling the effects of pumping (see section 2.3.5.5), but more general distribution functions may be used also (see section 2.4.1).

Because of the increasing physical orientation, also in the hydrologic sciences the focus thus shifts from time series analysis to system identification, or from analyzing the correlations in and between time series to identifying the key parameters in a system and the effects of excitations imposed to it. Time series analysis, as such, then becomes a tool in the more general area of system identification (analogous to the way it is presented in e.g., [*Dooge,* 1973]). In the next sections, first the statistical or black-box viewpoint on time series models is treated. Second, the physical viewpoint and response of several elementary groundwater systems are discussed. Third a 'mix' between these worlds is presented, which involves the use of (distribution) functions that do not have an apparent physical meaning or whose physical meaning is relaxed.

## 2.2    Statistical, 'black box' viewpoint

### 2.2.1    ARMA models as linear regression equation

From a statistical viewpoint, a time series model can be regarded as a variant of linear regression. The equation defining a simple regression line can be written as:

$$y = \omega x + \mu + n \tag{2.1}$$

where

$y$  : observations of the explained variable (e.g., groundwater head)

$x$  : observations of the explanatory variable (e.g., recharge)

$\omega$  : parameter defining the slope of the regression line

$\mu$  : parameter defining the intercept with the y-axis

$n$  : error or residual term

For now, we will ignore the residual term, on which we will elaborate in the next section. Of course, linear regression can be applied to all sorts of variables, including time series. In the latter case, the available observations are indexed with time $t$ [-], which gives the following dynamic regression equation:

$$y_t = \omega x_t + \mu \tag{2.2}$$

In many dynamic systems, however, the effect of one variable onto the other is not (or not only) instantaneous. In such cases, the system is said to have a memory. The rise in water level in an arbitrary hydrologic system caused by a rain shower, for instance, will not disappear immediately when the rain stops. Consequently and *vice versa*, the current water level in hydrologic systems is a function of present ánd previous rainfall events. Mathematically, the effects of previous states of the explanatory variable can be simply added to the regression equation in the following manner:
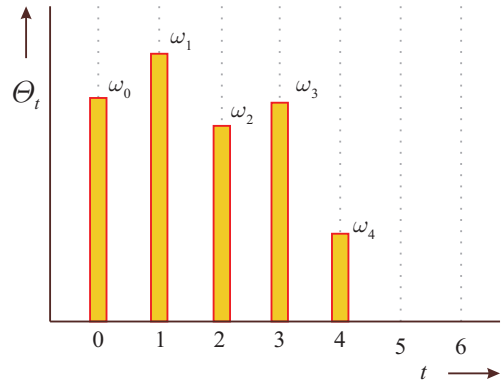


**figure 2.7:** *Transfer function of a hypothetical MA(4) system.*

$$y_t = \omega_0 x_t + \omega_1 x_{t-1} ..... + \omega_{ns} x_{t-ns} + \mu \tag{2.3}$$

The parameters $\omega_{ns}$ are known as moving average (MA) parameters (as (2.3) resembles the calculation of a weighted moving average of $x$), and the resulting model is said to have order MA($ns$). If we plot $\omega$ versus $ns$ (figure 2.7), the result

shows what is known as the transfer or impulse response function of the system under consideration, which we denote by $\Theta_t$. The name impulse response function stems from the fact that, after a sole 'impulse' of unit value of $x$ at time $t$, $y$ will respond with the subsequent values of $\omega$ if we move forth in time. Please note, however, that what is taken to be an impulse of $x_t$ in discrete-time notation, often is a mathematical abstraction of continuous time reality in which $x_t$ is perhaps a sample, time average or sum of $x(t)$. Because of that, we will avoid the term 'impulse response' when discussing discrete-time mathematics and use 'transfer function' in stead. In theory, the number of steps $ns$ with which we can look back in time in this way may be infinite. In practice, however, we cannot deduce an infinite number of parameters from a finite data set. Therefore, especially when the time steps are small as compared to the response time of the system, the direct use of equation (2.3) as a regression model has practical limitations.

A second option for dealing with memory is to model a variable as a function of its value at a previous time step, in the following manner:

$$y_t - \mu = \delta(y_{t-1} - \mu) + \omega x_t \tag{2.4}$$

Here, $\delta$ is known as an autoregressive (AR) parameter, and the model is referred to as having order AR(1). If we look at the transfer function of such a system (figure 2.8), we find that it shows exponential decay, given by:

$$\Theta_t = \omega \delta^t \tag{2.5}$$

Also here, we can choose to look further back in time and add more AR terms to equation (2.4). By doing so, the transfer function will take the shape of a mixture of exponentials (see section 2.3.1). Summarizing, a limitation of using MA terms in dynamic regression equations is the fact that they can only model the response of a system for a limited period back in time. An advantage of MA parameters is that they can take on any value, so the shape of the transfer function is free in the MA part. A limitation of using AR terms, on the other hand, lies in the fact that they limit



*figure 2.8: Transfer function of a hypothetical AR(1) system.*

the shape of the transfer function to one or more exponentials. A logical solution to this problem is to use both AR and MA terms in one model, which results in the general equation of a transfer function model of order $(nr, ns)$:

$$y_t - \mu - \delta_1(y_{t-1} - \mu) - ... - \delta_{nr}(y_{t-nr} - \mu) = \omega_0 x_t + \omega_1 x_{t-1} + ... + \omega_{ns} x_{t-ns} \qquad (2.6)$$

When we introduce the backward shift operator $\mathrm{B}$, defined by:

$$\mathrm{B}^b y_t = y_{t-b} \qquad (2.7)$$

and autoregressive and moving average operators of order $nr$ and $ns$, defined respectively by:

$$\begin{cases} \delta(\mathrm{B}) = 1 - \delta_1 \,\mathrm{B} - \delta_2 \,\mathrm{B}^2 - ... - \delta_{nr} \,\mathrm{B}^{nr} \\ \omega(\mathrm{B}) = 1 + \omega_1 \,\mathrm{B} + \omega_2 \,\mathrm{B}^2 + ... + \omega_{ns} \,\mathrm{B}^{ns} \end{cases} \qquad (2.8)$$

equation (2.6) can be written more economically as:

$$\delta(\mathrm{B})\tilde{y}_t = \omega(\mathrm{B})x_t \qquad (2.9)$$

or:

$$\tilde{y}_t = \delta^{-1}(\mathrm{B})\omega(\mathrm{B})x_t \qquad (2.10)$$

where $\tilde{y}_t$ are deviates of $y_t$ from $\mu$, and

$$\Theta(\mathrm{B}) = \delta^{-1}(\mathrm{B})\omega(\mathrm{B}) \qquad (2.11)$$

is the transfer function. When a system has a delayed response, or dead time of duration $b$, the transfer function becomes:

$$\Theta(\mathrm{B}) = \delta^{-1}(\mathrm{B})\omega(\mathrm{B})\mathrm{B}^b \qquad (2.12)$$
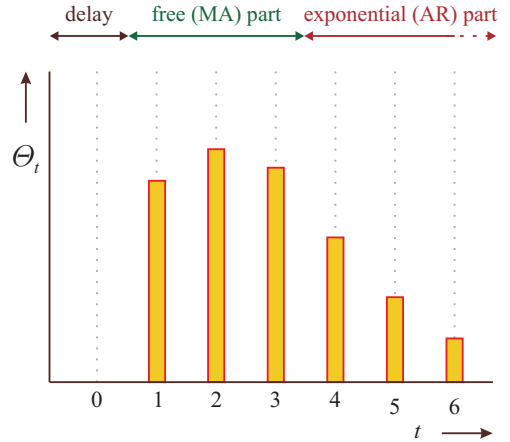


*figure 2.9: Transfer function of a hypothetical ARMA(1,2) system with delay 1.*

Alternatively, (2.10) may be written as a so-called discrete convolution product (see also section 2.3.2):

$$\tilde{y}_t = \sum_{i=-\infty}^{t} \Theta_{t-i} x_i \equiv \sum_{i=0}^{\infty} \Theta_i x_{t-i} \equiv \Theta(\mathrm{B})x_t \equiv (\Theta * x)_t \qquad (2.13)$$

where $\Theta_t$ is the transfer function (figure 2.9). ARMA transfer function models can serve as close approximations for many quite complicated dynamic systems. Apart from the ARMA structure, however, two basic assumptions are made. The first is the

fact that transfer function models assume that the system is linear, or can be linearized. In short, this implies that the effects of all pulses and explanatory variables can be added, regardless of the state of the explained variable. Second, the model assumes the transfer function to be time-invariant or remain constant over time.

### 2.2.2 Effects of noise and noise model

If we look at the real world, there are several reasons why mathematic equations will never be able to match observations from dynamic systems exactly. Observations are bound to be affected by noise, if only because of errors and uncertainties in the measurement process. One of the main problems in signal processing or system identification is to separate the signal or useful information from the noise or disturbances in a given time series. When creating a model of a dynamic system, possible sources of noise, each with a different character and effect, are in general (figure 2.10):

- Errors in input measurements
- Errors in model concept or parameters
- Unknown system disturbances
- Errors in output measurements

In model terms, the output of the (deterministic part of the) equations is called the model prediction. The differences between model predictions and the observed output form a time series of their own, called the 'residuals'. In general, a residual series cannot simply be modeled or taken to be a set of independent Gaussian deviates. System disturbances, model errors and errors in the input will, just as the input itself, have an effect that is not or not only instantaneous. Because effects linger on, the value of a model residual at a certain point in time will be correlated with its value at previous times. This phenomenon is called autocorrelation. Autocorrelation
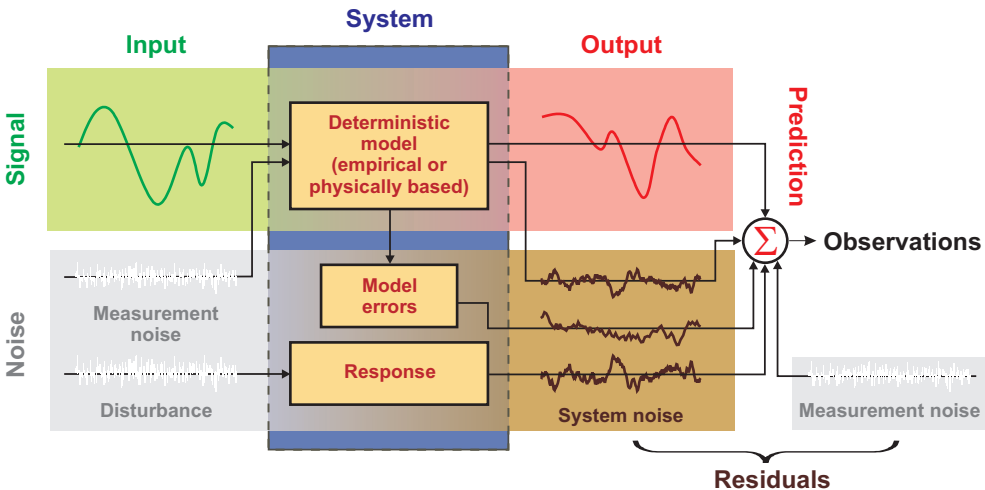


**figure 2.10:** *Sources and effects of noise in models of (linear, open) systems.*

causes individual residuals to be colored or have a lesser, non-unique information content, depending on the observation frequency and the response time of the system. If it is not properly taken into account, it causes parameter estimation algorithms to be inefficient and the parameter covariance matrix to be underestimated.

The fact that the response of a system introduces autocorrelation creates the possibility of regarding the residual series as the outcome of a stochastic process. Because the autocorrelation is accounted for by the (stochastic part of the) system, the stochastic input 'causing' the residuals can be modeled as uncorrelated or white noise (figure 2.11). The noise model used in traditional transfer function-noise (TFN) models follows this concept. Explicit modeling of the behavior of model residuals can have several purposes:

- First, when a noise model is fitted to the residuals, it can be used for stochastic simulation. This is especially useful in case one is interested in the probability of extremes, as such probabilities are underestimated when only the deterministic model is used [e.g., *Knotters and Van Walsum,* 1997]).
- Second, because of the autocorrelation in the signal, the noise model can also be used to yield predictions of the residuals at unobserved time steps, either for smoothing, forecasting or updating purposes. Such applications are widely used in the meteorologic sciences and are often referred to as data assimilation [*McLaughlin,* 1995*; Kalnay,* 2002]. They make optimal use of both model prediction and observations and can significantly improve the accuracy of the predictions of the combined model.
- Third, correlations between the input, predicted output and the residual series hamper an independent estimate of the model parameters, whereas autocorrelation in the residuals causes the variance of the parameters to be underestimated. 'Whitening' of the residuals with a noise model can thus improve estimates of the parameters of transfer and other deterministic models and their covariance [*Bryson and Henrikson,* 1965*; Te Stroet,* 1995].

In the ARMA-type TFN models as introduced by [*Box and Jenkins,* 1970], the sources that contribute to the noise in observations of the explained variable are modeled together as a single ARMA process. If the response of the system to noise is assumed to be the same as the deterministic response (e.g., situations where the effect of errors in the input series are dominant) equation (2.10) becomes:

$$\tilde{y}_t = \delta^{-1}(\mathrm{B})\omega(\mathrm{B})(x_t + a_t) \qquad\qquad (2.14)$$

where $a_t$ is white or uncorrelated noise which is assumed to be normally independently distributed with mean zero and variance $\sigma_a^2$, abbreviated as NID( $0, \sigma_a^2$ ). Several tests, however, are available and recommended for testing if these assumptions hold. Equation (2.14) is known as an ARMAX model (ARMA model with eXogenous variable). If however, the response to noise is different, and there is also a
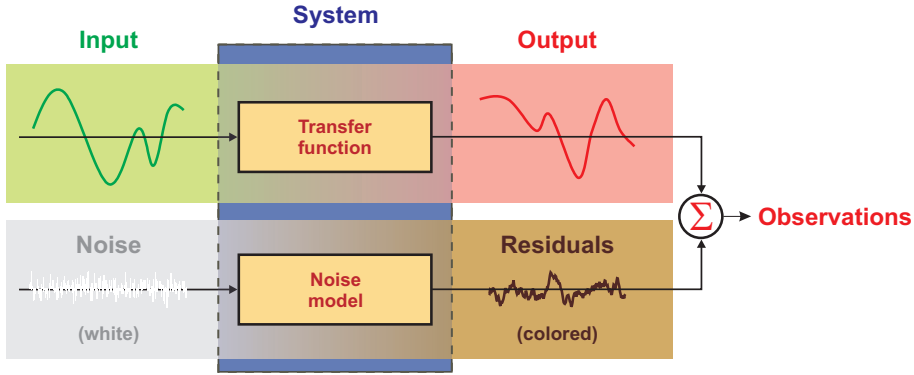
*figure* **2.11**: *Scheme of a combined single input, transfer function-noise model.*

delay, we arrive at the general equation of a combined transfer function and noise model:

$$\tilde{y}_t = \delta^{-1}(\mathrm{B})\omega(\mathrm{B})\mathrm{B}^b\, x_t + \varphi(\mathrm{B})^{-1}\theta(\mathrm{B})a_t \tag{2.15}$$

where $\varphi(\mathrm{B})$ and $\theta(B)$ are the autoregressive and moving average operators for the noise process, with orders $np$ and $nq$ respectively. A more general solution to this problem is provided by the so called Kalman filter [*Kalman,* 1960]. The Kalman filter allows for the distinction between system or correlated and measurement or uncorrelated noise, and can even be used for what is called adaptive filtering, when the system parameters are allowed to change over time. ARMA-type transfer functions and noise models can be embedded in a Kalman-filter [e.g., *Bierkens et al.,* 1999*; Berendrecht,* 2004], which serves as a solution to the limitation of ARMA-models of handling irregular data (see also next section).

## 2.2.3   Limitations of ARMA models

In spite of the general applicability, wide spread use and often accurate predictions of ARMA-type time series models, there are also several limitations to their use. Perhaps foremost important, the literature and theory on time series analysis have a strong statistical focus and use a specific mathematical notation, jargon and viewpoint. Consequently, it takes a long time for scientists, engineers and others who mostly received intensive training in their own discipline, but not in statistics let alone time series analysis, to understand the ins and outs. Once the theory is adopted, however, there are still several practical limitations and difficulties to overcome:

- First of all, there is the problem of identifying the correct model order. As follows from the theory addressed in short in sections 2.2.1 and 2.2.2, a time series modeler has to identify the so called model order or value of $[nr, ns, np, nq, b]$, i.e. the number of parameters in the autoregressive and moving average parts of the transfer and noise model respectively, and the delay time, before the actual parameter values can be estimated from the data. Next to that, a modeler has to apply a suitable degree of differencing to the data to induce stationarity on non-stationary series (e.g., series that do not remain in



figure 2.12: *Iterative model identification procedure as proposed by Box and Jenkins.*

equilibrium about a constant mean or show a trend). Also, an appropriate model frequency has to be selected in real time, as time in ARMA models is dimensionless. [*Box and Jenkins,* 1970] devised a specific iterative procedure for this task based on statistical criteria, known as the model identification procedure (figure 2.12). A general disadvantage of this procedure, however, is that its results can be ambiguous [*Hipel and McLeod,* 1994] and the process itself is rather heuristic and can be very knowledge and labor intensive [*De Gooijer et al.,* 1985]. Those who have experience with manual model calibration will probably also know to what task a modeler is up when he has to manually identify six or seven parameters at the same time from a given data set.

- Second, there is the matter of handling irregular or high-frequency data. In traditional time series literature, relatively little attention is paid to the fact that time is a continuous phenomenon. ARMA models are specified in discrete-time mathematics, and it is implicitly assumed that data are available without restrictions at the mostly unspecified but fixed frequency used in practice. For applications in for instance electro-engineering, this does not pose a serious problem, as the frequency of the measuring equipment can be easily adapted. In hydrologic practice however, the analyst will often have to make do with the often messy data available. Time series of groundwater levels, for example, are often collected manually and will tend to be non-equidistant and have missing records, as the observer will probably not work in weekends and is ill occasionally. In such cases, the fact that ARMA models can not deal with missing records and a varying frequency poses a serious problem. Furthermore, when the data and model frequency are high as compared to the response time of the system, the duration of the MA-part of the transfer function is limited in real time (see figure 2.9) so the model order is in practice restricted effectively to AR($nr$). High frequency data can also cause numeric problems as AR-parameters asymptotically approach the value

of one when the frequency increases, in which case the autocorrelation in subsequent values of the output will also be very high.

- Third, there is the problem of extrapolating results. Because of the data-based nature of time series models, and for any model for that matter, there is a risk in extrapolating time series analysis results to situations for which no data are available. Care has to be taken in this respect with scenario analysis, in which the level of one or more inputs or the output reaches unprecedented values. Even more problematic, and without additional knowledge even impossible, is extrapolation of results to situations where the system itself is changed and the assumption of time-invariance no longer holds. Another variant of this problem is extrapolation of results to other points in space, for which no data are available.
- Fourth, ARMA models are generally over-parameterized. When a system is affected by many inputs, when the input(s) do not change much over time and/or are cross-correlated, it will generally be difficult to separate the different effects and get correct, reliable and unique parameter estimates. The fact that the MA-part of ARMA models is parameter-inefficient, especially when dealing with high-frequency data, adds to this problem.

In light of all the above mentioned problems, i.e. whenever purely statistical methods fail, a possible solution is to use physical knowledge on the functioning of a specific system, or the type of systems addressed in general. Therefore, we will focus on the physical basis of time series models in the next sections.

## 2.3    Physical, 'white box' viewpoint

### 2.3.1    AR models as linear reservoir system

In [*Box and Jenkins,* 1970], the example of a series of water reservoirs (figure 2.13) is used to illustrate the fact that ARMA models can be regarded as models of physical systems that obey a temporal difference equation. Here, two linear reservoirs are considered with water volumes $y_1(t)$ and $y_2(t)$, and a zeroth reservoir whose volume $x_0(t)$ is manipulated and drives the others. Volumes are used in stead of levels to allow for systems with a zeroth moment or gain (see section 2.3.4) that differs



**figure 2.13:** *Series of coupled, linear reservoirs according to [Box and Jenkins, 1970].*

from one. Also, $x$ is used for denoting the volume of the zeroth reservoir because that is the forcing or explanatory variable (or in hydrologic terms the boundary condition). By doing so, the notation is kept comparable to that of section 2.2.1, where ARMA models were introduced as linear regression models.

In a steady state situation, logically the level in all reservoirs is the same. In that case, the relationship between the levels is:

$$y_i = \frac{A_i}{A_0} x_0 \tag{2.16}$$

where $i$ denotes the reservoir number and $A_i$ its basal area. As the levels are equal, it is the ratio between the basal areas that determines the gain. If the levels are not in equilibrium, the difference in level is $\frac{x_0}{A_0} - \frac{y_i}{A_i}$. In case of linear reservoirs, the rate of flow through the connecting pipe is proportional to the difference in level. For the first reservoir, this yields the following differential equation:

$$\frac{d\,y_1}{d\,t} = \frac{1}{c_1}(\frac{x_0}{A_0} - \frac{y_1}{A_1}) - \frac{1}{c_2}(\frac{y_1}{A_1} - \frac{y_2}{A_2}) \tag{2.17}$$

where $c_1$ and $c_2$ are the resistances [T] of the respective pipes. For the second reservoir, the differential equation is given by:

$$\frac{d\,y_2}{d\,t} = \frac{1}{c_2}(\frac{y_1}{A_1} - \frac{y_2}{A_2}) \tag{2.18}$$

Solving (2.18) for $y_1$ yields:

$$y_1 = A_1 c_2 \frac{d\,y_2}{d\,t} + \frac{A_1}{A_2} y_2 \tag{2.19}$$

Next, eliminating $y_1$ from (2.17) using (2.19) yields:

$$\frac{d^2\,y_2}{d\,t^2} = -\frac{(A_1 + A_2)c_1 + A_2 c_2}{A_1 A_2 c_1 c_2} \frac{d\,y_2}{d\,t} - \frac{1}{A_1 A_2 c_1 c_2} y_2 + \frac{1}{A_0 A_1 c_1 c_2} x_0 \tag{2.20}$$

A second order temporal differential equation like (2.20) can be written as a temporal difference equation in discrete time, and in turn as an AR(2) model. See [*Box and Jenkins,* 1970, page 343] for details in this respect. Please note, however, that the second term on the right side of equation (2.17) is forgotten by Box and Jenkins, when they take the step of deriving their second order temporal differential equation. They place the differential equation of the third reservoir in that of a two reservoir system, resulting in equations with erroneous coefficients onwards.

Equation (2.20) can be solved by applying a Laplace transform, under initial condition $y_2(0) = 0$. However, especially for higher order systems, matrix differential calculus is more convenient. We may combine equations (2.17) and (2.18) in one matrix equation as:

$$\frac{d\mathbf{y}}{dt} = -\mathbf{A}\mathbf{y} + \frac{1}{c_1 A_0}\mathbf{x} \tag{2.21}$$

where:

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}, \quad \mathbf{A} = \begin{bmatrix} \dfrac{1}{c_1 A_1} + \dfrac{1}{c_2 A_1} & -\dfrac{1}{c_2 A_2} \\ -\dfrac{1}{c_2 A_1} & \dfrac{1}{c_2 A_2} \end{bmatrix} \tag{2.22}$$

If the input $\mathbf{x}$ is stationary and equals:

$$\mathbf{x} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \tag{2.23}$$

and we use as initial condition:

$$\mathbf{y}(0) = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{2.24}$$

the solution to (2.21) will yield a vector of step response functions $\mathbf{y}$. Next, we write (2.21) as:

$$\frac{d\mathbf{z}}{dt} = -\mathbf{A}\mathbf{z} \tag{2.25}$$

where:

$$\mathbf{z} = \mathbf{y} - \frac{1}{c_1 A_0}\mathbf{A}^{-1}\mathbf{x} \tag{2.26}$$

The solution to (2.25) is:

$$\mathbf{z} = \exp(-\mathbf{A}t)\mathbf{z}(0) \tag{2.27}$$

Background and theory

or:

$$\mathbf{y} - \frac{1}{c_1 A_0} \mathbf{A}^{-1}\mathbf{x} = \exp(-\mathbf{A}t)\{\mathbf{y}(0) - \frac{1}{c_1 A_0}\mathbf{A}^{-1}\mathbf{x}\} \tag{2.28}$$

Using (2.24), this reduces to:

$$\mathbf{y} = \frac{1}{c_1 A_0}\{\mathbf{I} - \exp(-\mathbf{A}t)\}\mathbf{A}^{-1}\mathbf{x} \tag{2.29}$$

where $\mathbf{I}$ is the identity matrix. By definition, the derivative of the step response functions $\mathbf{y}$ to $t$ is a vector of impulse response functions $\boldsymbol{\theta}$, which equals, using (2.23):

$$\boldsymbol{\theta} = \frac{d\mathbf{y}}{dt} = \frac{1}{c_1 A_0}\exp(-\mathbf{A}t)\begin{bmatrix} 1 \\ 0 \end{bmatrix} \tag{2.30}$$

The exponential in (2.30) is a so-called matrix exponential that can be evaluated using a program like Matlab. In case of identical reservoirs, the solution or impulse response of the individual reservoirs is (see appendix A):

$$\theta_i = \frac{1}{c_1 A_0}\sum_{j=1}^{nr} e_{ij}e_{1j}\exp(-\lambda_j t) \tag{2.31}$$

where $i$ is the position of the reservoir, $nr$ is the number of reservoirs and $e_i$ the $i^{th}$ value in eigenvector $j$ with eigenvalue $\lambda$. In case of non-identical reservoirs, the solution reads (see appendix A):

$$\theta_i = \frac{1}{c_1 A_0}\sum_{j=1}^{nr} \alpha_{ij}\exp(-\lambda_j t) \tag{2.32}$$

where the $\alpha$'s weigh the exponential functions. In [*Sahuquillo*, 1983], a method is presented for solving the groundwater flow differential equation using a related approach [see also *Sloan*, 2000; *Pulido-Velazquez et al.*, 2005; *Bidwell*, 2005]. In case of linearity, the cells in a finite difference or finite element model can be considered to be linear reservoirs. Based on this, impulse response functions (or influence functions in Sahuquillo's terms) can be derived directly from the matrix of cells in a groundwater model (1, 2 or 3D), and be stored for simulation purposes. In this way, the solution to the differential equation is exact in the time domain, and it is not necessary to solve the entire matrix for every time step. Next to the fact that this method is useful in itself, especially when results from only a few locations are needed, it also nicely

shows that an AR model (or a sum of exponential functions) in principle can form a solution to an entire groundwater model.

### 2.3.2 Differential equations and convolution

In the previous section, we have regarded AR models as temporal difference equation of a set of linear reservoirs. The driving forces and state variables in groundwater systems, however, are a function of both time and space, and space is not explicitly defined in an ARMA model. In order to gain better insight into the physical meaning of time series models and response functions, here we start from the beginning. The construction of a physically-based model for the dynamics of an arbitrary variable starts with the derivation of an appropriate mathematical expression, commonly a differential equation, for it. Differential equations are derived from two types of equations [e.g., *Bear,* 1972]:

- The continuity equation(s)
- The constitutive equation(s)

For groundwater, the continuity equation is found by applying the principle of mass conservation (while in other cases, also other conservation laws may apply, like those for energy or momentum). In simple words, this principle states that no mass can disappear without reason and it is valid for all types of masses. In case of water, it is also known as the water balance equation. The constitutive equation specifies a property of the specific type of mass under consideration, which in case of groundwater is Darcy's law. This equation treats a representative elementary volume (REV) of a porous medium as a continuum, thereby disregarding the spatial variability below this scale (consisting of grains and pores, or molecules on an even smaller scale).

A standard technique for solving differential equations of linear dynamic systems exactly is to determine their solution for a Dirac delta function $\delta(t)$ [*Dirac,* 1947] (not to be confused with $\delta$, the autoregressive parameters of ARMA models) as input $p(t)$. The delta function has the following properties:

$$\begin{cases} \delta(t) = 0, & t \neq 0 \\ \int\limits_{-\infty}^{\infty} \delta(t)\mathrm{dt} = 1 \end{cases} \tag{2.33}$$

and can be seen as the limit of (for instance) the Gaussian distribution function with mean zero and a variance that also approaches zero (which, in mathematical terms, is a way of 'mollifying' $\delta(t)$). In words, the delta function is an instantaneous pulse or impulse of unit area. The impulse response (IR) function $\theta(t)$ is then defined as the effect of $\delta(t)$ as input $p(t)$ on the state $h(t)$ of a system, i.e. the deviation in time from an otherwise steady state $d$. In terms of groundwater head and precipitation,

$\theta(t)$ can be thought of as the response to a very short shower of unit height, when the head is otherwise constant or equals the local drainage base. In mathematical terms, $\theta(t)$ is defined by the following conditions:

$$\begin{cases} \theta(t) = h(t) - d \\ h(t) = d, \qquad t < 0 \\ p(t) = \delta(t) \end{cases} \tag{2.34}$$

If $\theta(t)$ and $d$ are known, $h(t)$ can be obtained for an input $p(t)$ that varies arbitrarily in time through convolution (Duhamel's principle [*Duhamel,* 1833]):

$$h(t) - d = \int_{-\infty}^{t} \theta(t-\tau)p(\tau)\mathrm{d}\tau \equiv \int_{0}^{\infty} \theta(t)p(t-\tau)\mathrm{d}\tau \equiv (\theta * p)(t) \tag{2.35}$$

Equation (2.35) implies that the dynamics of an arbitrary, linear dynamic system at a certain location are completely governed by the dynamics of the input and the impulse response function. The impulse response function completely characterizes the dynamically relevant, physical properties of the system at that location, and is as such an integral property thereof.

### 2.3.3 Evaluation of a convolution integral

Of course, the first step in applying a model is evaluation of the equations. For evaluating the convolution integral in (2.35), first the input data has to be transformed to continuous time, as most data are discrete due to the observation and digital storage process. In case of groundwater systems, common excitations are either levels (surface water or other groundwater levels) or fluxes (precipitation, evaporation, pumping). For fluxes, the discrete data may often be regarded as changes in the primitive function of the underlying continuous flux or intensity, as is the case with precipitation amount series $P_t$ [L]:

$$P_{t_i} = P(t_i) - P(t_i - \Delta t_i) = \int_{t_i - \Delta t_i}^{t_i} p(\tau)\mathrm{d}\tau \tag{2.36}$$

Of course, the continuous precipitation intensity series $p(t)$ [LT$^{-1}$] cannot be reconstructed exactly, but it can be approximated by assuming that the flux is constant over the time step $\Delta t$ and equals the average. Equation (2.36) may then be written as:

$$p(\tau) = \frac{P_{t_i}}{\Delta t_i}, \quad t_i - \Delta t_i < \tau < t_i \tag{2.37}$$
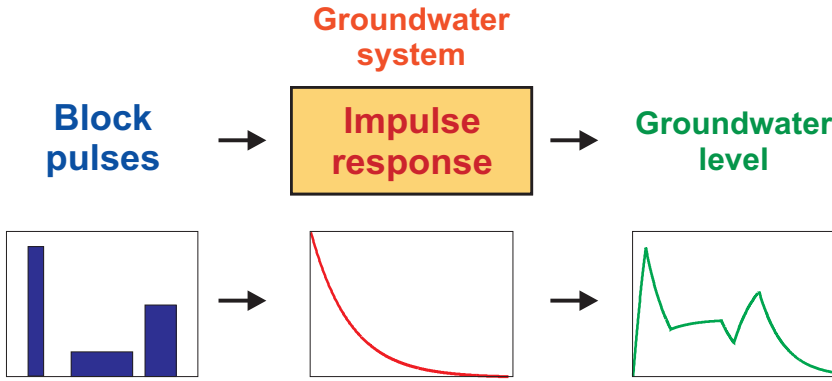
**figure 2.14:** *Transformation of input to output in case of an exponential impulse response and three block pulses of different height and duration.*

The higher the observation frequency, the smaller the time step and the better (2.37) will approximate $p(t)$. For stresses that are levels, we simply assume that the level is constant over the time step. Note that the error made by these assumptions will vary in time when input series with an irregular frequency are used, which will cause the model residuals to have a varying variance. When the time steps are not too large and irregular, however, this effect is small as compared to the other sources of model error and will therefore be neglected.

The transformation described above results in continuous input series that take the shape of consecutive series of blocks. As a next step, we can evaluate the effect of a single block pulse using the block response function $\Theta(t)$. The block response is obtained when the impulse response is convoluted with a block pulse of unit height over a period $\Delta t$, which equals:

$$\Theta(t, \Delta t) \; = \; \int_{t-\Delta t}^{t} \theta(\tau) \mathrm{d}\tau \tag{2.38}$$

The block response function is equivalent to a discrete-time transfer function when it is divided by $\Delta t$ in order to scale it to the effect of a discrete instantaneous input of unit height. In case of an irregular frequency, however, the block response isn't constant but a function of the time step $\Delta t_i$. Because convolution is a linear operation, its result or groundwater level $h(t)$ may be obtained by adding the responses to all individual blocks of the input series. The transformation of input to output is illustrated in case of an exponential impulse response and three block pulses of different height and duration in figure 2.14. Because $\Theta(t)$ and $p(t)$ are continuous functions, this procedure results in a definition of $h(t)$ that is continuous also. In contrast to the equations of ARMA models, no explicit reference is made to a model

or data frequency. Problems related to the frequency, such as handling irregular or high-frequency data, are consequently circumvented. In continuous time, there is no such thing as a gap or 'missing' observation of $h(t)$.

In the convolution integral of equation (2.35), time starts at minus infinity. As a consequence, we have to define an initial value for the input series from $t = -\infty$ to $t = 0$, the time of the first available observation, in order to be able to evaluate it. For stresses that are stationary or in general fluctuate around some average level, like precipitation, we can define the initial value $\bar{p}$ to be the temporal average of the available observations (a historical average, if available, may also be used). For non-stationary stresses, e.g., a pumping well that was constructed in a certain year and gradually increased in rate, the temporal average is inappropriate as initial value. In that case, the initial value can either be taken to be zero before pumping started, or to be some average historic rate prior to the first available recorded discharge. Also, a run-in period is needed from $t = 0$ to $t = t_h$, the start of the calibration period or time of the first groundwater level observation, to obtain a reasonable estimates of the groundwater levels at early time. The actual outcome of the convolution integral is the sum of the effect of the initial value and the available observations of the input series, in both the run-in and calibration period. Together with local drainage base $d$, this yields a prediction for $h(t)$ that can be compared with the available observations (figure 2.15).

Convolution is a very practical method that unfortunately is missing in many textbooks on groundwater hydrology [*Olsthoorn,* 2008]. It can not only be used for time series analysis, but also for efficiently generating high-resolution time series using response
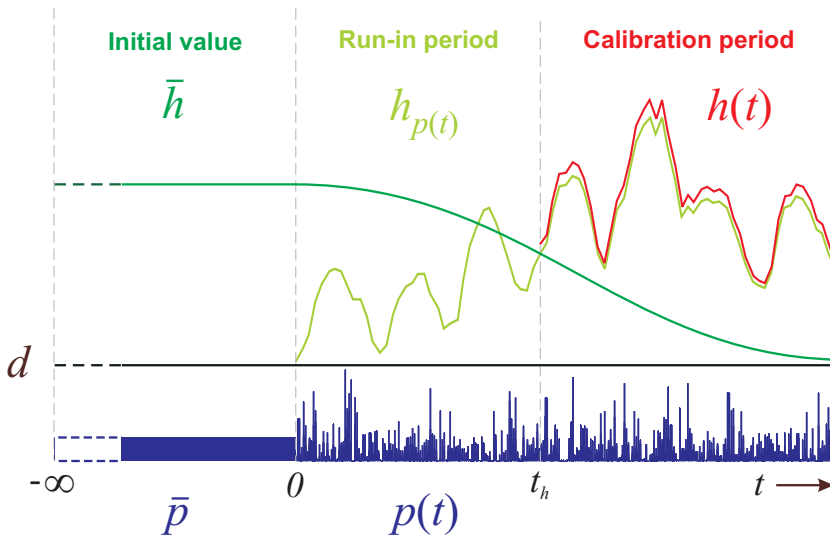


*figure* 2.15: *Role of the initial value, run-in and calibration period in the evaluation of a convolution integral.*

functions obtained from more computationally demanding deterministic models [*Sahuquillo,* 1983].

### 2.3.4 Impulse responses and their characteristics

In groundwater hydrology, a typical impulse response takes the shape of a skewed distribution function (figure 2.16). In the field of statistics, distribution functions are commonly characterized by their moments. Here, we also use the concept of moments to characterize the properties of impulse responses. The $n^{\text{th}}$ -order temporal moment $M_n$ of an impulse response function is defined by:



$$M_n = \int_{-\infty}^{\infty} t^n \theta(t)\mathrm{d}t \qquad (2.39)$$
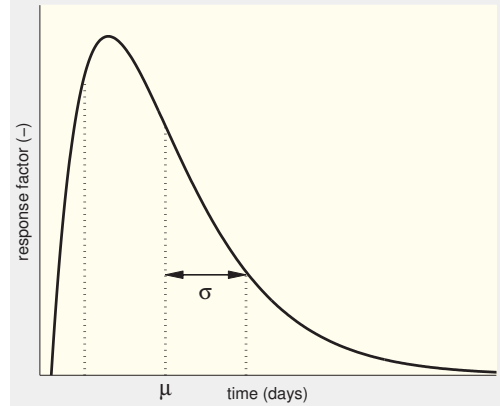
*figure 2.16: Skewed impulse response function with mean $\mu$ and variance $\sigma^2$.*

Higher order moments are generally scaled by the zeroth moment (although for frequency or probability distributions by definition $M_0 = 1$, so there it does not make a difference) and centralized about their mean $\mu$. Consequently, central moments are given by:

$$\overset{c}{M}_n = \frac{\int_{-\infty}^{\infty} t^n \theta(t-\mu)\mathrm{d}t}{\int_{-\infty}^{\infty} \theta(t)\mathrm{d}t} \qquad (2.40)$$

The zeroth through second central moments are also commonly known by the terms area, mean $\mu$ and variance $\sigma^2$. Other characteristics of $\theta(t)$ are its skewness, defined by:

$$\overset{1}{\gamma} = \frac{\overset{c}{M}_3}{\sigma^3} \qquad (2.41)$$

and its kurtosis or excess kurtosis:

$$\overset{2}{\gamma} = \frac{\overset{c}{M}_4}{\sigma^4} - 3 \tag{2.42}$$

The effect of differences in kurtosis is instantly clear if one compares well-known distributions with the same mean, variance and skewness, but a different kurtosis (figure 2.17). These distributions can be written as special cases of the Skew Exponential Power distribution [*Fernandez et al.,* 1995*; Schoups and Vrugt,* 2010], which has the added advantage that its skewness and kurtosis can be varied gradually. An alternative, and in some cases more convenient way of characterizing distribution functions is by their cumulants. From a hydrologic perspective, $M_0$ defines the effect of a stationary excitation of unit strength (e.g., stationary precipitation). In time series analysis literature, $M_0$ is called

*figure* **2.17:** *Well-known distributions with different kurtosis (resp. uniform, Wigner semicircle, normal, hyperbolic secant, Laplace double exponential (source: Wikipedia)).*

the gain. $\mu$, on the other hand, is the mean delay or response time to an instantaneous input (e.g., a sudden shower of precipitation). The standard deviation $\sigma$ is a measure for the temporal 'dispersion' of an impulse.

### 2.3.5    Responses of elementary groundwater systems

#### 2.3.5.1    Recharge on a system with a uniform internal head

A consequence of the law of mass conservation is that the amount of water in a volume must increase when the inflow at a certain time step exceeds the outflow. When we simplify the functioning of a soil column to a simple linear reservoir with inflow at the top and outflow at the bottom, analogous to [*Knotters and Bierkens,* 2000], the water balance is given by (figure 2.18):

$$r - q = S\frac{\mathrm{d}h}{\mathrm{d}t}$$

(2.43)

where

$q$   : discharge per unit area and time
     $[LT^{-1}]$
$h$   : groundwater head at location $z$ [L]
$S$   : storativity [-]
$r$   : recharge per unit area and time $[LT^{-1}]$

When considering surface water, the storativity $S$ equals 1 as one millimeter of precipitation causes a rise in water level of one millimeter. In groundwater systems, however, $S$ is smaller because only the free pore space is available for additional storage of water (in phreatic systems), or because the pore space can only change a little by compression or expansion of the grain skeleton, water, or included gas bubbles (in confined aquifers). Equation (2.43) contains two different variables, $q$ and $h$, that are normally both



***figure* 2.18:** *Water balance of a linear reservoir system.*

unknown, while the third ($r$) is normally given. In order to solve the equation, we need a second, independent relationship between $q$ and $h$, which we find from Darcy's law:

$$q = \frac{h - d}{c}$$

(2.44)

where $c$ is the drainage resistance [T] and $d$ the drainage level [L]. Elimination of $q$ from (2.43) yields the following first order differential equation:

$$S\frac{\mathrm{d}h}{\mathrm{d}t} = \frac{d-h}{c} + r \qquad (2.45)$$

It is noted that equation (2.45) does not contain any reference to a spatial dimension. The internal head of the system is uniform, and only a difference in head with the outside world is specified, without a direction. Graphically (as in figure 2.18) and intuitively, however, it may be convenient to picture the head difference as a vertical column of water, which of course requires the vertical dimension. If we solve (2.45) for an impulse of recharge and use (2.34) , we find that the impulse response function [-] is an exponential and equals that of a single, linear reservoir (see also section 2.3.1):



**figure 2.19:** *Example IR functions of the reservoir system of figure 2.18 for different values of $cS$ .*

$$\theta(t) = \frac{1}{S}\exp(-\frac{t}{cS}) \qquad (2.46)$$

As for any linear system, also in this case the groundwater level fluctuations can now be found through convolution. The linear reservoir, however, has the remarkable and useful property that the fluctuations in level may also be computed recursively. In other words, the value of $h(t)$ at any time step may be computed from $h(t-\Delta t)$ and the total input over the time step $R_t$ , as [*De Zeeuw and Hellinga,* 1958]:

$$h(t) = \{h(t-\Delta t) - d\}\exp(-\frac{\Delta t}{cS}) + R_t c\{1 - \exp(-\frac{\Delta t}{cS})\} + d \qquad (2.47)$$

Equation (2.47) is equivalent to an ARX(1) model (equation (2.4) ) in discrete time [*Knotters and Bierkens,* 2000]. The fact that $h$ can be computed recursively, allows the response parameters [ $c,S$ ] to change piecewise over time. Consequently, equation (2.47) can also be applied to non-stable or non-linear systems (where $c$ and/or $S$ are a function of $h$ ).

### 2.3.5.2 Recharge on a system with a one-dimensional head gradient

In a system with a one-dimensional, horizontal head gradient, the water balance equation is given by:

$$-\frac{\partial q_x}{\partial x} = S\frac{\partial h}{\partial t} - r \qquad (2.48)$$

**figure 2.20:** *Water balance of a system with a one-dimensional, horizontal head gradient and parallel surface waters.*

Under phreatic conditions, the saturated thickness $H$ [L] of an aquifer depends on $h$ (and therefore on $t$ and $x$). If $H - h$ is small relative to $H$, however, we can approximate $H$ as constant. In case of an impermeable base, fully penetrating ditches and a distance between the surface waters $L$ [L] that is much larger than $H$, we can also approximate the flow as horizontal [Dupuit-Forchheimer approximation, *Dupuit, 1863; Forchheimer, 1901*]. Darcy's law may in such a case be written as:

$$q_x = -KH\frac{\mathrm{d}h}{\mathrm{d}x} \tag{2.49}$$

where
$KH$ : transmissivity [$L^2T^{-1}$]
$K$ : permeability [$LT^{-1}$]

Elimination of $q$ from (2.48) and (2.49) yields the following second order differential equation:

$$KH\frac{\partial^2 h}{\partial x^2} = S\frac{\partial h}{\partial t} - r \qquad (2.50)$$

Here, the differential equation contains one spatial dimension, allowing the heads in the system to depend on the location. Again, we use a second, vertical dimension to graphically represent the system (figure 2.20). Equation (2.50) can be solved analytically for the effects of recharge. For groundwater problems, the solution has been found first by Glover and was published in [*Dumm,* 1954] and also discussed in [*Kraijenhoff van de Leur,* 1958]. When the water level at the



*figure* **2.21:** *Example IR functions of the system of figure 2.20 for different values of* $C$.

boundaries is constant $(h_{x=0} = h_{x=L} = d)$, the recharge is an impulse and (2.34) is used, the impulse response function [-] may be found and be written as:

$$\theta(t) = \frac{1}{S}\frac{4}{\pi}\sum_{n=1,3,5,\dots}^{\infty}\frac{1}{n}\exp(\frac{-n^2\pi^2 KHt}{SL^2})\sin\frac{n\pi x}{L} \qquad (2.51)$$

where $x$ [L] defines the position of the location under consideration. The response of such a system is not only a function of the storativity and resistance, as in the linear reservoir case, but also of the relative position or centrality $C$ [-] of the location under consideration, which we define by:

$$C = \begin{cases} \dfrac{2x}{L}, & x \leq \dfrac{L}{2} \\ \dfrac{2(L-x)}{L}, & x > \dfrac{L}{2} \end{cases} \qquad (2.52)$$

In this way, the centrality is 1 if the location is in the center of the system and 0 if it lies on its boundaries. In figure 2.21, example IR functions are plotted for different values of $C$, which shows the difference in behavior depending on the location. In short: in the center of the system, the level can only drop after the levels near the boundaries have dropped first. Near the boundaries, the water level drops quickly at first but stabilizes later on because of water flowing in from the center.

### 2.3.5.3    From infiltration to recharge in the unsaturated zone

Up until this point, we have neglected the fact that in reality recharge is difficult to measure and generally unknown. Recharge, however, can be seen as the outcome of the processes in the unsaturated zone, of which the driving forces are precipitation $p$ and (potential) evaporation $e$. The latter are generally available from meteorologic stations, and can therefore be used as input if we are able to model or mimic the unsaturated zone processes. In case of unsaturated flow, the water balance equation and Darcy's law again take different shapes, because of the varying water content and permeability. If we consider one-dimensional, vertical flow only, the water balance equation is given by:
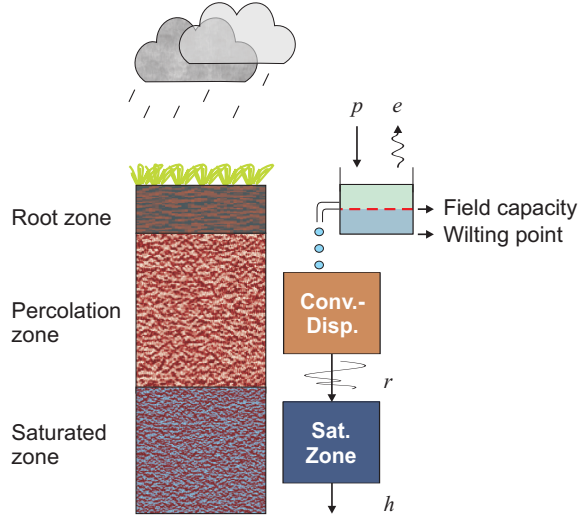


*figure 2.22: Schematic representation of the unsaturated zone.*

$$\frac{\partial q}{\partial z} = \frac{\partial \vartheta}{\partial t} \qquad (2.53)$$

and Darcy's law by:

$$q = K(\vartheta)\frac{\mathrm{d}h}{\mathrm{d}z} \qquad (2.54)$$

where
$z$    : depth [L]
$\vartheta$    : volumetric water content [-]
$K(\vartheta)$ : unsaturated permeability [LT$^{-1}$]

By combining the equations above, we find the following differential equation for flow in the unsaturated zone:

$$\frac{\partial}{\partial z}\left[ K(\vartheta)\frac{\partial h}{\partial z} \right] = \frac{\partial \vartheta}{\partial t} \qquad (2.55)$$

which is a modified form of Richards'
equation [*Richards,* 1931]. The
unsaturated zone is generally divided
into a) the root zone, where there is
direct input and uptake of water, b) the
percolation zone, where water primarily
percolates downwards due to
gravitational force and c) the capillary
zone, in which there is capillary contact
and possible feeding from the saturated
zone. As it is nearest to the driving
forces, the water content in especially
the root zone is highly dynamic and
consequently, because of the varying
permeability $K(\vartheta)$, its functioning is
highly non-linear. In a wet state, when
the soil is at field capacity, permeability
is high, evaporation is at its potential



*figure* 2.23: *Example IR functions of a linearized unsaturated zone (convection-dispersion equation) for different values of* $\sigma$.

rate and precipitation can infiltrate and percolate freely. In a dry state, capillary
pressure is high, the permeability of the soil is greatly reduced and evaporation and
percolation are strongly limited. The effect of variable storage of water in the root
zone is sometimes modeled conceptually as a reservoir in which uptake or evaporation
stops at the so-called wilting point, and from which water only percolates at field
capacity (figure 2.22). The fact that such non-linear behavior cannot be described with
a linear model can account for the ubiquitous seasonal patterns in the residual series
of linear time series models. It can also be a reason for combining time series models
with more or less complex models of unsaturated zone processes [e.g., *De Keizer,*
2003; *Berendrecht et al.,* 2006] .

In the percolation zone, in situations with deeper water tables, the water content is
less variable and its main effect is retardation and dispersion of infiltrating water
[*Besbes and de Marsily,* 1984; *Parlange et al.,* 1992]. When Richards' equation is
linearized around some constant, 'average' water content $\vartheta_c$ using Taylor series
expansion, it transforms into the convection-dispersion equation [*Jury and Roth,* 1990;
*Zwamborn,* 1995]. The convection-dispersion equation can consequently be used to
describe the 'average' effect of the unsaturated zone on the damping and retardation
of precipitation and evaporation pulses. Here, we define $z$ to be zero at the top of the
percolation zone. If we apply the following initial conditions:

$$\begin{cases} \vartheta(z < 0,\ t = 0) - \vartheta_c = 1 \\ \vartheta(z > 0,\ t = 0) - \vartheta_c = 0 \end{cases}$$

(2.56)

and boundary conditions:

$$\begin{cases} \vartheta(z = -\infty, t) - \vartheta_c = 1 \\ \vartheta(z = \ \ \infty, t) - \vartheta_c = 0 \end{cases} \tag{2.57}$$

and the step response function $\Omega$ [-] to be:

$$\Omega(z,t) = \vartheta(z,t) - \vartheta_c \tag{2.58}$$

the following solution to the convection-dispersion equation may be found [*Crank, 1957*]:

$$\Omega(z,t) = \frac{1}{2}\mathrm{erfc}\{\frac{z - vt}{2Dt}\} \tag{2.59}$$

where
$D$ : (effective) diffusion-dispersion coefficient [$L^2\ T^{-1}$]
$v$ : (effective) propagation velocity [$LT^{-1}$]

The impulse response $\theta$ [$T^{-1}$] is the derivative of (2.59) or [e.g., *Jury and Sposito, 1985*; *Maas, 1994*]:

$$\theta(z,t) = \frac{\mathrm{d}\Omega(z,t)}{\mathrm{d}t} = \frac{z}{4\sqrt{\pi Dt^3}}\exp\{-\frac{(z - vt)^2}{4Dt}\} \tag{2.60}$$

A shortcoming of (2.59) is that it also predicts upward dispersion beyond $z = 0$. When the boundary conditions are chosen such that upward dispersion is prevented $(\vartheta(z = 0,t) - \vartheta_c = 1)$, a second term is introduced in the step response function [*Rifai et al., 1956*]. The contribution of this second term, however, rapidly decays with distance, so that only (2.60) remains at some distance from the boundary [see *Maas, 1994, page 61*]. $\theta$ in this case is not a response in level but in water content, and its area or zeroth moment equals one because of the principle of mass conservation. Please note also that the propagation velocity $v$ is the velocity at which the pulse or pressure wave propagates. It is not the (effective) velocity of water itself in the unsaturated zone. The velocity of water is much smaller than that the pressure propagation velocity, and equals the recharge rate in a linearized situation. The first ($\mu$ [$T^{-1}$]) and second ($\sigma^2$ [$T^{-2}$]) central moments of $\theta$ are given by:

$$\begin{cases} \mu = \dfrac{z}{v} \\ \\ \sigma^2 = \dfrac{2zD}{v^3} \end{cases} \tag{2.61}$$

In figure 2.23, some example IR functions are plotted that have the same, average arrival time $\mu$, but a different standard deviation $\sigma$. An outcome of the convection-dispersion equation is that the spatial (vertical) distribution of water pulses is symmetric, which is a consequence of the linearization of Richards' equation around the average water content. In dynamic, non-linear reality, the 'wetting' front will be sharper because of the increased permeability when the water content is high. On the other hand, macro pores and preferential flow will increase the flux at relatively early arrival times and consequently counteract the sharpness.

#### 2.3.5.4 Convolution of response functions

In section 2.3.5.3, we have discussed the effect of the unsaturated zone in delaying and damping precipitation pulses. As recharge is generally unknown, however, we can not obtain a separate estimate of the properties or response of the unsaturated zone through time series analysis. Instead, if we use precipitation as input and estimate its effect on the groundwater level, we obtain an estimate of the response of the combined saturated and unsaturated zone. If linear systems are placed in series, the output of one is input to the other, so their combined effect equals:

$$h(t) = \int_{-\infty}^{t} \theta_{sz}(t-\tau_2)\int_{-\infty}^{t}\theta_{uz}(t-\tau_1)p(\tau_1)\mathrm{d}\tau_1\mathrm{d}\tau_2 \tag{2.62}$$

where $\theta_{sz}$ and $\theta_{uz}$ are the response functions of the saturated and unsaturated zone, respectively. Mathematically, equation (2.62) equals a single convolution integral with
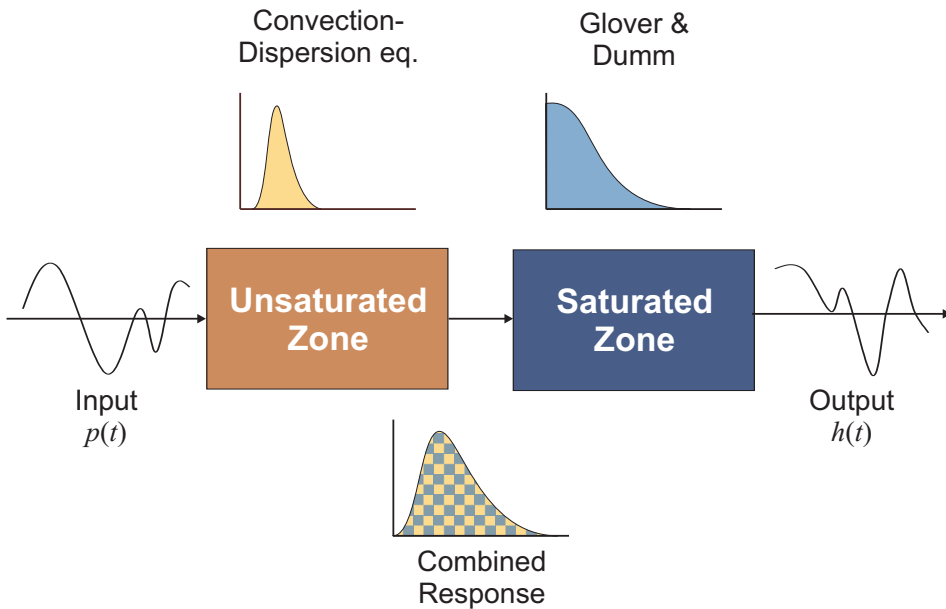


*figure* 2.24: *Combination of the responses of the unsaturated and saturated zones.*

a combined response that is found by convoluting the separate responses:

$$\theta_c(t) = \int\limits_{-\infty}^{t} \theta_{sz}(t-\tau)\theta_{uz}(\tau)\mathrm{d}\tau \qquad (2.63)$$

In figure 2.24 the process and example responses are illustrated. Depending on the situation, the combined response can take the shape of a function that rises gradually at first, due to the effect of the unsaturated zone, until the peak is reached. After that, the response slowly decays as recharge diminishes and groundwater drains towards the surface waters or drainage means. In short, the combined response takes the shape of a skewed distribution function. When a signal is transferred through $n$ linear systems placed in series, their $k^{\text{th}}$-order impulse response moments combine in the following way [e.g., *Maas,* 1994]:

$$\begin{cases} M_{0,total} = \prod\limits_{i=1}^{n} M_{k,i}, & k = 0; \\ \overset{c}{M}_{k,total} = \sum\limits_{i=1}^{n} \overset{c}{M}k_{k,i}, & k > 0. \end{cases} \qquad (2.64)$$

Although there is no closed mathematical expression, for time series analysis purposes it is well possible to use for instance the convection-dispersion and Glover's equation as separate solutions for the unsaturated and saturated zone and combine them numerically via convolution [see also *Bierkens and Walvoort,* 1998*; Bierkens and Bron,* 2000]. When doing so, however, the parameters of both equations prove to be highly correlated and both responses are consequently difficult to separate [*Kruithof,* 2001].

### 2.3.5.5    Other situations and excitations

The previous sections served to illustrate the shapes that responses of groundwater systems take from a physical point of view. As we only discussed elementary groundwater systems, the text can hardly serve as an overview of the possible variability in responses of groundwater systems. For a thorough treatment on the existing analytic solutions in various geohydrologic situations, we refer to [*Bruggeman,* 1999]. In the coming sections, however, we will step away from the traditional, physical viewpoint and introduce the use of behavioral response functions that have general usability but no (apparent) physical meaning, and discuss its consequences. A point that remains here is the fact that we only discussed the effects of precipitation and evaporation. As the dimensionality of a system can have a pronounced effect on its response, so does the dimensionality or spatial distribution of the excitations. In contrast to precipitation that has areal coverage, rivers, ditches and other surface waters can be seen as line features in the horizontal plane, and pumping wells as points. In figure 2.25, schematizations are given of a pumping well according to Hantush' well function [*Hantush,* 1956*; Veling and Maas,* 2010] and a river or canal according to the polder function [*Bruggeman,* 1999]. In both schematizations, there is
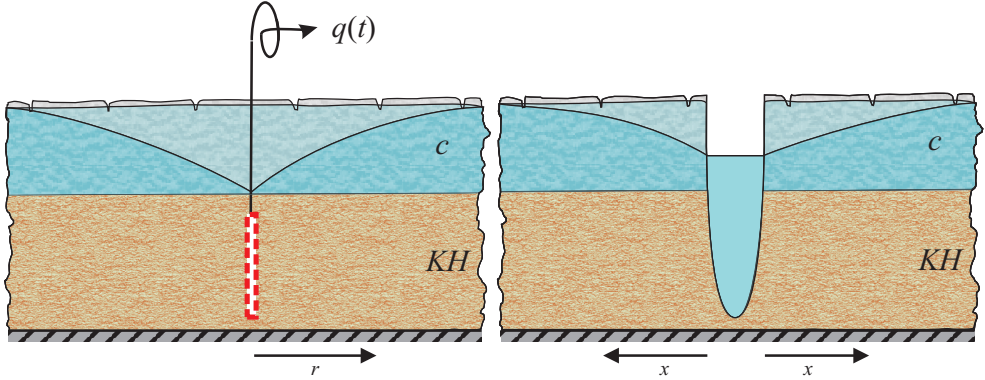
*figure* **2.25:** *Schematization of a pumping well according to Hantush (left), and a river according to the Polder function of Bruggeman (right).*

an aquifer with transmissivity $KH$ and storativity $S$ and an aquitard with resistance $c$, without storage. Furthermore, both the pumping well and the river fully penetrate the aquifer. A difference, however, is that in the differential equation for which Hantush' well function is derived a radial symmetric head gradient is assumed, whereas in the situation of the polder function the head gradient has bilateral symmetry. As a standard well test yields a step response, here we seek the impulse response or derivative of Hantush' well function with respect to time. The impulse response function describes the response to an instantaneous extraction of a unit volume of water and equals:

$$\theta(t) = -\frac{1}{4\pi KHt}\exp(-\frac{r^2 S}{4KHt} - \frac{t}{cS}) \tag{2.65}$$

In the polder function, the excitation is not a pumped amount but a rise in surface water level. Also here, the polder function is a step response, whose derivative equals:

$$\theta(t) = -\frac{1}{\sqrt{\dfrac{4\pi KHt^3}{x^2 S}}}\exp(-\frac{x^2 S}{4KHt} - \frac{t}{cS}) \tag{2.66}$$

## 2.4    Distribution functions as response models

### 2.4.1    The skew-Gaussian convolutional limits

In statistics and probability theory, assumptions about the probability distribution or 'behavior' of variables are commonly made without physical justification or physical analysis of the system to which they are related. An appropriate distribution is simply selected from a limited set of candidates, based on their match with the available

observations. On the one hand, such a practice is formed out of necessity, as probability theory is generally applied to variables generated by 'systems' whose physical properties and processes are not or only incompletely known (and some processes are intrinsically of a stochastic nature). On the other hand, this practice is also adequate, as a limited set of probability distributions prove to match a wide variety of cases, the Gaussian distribution being the most prominent example. For hydrologic applications, the use of models that do not require the explicit definition and parameterization of the physical, spatial structure of a system is very appealing. By excluding the spatial structure, the degree of freedom in the modeling problem is strongly reduced, and so is the problem of equifinality described in section 2.1.3. Next to that, only for a limited number of schematizations exact, analytic solutions are known. In practice, for use in a time series model, therefore there is often no response function that adequately matches the physical structure of the system, known or unknown, under consideration. Here, we therefore follow on the path laid down by Jury and Maas in trying to describe the behavior of groundwater systems with simple functions, without assuming or explicitly referring to a physical structure. In solute transport modeling, the problem is restricted to matter going in and coming out of a system, while the travel time of individual particles differs because of the heterogeneity of the porous medium. Here, it may seem logical to model the transfer of solutes as a (travel time) distribution function. The use of distribution functions in modeling groundwater level responses, however, is perhaps less obvious and to this date remains somewhat heuristic. All in all, there are three main arguments for assuming that (skewed) distribution functions can also be used for modeling the behavior of systems in case of groundwater heads:

- The combined, physically-based response of elementary versions of the saturated and unsaturated zone takes the shape of a skewed distribution function, as shown in section 2.3.5. Also, Hantush' well function and Bruggemans' polder function are essentially skewed distribution functions (see later on).
- As moments of higher order are of decreasing significance in characterizing probability distributions, so does the effect on the performance of a time series model decrease with the order. Consequently, a given function does not have to match the 'true' response of a system exactly in order to yield acceptable results, as long as the first few moments match.
- When considering the transfer of a signal through a porous medium, the latter can be thought to consist of a series of $n$ sections or subsystems. In [*Maas, 1994*] it is shown that transferring a signal through $n$ linear, independent systems is mathematically equivalent to adding up $n$ random, independent variables. In both cases, the combined response or probability distribution is a convolution of the individual distributions (see also section 2.3.5.4). Consequently, the central limit theorem of statistics also applies, under fairly general conditions, to the transfer of signals. This means that like probability distribution functions, also impulse responses will approach a Gaussian distribution with mean $n\mu$ and variance $n\sigma^2$ when $n$ approaches infinity and the systems are independent and identical (in line with equation

(2.64)). Maas also shows that before reaching this limit, impulse responses will tend to a skew-Gaussian shape. Furthermore, he shows that these convolutional limits also hold for parallel series of systems with lateral interactions.

The central limit theorem and skew-Gaussian convolutional limit of [*Maas, 1994*], however, are not directly applicable to the case of groundwater heads. If we, analogous to Maas, consider a porous medium as a series of identical sections or subsystems, we can distinguish three types of input-output relationships (figure 2.26). In the first case, the response of individual sections is independent of the others. This is approximately the case for transport of solutes and percolation in the unsaturated zone. Only in this case, the response of the combined system is simply a convolution of the separate responses and the convolutional limits directly apply. In the second case, the response of individual sections depends on the properties of the system in total (or $n$ in case of identical sections of fixed size), which is generally the case in saturated flow problems. When, for example, hypothetically 'adding' sections to a system at whose boundaries a constant head difference $\Delta \overline{h}$ is imposed, the length

$\Delta x$ of the system changes and therefore so do $\dfrac{\Delta \overline{h}}{\Delta x}$ or the $M_{0,i}$ of individual sections.
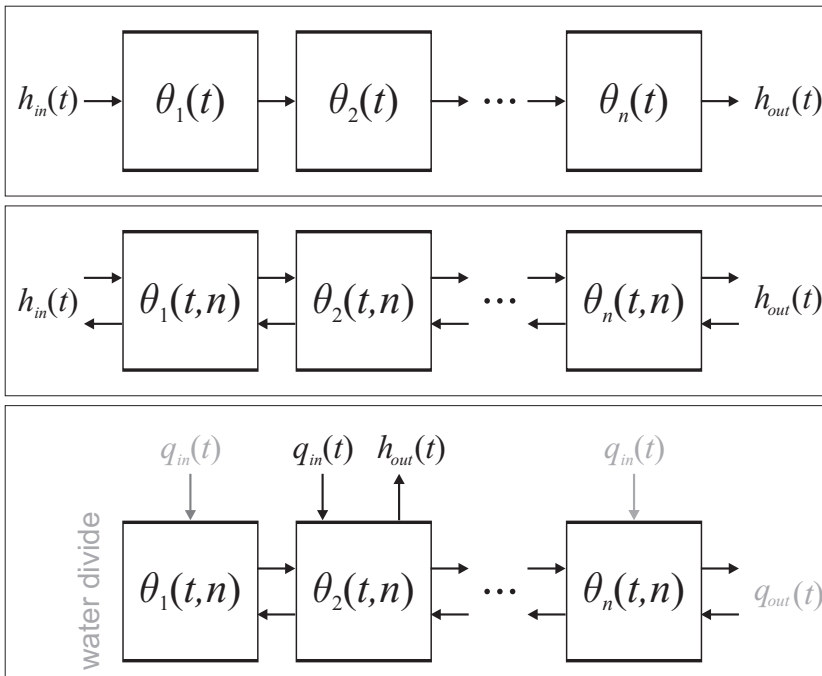


*figure* **2.26:** *Three types of input-output relationships, when considering a porous medium as a series of identical sections: a) signal $h(t)$ travels through sections with independent $\theta(t)$, b) as a), but with dependent $\theta(t,n)$ , c) as b) but with $h_{out}(t)$ and $q_{in}(t)$ (partly) at the same location.*

Also, physically identical sections do not necessarily have the same $M_{0,i}$ in such cases.

As the individual responses are not independent and identical, the central limit theorem and skew-Gaussian limit are not directly applicable. Remarkably, however, impulse responses of several systems that fall in this category, like the Polder function, Hantush' Well function and convection-dispersion equation, are also of skew-Gaussian nature and are special cases of one distribution function (see later on). In the third case, $\theta_i(t,n)$ depends on the properties of the system in total also. Here, however, the output signal $h_{out}(t)$ is obtained from the same location as where the input $q_{in}(t)$ (at least partly) enters the system. This is for example the case with inputs that have areal coverage, like recharge or seepage, and their effect on the phreatic groundwater level. Because in this case, the



*figure* **2.27**: *Cascade of reservoirs as physical analog of the gamma distribution.*

signal does not have to travel first, there is no delay in the initial response $\theta(0)$ and in principle $\theta(0)$ is also independent of the surrounding system. In case of recharge, $\theta(0)$ equals $\dfrac{1}{S}$, in line with equations (2.46) and (2.51). Consequently, in such cases, the response does not attain a (skew-)Gaussian shape with increasing system size, as $\theta(0)$ does not approach zero.

### 2.4.2 The Pearson type III, scaled gamma and generalized moving Gaussian distribution

An example of a distribution function that matches the skew-Gaussian convolutional limits of [*Maas,* 1994] (i.e. it can be seen as the output of a series of independent, identical systems) is the Pearson type III distribution function. This function is also used by Jury (with $b=0$) and Maas for modeling solute transport, and is given by:

$$\begin{cases} \theta(t) = \dfrac{a^n (t-b)^{n-1} \exp\{-a(t-b)\}}{\Gamma(n)}, & t \geq b; \\ \theta(t) = 0 & , & t < b. \end{cases} \qquad (2.67)$$

where $a$, $b$ and $n$ are parameters. When $b$ equals zero (no delay), equation (2.67) reduces to what is known as the gamma distribution. The gamma distribution does have a physical analog, as it describes the transfer function of a series of coupled linear reservoirs (figure 2.27), also known as the Nash cascade in surface water
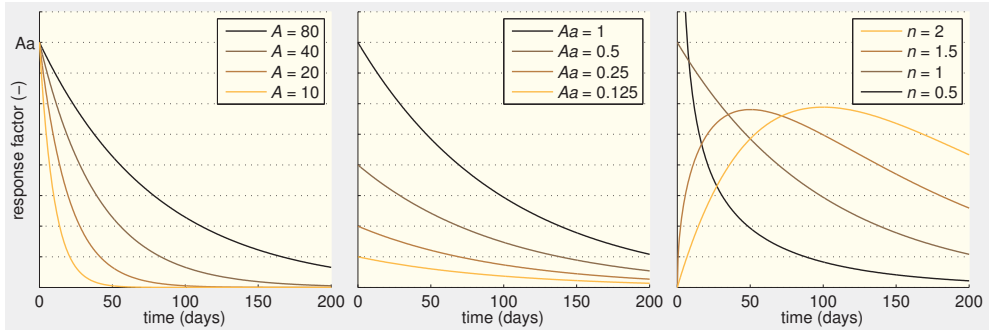
*figure* 2.28: *Example curves of the scaled gamma distribution, illustrating its variability. In a) the drainage resistance ( $Aa = 1, n = 1$ ), in b) the 'storage coefficient' ( $a = 0.01, n = 1$ ) and in c) the 'centrality' of the location and/or 'influence' of the unsaturated zone ( $A = 400 \cdot n, a = 0.01$ ) were varied.*

hydrology [*Nash,* 1958]. In that respect, the parameter $n$ denotes their (not necessarily integer) number and $a$ equals the inverse of the reservoir coefficient. In case of groundwater level responses, the response does not describe the transformation of one flux into another, but the transformation of a level or flux into a level. In that case, a conservation law does not necessarily apply and (2.67) has to be multiplied with a factor $A$ in order to allow the area to differ from one. In the absence of a delay $b$, the result is given by:

$$\begin{cases} \theta(t) = A \dfrac{a^n t^{n-1} \exp(-at)}{\Gamma(n)}, & t \geq 0; \\[2mm] \theta(t) = 0 & , \quad t < 0. \end{cases} \tag{2.68}$$

In the following, we shall refer to (2.68) as the scaled gamma distribution function (SG df). In figure 2.28, example curves of the SG df are plotted for different values of $A$, $a$ and $n$, to illustrate its flexibility and the effects of the different parameters. When used for modeling groundwater head series, the parameters in principle merely define the shape of the SG df and have no direct physical meaning. In case of the response to precipitation and evaporation, however, the parameters can be related to geohydrologic parameters in general terms, in the following way, to illustrate the range of physical variability the SG df can cover:

$A$ –  equals the area of the SG df. It also is the ratio of the mean convexity of the groundwater head above the local drainage base to the mean groundwater recharge. $A$ therefore by definition equals the (local) drainage resistance. With increasing drainage resistance, the decay rate $a$ of the response decreases ( $a \approx 1/A$ ). In simple words, the water table will drop slowly when the drainage resistance is high.

*Aa* - As outlined in section 2.3.5.1, a single linear reservoir (a SG df where $n=1$) equals a simple physical model of a one dimensional soil column, discarding lateral flow and the functioning of the unsaturated zone. Such a system has an exponential response, whose y-intercept or maximum *Aa* is determined by the storage coefficient or porosity and moisture content of the soil. The lower the storage coefficient, the higher the water table will rise due to a unit impulse of precipitation, but the higher the drainage resistance A when the decay rate a is kept constant. For $n \neq 1$, however, the y-intercept of the SG df is either zero or infinite, in which case the storage coefficient cannot be determined exactly.

*n* – as stated denotes the (non-integer) number of linear reservoirs, and allows the SG df to gradually shift from a function that is steeper than exponential towards a Skew-Gaussian and finally a Gaussian shape (figure 2.28c). In this behavior, two physical properties are combined. First, this behavior resembles the effect of the relative position of a location within a system (see section 2.3.5.2). Second, and especially for $n>1$, this behavior resembles the effect of convection and dispersion of water in the unsaturated zone (see sections 2.3.5.3 and 2.3.5.4). With increasing $n$, the influence of the unsaturated zone increases and pulses of precipitation and evapotranspiration are increasingly dispersed and delayed.

In spite of the fact that the variation described above is of course still very limited as compared to the variation found in reality, the SG df generally proves to capture the main dynamic properties of geohydrologic systems well. In [*Von Asmuth et al.,* 2002], the performance of the SG df in modeling the groundwater heads in a dune area at 15 locations was tested, and compared to ARMA(1,s) transfer functions of optimal order. Both approaches gave highly comparable model performances and shapes of the response functions. The performance of the SG df was even slightly better when validation in stead of calibration results were compared, confirming its usability. After this study, the so-called PIRFICT time series model containing the SG df as impulse response was successfully applied to tens of thousands of groundwater head series by von Asmuth and many others.

In [*Bakker et al.,* 2008], it was pointed out that impulse responses according to the scaled gamma distribution (equation (2.68)), Hantush' well function (2.65) and Bruggeman's polder function (2.66) are all special cases of the following function:

$$\theta(t) = At^{\upsilon} \exp(-at - \frac{b}{t}) \tag{2.69}$$

with $\upsilon$ as a fourth parameter. In fact, this holds for the solution to the convection-dispersion equation of equation (2.60) as well, indicating that (2.69) matches the behavior of a variety of systems with convection-dispersion or diffusion type processes. In [*Veling,* 2010], equation (2.69) is named the generalized moving Gaussian distribution, its mathematical properties (moments, cumulants, derivatives) are given, and its performance in approximating responses according to the convection–dispersion equation for several boundary conditions, along with a measured response

curve, is discussed and compared to expansions based on the pure Gaussian distribution. Glover's equation (2.51), however, is not an exact, special case of (2.69), in line with the fact that as discussed before, here the initial response is fixed to the inverse of the storativity. Perhaps, in case of shallow water tables, (2.68) and (2.69) may therefore not be optimal for modeling the early time effects of e.g., precipitation using high frequency data. In [*Veling,* 2010] and elsewhere, distribution functions like (2.68) or (2.69) are referred to as approximate. The term approximate, however, seems not completely adequate, as (2.69) exactly matches the physically-based solution for some schematizations, whereas for others, it is approximate. Next to that, our reference is not the exact solution to some approximate schematization, but reality itself, which any model will always only approximate. In principle, an 'approximate' response function may better fit the real world behavior than an 'exact', physically-based solution. Use of the term behavior, on the other hand, stresses that the spatial structure of the system under consideration does not have to be explicitly defined to model its dynamics. The term behavior is also used in such a sense in the context of systems theory, system identification and equifinality [e.g., *Whitehead and Young,* 1979; *Polderman and Willems,* 1998; *Beven,* 2006].

### 2.4.3   Matching temporal moments with spatial models

A direct consequence of the fact that the spatial structure of a system does not have to be defined in a time series model (which is comparable to the statement made by [*Jury and Roth,* 1990], discussed in section 2.1.3), is that time series models can not be used to make predictions about locations other than those observed. This of course is an important limitation. However, in the following we shall see that the outcomes of time series models can be matched with those of spatial models by transforming ordinary spatiotemporal differential equations into moment-generating differential equations. The concept of moment-generating differential equations has been introduced to subsurface hydrology first by [*Harvey and Gorelick,* 1995], who used it for analyzing solute transport, and independently found by [*Maas,* 1995]. An elaborate treatment of using moments in the field of solute transport can be found in [*Govindaraju and Das,* 2007]. [*Li et al.,* 2005] used moments to characterize the drawdown in pumping tests. The use of moment-generating differential equations and moment matching for groundwater heads, or systems with multiple excitations in general, requires the use of a time series model to estimate and separate the impulse response functions. This was proposed by [*Lankester and Maas,* 1996] and [*Von Asmuth and Maas,* 2001]. Actual practical testing of the method was performed in [*Bakker et al.,* 2008], whereas in [*Bakker et al.,* 2007] moment matching was used for transient modeling of groundwater heads using analytic elements.

In section 2.3.5.2, we have derived a differential equation for the case of one dimensional, horizontal flow. Application of the same procedure to a system with two dimensional flow and a homogeneous, isotropic aquifer yields the following differential equation:
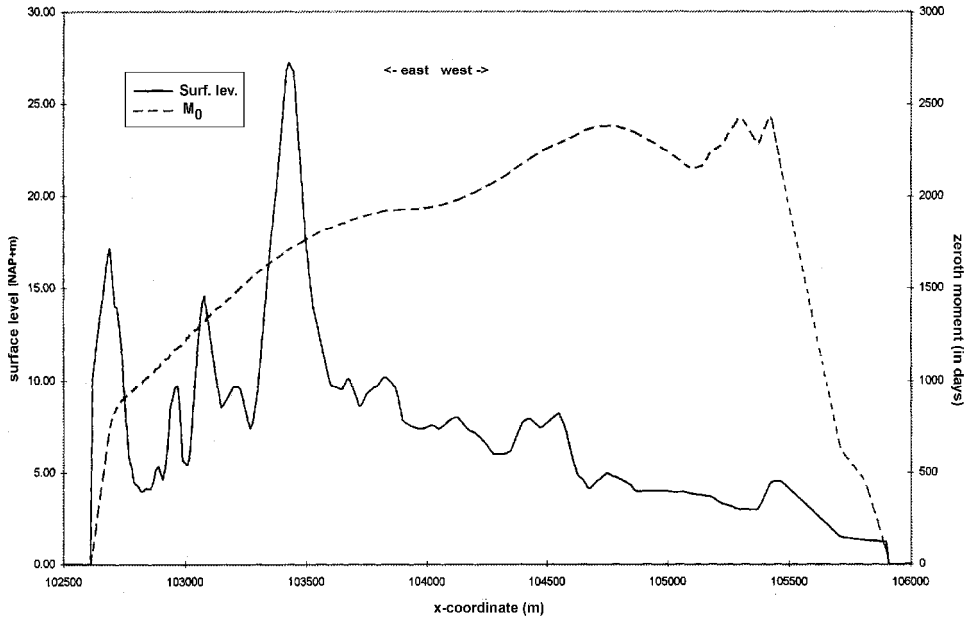
*figure* **2.29:** *East to west cross-section of the dune area near Wijk aan Zee (The Netherlands), showing the surface level and the course of $M_0$ [Van de Vliet & Boekelman, 1998].*

$$KH\nabla^2 h = S\frac{\partial h}{\partial t} - r \qquad (2.70)$$

where $\nabla^2$ is the Laplacian $\dfrac{\partial^2}{\partial x^2} + \dfrac{\partial^2}{\partial y^2}$. If we replace the recharge in (2.70) by a Dirac delta function, the head by definition equals the impulse response function $\theta$:

$$KH\nabla^2 \theta(t) = S\frac{\partial \theta(t)}{\partial t} - \delta(t) \qquad (2.71)$$

In order to obtain the $n^{\text{th}}$ - temporal moment of $\theta$, both sides of equation (2.71) are multiplied with $t^n$ and integrated with respect to time from $-\infty$ to $+\infty$, which gives:

$$\nabla^2 \int_{-\infty}^{\infty} t^n \theta(t)\mathrm{d}t = \int_{-\infty}^{\infty} t^n \frac{S}{KH}\frac{\partial \theta(t)}{\partial t}\mathrm{d}t - \frac{P_0}{KH}\int_{-\infty}^{\infty} t^n \delta(t)\mathrm{d}t \qquad (2.72)$$

This yields, using the definition of temporal moments given in (2.39):

$$\begin{cases} \nabla^2 M_0 = -\dfrac{1}{KH} & , \qquad n = 0 \\[4mm] \nabla^2 M_n = -\dfrac{S}{KH} M_{n-1}, & \qquad n > 0 \end{cases} \tag{2.73}$$

Note that the factor time is not present in equation (2.73). For $n = 0$, equation (2.73) is mathematically identical to the equation of stationary groundwater flow and can be solved using an ordinary groundwater model using the finite difference, finite element or analytic element method. For higher $n$, $M_n$ can be solved using the moments of order $n-1$ as input. [*Van de Vliet and Boekelman,* 1998] used the finite element code TRIWACO to obtain $M_0$ and $M_1$ from a dune area in The Netherlands. In figure 2.29, an east to west cross-section of the modeled dune area is given. The values for $M_0$ obtained from TRIWACO proved to be highly comparable (most differences were less than 5%) to the values obtained from analysis of time series generated by the same groundwater model. The results for $M_1$ were not comparable, perhaps because the retardation in the unsaturated zone was not adequately treated. An advantage of the analytic element method is that its continuous spatial mathematics combine elegantly with the continuous time PIRFICT method of time
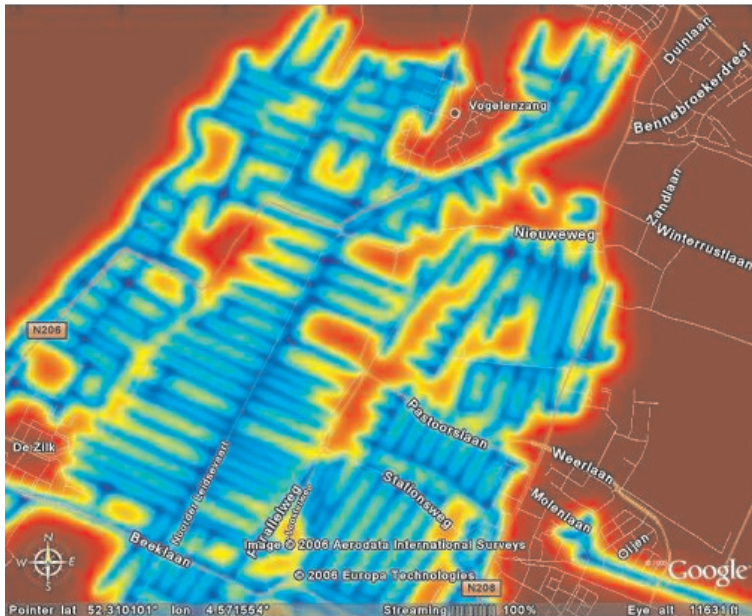
*figure **2.30**: Color image of $M_0$ created with the analytic element method [Bakker et al., 2007] in an area with shallow water tables and a detailed drainage pattern (created with TimML by Frans Schaars, Artesia (www.artesia-water.nl)*

series analysis developed in this thesis. As it is a gridless method, the analytic element method allows for the precise placement of ditches, streams and pumping wells. Furthermore, moments can be computed analytically at the exact location of an observation well, which circumvents scale and discretization problems. The power of the analytic element method becomes especially clear in areas with shallow water tables and a detailed drainage pattern, as is the situation in large parts of the Netherlands (figure 2.30).

# Chapter

**3**

# Time series modeling using continuous response functions

**Adopted from:**

Von Asmuth, J.R., M.F.P. Bierkens and K. Maas (2002) Transfer function noise modeling in continuous time using predefined impulse response functions.
*Water Resources Research*, 38(12), 23_1-23_12.

**Abstract:**

In this chapter, a method of transfer function-noise (TFN) modeling is presented that operates in continuous time and uses predefined impulse response (IR) functions. The resulting class of models is referred to as PIRFICT (Predefined IR Function In Continuous Time). It provides a useful tool for standardized analysis of time series, as it can be calibrated using irregularly spaced data and does not require a model identification phase prior to calibration. In the methodological section, the discrete ARMA (AutoRegressive-Moving Average)-type TFN model of Box and Jenkins [1970] is presented and transformed into continuous time to obtain the PIRFICT model. The discrete ARMA transfer function, which is made up of a variable number of parameters, is replaced by a simple analytic expression that defines the IR function. From the IR function, block response functions are derived that enable the model to handle irregularly spaced data. In the example application, the parameter estimates and performance of the ARMA and PIRFICT model are compared using a data set of 15 piezometers and a simulated series. It was found that the estimated transfer and BR functions of both models follow the same general pattern, although the ARMA transfer functions are partly irregular. The performance of both models proves to be highly comparable for all piezometers.

## 3.1    Introduction

Time series models can be used to model the dynamic behavior of a wide range of variables. Once a time series model is calibrated to a limited set of observations, it is thought to represent the dynamic properties of the system that generated the observations. The model can then be used to make real time forecasts of future values of the output variable and its uncertainty, to predict values at non-observed periods and to quantify and separate the influence of different variables on the output signal. An extensive overview of time series modeling theory was first published by Box and Jenkins [1970]. In univariate time series models, the modeled time series is assumed to be generated by a linear transformation of a random input signal. The input signal is considered to be a discrete white stochastic process, while the characteristic response of the analyzed system is estimated by minimizing the likelihood function, or a least squares criterion, of the noise series. Transfer Function-Noise (TFN) time series models are used whenever a time series can be modeled by linearly transforming one or more deterministic input series, while the residuals of the transfer model are auto-correlated. As these conditions are often met in hydrology, TFN models have been used on many hydrologic variables, including time series of groundwater head [*Tankersley et al.,* 1993*; Gehrels et al.,* 1994*; Van Geer and Zuur,* 1997]. Most time series can also be modeled by mechanistic models that operate according to the physical laws of the analyzed system, for example by a transient distributed groundwater model such as MODFLOW in the case of time series of groundwater head. The use of TFN models is, however, often preferred over the use of mechanistic models, because TFN models often yield more accurate predictions and are less complicated than mechanistic models [*Hipel and McLeod,* 1994]. In addition, because of their stochastic nature, TFN models are well equipped to model the behavior and uncertainty of phenomena that are not well explained by physical laws only. Apart from these favorable properties, there are also drawbacks to the use of TFN models as described by Box and Jenkins. First, TFN models require very specific knowledge from the analyst, as he or she has to be able to perform an iterative model identification procedure. Furthermore, traditional time series models can only operate on time series that are equally spaced in time and are non-interrupted, while the frequency of all input and output variables is coupled and has to be equal.

In time series literature [e.g., *Box and Jenkins,* 1970*; Hipel and McLeod,* 1994*; Ljung,* 1999] often little attention is paid to the fact that time is a continuous phenomenon. Most time series models simply divide time into fixed portions, and only perform calculations with time series that are observed in accordance with their time discretization. For many applications, this does not pose a serious problem, as the frequency of the measuring equipment can easily be adapted to match the demands of the time series model. In hydrologic practice, however, the analyst will often have to make do with the data available. Time series of the groundwater level, for example, are often collected manually and tend to be non-equidistant and contain missing data. [*Bierkens et al.,* 1999] published a possible solution to this problem. In their article they show that an ARX(1,0) model (i.e. an AutoRegressive-eXogenous variable model [*Hipel and McLeod,* 1994]) that operates on a daily frequency can be fitted on irregularly observed time series with the aid of a Kalman filter, because the Kalman

filter estimates the variance of the one-step ahead prediction error or innovation series at non-observed time steps. Although this 'KALMAX' approach to a large extent alleviates this problem for simple exponential systems, it does not offer a satisfactory solution for time series of slow systems with a non-exponential response, e.g., large groundwater systems with thick unsaturated zones [*Gehrels et al.,* 1994]. An adequate daily frequency TFN model for such complex, slow systems would require an ARMAX (i.e. an AutoRegressive-Moving Average-eXogenous variable [*Hipel and McLeod,* 1994]) or full scale TFN model with too many parameters to be calibrated smoothly, as the number of parameters is coupled to the model frequency. Moreover, using the KALMAX approach, all input series still have to be available with equidistant, but small, time steps. Another consequence of time discretization is that the model order and parameter values depend on the frequency of observation. Because of this dependency, a transformation is necessary to couple time series models with different frequencies [*Koutsoyiannis,* 2001]. Problems can also occur with the scaling of possible autoregressive (AR) parameters, as they asymptotically approach a value of 1 when the frequency increases.

The topic of model identification has received much attention in time series literature, as it is an important and difficult part of time series analysis. In general, the approaches and methods proposed by several authors are very diverse. Following [*Knotters,* 2001], they can be subdivided into three main categories. The first category is formed by iterative procedures of identification, calibration and diagnostic checking. Such procedures are normally based on stochastic realization theory and/or statistical hypothesis testing, and include the traditional and still much practiced procedure that was proposed by Box and Jenkins [1970]. A general disadvantage of this approach is that the results of the model identification procedure can be ambiguous [*Hipel and McLeod,* 1994] and the process itself is rather heuristic and can be very knowledge and labor intensive [*De Gooijer et al.,* 1985]. Second, the model order can be identified using automatic model selection procedures, using information measures, Bayesian methods and/or the one-step-ahead prediction error. Among the most used criteria are Akaike's final prediction error (FPE) criterion, information criterion (AIC) and Bayes information criterion (BIC) [*Akaike,* 1970*; Akaike,* 1974*; Akaike,* 1979]. Both AIC and FPE, however, are not consistent but asymptotically overestimate the true order of time series models [*Shibata,* 1976]. Although many of the available measures have desirable characteristics, there is not one definitely superior method, and a coherent and systematic framework which indicates when to use a particular method is lacking [*De Gooijer et al.,* 1985]. A third category of model identification procedures is formed by procedures that use physical insight into the system, in contrast to the above mentioned methods that are based only on the data itself. In the data-based mechanistic modeling methodology [e.g., *Young and Beven,* 1994*; Price et al.,* 2000] physical analysis is combined with statistical model identification. TFN models are identified directly from the data but are only accepted as a reasonable representation of the system if they have a valid physical interpretation. [*Knotters and Bierkens,* 2000] went one step further and used physical analysis to limit the set of candidate TFN models to one. In their study, the ARX(1,0) model is chosen as the most appropriate linear time series model for time series of the groundwater level on the basis of a water balance of the phreatic groundwater zone.

In this paper we present a method of transfer function-noise modeling in continuous time, which uses predefined impulse response (IR) functions. The resulting class of TFN models is referred to as PIRFICT models (PIRFICT stands for Predefined IR Function In Continuous Time), and circumvents a number of limitations of discrete TFN models linked to time discretization and model identification. The paper is organized as follows. First, the theory and basic equations of discrete TFN models are given and subsequently transformed into continuous time to obtain the PIRFICT model. Next, the implementation of the PIRFICT model is elucidated step-by-step, and is summarized at the end of the methodological section. In the example application, the performance of the PIRFICT model is compared to that of traditional ARMA-type TFN models and finally, discussion and conclusions are given. For reasons of simplicity, the theory is developed and illustrated using a single input TFN model with the groundwater level as output series and precipitation surplus as input series, but the equations can be readily extended to include multiple input series.

## 3.2 Methods and theory

### 3.2.1 The discrete ARMA TFN model

For linear, undisturbed phreatic systems that are influenced by precipitation surplus only, the following single input discrete TFN model can be used to model groundwater level fluctuations:

$$h_t = \hat{h}_t + n_t + d \tag{3.1}$$

$$\hat{h}_t = \Theta(\mathrm{B})p_t = \sum_{i=0}^{\infty} \Theta_i p_{t-i} \tag{3.2}$$

$$n_t = \Phi(\mathrm{B})a_t = \sum_{i=0}^{\infty} \Phi_i a_{t-i} \tag{3.3}$$

where

$t$    : discrete time step ($t \in \mathbb{N}$, [-])
$h$    : observed groundwater level, relative to some reference level [L]
$\hat{h}$    : predicted groundwater level, attributable to $p$ and relative to $d$ [L]
$p$    : precipitation surplus [L]
$d$    : level of $h$ without precipitation, or in other words the local drainage level [L]
$n$    : residual series [L]
$\Theta$    : deterministic transfer function [-]
$\mathrm{B}$    : backward shift operator [-], defined as $\mathrm{B}^i p_t = p_{t-i}$
$\Phi$    : noise transfer function [-]
$a$    : discrete white noise process with zero mean [L]

In ARMA TFN models the transfer function $\Theta(B)$ is defined as a fraction, where the numerator is a so-called moving average (MA) function $\omega(B)$, and the denominator an autoregressive (AR) function $\delta(B)$, so that $\Theta(B) = \omega(B) / \delta(B)$ [1]. The weights $\Theta_0, \Theta_1, \ldots \Theta_\infty$ of the transfer function are normally referred to as the impulse response (IR) function, but from a continuous point of view this response is actually the response to an input series with the shape of a block (see section 2.3). To avoid confusion, we will use the term transfer function in the discrete case and reserve the term IR function for the response to an actual instantaneous impulse. In the model identification phase of discrete TFN models, the model order is defined by the choice of a delay time and the number of MA and/or AR parameters, which together form the nodes of the transfer functions in equations (3.2) and (3.3), and define their general structure:

$$
\begin{cases}
\Theta(B) = B^b \dfrac{\omega(B)}{\delta(B)} = B^b \dfrac{\omega_0 + \omega_1 B + \omega_2 B^2 + ..\omega_{nr-1}B^{nr-1}}{1 + \delta_1 B + \delta_2 B^2 + ..\delta_{ns}B^{ns}} \\[4mm]
\Phi(B) = \dfrac{\theta(B)}{\varphi(B)} = \dfrac{1 + \theta_1 B + \theta_2 B^2 + ..\theta_{np}B^{np}}{1 + \varphi_1 B + \varphi_2 B^2 + ..\varphi_{nq}B^{nq}}
\end{cases}
\tag{3.4}
$$

where
$nr$ : number of MA parameters of the transfer model
$ns$ : number of AR parameters of the transfer model
$np$ : number of MA parameters of the noise model
$nq$ : number of AR parameters of the noise model
$b$ : delay time

Thus, the order of an ARMA TFN model is usually specified as $[nr, ns, np, nq, b]$. Although it is often not explicitly mentioned, the modeler also has to specify the time discretization $\Delta t_h$ of the groundwater level series and $\Delta t_p$ of the precipitation series, yet the options are normally restricted in practice, because of data availability and because $\Delta t_h = \Delta t_p$ for ARMA models. We therefore specify the order of an ARMA TFN model as $[nr, ns, np, nq, b][\Delta t_h = \Delta t_p]$. The overall effect of moving-average parameters on the model structure is that they provide the transfer function freedom of shape with respect to real time from $b \cdot \Delta t_p$ to $(b + nr) \cdot \Delta t_p$, while an additional AR parameter causes the tail of the IR function to be exponential (see figure 3.1).

---

[1] Although other authors also use this interpretation and notation, it is incorrect. The correct notation is $\Theta(B) = \delta^{-1}(B)\omega(B)$ (see also paragraph 2.2), which is not really a fraction and only has a symbolic meaning. This point, however, only concerns the notation and not the contents of this chapter.

 In the model identification procedure as proposed by Box and Jenkins, the analyst identifies, by iteratively adding or removing MA parameters, the point from which the remainder of the transfer function can be adequately approximated by an exponential function, which often lies just beyond the peak of the transfer function. The number of parameters necessary to reach this point, however, is dependent on the observation frequency, as an MA parameter is necessary for every time step in between. Therefore, the modeler also has to balance the observation frequency against the number of parameters, in order to get the best model result.
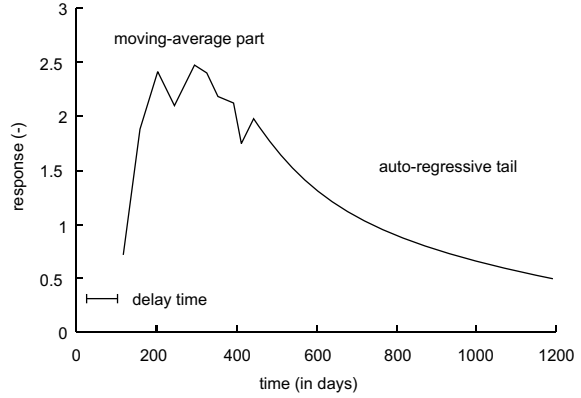


**figure 3.1:** *Example ARMA transfer function of a model with order [10 1 0 1 3][30.4], illustrating the portions of the transfer function that are determined by the MA and AR parameters.*

It can be readily seen that the limitations of discrete TFN models with regard to irregularly observed time series follow directly from Eqs. (1-3). In these equations, time is considered to be a dimensionless index $t \in \mathbb{N}$ , such that each time step is equal to one, regardless of the discretization in real time. Furthermore, the same index is used for $h$ and $p$ , so the sample intervals of the input and output series have to be identical. Consequently, the analyst is forced to lower the frequency of all time series to the lowest one available and thus disregard relevant information about the distribution of the input series in between the time steps.

### 3.2.2 The continuous time PIRFICT model

Most processes modeled with TFN models, such as precipitation and groundwater level fluctuations, are in reality not discrete but continuous. In continuous time, the transformation of a time series with a linear, time-invariant transfer function is given by a convolution integral [*Quimpo,* 1971]. Thus, equations (1-2) can be written in continuous time as:

$$h(t) = \hat{h}(t) + n(t) + d \tag{3.5}$$

$$\hat{h}(t) = \int_{-\infty}^{t} \theta(t - \tau) p(\tau) \mathrm{d}\tau \tag{3.6}$$

Unlike the transfer function of discrete TFN models, the IR function $\theta(t)$ of a convolution integral does not depend on the observation frequency of the input series. It describes the dynamic response of a system to an instantaneous impulse and is time-invariant and an integral property of a specific system [*Jury and Roth,* 1990;

*Maas,* 1994]. The IR function, in hydrologic terms, describes the way in which the water table responds to an instantaneous impulse of precipitation surplus. In that respect it is similar to the instantaneous unit hydrograph used in surface water hydrology [e.g., *Dooge,* 1973]. Equation (3.3) can be transformed to continuous time by replacing the discrete white noise process $a_t$ by a continuous white noise process. The residual series $n(t)$ can then be modeled as a colored stochastic process, which is given by the following stochastic integral:

$$n(t) = \int_{-\infty}^{t} \phi(t - \tau) \mathrm{d}W(\tau)$$

(3.7)

where $W(t)$ is a continuous white noise (Wiener) process [L], with properties $\mathrm{E}\{\mathrm{d}W(t)\} = 0$, $\mathrm{E}[\{\mathrm{d}W(t)\}^2] = \mathrm{d}t$, $\mathrm{E}[\mathrm{d}W(t_1)\mathrm{d}W(t_2)] = 0, t_1 \neq t_2$. To allow for discretely available data in continuous time with non-equidistant intervals we use $t$ in real time, and index $t$ with $i$ instead. A series of $N$ discretely available observations of a continuous process, such as groundwater level fluctuations, can then be written as:

$$h(t_i) = [h(t_0), h(t_1), h(t_2), ........h(t_N)]$$

(3.8)

The drainage level $d$ in equation (3.5) is usually unknown, but can be eliminated from the equations by summing both sides of equation (3.5) for all time steps (from $t = t_0$ to $t_N$), dividing its result by $N$, and assuming that the drainage level is constant, which gives:

$$d = \frac{\sum_{i=0}^{N} h(t_i)}{N} - \frac{\sum_{i=0}^{N} \hat{h}(t_i)}{N} - \frac{\sum_{i=0}^{N} n(t_i)}{N} = \overline{h} - \overline{\hat{h}} - \overline{n}$$

(3.9)

Combining equations (3.5) and (3.9) gives an overall equation for a continuous TFN model of which the components are all centered about their respective temporal averages (the overbar symbols), and in which $d$ is no longer present:

$$\{h(t) - \overline{h}\} = \{\hat{h}(t) - \overline{\hat{h}}\} + \{n(t) - \overline{n}\}$$

(3.10)

In the continuous case, the model order is defined by choosing continuous mathematical functions to represent the IR functions. The mathematical functions are selected on physical grounds, by an iterative procedure of model identification, estimation and diagnostic checking, or with the use of automatic model selection criteria. There are, however, several important differences with the discrete model identification procedure. First of all, when chosen carefully, a continuous IR function

can have a flexible shape and be equivalent to a series of ARMA transfer functions. Second, the model identification procedure is simplified, because the model frequency does not interfere with the model order. Third, a continuous IR function can be objectively chosen as the function that represents the physics of the analyzed system best. A physically-based IR function on the one hand reduces the sensitivity of the model to coincidental correlations in the data, but on the other hand it can reduce the fit if for some reason the physical assumptions prove incorrect. In ARMA TFN models the model order can be chosen on physical grounds [*Young and Beven,* 1994], but with models that contain MA parameters the exact shape of the transfer function cannot.

Here we choose the Pearson type III distribution function (PIII df), with an extra parameter A that adjusts the area, to describe the response of the water table to precipitation surplus:

$$\theta(t) = A \frac{a^n (t-b)^{n-1} \exp\{-a(t-b)\}}{\Gamma(n)} \tag{3.11}$$

with $[a,n,b,A]$ being parameters. The physical basis of the PIII df lies in the fact that it describes the transfer function of a series of coupled linear reservoirs [*Nash,* 1958], the parameter $n$ denoting their number, $a$ equaling the inverse of the reservoir coefficient normally used, and $b$ being the delay time. The extra parameter $A$ is necessary because in the case of equation (3.6), where a precipitation surplus series is transformed into a groundwater level series, the law of conservation of mass does not apply. Mathematically, $n$ is not restricted to integer values, which further increases the flexibility of the PIII df IR function. The PIII df can take shapes gradually ranging from steeper than exponential, via exponential to Gaussian (see figure 3.2). In discrete terms, it can be equivalent to an AR(1) model, but also to all ARMA models of which the weights of the transfer function together form one of the curves in the PIII df family. Knotters and Bierkens [2000] showed that an ARX(1,0) model is a discretized version of a single linear reservoir, and subsequently identified it as the most appropriate linear time series model for time series of groundwater level. Their choice of the ARX model, however, was
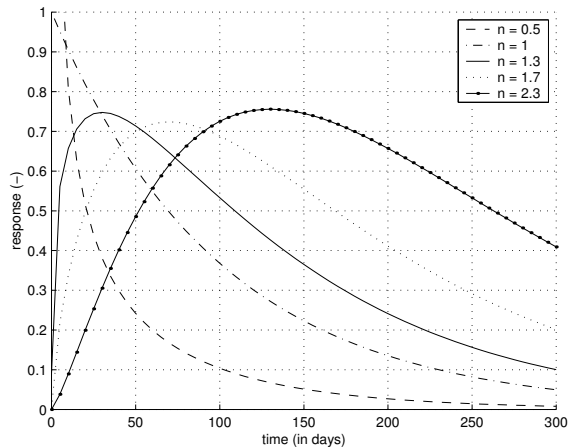


*figure* **3.2:** *A selection of curves of the Pearson type III df* ($n$ = [0.5 1 1.3 1.7 2.3], $A = n \cdot 100$, $a = 0.01$, $b = 0$), *giving examples of the range of shapes the PIII df can take.*

based on a simple physical model of a one dimensional soil column, discarding lateral flow and the functioning of the unsaturated zone. In this sense the PIII df forms an extension to their method, as it includes the ARX model but can for example also describe the combined response of a saturated and layered unsaturated zone.

As the residuals, which are thought to be the output of the noise model, are the result of a variety of causes (e.g., errors in the observations of the input and output series, errors in the model parameters, simplifications or errors in the model concept), it is difficult to make a clear choice of the noise IR function on physical grounds. We therefore choose a simple AR(1) noise model, and rely on diagnostic checks to test its adequacy. The choice of an AR(1) model equals the choice of an exponential IR function in continuous time [e.g., *Box and Jenkins,* 1970*; Chatfield,* 1989]. In order to get an exponential noise model with an appropriate innovation variance function, here we use an IR function of the following form [*Von Asmuth and Bierkens,* 2005]:

$$\phi(t) = \sqrt{2\alpha\sigma_n^2}\,\exp(-\alpha t) \tag{3.12}$$

with the parameter $\alpha$ determining the decay rate of $\phi$ and $\sigma_n^2$ denoting the variance of the residuals.

### 3.2.3 Evaluation, parameter estimation and diagnostic checking

When the IR functions have been chosen, the PIRFICT model is identified and can be evaluated. In the following, we will describe the different steps in the implementation of the PIRFICT model. First, the available time series have to be transformed to continuous series, as most time series are not available as continuous series due to the discrete observation process. Many collected time series, however, can be regarded as the change in the primitive function of some underlying continuous process, as is the case with precipitation surplus:

$$p_t = P(t_i) - P(t_{i-1}) = \int_{t_{i-1}}^{t_i} p(\tau)\mathrm{d}\tau \tag{3.13}$$

When precipitation surplus data are only available at discrete intervals, the continuous series $p(\tau)$ cannot be reconstructed exactly, but it can be approximated by assuming that the distribution of $p$ is uniform during the period $t_{i-1}$ to $t_i$ [*Ziemer et al.*, 1998]. Equation (2.36) can then be written as:

$$p(\tau) = \frac{p_t}{t_i - t_{i-1}}, \quad t_{i-1} \geq \tau > t_i \tag{3.14}$$

In this way, the higher the frequency of observation, the more $t_i - t_{i-1}$ approaches

zero and the better the approximation of $p(\tau)$ (with units $[LT^{-1}]$) will be. Note that the error made by this assumption will vary in time when non-equidistant precipitation surplus observations are used, which will introduce non-stationarity in the model residuals. However, we assume that, when the time steps are not too large and irregular, this effect is small when compared to the other sources of model error and can therefore be neglected.
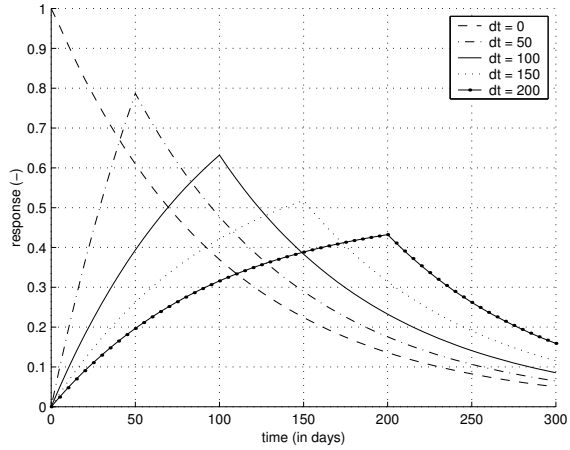


figure 3.3: Example BR functions of a single linear reservoir ( $\Delta t = [0, 50, 100, 150, 200]$ , $n = 1$ , $A = 100$ , $a = 0.01$ ), illustrating the effect of different time steps on the shape of the BR function.

Second, as equation (3.6) and (3.7) contain time from $-\infty$ , we have to define an initial value to be able to evaluate the equations with observations starting from $t = 0$ . Here we define the initial value of $p$ from $t = -\infty$ to $t = 0$ to be $\overline{p}$ , the temporal average of the precipitation surplus series used for the simulations (but when available a historical average could also be used). With the aid of equation (3.14), the transfer model (equation (3.6)) can now be evaluated using the block response (BR) function. The BR function $\Theta(t)$ can be obtained by convoluting the IR function with a 'block' of precipitation surplus with unit intensity over a period $\Delta t$ , which equals:

$$\Theta(t) \;=\; \int\limits_{t-\Delta t}^{t} \theta(\tau) \mathrm{d}\tau \tag{3.15}$$

To make the BR function equivalent to the discrete transfer function, it has to be divided by $\Delta t$ in order to scale it to a block input with unit area. In figure 3.3 some example BR functions of a single linear reservoir are plotted for different time steps. The predicted groundwater level series $h^*$ can be obtained by adding the responses of all 'blocks' of precipitation. Because $\Theta$ is a continuous function, $h^*$ is also continuous and for every observation of $h$ a sample of the residual series $n$ is obtained.

Next, the noise model (equation (3.7)) is evaluated in order to obtain a series of innovations $\nu$ . To evaluate the noise model without having to use a Kalman Filter (which is computationally expensive) we will derive a direct relation between the residuals $n$ and the innovations $\nu$ . Consider the innovation series $\nu$ as the non-equidistantly sampled change in the solution to the stochastic integral describing the

residual series:

$$v(t) = \int_{t-\Delta t}^{t} \phi(t-\tau)\mathrm{d}W(\tau) \tag{3.16}$$

When the noise IR function $\phi$ equals equation (3.12), equation (3.7) can be written as:

$$n(t) = \exp(-\alpha\Delta t)n(t-\Delta t) + \int_{t-\Delta t}^{t} \sqrt{2\alpha\sigma_n^2}\exp\{-\alpha(t-\tau)\}\mathrm{d}W(\tau) \tag{3.17}$$

which is known as an Ornstein-Uhlenbeck process [*Uhlenbeck and Ornstein,* 1930]. By combining equations (3.16) and (3.17), $v$ can be calculated from the available data as:

$$v(t) = n(t) - \exp(-\alpha\Delta t)n(t-\Delta t) \tag{3.18}$$

Subsequently, the parameter set $\Psi = [A, a, n, b, \alpha]$ contained in the IR functions has to be estimated from the data. By adjusting the value of the parameters the time series model can be calibrated on a set of observations of the groundwater level series $h(t_i)$ by minimizing a certain objective function. Bierkens et al. [1999] use a log-likelihood function [*Schweppe,* 1973] as objective function for the innovations of a Kalman filter with an observation error variance of 0, to get a maximum likelihood estimate of the model parameters (under the assumption that the innovations are Gaussian). However, for reasons of efficiency, we seek an objective function that can be expressed in terms of individual innovations. From (3.12) and (3.16) we have:

$$\sigma_v^2(\Delta t, \Psi) = \{1 - \exp(-2\alpha\Delta t)\}\sigma_n^2 \tag{3.19}$$

With (3.19) the likelihood function can be approximated by the following weighted least squares criterion [*Von Asmuth and Bierkens, 2005*]:

$$S\{\Psi \mid h(t_i)\} = \sum_{j=1}^{N} \frac{\sqrt[N]{\prod_{i=1}^{N}\{1 - \exp(-2\alpha\Delta t_i)\}}}{1 - \exp(-2\alpha\Delta t_j)} v^2(t_j, \Psi) \tag{3.20}$$

with $v^2(t_j, \Psi)$ calculated from the residual series using equation (3.18). From equation (3.20), a Jacobian matrix can be easily obtained, so it can be minimized with respect to $\Psi$ using a Levenberg-Marquardt method. This makes the parameter

estimation problem much more efficient than using a Kalman filter in conjunction with a log-likelihood function and some global optimization algorithm.

Finally, the accuracy and validity of the model results is checked by examining the auto- and crosscorrelation functions of the innovations, the covariance matrix of the model parameters and the variance of the IR functions. The covariance matrix of the parameters $C(\Psi)$ is estimated using the Jacobian matrix obtained from the calibration routine and $\sigma_v^2$. From the covariance matrix, also the correlation between the parameters is calculated. The variance of the IR function $\theta$ can be obtained by:

$$\sigma_{\theta(t)}^2 = (\frac{\partial \theta(t)}{\partial A})^2 \sigma_A^2 + (\frac{\partial \theta(t)}{\partial a})^2 \sigma_a^2 + (\frac{\partial \theta(t)}{\partial n})^2 \sigma_n^2 + 2(\frac{\partial \theta(t)}{\partial A})(\frac{\partial \theta(t)}{\partial a})C(A,a) + ....$$
$$2(\frac{\partial \theta(t)}{\partial A})(\frac{\partial \theta(t)}{\partial n})C(A,n) + 2(\frac{\partial \theta(t)}{\partial a})(\frac{\partial \theta(t)}{\partial n})C(a,n)$$

(3.21)

Assuming a normal distribution, a confidence interval for $\theta$ can be plotted as +/- $2\sigma$. As in discrete TFN models, serious model inadequacy can be detected by examining the autocorrelation function of the innovation series $v$ (which indicates whether the white noise assumption holds) and the crosscorrelation function of $v$ and the input series $p$ (which indicates whether there are still patterns left in the innovation series that could be explained by the input series). The autocorrelation and crosscorrelation functions at lag $k$ are defined in the same way as in discrete TFN models, but because of the non-equidistant sampling a tolerance around lag $k$ of $\pm 0.5 k$ is implied. A similar approach is used in the field of geostatistics to obtain spatial variograms [*Journel and Huijbregts,* 1978]. An example of the autocorrelation functions of an ARMA and a non-equidistant PIRFICT model is given in figure 3.4.
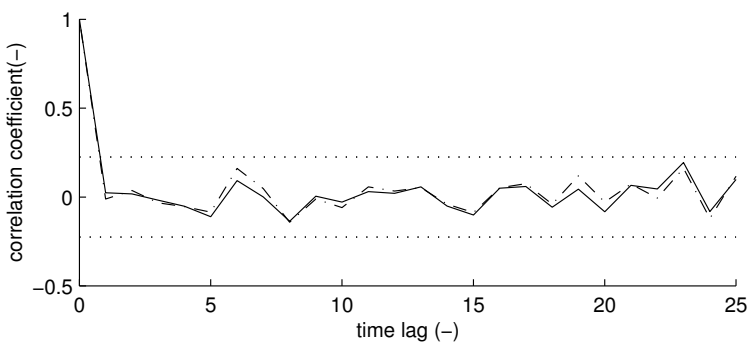


*figure 3.4:* *Autocorrelation functions of an [10 0 1 1 0][30.4] ARMA TFN model and a [30.4 30.4] PIRFICT model. The dotted lines denote the 95% confidence interval. From the figure it can be seen that for both models, the autocorrelation functions are very similar and the white noise assumption is valid.*

### 3.2.4    Summary of method

In summary, the method described above consists of the following steps. First, for every input series an IR function is chosen, which in principle can be any continuous function or combination of functions, but in practice will often be based on the physical laws of the analyzed system. The input series are assumed to be uniformly distributed in between the time steps and transformed into continuous series using (3.14). The transfer convolution integral (equation (3.6)) can now be evaluated using block response functions for every block pulse, to obtain a continuous prediction of the output series. Using (3.18), a sample of the innovation series is obtained for every observation of the output series, whether or not equidistant. An estimate of the model parameters is made with the aid of a Levenberg-Marquardt algorithm, which numerically minimizes weighted least squares criterion (3.20) that is based on the likelihood function of the innovations. Finally the accuracy and validity of the model is checked using the auto- and crosscorrelation functions of the innovations, the covariance matrix of the model parameters and the variance of the IR functions.

## 3.3    Example application

### 3.3.1    Set-up and data set

The example application is devised to illustrate the performance of the PIRFICT model in practice by comparing its calibration and validation results and parameter estimates with those of ARMA TFN models fitted on the same data. For this purpose, groundwater level series from 15 piezometers are selected, all lying in a dune reserve in the province North-Holland, The Netherlands, near the town of Egmond. Piezometers are selected at locations that are little disturbed by the groundwater abstraction in the area [*Rolf and Lebbink,* 1998], thus allowing the groundwater level series to be modeled with precipitation surplus as a single input series. The groundwater level in all piezometers is well observed in the same period, with observations taken manually about the 14th and 28th of every month, so that the model results are little influenced by a difference in length of the calibration period or the number of observations in that period. For the calibration process, observations of the groundwater level from 1-1-1990 until 1-1-2001 are used and observations of the precipitation surplus (precipitation minus potential evaporation) starting from 1-1-1987. The precipitation series is available on a daily basis and is observed by the Provincial Water Company of North-Holland in the dunes near the town of Castricum, whereas the daily potential evaporation series originates from a station of the Royal Dutch Meteorologic Institute near the town of De Kooy.

As the performance of models may be dependent on the properties of a data set, the example application is not solely focused on assessing the performance of both models for this specific data set, but also on clarifying the mechanisms which influence model performance. First of all, we will illustrate the performance of a range of ARMA TFN models, their dependence on user defined choices such as model order and time discretization, and the dependence of the results of PIRFICT models fitted on the same data from an example series. For this purpose, piezometer 19AZW246_1 is selected. At this location the water table showed the slowest response to precipitation

surplus, which makes it especially suited for illustrating the beneficial properties of the PIRFICT model. Next, we will identify a single ARMA model order for all 15 piezometers, and compare the calibration results with those of the PIRFICT model. For six of the piezometers, observations that were taken before the calibration period are available, and for these series a validation is also performed. Unfortunately, the observations in this period are only available on a quarterly basis and not suited for a cross-validation. Finally, a small simulation experiment is performed to confirm some hypotheses based on results from the real world data.

### 3.3.2    Single series

In order to be able to calibrate an ARMA model on a specific time series, the model order has to be identified first. As Knotters and Bierkens [2000] have shown, the response of the saturated zone to precipitation, under the assumption of vertical flow, shows exponential decay and therefore equals an ARX model. However, the first ordinates of the response function can also be non-exponential and the response can be delayed when horizontal flow and the functioning of the unsaturated zone are taken into account.  Therefore, we only consider ARMA models with order $[nr,1,0,1,b][\Delta t_h = \Delta t_p]$, i.e. models in which the transfer model contains one auto-regressive and a variable number of moving-average parameters, in which the noise model one auto-regressive parameter and there is a delay time. The number of moving average parameters $nr$ is chosen to be 5, 10 and 20, also dependent on the time discretization, which covers the range of probable model orders. The time discretization is chosen in accordance with the observation frequency and such that the response of the system can be described adequately with a reasonable number of parameters. In the case of these time series this means that $\Delta t_h$ = 15.2, 30.4 or 60.8 days, so that exactly 24, 12 or 6 groundwater level observations are available every year. As ARMA models require equidistant time series for their calibration and the groundwater level measurements are taken approximately the 14th and 28th of every month, the groundwater level series has to be made equidistant. In this application, an equidistant series starting from 1-1-1990 is obtained by linearly interpolating the groundwater level between the two nearest observations. Observations of the precipitation surplus are available on a daily basis, and an equidistant series is obtained by taking the sum of the daily series for every time interval.

As both the input and output time series are modified by the interpolation and resample operations required for ARMA models, PIRFICT models are fitted for every data set obtained this way. To make a distinction between the results thus obtained, the PIRFICT models are tagged as $[\Delta t_h, \Delta t_p]$ and as (non)equidistant, which in this case does not indicate a difference in model order but a difference in the data set used. The calibration results of the different ARMA models and the PIRFICT model are given in table 3.1, expressed in the form of a number of criteria. In the table, first the root mean squared error (RMSE) and root mean squared innovation (RMSI) are given, which form a measure for the error of the transfer model and the variance of the noise process, respectively. Because the variance of the groundwater level series and

*table 3.1:* *Calibration results for ARMA TFN models of different order and PIRFICT models calibrated on the same data (piezometer 19AZW246_1).*

| | ARMA TFN models | | | | |
|---|---|---|---|---|---|
| $[nr,ns,np,nq,b]$ $[\Delta t_h = \Delta t_p]$ | [5 1 0 1 0] [30.4] | [10 1 0 1 0] [30.4] | [20 1 0 1 0] [30.4] | [20 1 0 1 0] [15.2] | [5 1 0 1 0] [60.8] |
| RMSE (cm) | 17.82 | 16.39 | 17.08 | 16.58 | 17.33 |
| RMSI (cm) | 8.61 | 8.42 | 8.06 | 7.04 | 10.41 |
| EVP (%) | 82.35 | 85.08 | 83.79 | 84.83 | 83.59 |
| | PIRFICT model | | | | |
| RMSE (cm) | | 16.25 | | 16.41 | 17.43 |
| RMSI (cm) | | 9.03 | | 7.27 | 10.52 |
| EVP (%) | | 85.32 | | 85.14 | 83.39 |

the residual series can be influenced by interpolation and resampling operations, also the explained variance percentage (EVP) is given. The EVP is defined as:

$$EVP = \frac{\sigma^2_{h(t)} - \sigma^2_{n(t)}}{\sigma^2_{h(t)}} *100\% \tag{3.22}$$

A logical way of comparing model results seems to be the use of automatic model order selection criteria such as AIC and FPE. However, both criteria use the innovation variance or their likelihood for determining the best model order, which are influenced by the sample frequency, so these criteria cannot be used to compare models with different sample frequencies or data sets. From the results in table 3.1 it can first of all be seen that the differences between the fit of both models are small. Although the PIRFICT model does show the lowest RMSE and highest EVP, the variation caused by differences in observation frequency and model order is larger than the differences between both model types. The optimal result for the ARMA model appears to be given by the [10 1 0 1 0] [30.4] model, so we can identify this for the moment as the optimal model order. As expected, the RMSI of both models varies with the time lag between the observations of the output series. When $\Delta t_h = 15.2$, the RMSI lies in the order of 7 to 8 cm, whereas the RMSI is 9 to 10 cm when $\Delta t_h = 60.8$. Furthermore, the results show that the RMSI of the ARMA TFN model decreases when the number of parameters increases, while the RMSE shows an optimum for the [10 1 0 1 0] model for $\Delta t_h = 30.4$. This phenomenon could well be attributed to the fact that the models are calibrated by minimizing the RMSI (which is linked to the noise part of the model) rather than the RMSE (which is linked to the transfer model). Because of this, adding extra parameters and thereby overfitting the data will result in a gradually improving fit of the noise model, but can at the same have a negative effect on the fit of the transfer model. Overfitting behavior, or the generally higher
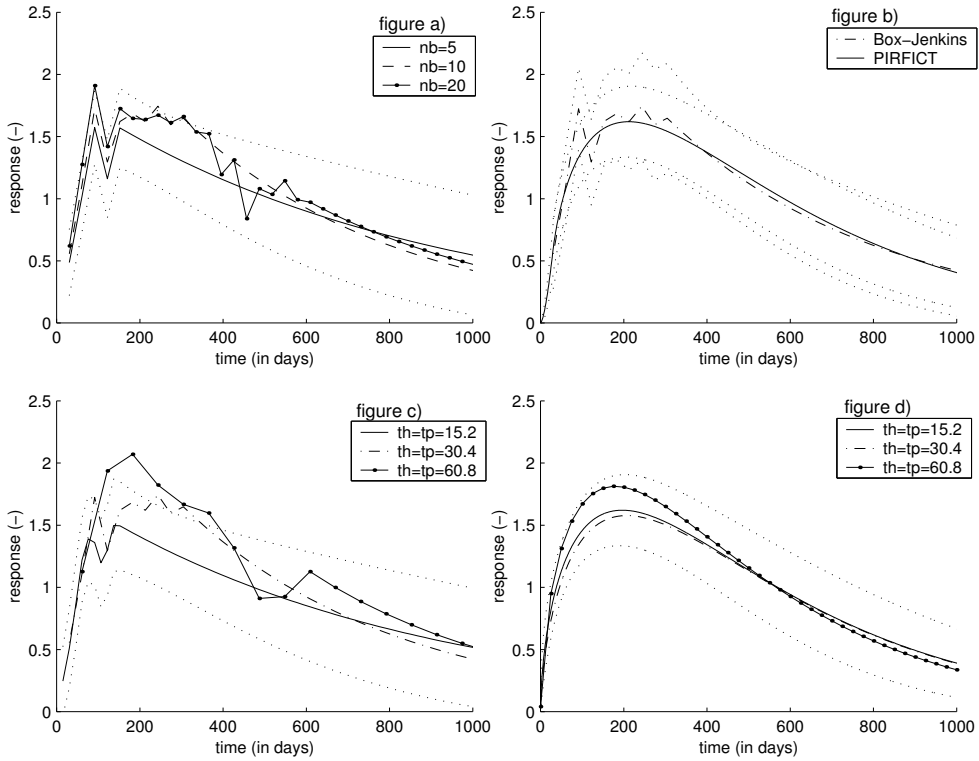
*figure* **3.5:** *BR functions of several TFN models of piezometer 19AZW246_1. a) Three [nr 1 0 1 0] [30.4] ARMA models with nr = 5, 10 and 20. b) An [10 1 0 1 0][30.4] ARMA model and a PIRFICT model using the same data. c) three [10 1 0 1 0][ $\Delta t$ ] ARMA models with $\Delta t$ = 15.2, 30.4 and 60.8. d) Three PIRFICT models with $\Delta t_h = \Delta t_p$ = 15.2, 30.4 and 60.8. The dotted lines denote the 95% confidence interval of the nb = 5 model in a), both models in b), and the $\Delta t$ = 30.4 model in c) and d).*

number of parameters of an ARMA model, could also explain the fact that the RMSI of the ARMA TFN model is lower than that of the PIRFICT model, while its RMSE is higher.

The parameter estimates of the different models can best be compared by plotting the estimated BR functions. In figure 3.5, the transfer functions of three [ $nr$ 1 0 1 0][30.4] ARMA models are plotted, with $nr = 5$, 10 and 20. The results appear to be significantly influenced by the model order, as the transfer functions of the $nr = 10$ and 20 model lie partly outside the confidence interval of that of the $nr = 5$ model. The number of MA parameters for the case $nr = 5$ is apparently too low to model the slow response of the system well. According to this figure and the RMSE, in this case 10 MA parameters just about suffice to model the first part of the response of the system, while the remainder of the response function can be described adequately by a single AR parameter. From figure 3.5b it can be seen that the response of the

PIRFICT model follows the response of the [10 1 0 1 0] ARMA model rather closely. The results of both models should therefore be highly comparable, as the only difference in the transfer functions is the irregular pattern of the ARMA model around the smooth curve of the PIRFICT model. In figure 3.5c the order of the ARMA model, in the traditional sense, is kept constant while the sample frequency is varied, resulting in three $[10\ 1\ 0\ 1\ 0][\Delta t]$ ARMA models with $\Delta t$ = 15.2, 30.4 and 60.8. As expected, the parameter estimates of the ARMA model prove to be also significantly influenced by the sample frequency, which interferes with the model order when time is used as a dimensionless index. Finally, in figure 3.5d, the BR functions of three PIRFICT models are plotted, calibrated on the same data as the ARMA model (i.e. $\Delta t$ = 15.2, 30.4 and 60.8). From the results, the first two BR functions prove to be almost identical, while the model apparently has difficulties estimating the first part of the BR function correctly for $\Delta t$ = 60.8, although it does not differ significantly. This effect is probably caused by the increasing time interval and/or decreasing number of observations, as the ARMA [60.8] model estimates the transfer function in about the same way.

### 3.3.3 Multiple series and validation study

For a broader comparison between the ARMA and PIRFICT TFN model, both models are calibrated on time series from 15 piezometers with observations ranging from 1-1-1990 until 1-1-2001. For reasons of comparability and objectivity, no Box and Jenkins style iterative model identification procedure is performed for each separate
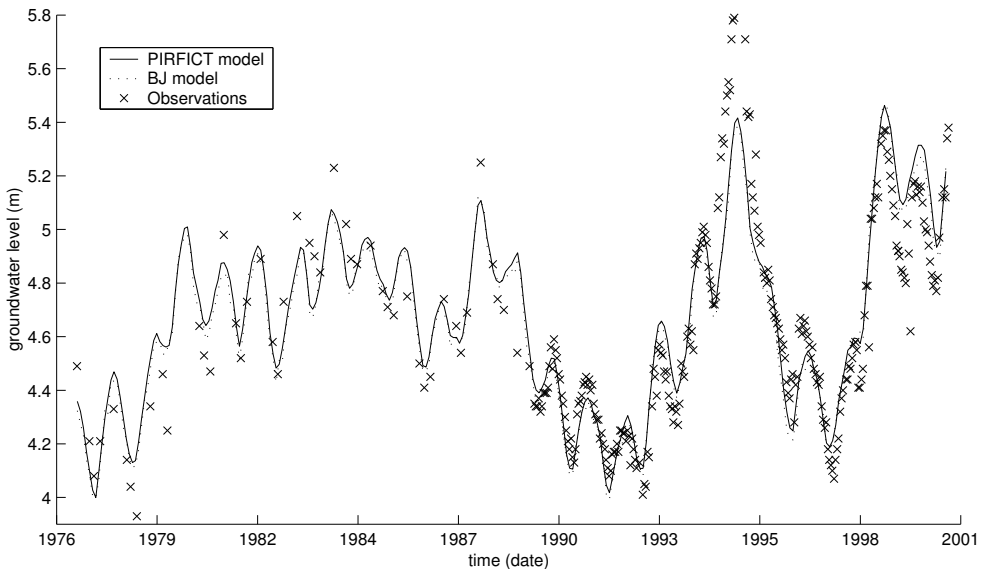


*figure 3.6: Simulations from a [10 1 0 1 0][30.4] ARMA model and a [30.4 30.4] PIRFICT model calibrated on observations of the water table depth of piezometer 19AZW246_1 between 1990 and 2001. From the figure it can be seen that the simulation results in both the calibration and validation period are similar.*

piezometer to obtain the most appropriate model order (in our case defined by $nr, b$ and $\Delta t_h$). Instead, a [10 0 1 1 $b$][30.4] ARMA model is calibrated on all piezometers, which proves to give the transfer function just about enough MA parameters to model the slowest response of the 15 piezometers well (see section 3.2). For all models, a delay time is applied, which has shown to improve both the calibration and validation results. As expected, the improvements are greatest for the PIRFICT model because of its predefined shape. On the basis of a manual model identification procedure, the delay time for the 15 piezometers is chosen to be $b = $ [12 6 0 0 0 3 1 13 13 6 6 8 1 5 0] days. It is applied by shifting the precipitation series along the time axis. Shifting the input series rather than the response function makes it possible to apply delay times which are smaller than the discrete-time interval, and by doing so the delay time could be kept equal for the ARMA and PIRFICT model. The PIRFICT model is first calibrated using the same data as the ARMA model, which is therefore denoted as equidistant [30.4 30.4], and second using non-equidistant [30.4 1] data. The results should therefore show the combined effect of the interpolation operations needed to make the data equidistant and the effect of summing the daily precipitation surplus series into 30.4 day totals. In addition, for six of the piezometers, a validation is performed on the observations taken before 1-1-1990. A time plot of the available observations in both the calibration and validation period for piezometer 19azw246_1 is shown in figure 3.6, along with predictions of a [10 1 0 1 0][30.4] ARMA model and a PIRFICT model using the same data. The figure clearly shows that the observation frequency in this period is not equal to that of the calibration period, but has been changed at the end of 1989 from 4 times a year into 24 times a year, on average. Because the frequency of the observations is much lower than the frequency of the predictions, the predicted values are linearly interpolated to match the dates of the observations, and not vice versa as in the calibration routine. For the PIRFICT model, the simulated values do not have to be interpolated, as the simulated water table depth is continuously defined.

The average results of all 15 piezometers are summarized in table 3.2. As in table 3.1, the RMSE, RMSI, and EVP are given, along with the validation RMSE. Again, the differences are small as the average RMSE of the [10 1 0 1 $b$][30.4 30.4] ARMA model, with a difference of only 0.8 millimeter, almost equals the RMSE of the non-equidistant [30.4 1] PIRFICT model. In 7 out of the 15 piezometers the fit is

*table 3.2: Average calibration and validation results of the ARMA and PIRFICT model for 15 and 6 piezometers respectively.*

|  | **ARMA** [10 1 0 1 0] [30.4] | **PIRFICT** [30.4 30.4] equi. | **PIRFICT** [30.4 1] non-equi. |
|---|---|---|---|
| RMSE (cm) | 12.42 | 12.56 | 12.50 |
| RMSI (cm) | 7.20 | 7.65 | 7.89 |
| EVP (%) | 88.66 | 88.37 | 88.36 |
| V-RMSE (cm) | 15.76 | 15.64 | 15.38 |

slightly better when the PIRFICT model is used, whereas 8 piezometers show a better result with the ARMA model. At the same time, however, the RMSI of the ARMA model is lower for all piezometers, whereas the opposite is true for the validation RMSE. The lower RMSI is probably to a certain extent caused by the interpolation
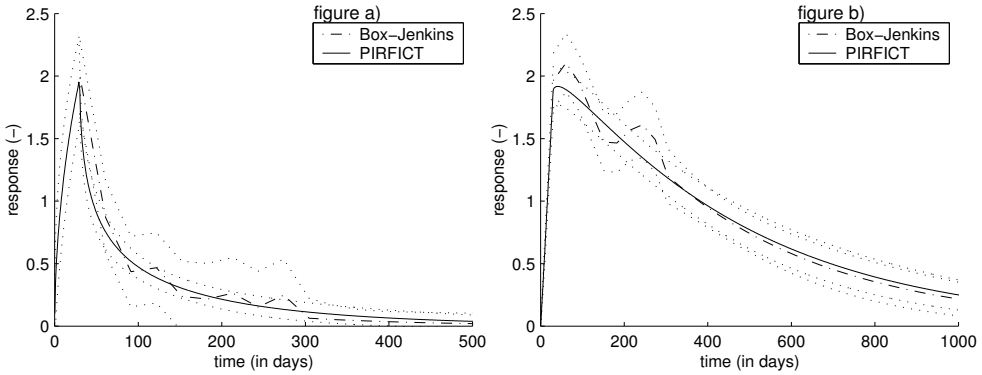
*figure 3.7:* *Transfer functions of a [10 1 0 1 b][30.4] ARMA TFN model and a PIRFICT model calibrated on the same data from piezometers 19AZL5038_1, shown in a), and 19AZW195_1 shown in b). The dotted lines denote the 95% confidence intervals.*

operations that were carried out on the data to make it equidistant, which tends to smooth high-frequent variations of a signal. This is corroborated by the fact that the PIRFICT model yields a RMSI that is 2.4 millimeter smaller for the equidistant data than for the non-equidistant data. However, the RMSI of the ARMA model is on average 4.5 millimeter smaller than that of the [30.4 30.4] PIRFICT model while the VRMSE of the ARMA model is higher. This, again, points to overfitting behavior of ARMA models. The ARMA model is able to produce a lower value of the objective function using the same data, but apparently the correlations fitted are to some extent coincidental as they are accompanied by a higher VRMSE. When the response functions of both models are compared, the smooth continuous BR functions prove to fit the general shape of the more irregular discrete transfer functions well for all piezometers, indicating that the PIRFICT model can adequately model phreatic systems with a relatively fast, intermediate and slow response (figure 3.7a and b and figure 3.5b).

### 3.3.4 A small simulation study

Because we were not able to perform a cross-validation and cannot determine the 'true' response of the real world phreatic systems, there is only indirect evidence that the irregular shape of the ARMA transfer function is to some extent coincidental or due to overfitting behavior. In order to have more direct evidence, we performed a short study in which we simulated a groundwater level series by transforming a daily frequency precipitation surplus series in a synthetic system with a known response (a PIII IR function with $A = 1500$, $a = 0.002$, $n = 1.5$, $b = 0$). To the series thus obtained, we added a residual series that was generated by convoluting a 'discrete white noise' series of normally distributed random numbers with a known, exponential noise IR function ($\alpha = 25$, $\sigma_\nu = 2.5$ cm on a daily frequency). From the series thus obtained the parameters that we should have retrieved using the TFN models are known exactly. The parameter estimates of a [10 1 0 1 0][30.4] ARMA model and a PIRFICT model are depicted in figure 3.8 as BR functions, along with the true BR function of the synthetic

system. The estimated parameters of the noise model are [ $\alpha = 27.9$, $\sigma_v = 2.43$ cm] and $\alpha = 29.6$, $\sigma_v = 2.39$ cm] for the PIRFICT and ARMA model respectively. As was the case with the real world data, the ARMA transfer function shows a partly irregular pattern and in this case even falls partly outside the confidence interval of the PIRFICT response. When comparing the estimates of both models with the true response function the deviation of the continuous BR function is less than that of the discrete transfer function, as is the case for the parameters of the noise model.



*figure 3.8: Transfer functions of a [10 0 1 1 0] [30.4] ARMA model and a PIRFICT model calibrated on the same simulated series. The dotted lines denote the 95% confidence intervals. The figure illustrates that the partly irregular shape of the ARMA transfer function is due to overfitting the data with too high a model order and not conform the 'true' response of the synthetic system.*

## 3.4    Discussion and conclusions

The method presented has shown to circumvent a number of the limitations of discrete ARMA Transfer Function-Noise models. First, the PIRFICT model can be calibrated on data at any frequency available, because it operates in a continuous time domain and the time steps of the output variable are not coupled to the time steps of the input variables. Thus, also the frequency of the input series can be irregular. Second, compared to the combined ARX model and Kalman filter, the PIRFICT model offers a further extension of the possibilities of calibrating TFN models on irregularly spaced time series, because the shape of the transfer function is not restricted to an exponential. Third, using the PIRFICT model, the model identification process is simplified because the model frequency does not interfere with the model order and parameter values, and the flexibility of a single continuous IR function can be such that it comprises a range of ARMA transfer functions. Furthermore, the model can be readily identified using physical insight. Although the method is presented in the form of a single input TFN model for time series of groundwater head, it is in fact quite general, and could be used for a variety of single or multi input, (non) hydrologic problems (i.e. for the same purposes that ARMA TFN models are used, see introduction). The continuous time approach probably offers most advantages in cases where the above mentioned limitations occur, where the analyst is interested in the functioning of a system on different or small time scales, in the automated analysis of large quantities of time series, or in assessing time invariant response characteristics of systems. While the PIRFICT model probably performs best when the dynamic behavior of a system can be expressed in the form of a simple analytic formula, this should not pose a restriction to the use of the model, as the shape of the IR function can also be generalized by using sums of  PIII df (or other) functions.
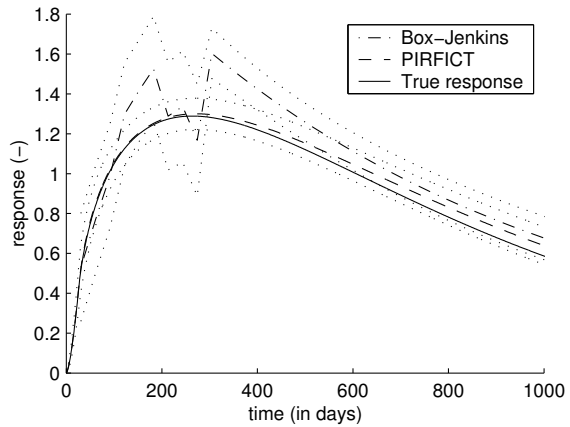
It was shown in the example application that the PIRFICT model yielded comparable parameter estimates when the sample interval was varied, as long as the series of observations was sufficiently long and dense. As expected, the results of the ARMA TFN model proved to be influenced by user defined choices such as model order and time discretization. As for model performance, the PIRFICT model gave results comparable to those of the discrete ARMA models for all series analyzed. A first indication of this can be found in the fact that the estimated BR and transfer functions of both models show the same general behavior (but the ARMA transfer function is partly irregular while the PIRFICT BR function is smooth). Second, the average calibration RMSE of the non-equidistant [30.4 1] PIRFICT model of the 15 piezometers was on average only slightly higher than that of the ARMA TFN model, whereas the validation RMSE was actually slightly lower. Although the differences are small, they could well be explained by the different structure of the transfer function of both models, which influences their overfitting behavior. As the shape of the transfer function of ARMA models is partly free, the model is also free to fit coincidental crosscorrelations between the input and output series, which will result in a lower calibration RMSE, a higher validation RMSE and a partly random pattern of the transfer function. Because of the use of a predefined IR function, the estimates of the PIRFICT model are forced to follow a certain physical behavior. This on the one hand makes the model less sensitive to overfitting behavior (next to its generally lower number of parameters) and can therefore yield better estimates, but can on the other hand also negatively influence the results when the physical assumptions prove incorrect or are to rough. These hypotheses were corroborated by the results of the simulation experiment, in which the response of the synthetic system was a PIII distribution and therefore equal to that of the predefined IR function in the PIRFICT model. For all series analyzed, the PIII df, or a system of serially coupled linear reservoirs, has shown to model the response of the water table to precipitation surplus adequately.

# Chapter

4

# A continuous noise model for autocorrelated data of irregular frequency

**Adopted from:**

**Abstract:**

In this chapter, the background and functioning of a simple but effective continuous time approach for modeling irregularly spaced residual series are presented. The basic equations were published earlier in [*Von Asmuth et al.,* 2002] where they were used as part of a continuous time transfer function noise (TFN) model. It is shown that the methods behind the model are build on two principles: first, the fact that the equations of a Kalman filter degenerate to a form that is equivalent to 'conventional' autoregressive moving average (ARMA) models when the modeled data are considered to be free of measurement errors. This assumption, in comparison to the 'full' Kalman filter also yields a better prediction efficiency [*Ahsan and O'Connor,* 1994]; second, the mathematical equivalence between discrete-time AR parameters and continuous exponentials and the point that continuous time models provide an elegant solution for modeling irregularly spaced observations [e.g., *Harvey,* 1989]. Because simple least-squares methods do not apply in case of modeling irregular data, a sum of weighted squared innovations (SWSI) criterion is introduced and derived from the likelihood function of the innovations. In an example application, it is shown that the estimates of the SWSI criterion converge to maximum likelihood estimates for larger sample sizes. Finally, we propose to use the so-called innovation variance function as an additional diagnostic check, next to the well-known auto and crosscorrelation functions.

## 4.1    Introduction

One of the immediate consequences of the stochastic nature of natural processes is that the predictions of a deterministic model will never match a set of observations completely. The difference between the model predictions and the observed time series forms a time series of its own, called the 'residuals'. Residuals are 'caused' by errors in the observation process, errors in the model parameters, simplifications or errors in the model concept and/or numeric errors when evaluating the model equations. Furthermore, the value of a model residual at a certain time instant is often correlated with its value at earlier time instants, so residuals cannot simply be modeled as a set of independent Gaussian deviates.  Explicitly modeling the behavior of the residuals of a transfer or deterministic model can have several purposes. First, when a noise model is fitted to the residuals, it can be used for stochastic simulation. This is especially useful in case one is interested in the probability of extremes, as such probabilities are underestimated when using only the deterministic model [e.g., *Knotters and Van Walsum, 1997*]. Second, because of the autocorrelation in the signal, the noise model can also be used to yield predictions of the residuals at unobserved time steps, either for smoothing, forecasting or updating purposes. Such applications, which are widely used in the meteorologic sciences and are often referred to as data assimilation [*McLaughlin, 1995; Kalnay, 2002*], make optimal use of both model prediction and observations and can significantly improve the accuracy of the predictions of the combined model. Third, optimization algorithms for estimating the model parameters and their covariance matrix often assume that model errors are uncorrelated. Correlations between the input, predicted output and the residual series hamper an independent estimate of the model parameters, whereas autocorrelation in the residuals causes the variance of the parameters to be underestimated. 'Whitening' of the residuals with a noise model thus improves the parameter estimates of transfer or physical-deterministic models [*Bryson and Henrikson, 1965; Te Stroet, 1995*].

In light of the problem of adequately dealing with residuals or more in general noise corrupted signals, one can discern two directions in the methods used. First, there are statistical time series analysis methods, which are popular in fields of science like econometrics or social sciences. Their application developed rapidly after the publication of the comprehensive text book by Box and Jenkins [1970]. In the univariate version of these models, which are classified under the name autoregressive integrated moving average or (X)ARIMA, a series of observations of a given variable is modeled using only the temporal correlation structures in the data itself. In the multivariate case, also data of other, explanatory variables is used (transfer part of the model), while the errors of the transfer model are modeled with a separate univariate or so-called noise model. The combined model is referred to as a transfer function noise (TFN) model. Second, there are the filtering methods, of which Wiener can be considered the founder [e.g., *Wiener,* 1949]. Such methods were first applied in the field of control, navigation and communication engineering, where the problem of extracting useful information from signals that are noise corrupted arises most naturally. Here, the introduction of the Kalman filter [*Kalman,* 1960] was an important factor in the rapid spreading of its application across the more exact sciences, owing
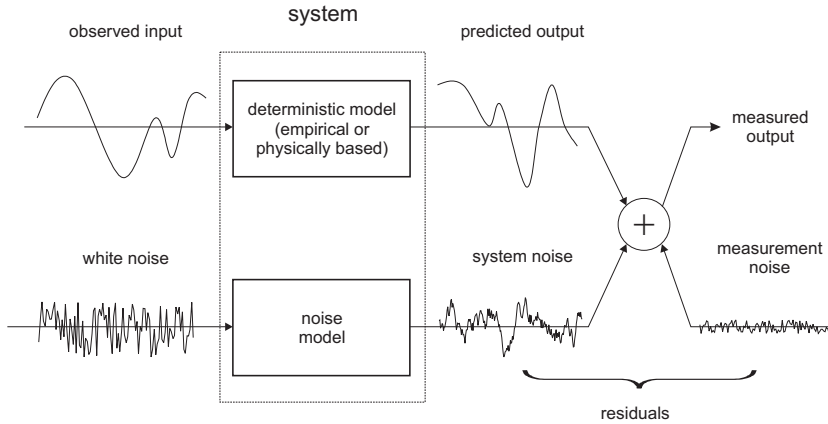
*figure 4.1: Schematic representation of a combined deterministic-stochastic model of a system under the influence of system noise, measurement noise, and an input signal.*

to its generality, simplicity, and ease of application. Nowadays, the Kalman filter is applied across a wide range of disciplines for modeling the behavior and uncertainty of phenomena that are not well explained by physical laws alone.

The distinction made above, however, is somewhat artificial, because both methods are closely related. Apart from the fact that their mathematical treatment and notation differs, there are two important differences. First, in the Kalman filter a distinction is made between noise that perturbs the state of the system itself, and noise perturbing the process of measuring that state (figure 4.1). In ARIMA models, on the other hand, there is only one noise term uniting all sources of uncertainty in the modeled signal. However, the Kalman filter is often also applied for pure prediction problems where measurement error is assumed to be absent. In Ahsan and O'Connor [1994], it is shown that in that case, the Kalman gain becomes redundant, and the filter equations reduce to a simpler form equivalent to that of an ARMA model. A second important difference, one that was not discussed by Ahsan and O'Connor, lies in dealing with irregular data. On this point, 'conventional' discrete-time ARMA models have an important practical drawback, as they cannot be readily applied to data with missing observations, while also the frequency of the input and output variables is coupled and has to be equal. Over the years, there has been a lot of effort in solving this problem, in the time series literature [*Jones,* 1980*; Harvey and Pierse,* 1984*; Little and Rubin,* 1987] but more recently also specifically focused on hydrologic problems [*Bierkens et al.,* 1999*; Koutsoyiannis,* 2001*; Berendrecht et al.,* 2003*; Yi and Lee,* 2004]. Mostly, these solutions involve a state space representation (SSR) of the model equations embedded in a Kalman filter. This way, missing observations can be easily handled by simply omitting the updating equations while retaining the prediction equations. Furthermore, to estimate the ARMA parameters, a likelihood function is constructed using the prediction errors and their variances, as simple least-squares methods no longer apply.

In this paper, we specifically focus on dealing with irregular data, and argue that also in that case the Kalman filter can be reduced to a more simple form. Furthermore, we will solve the problem in a continuous time domain using a simple Ornstein-Uhlenbeck based (OUB) noise model. Continuous time models are often considered to be more fundamental than their discrete-time counterparts, and can provide an elegant solution for modeling irregularly spaced observations and data with mixed frequencies [*Harvey,* 1989]. For many variables, the process generating the observations can be regarded as a continuous one even though the observations themselves are only made at discrete intervals. In the economic sciences, already a good deal of the theory is based on continuous time models [*Khabie-Zeitoune,* 1982*; Bergstrom,* 1990*; Brockwell,* 2001].  The basic equations of the OUB noise model were published earlier in [*Von Asmuth et al.,* 2002] where they were used as part of a continuous time transfer function noise (TFN) model. In that paper, we mainly restricted ourselves to describing how the transfer part of the model transforms irregularly spaced input series. The noise model, however, deals with the prediction errors in the output, the parameter optimization process, and stochastic simulation applications. Because of its stochastic nature, its properties are fundamentally different from those of the transfer model, and certainly as important in light of dealing with irregular data. Therefore, its background and functioning are described separately in the current paper. The OUB model is equivalent to an AR(1) model, which is often used to model the residuals in hydrologic applications, and we will show that it also in this case suffices to effectively whiten the residuals.

The paper is organized as follows: first, the forecasting mode of an AR(1) model, the equations of the time update in the Kalman filter, and the OUB model are discussed, and it is shown that they are mathematically equivalent. Next, specific attention is paid to the methods for estimating the model parameters. As an alternative for the maximum likelihood (ML) function, a sum of weighted squared innovations (SWSI) criterion is derived and introduced. In an example application, both optimization criteria are compared and it is shown that the SWSI criterion converges to the ML function for larger sample sizes. Finally, the way the prediction error or innovation variance varies with the time step is investigated, because that is a crucial part in handling data with irregular time steps correctly. In the appendix, the main derivations are given in detail.

## 4.2    Methods and theory

### 4.2.1    Irregular data and the AR(1) model

For a general treatment on the functioning of ARMA models we refer to [Box and Jenkins, 1970]. Here we will restrict ourselves to an AR(1) model and aspects that concern handling irregular data, because of the equivalency with the OUB model discussed later on. The AR(1) model is given in mathematical terms by:

$$\tilde{n}_t = \varphi \tilde{n}_{t-1} + a_t \tag{4.1}$$

where

$\tilde{n}$ : deviations of the residual series $n$ from its mean $E(n_t)$

$\varphi$ : autoregressive parameter

$a$ : discrete white noise process with properties $E\{a_t\} = 0$ (so that $a_t = \tilde{a}_t$),

$\quad E\{a_t^2\} = \sigma_a^2$, $E[a_t a_{t+i}] = 0$ for $i \neq 0$

In ARMA models in general, a white noise series at equals the series of one-step-ahead prediction errors, or innovations. For the AR(1) model, an estimate of at is obtained by subtracting the one step ahead prediction $\hat{\tilde{n}}_{t|\tilde{n}_{t-1}} = \varphi \tilde{n}_{t-1}$ from the observed value of $\tilde{n}_t$. Because of this, a missing observation of $\tilde{n}$ at time $t$ implies that $a_t$ and $a_{t+1}$ can not be calculated. Prediction of $\tilde{n}_t$ using $\tilde{n}_{t-2}$ or more time steps back does not offer a straightforward solution, because the lead time influences the prediction error or variance of $a_t$, which should be stationary. Therefore, a time series must be complete and given at regular intervals in order to be able to fit an ARMA models on the data. The forecasting mode of ARMA models, however, does provide the necessary equations to predict over a variable lead time and to quantify the accompanying prediction error [see *Box and Jenkins,* 1970]. For the AR(1) model, the prediction and its variance are a function of the time lag l and given by:

$$\begin{cases} \hat{\tilde{n}}_{t+l} = \phi^l \tilde{n}_t \\ \sigma_{e_{t+l}}^2 = \sigma_a^2 \dfrac{(1-\varphi^{2l})}{(1-\varphi^2)} \end{cases} \qquad (4.2)$$
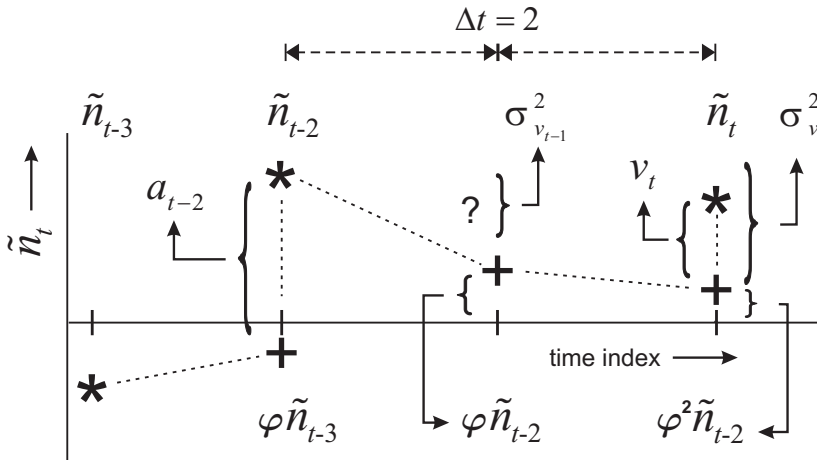


*figure 4.2: Schematic representation of the functioning of an AR(1) model combined with a 'degenerate' Kalman filter. The (\*) denotes a residual, whereas the (+) denotes a predicted residual. The model predicts missing values of the residual series at every time step, along with the prediction error variance.*

where

$\hat{\tilde{n}}_{t+l}$     : unbiased prediction of $\tilde{n}_{t+l}$ given the available observations up to time step $t$

$\sigma^2_{e_{t+l}}$     : variance of prediction error $e_{t+l}$

In other words, the main reason that ARMA models can not cope with irregular data are that parameter estimation and estimation of missing values can not be done simultaneously. In [*Brubacher and Wilson,* 1976], this was already recognized and they devised a technique that regenerates the residuals using the forecasting and backcasting mode of ARMA models, but according to [*Hipel and McLeod,* 1994, p. 695] in practice it was not very convenient.

### 4.2.2    The combined AR(1) model and 'degenerate' Kalman filter

We refer to Ahsan and O'Connor [1994] for an extensive treatment on how the Kalman can be 'degenerated' to a simpler form in the pure prediction scenario, and how the state space representation in general relates to ARMA models in standard notation. To facilitate comparison, here we will use standard notation. In figure 4.2, a schematic representation of a combined AR(1) model and the 'degenerate' Kalman filter is given. When observations are missing or scarce, the value of $a$ cannot be determined for every time step. Instead, an irregularly spaced innovation series $v$ is estimated. While $\sigma^2_a$ is a constant, $\sigma^2_{v_t}$ depends on the time lag between two observations. In this approach, fitting the model to an irregular series yields an estimate of $\varphi$, $\sigma^2_a$, $v$ and a specific variance $\sigma^2_{v_t}$ for every $v$. The parameters are estimated by optimizing a ML function, made up of the innovations and their variances. The equations of the Kalman filter are evaluated recursively, while different actions are taken depending on whether or not $n$ is available at a time step. Starting with initial conditions $\breve{n}_0$ and $\sigma^2_{e_0}$, the following equations are evaluated in the so-called time update [after *Bierkens et al.,* 1999]:

$$\begin{cases} \hat{\tilde{n}}_t = \varphi \breve{\tilde{n}}_{t-1} \\ \sigma^2_{e_t} = \sigma^2_a + \varphi^2 \sigma^2_{e_{t-1}} \end{cases} \qquad (4.3)$$

where

$\hat{\tilde{n}}_t$     : prediction of $\tilde{n}_t$ in the time update

$\breve{\tilde{n}}_t$     : prediction of $\tilde{n}_t$ in the measurement update

$\sigma^2_{e_t}$     : variance of the error in the time update

When an observation is available, the measurement update is evaluated:

$$\begin{cases} v_t = \tilde{n}_t - \hat{\tilde{n}}_t \\ \sigma_{v_t}^2 = \sigma_{e_t}^2 \\ \breve{\tilde{n}}_t = \tilde{n}_t \\ \sigma_{e_t}^2 = 0 \end{cases} \qquad (4.4)$$

If, however, no observation is available:

$$\breve{\tilde{n}}_t = \hat{\tilde{n}}_t \qquad (4.5)$$

The recursive application of equations (4.3),(4.4) and (4.5) mathematically equals the forecasting mode of an AR(1) model as described by (4.2), because after a time lag $l$ the error in the time update is:

$$\sigma_{e_{t+l}}^2 = \sigma_a^2(1 + \varphi^2 + ... + \varphi^{2l-2}) = \sigma_a^2 \frac{(1 - \varphi^{2l})}{(1 - \varphi^2)} \qquad (4.6)$$

Hence, in the SSR 'the forecasting mode' of the AR(1) model that handles predictions over variable lead times is part of the basic model equations. However, even in the 'degenerate' form, the recursive evaluation of the equations in SSR becomes computationally increasingly inefficient with an increasing model frequency. Even more so when it is used in combination with a deterministic model, because also that has to operate on the same frequency. Next to that, the autoregressive parameter $\varphi$ becomes badly scaled as it asymptotically approaches the value of 1 when the time step $\Delta t$ (in real time units) approaches 0. Finally, operating on a high frequency also
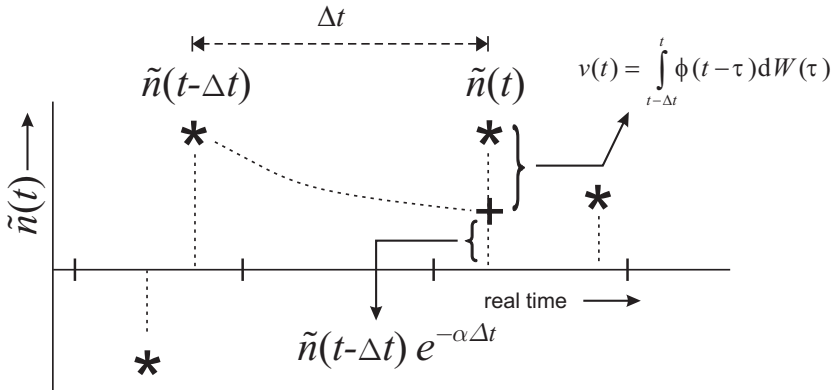


*figure 4.3:* *Schematic representation of the functioning of the Ornstein-Uhlenbeck based model. The (*) denotes a residual, whereas the (+) denotes a predicted residual. The model directly predicts the value of the residual series for the next observation available.*

poses a problem for ARMA type transfer models, because the number of MA parameters increases linearly with the frequency [*Von Asmuth et al.,* 2002].

### 4.2.3    The Ornstein-Uhlenbeck based model

In figure 4.3, a schematic representation of the functioning of the OUB noise model is given. In this approach all time series and functions, including the noise process, are regarded as being continuous. Equation (4.1) can be transformed to continuous time by writing it in moving-average form:

$$\tilde{n}_t = \sum_{i=0}^{\infty} \varphi^i a_{t-i} \tag{4.7}$$

and replacing at by a continuous white noise process $dW(t)$. The residual series $n$ can then be modeled as a continuous stochastic process, given by:

$$\tilde{n}(t) = \int_{-\infty}^{t} \phi(t - \tau) dW(\tau) \tag{4.8}$$

where

$\phi$   :   noise impulse response function

$W(t)$ :  Wiener process [L], with properties $\mathrm{E}\{dW(t)\} = 0$, $\mathrm{E}[\{dW(t)\}^2] = \beta dt$,

$\quad\quad\ \mathrm{E}[dW(t_1)dW(t_2)] = 0$ for $t_1 \neq t_2$.

$\beta$   :   a parameter [L$^2$T$^{-1}$]

As in the Kalman-filter approach, the model does not yield an estimate of the white noise series $dW(t)$, but instead gives an estimate of the innovation series $v$. $v$ is modeled as the irregularly sampled effect of the noise process on the residual series between time steps $t - \Delta t$ and $t$, which is given by:

$$v(t) = \int_{t-\Delta t}^{t} \phi(t - \tau) dW(\tau) \tag{4.9}$$

For the noise impulse response (IR) function the following exponential is chosen, so that (4.8) reduces to an AR(1) model when it used on data with regular time steps:

$$\phi(t) = \sqrt{\frac{2\alpha\sigma_n^2}{\beta}} \exp(-\alpha t) \tag{4.10}$$

with $\alpha$ defining the decay rate of the noise, and $\sigma_n^2$ denoting the variance of the residuals. With the choice of an exponential IR-function, equation (4.8) can be written as [see e.g., *Gardiner, 1994*]:

$$\tilde{n}(t) = \exp(-\alpha\Delta t)\tilde{n}(t-\Delta t) + \int_{t-\Delta t}^{t} \sqrt{\frac{2\alpha\sigma_n^2}{\beta}} \exp\{-\alpha(t-\tau)\}\mathrm{d}W(\tau) \qquad (4.11)$$

which is known as an Ornstein-Uhlenbeck process [*Uhlenbeck and Ornstein,* 1930; *Gardiner,* 1994]. By combining equation (4.9) and (4.11), the innovation series $v$ can be calculated from the available residuals using simply:

$$v(t) = \tilde{n}(t) - \exp(-\alpha\Delta t)\tilde{n}(t-\Delta t) \qquad (4.12)$$

where the last term on the right side equals $\hat{\tilde{n}}_{t|\tilde{n}_{t-\Delta t}}$ (in figure 4.3 illustrated by the dotted exponential). Because of its continuous formulation, (4.12) gives an exact solution and only needs to be evaluated once for every observation of $n$. This can reduce computation times substantially, compared to a recursive algorithm which discretizes $\Delta t$ in small time steps.

### 4.2.4    Parameter estimation and SWSI criterion

Next to the equations discussed in the previous section, an important aspect of dealing with irregular data lies in the optimization methods used. For fixed time steps, $\sigma_a^2$ is constant and minimizing the sum of squares of $a$ yields exact ML estimates, under the assumption of a Gaussian distribution. This is, however, also conditional on the choice of the starting values $a_0$ and $n_0$, and is therefore called the conditional sum of squares [*Box and Jenkins,* 1970]. With a variable $\sigma_{v(t)}^2$, simple least-squares methods are no longer straightforwardly applicable. Instead, a likelihood function is constructed and maximized using the Kalman filter estimates of $v$ and $\sigma_{v(t)}^2$ at every time step [see e.g., *Schweppe, 1973; Mélard, 1984*]. The likelihood function, however, is a rather complicated expression and flexible algorithms are needed in order to maximize it [*Hipel and McLeod, 1994*] . In the following, we will show that also for irregular data, a simple least squares criterion can be derived using the innovation variance function (IVF) that defines the general relationship between $\sigma_{v(t)}^2$ and $\Delta t$. For ARIMA models, when the set of parameters that is estimated is $\Psi = [\psi_1, \psi_2, .., \psi_p]$, and assuming that $a$ is a series of $N$ random variables which are normally independently distributed (NID(0, $\sigma_a^2$)), then their joint pdf can be written as:

$$P(a_t \mid \Psi, a_0, n_0) = (2\pi\sigma_a^2)^{-\frac{N}{2}} \exp(\sum_{i=1}^{N} \frac{-a_t^2}{2\sigma_a^2})$$ (4.13)

in which $a_0$ and $n_0$ denote the initial conditions of $a$ and $n$. Using (4.13) the log likelihood function of $\Psi$ is given by:

$$\Lambda(\Psi \mid a_t, a_0, n_0) = -0.5N\ln(2\pi) - 0.5N\ln(\sigma_a^2) - 0.5(\sum_{i=1}^{N} \frac{a_t^2}{\sigma_a^2})$$ (4.14)

For the likelihood function of continuous time innovations, we first adapt the discrete notation to allow for discretely sampled, irregularly spaced observations. For this reason we use $t$ in real time, and index that with $i$ instead. A data set $O$ of $N$ observations of a continuous process like $n$ is then given by:

$$O = [n(t_1), n(t_2), \ldots\ldots, n(t_N)]$$ (4.15)

As stated earlier, the innovation variances $\sigma_{v(t)}^2$ that are otherwise individually estimated and stored for every time step, can be jointly described by the IVF as a function of the time step $\Delta t$ (and the parameter set $\Psi$). In the example application, we will pay more attention to the IVF as it plays a crucial function in handling irregular data. The likelihood function can now be written as [*Schweppe, 1973*]:

$$\Lambda(\Psi \mid O) = -0.5N\ln(2\pi) - 0.5\sum_{i=1}^{N} \ln\{\sigma_v^2(\Delta t_i, \Psi)\} - 0.5\sum_{i=1}^{N} \frac{v^2(t_i, \Psi)}{\sigma_v^2(\Delta t_i, \Psi)}$$ (4.16)

In order to reduce the amount of parameters that has to be numerically optimized, we will start with eliminating $\mathrm{E}\{n(t)\}$. For sample sizes usually considered, the mean of a time series can be adequately estimated as [*Box and Jenkins, 1970*]:

$$\mathrm{E}\{n(t)\} = \frac{\sum_{i=1}^{N} n(t_i)}{N}$$ (4.17)

Hence, we can estimate $\tilde{n}$ directly from data set $O$. Because $\sigma_v^2(\Delta t, \Psi)$ is a function of the time step, it cannot be straightforwardly estimated from the available innovations. However, using (4.9) and (4.10), $\sigma_v^2(\Delta t, \Psi)$ can be written as a function of the residual variance as [*Gardiner, 1994*]:

$$\sigma_v^2(\Delta t, \Psi) = \{1 - \exp(-2\alpha\Delta t)\}\sigma_n^{2}(\Psi)$$ (4.18)

which in turn can yield an estimator for $\sigma_n^{2}(\Psi)$ using the individual innovations (see appendix):

$$\sigma_n^2(\Psi) = \frac{\sum\limits_{i=1}^{N}\left\{\dfrac{1}{1-\exp(-2\alpha\Delta t)}\right\}\nu^2(t_i,\Psi)}{N} \tag{4.19}$$

With (4.18) and (4.19) in (4.16), we can now eliminate $\sigma_v^2(\Delta t, \Psi)$ form the equations, and the likelihood function can be written as (see appendix):

$$\Lambda(\Psi\,|\,O) = -0.5N\ln(2\pi) - 0.5N\ln[\frac{\sum\limits_{j=1}^{N}\dfrac{\sqrt[N]{\prod\limits_{i=1}^{N}\{1-\exp(-2\alpha\Delta t_i)\}}}{1-\exp(-2\alpha\Delta t_j)}\nu^2(t_j,\Psi)}{N}] - ...\\ 0.5N \tag{4.20}$$

Thus, the only parameter that has to be numerically optimized for the noise model is $\alpha$, defining the decay rate of the noise. Because the first and last terms in the likelihood function are constant, an estimate of $\Psi$ can be obtained by minimizing the following criterion:

$$S^2(\Psi\,|\,O) = \sum\limits_{j=1}^{N}\frac{\sqrt[N]{\prod\limits_{i=1}^{N}\{1-\exp(-2\alpha\Delta t_i)\}}}{1-\exp(-2\alpha\Delta t_j)}\nu^2(t_j,\Psi) \tag{4.21}$$

which may be referred to as the Sum of Weighted Squared Innovations (SWSI) criterion. The SWSI criterion is similar to other weighted least squares criteria, as also here the weights reflect the variances of the innovations. For optimization and parameter estimation, we can now use standard nonlinear least squares regression methods [see e.g., *Snedecor and Cochran, 1967*]. First, the partial derivatives of (4.21) are obtained, either numerically or analytically, and used to construct a Jacobian matrix:

$$J = \begin{bmatrix} \dfrac{\partial S_1}{\partial \beta_1} & \cdots & \dfrac{\partial S_1}{\partial \beta_1} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial S_N}{\partial \beta_1} & \cdots & \dfrac{\partial S_N}{\partial \beta_p} \end{bmatrix} \quad (4.22)$$

Using $J$ the model is calibrated on a set of observations with a Levenberg-Marquardt optimization algorithm [*Marquardt, 1963*] that adjusts $\Psi$, while minimizing $S\{\Psi|O\}$. Then, the covariance matrix of the parameters can be estimated with:

$$\sigma_\beta{}^2 = \frac{\min(S^2)}{N-p}(J'J)^{-1} \quad (4.23)$$

Next, $\sigma_n{}^2$ is estimated using (4.19). The model results can be checked by examining the covariance matrix and the variance of the IR-function. The variance of the IR-function $\phi$ equals:

$$\text{var}\{\phi(t)\} = \{\frac{\sqrt{2\sigma_n^2}}{2\sqrt{\alpha}}\exp(-\alpha t) - t\sqrt{2\alpha\sigma_n^2}\exp(-\alpha t)\}^2 \, \text{var}(\alpha) \quad (4.24)$$

A confidence interval for $\phi$ can be plotted as +/- $2\sigma$, when assuming a normal distribution for $\alpha$. As in the ARIMA approach, serious model inadequacy can be detected by examining the autocorrelation function of the innovation series $v$ (which gives an indication on whether the white noise assumption holds) and the crosscorrelation function between $v$ and the transfer model input series $p$ (which indicates whether there are still patterns left in the innovation series that could be explained by the input series). The autocorrelation and crosscorrelation functions at lag $k$ are defined in the same way as in discrete TFN models, but because of the irregularity of the time steps, a tolerance around lag $k$ of $\pm 0.5 k$ is implied. A proof for the fact that, like AR(1) models, the innovations of the continuous model are not autocorrelated, nor crosscorrelated with observations of the process $n$ itself, is given in the appendix. This makes the model suitable for simulation purposes, and also is a prerequisite for the parameter estimation process.

### 4.2.5    Summary of method

When the parameters of a combined deterministic or transfer model and noise model are estimated simultaneously, the procedure is as follows (see also figure 4.4). First, using the initial values of $\Psi$, the deterministic model is evaluated in order to get a time series of residuals. Under the assumption of exponential noise decay (equation (4.10)), the innovation series is obtained from the residual series using equation (4.12).

Next, the SWSI criterion is calculated by weighting the squared innovations according to their respective variances and relative to the geometrical average of the innovation variances for all time steps (equation (4.21)). Subsequently, the parameter set $\Psi$ is estimated by minimizing the criterion using a Levenberg-Marquardt algorithm. Finally the validity and performance of the model are checked using the auto- and crosscorrelation functions of the innovations, the correlation matrix of the model parameters and the variance of the IR function.
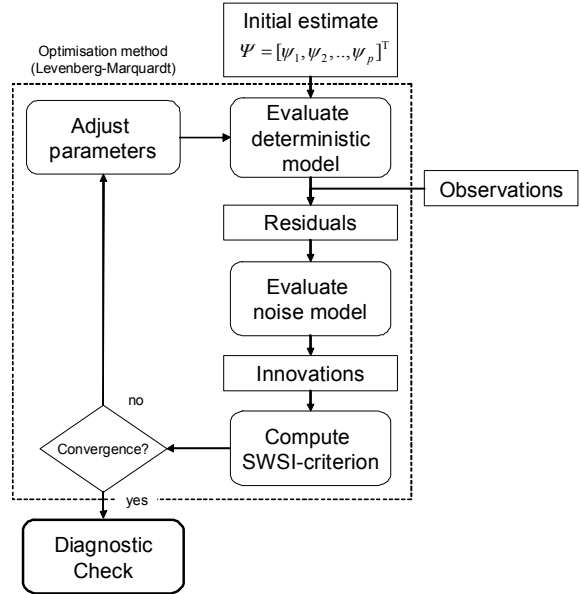


*figure 4.4:* *Procedure for applying the noise model in combination with a deterministic model.*

## 4.3 Example application

### 4.3.1 Set-up and data set

The effectiveness of a combined AR(1) model and Kalman filter in modeling irregularly observed hydrologic data has already been shown rather elaborately [*Bierkens et al.,* 1999*; Berendrecht et al.,* 2003*; Yi and Lee,* 2004]. Therefore, we will not focus on that here. The mathematical equivalency of the OUB model and the AR(1) model are straightforward, so logically its results will also be equivalent. Instead, we will first focus on comparing the SWSI criterion with the likelihood function. Logically, if the minima of both criteria are identical, they will yield identical parameter estimates and therefore also identical model predictions. Second, we will examine the innovation variance function, which is crucial for handling irregular data in both the OUB and the Kalman filter approach. We will illustrate that also in this case, the model effectively whitens the residuals and the assumption of exponential noise decay is therefore valid. As a test case, we will use the noise model in conjunction with a continuous transfer function model in the context for which it was developed, i.e. modeling the residuals of groundwater level observations. In the simple case where groundwater level fluctuations are influenced by precipitation surplus only, the combined transfer function noise model is given by:

$$h(t) = \int_{-\infty}^{t} \theta(t-\tau)p(\tau)\mathrm{d}\tau + n(t) + d \tag{4.25}$$

where
$t$     : time [T]

$h$   : observed groundwater level, relative to some reference level [L]
$p$   : precipitation surplus [LT$^{-1}$]
$\theta$   : impulse response function [-], for which a scaled gamma distribution is chosen,

i.e. $\theta(t) = A \dfrac{b^n t^{n-1} \exp(-bt)}{\Gamma(n)}$

$n$   : residual series [L], modeled conform equation (4.11)
$d$   : local drainage level, relative to some reference level [L]

More details on the background of the transfer function model can be found in [*Von Asmuth et al.,* 2002]. The TFN model is calibrated on a 15-year (1981-1996) groundwater level series observed with a daily frequency. However, this series is not totally complete but 8.75% of the observations are missing. The series originates from a piezometer located on the main meteorologic field of the Royal Dutch Meteorologic Institute at the town of De Bilt in the center of the Netherlands [see also *Bierkens et al., 1999*]. The precipitation surplus is obtained from daily averaged observations of precipitation and potential evapotranspiration at the meteorologic field. A time plot of the available groundwater level observations is given in figure 4.5.
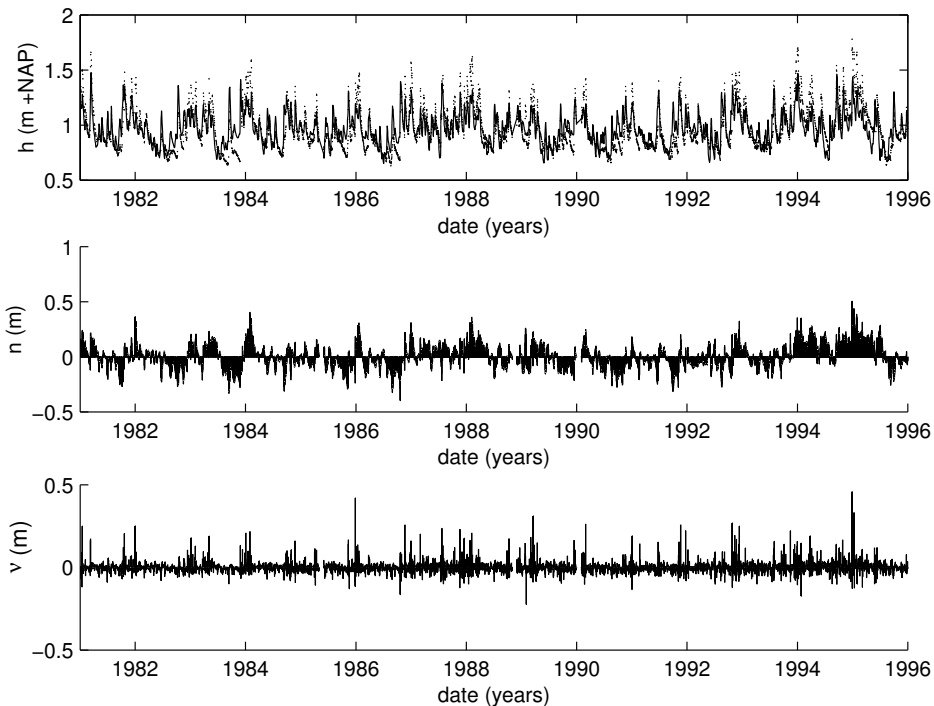


*figure 4.5:* *Time plot of the available groundwater level observations (dots) from piezometer 32cl0034 relative to the national reference level (h+NAP), the predictions of the transfer model (full line), the model residuals (n) and the innovation series (ν).*

The parameter set $\Psi$ of the combined model is integrally estimated using a numerically derived Jacobian matrix and the methods described in the previous section. As $d$ is eliminated from the equations, the parameters that have to be estimated are $A, b, n$ from the transfer model, along with $\alpha$ from the noise model. For this purpose, all available observations in the period were used. The results of the TFN model are given in figure 4.5, as time plots of the observations and predictions, and of the residuals and innovations. The

**table 4.1:** *Calibration results, estimated parameters and characteristics for piezometer 32cl0034.*

| Piezometer | 32cl0034 |
|---|---|
| $R^2$ | 0.665 |
| RMSE | 11,0 cm |
| RMSI | 3,6 cm |
| $A$  (+/- 2) | 61.6  (+/- 4.2) day |
| $b$  (+/- 2) | 0.28  (+/- 0.02) day$^{-1}$ |
| $n$  (+/- 2) | 1.88  (+/- 0.06) |
| $\alpha$  (+/- 2) | 18.2  (+/- 3.2) day$^{-1}$ |

parameter estimates and calibration results (coefficient of determination ($R^2$), root mean squared error (RMSE), root mean squared innovation (RMSI)) are listed in table 4.1. The autocorrelation function of the innovations, for which a time lag increment of one month is chosen in order to reveal seasonal patterns in the autocorrelation, indicates that the white noise assumption holds (figure 4.6). Because of the large number of available observations, the autocorrelation function is rather smooth and the accompanying confidence interval narrow.

### 4.3.2 Comparison of the likelihood function and the SWSI criterion

Plots of both functions were made by varying two of the parameters around their estimated value, while keeping the others constant (ceteris paribus). The parameters were varied +/- 0.5 half their estimated value. In order to make both functions comparable, they were normalized by respectively setting their maximum and minimum to zero, and dividing the result by the range. In figure 4.7, a contour plot of the normalized log likelihood function and SWSI criterion for the first two parameters of the transfer model ($A$ and $b$) is given. The figure shows that the contours of both objective functions are almost identical, which confirms that the SWSI criterion yields almost exact ML estimates. This can be expected, as the only difference between the
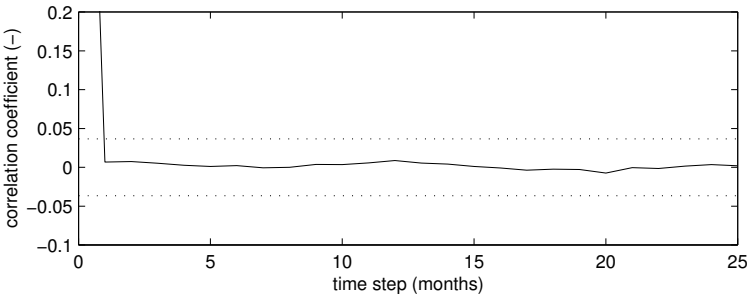


*figure 4.6: Autocorrelation function of the innovations of the OUB model. The dotted lines denote the 95% confidence interval. The autocorrelation function indicates that the white noise assumption holds.*

ML function and SWSI criterion is that in first case $\sigma_a^2$ (for the minimal time step) is part of the parameter set to be estimated, while in the SWSI criterion $\sigma_n^2$ is assumed known and implicitly approximated with Equation (4.19). This approximation converges to the true value of $\sigma_n^2$ for large $N$. Consequently, a comparison of $\sigma_a^2(\Psi)$ estimated as a parameter with the likelihood function or estimated afterwards using the SWSI approach will exemplify the likeness of both criteria for all parameters. Using the SWSI criterion $\sigma_a^2(\Psi)$ is estimated from the data in the following way:

$$\sigma_a^2(\Psi) = \frac{\sum_{i=1}^{n}\left\{\frac{1-\exp(-2\alpha)}{1-\exp(-2\alpha\Delta t_i)}\right\}v^2(t_i,\Psi)}{n} \tag{4.26}$$

which is equivalent to equation (4.6) for fixed $\Delta t$ when $\dfrac{\sum_{i=1}^{n}v^2(t_i,\Psi)}{n}$ is replaced by $\sigma_v^2(\Psi)$, and $\phi = \exp(-\alpha)$. In figure 4.8, the likelihood function of the innovations is plotted and compared to $\sigma_a^2(\Psi)$ obtained with equation (4.26) (i.e. 0.0013 m2).
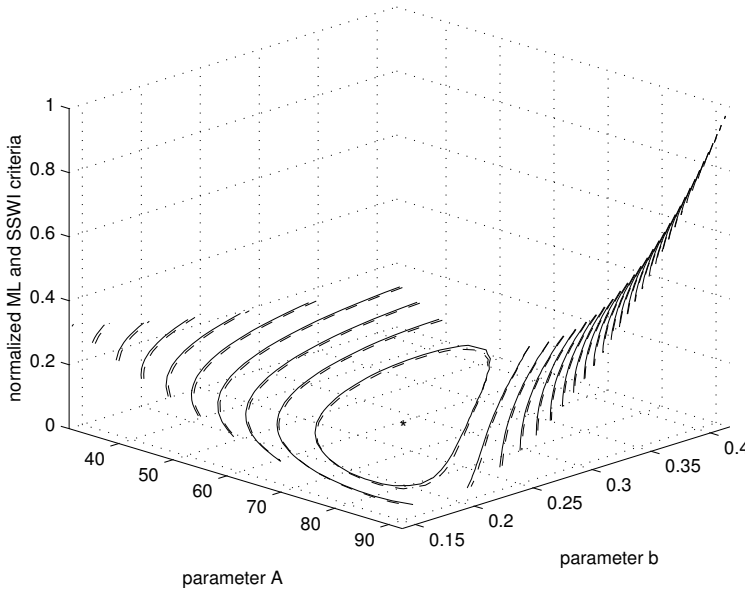
*figure 4.7:* Contour plot of the normalized log likelihood function of the innovations (dotted lines) and the SSWI criterion (full lines) as a function of the parameters A and b (ceteris paribus). The marker (*) denotes the estimated values. From the figure it can be seen that both functions are almost identical.
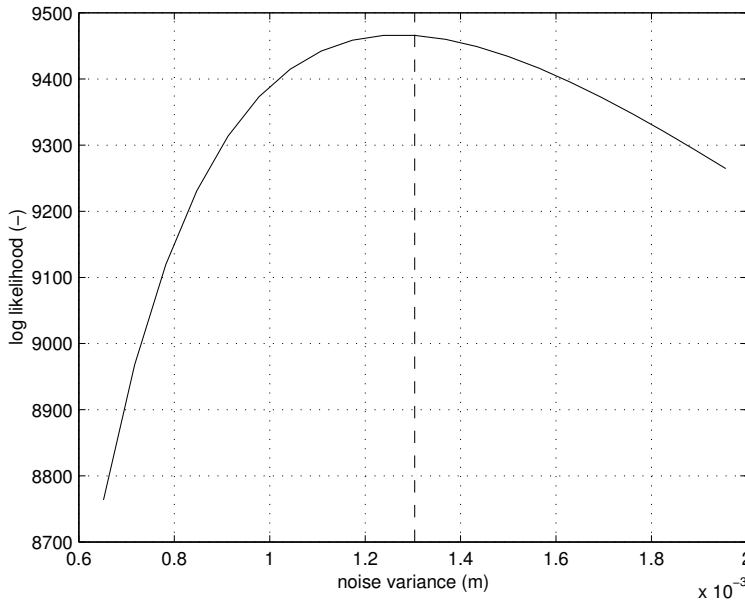
*figure 4.8:* *Plot of the log likelihood function of the innovations as a function of the parameter* $\sigma_a^2$.

*The dotted line denotes the value of* $\sigma_a^2(\Psi)$ *obtained with the SWSI criterion. The minimum of the log likelihood function proves to coincide well with the value obtained from the SWSI criterion.*

From this figure, the estimate of $\sigma_a^2(\Psi)$ proves to coincide well with the maximum of the likelihood function for this parameter.

### 4.3.3    A check on the innovation variance function

Because the IVF is a key factor in both the derivation of the SWSI criterion and in general for stochastic modeling at mixed frequencies, in the following we will examine it more closely using the data of section 4.3.1. For this purpose, the groundwater level series was resampled at frequencies $\Delta t = [1,2,3.....75]$ days. The transfer and noise model were evaluated using the estimate of $\Psi$ obtained with the daily series, which logically should be optimal. Thus, a residual and innovation series were obtained for every sample frequency. As stated earlier, the daily series was not complete and therefore also the resampled series were not. Because of this, only those innovations and residuals were selected from the resampled series for which

$\Delta t_i$ equaled the sampling frequency exactly. Both the innovation and residual variance of the resulting series were estimated and plotted in figure 4.9. In the same figure, equation (4.18) being the theoretical IVF, is plotted together with the 95% confidence interval of the estimates. The latter is given by [*Snedecor and Cochran, 1967*]:
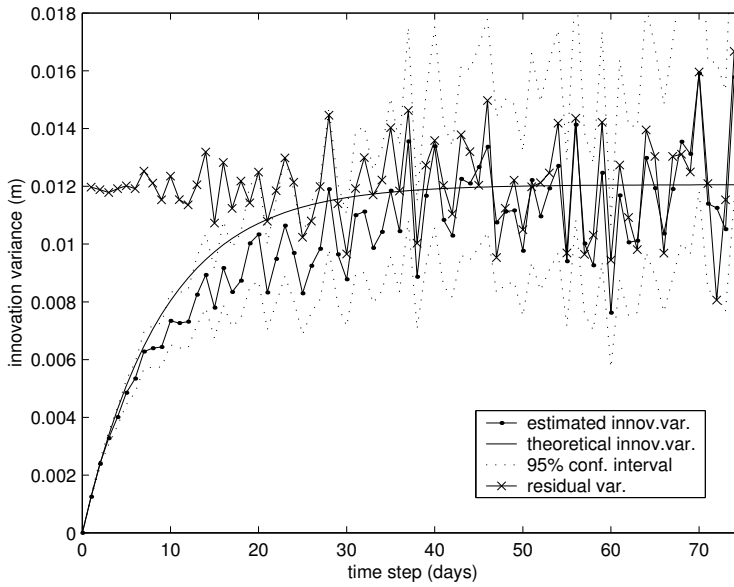
*figure 4.9:* *Plot of the estimated and theoretical IVF as a function of the time step, and of the estimated residual variance. In general, the estimated innovation variances do not differ significantly from the theoretical IVF, although the agreement is better for smaller time steps.*

$$\frac{\sigma_v^2(\Delta t_i, \Psi)*(N-1)}{\chi_{0.975}^2} \leq \sigma_v^2(\Delta t_i, \Psi) \leq \frac{\sigma_v^2(\Delta t_i, \Psi)*(N-1)}{\chi_{0.025}^2} \qquad (4.27)$$

Figure 4.9 thus allows for a comparison between the estimated and theoretical IVF. A check on whether the estimated variances differ significantly from the theoretical IVF could well be used as a further diagnostic check on the validity of the model, next to checking the auto- and crosscorrelation functions. From figure 4.9 it is concluded that the estimates of the innovation variance $\sigma_v^2(\Delta t, \Psi)$ do not significantly differ from the theoretical IVF, although the agreement is less for the intermediate time steps. This points to the fact that the parameter $\alpha$, defining the decay rate of the noise model, could be sensitive to the sampling frequency. Therefore, special care has to be taken when the assumption of exponential decay for the noise model does not, or not well, hold and/or when the model is used for simulations at a different frequency than that of the data with which it was calibrated. A standard diagnostic check like the autocorrelation function of the innovations will not provide enough information on this point.

## 4.4    Discussion and conclusions

A comparison of the AR(1) model, 'conventional' or embedded in a Kalman filter, and the Ornstein-Uhlenbeck based noise model shows that their equations for respectively the forecasting mode in the AR(1) model, those handling the predictions and their

variance in the time update of the 'degenerate' Kalman filter, and the continuous time innovations are mathematically equivalent. The continuous equations, however, are more general, give an exact solution and are computationally more efficient as they are not evaluated recursively. Because of its simplicity, the OUB model can be easily implemented for modeling irregularly spaced errors of various deterministic models. The continuous time approach, however, makes it especially suited for combination with analytic models or continuous time transfer models. A restriction of the OUB model in its present form, is that it is limited to processes that show exponential decay. Work is needed to further generalize and test the approach, for example using continuous time ARMA models in general [e.g., *Brockwell,* 2001], non-Gaussian OUB models [*Barndorff-Nielsen and Shephard,* 2001] or by superposition of OU type processes [*Barndorff-Nielsen,* 1999].

It was shown that from the likelihood function normally used for irregularly spaced innovations, a weighted least squares (SWSI) criterion can be derived. The main difference between both functions is that in the first case, the innovation variances are calculated by relating them to the innovation variance of the minimal time step, which itself is estimated as a parameter, and stored for every time step. In the latter case, the innovation variances are implicitly related to the variance of the residuals, and in turn approximated implicitly with equation (4.19). This approximation converges to the true value for large sample sizes, so parameter estimates that are based on the SWSI criterion should approximate the ML estimates well for sample sizes usually considered. With the aid of the SWSI criterion, the parameters and their covariance matrix can be estimated using standard nonlinear least squares regression methods. The use of such methods is more efficient than minimizing the log-likelihood function with some global optimization algorithm, and a Monte-Carlo approach to estimating the parameter variances. The SWSI criterion itself could also be used for optimizing AR(1) models, in the 'conventional' sense or embedded in a Kalman filter.

By comparing the estimated innovation variances for different time steps with the theoretical innovation variance function (IVF), the validity of exponential decay for the noise model can be checked. The impulse response function is reflected in the behavior of the IVF. Thus, a plot of both functions and the accompanying confidence interval is useful as a diagnostic check. Such a check provides additional information on the validity of the model, especially in light of its application at mixed frequencies.

# Chapter

5

# Modeling groundwater head series subjected to multiple stresses

**Abstract:**

In this chapter, the methods behind the PIRFICT (Predefined Impulse Response Function In Continuous Time) time series model are extended to cover more complex situations where multiple stresses influence groundwater head fluctuations simultaneously. In comparison to ARMA (Auto-Regressive Moving Average) time series models, the PIRFICT model is optimized for use on hydrologic problems. The objective of the paper is twofold. First, an approach is presented for handling multiple stresses in the model. Each stress has a specific parametric impulse response function. Appropriate impulse response functions for other stresses than precipitation are derived from analytic solutions of elementary hydrogeologic problems. Furthermore, different stresses do not need to be connected in parallel in the model, as is the standard procedure in ARMA models. Second, general procedures are presented for modeling and interpretation of the results. The multiple-input PIRFICT model is applied to two real cases. In the first one, it is shown that this model can effectively decompose series of groundwater head fluctuations into partial series, each representing the influence of an individual stress. The second application handles multiple observation wells. It is shown that elementary physical knowledge and the spatial coherence in the results of multiple wells in an area may be used to interpret and check the plausibility of the results. The methods presented can be used regardless of the hydrogeologic setting. They are implemented in a computer package named Menyanthes (www. menyanthes.nl).

## 5.1    Introduction

Transfer function noise (TFN) models are a convenient tool for modeling the evolution of a wide range of variables. The general theory of time series analysis [*Box and Jenkins,* 1970] originally stems from the statistical sciences. Because of their statistical background, the so-called ARMA (i.e. Auto-Regressive Moving Average) time series models can be applied to all sorts of data, as long as the behavior of the system to be modeled is sufficiently linear, or can be linearized by transforming the data. Time series models are especially useful for modeling systems whose behavior can not, or not easily, be described in terms of physical laws and properties (e.g., economical data). In addition, TFN models are often used in hydrology and other sciences, because they are relatively easy to construct and at the same time they can yield very accurate predictions.

When an ARMA type TFN model is applied to a data set, the so-called model order has to be specified. The model order includes the number of auto-regressive and moving average parameters in both the deterministic and stochastic parts of the model, and the delay time of the transfer function. Box and Jenkins devised an iterative model identification procedure to guide the modeler in finding the optimal model order. First, an initial model order is chosen based on statistical criteria like the crosscorrelation function between the explained and explanatory variables. Second, the parameters of the model are estimated by minimizing the variance of the 'innovations' or one-step ahead prediction error, using an optimization algorithm. Third, the adequacy of the model is checked diagnostically, using statistical criteria like the auto- and crosscorrelation functions of the innovations. If the model does not yet meet the criteria, the model order is updated, and the procedure is repeated until the modeler finds the results satisfactory. A disadvantage of this approach is that the results of the model identification procedure can be ambiguous [*Hipel and McLeod,* 1994] and the process itself is rather heuristic and can be knowledge and labor intensive [*De Gooijer et al.,* 1985].

Von Asmuth et al. [2002] presented the principles of a new type of TFN model that is optimized for use on hydrologic problems and operates in a continuous time domain. In this approach, the discrete transfer function used in ARMA models is replaced by a simple analytic expression that defines the impulse response function. The resulting class of models is referred to as PIRFICT (Predefined Impulse Response Function In Continuous Time). Von Asmuth et al. [2002] showed that PIRFICT models overcome a series of limitations of ARMA TFN models, including the use of irregular or high frequency data and the modeling of systems with a long memory. In addition, application of the PIRFICT model does not require a Box-Jenkins style model identification procedure. Since the transfer functions are confined a priori to physically plausible behavior, there is no need to identify the 'order' of the transfer functions on statistical grounds. Therefore, application of the model is standardized, which facilitates implementation in a computer package such as Menyanthes (www. menyanthes.nl). When the PIRFICT method was introduced [*Von Asmuth et al.,* 2002] we restricted ourselves to the case of a single input / output series. Here, we will extend the method to cover more complex, real world situations where multiple

stresses influence head fluctuations at one or multiple observation wells simultaneously.

Two important aspects of dealing with complex data sets are addressed in this paper. The first one is the treatment of different types of stresses within the model. Different stresses require different parametric impulse response functions. Von Asmuth et al. [2002] used a scaled gamma distribution for modeling the effect of precipitation surplus. Here we will introduce analytic solutions of elementary hydrogeologic schematizations as guides to develop appropriate impulse response functions for other stresses. . We will also show that from a physical point of view, different stresses do not always have to be connected in parallel and get a separate transfer function, as is the standard procedure in ARMA TFN modeling. The second aspect deals with the interpretation and checking of the plausibility of the results. While the time series literature commonly involves the analysis of individual time series, using the PIRFICT approach one can analyze and process all available series of heads in an area in batch. Given that they are part of the same hydrologic system, the results of neighboring observation wells may show a spatial coherence which yields valuable extra information as regards the properties of the system and the plausibility of the results. In this paper, however, all series are still modeled separately, and the resulting spatial patterns are analyzed a posteriori. Future research will include methods to impose spatial coherences a priori in the model.

This paper is organized as follows. First we discuss how different types of stresses are dealt with in the model. We illustrate the approach by analyzing a single series being influenced by precipitation, evaporation, groundwater withdrawal and river level fluctuations. Second, results are presented for a case where data are available from multiple observation wells. A discussion and conclusions are given at the end of the paper.

## 5.2 Methods and theory

### 5.2.1 From a single to a multiple input model

The basic equation of an ARMA model, which is discrete in time, is equivalent to the following convolution integral in continuous time [*Von Asmuth et al.,* 2002]:

$$\hat{h}_i(t) = \int_{-\infty}^{t} R_i(\tau)\theta_i(t-\tau)d\tau \tag{5.1}$$

where
$t$ : time [T]
$\hat{h}_i$ : predicted head [L] attributable to stress $i$ .
$R_i$ : value of stress $i$
$\theta_i$ : transfer or impulse response function of stress $i$ .

In ARMA time series models, multiple stresses are dealt with by connecting them in parallel and assigning a separate transfer function to each. The output series is then obtained by summing the separate effects of all stresses. For the case where a number of $N$ stresses influence the head, the equations of a continuous time TFN model may be written as:

$$h(t) = \sum_{i=1}^{N} \hat{h}_i(t) + d + n(t) \tag{5.2}$$

where
$h$   : observed head [L]
$d$   : local drainage level relative to some reference level [L]
$n$   : residual series [L].

Several main types of stresses can be distinguished. These types include precipitation ($p$), evaporation ($e$), groundwater withdrawal (or injection) ($w$), surface water level ($s$), barometric pressure ($b$), and (hydrologic) interventions ($m$). Please note that tidal fluctuations, on whose effect a large volume of papers is devoted, are included in the $s$-type. From a physical point of view, a groundwater system is likely to respond to different types of stresses differently, but there are also certain stresses that will cause a very similar response. In the latter case, separate stresses do not necessarily need separate response functions. For instance, the effect of evaporation $e$ on the head $h$ is essentially the same as that of precipitation $p$, but it is negative, and may be modeled as:

$$\hat{h}_e(t) = \int_{-\infty}^{t} -e(\tau) f \theta_p(t-\tau) d\tau \tag{5.3}$$

where $\theta_p$ is the response of the system to precipitation, and $f$ is a reduction of $e$ as compared to the reference evaporation series. The evaporation factor $f$ is a parameter that depends on soil and land cover and should not be confused with the crop factor [e.g., *Penman,* 1948], as the latter is used for a crop that is optimally supplied with water, while f also incorporates the average reduction of the evaporation due to actual soil water shortages. For sake of simplicity, we consider $f$ to be constant. Although it may actually be a function of time, the use of an average reduction factor is in fact already an improvement upon traditional ARMA modeling practice, where f is often simply ignored. In case of the barometric pressure $b$, which co-determines the loading on a (semi)confined aquifer, we propose to use the time derivative in the convolution:

$$\hat{h}_b(t) = -\int_{-\infty}^{t} \frac{db(\tau)}{d\tau} \theta_b(t-\tau)d\tau \qquad (5.4)$$

in order to be able to use an impulse response function $\theta_b$ that behaves similar to the other stresses, as a step change of $b$ will result in a quick increase of $h$, followed by a slow decay whenever the aquifer is not completely confined.

Interventions are defined as structural changes to a hydrologic system, like forest clearing, construction of ditches or drainage systems, etc. In general, the nature of an intervention determines the manner in which it should be modeled. For example, if the intervention causes a sudden change in actual evaporation on time $t_m$, such as a forest clearing, it may be modeled as:

$$\hat{h}_m(t) = \int_{-\infty}^{t} m(\tau)k\theta_p(t-\tau)d\tau \qquad (5.5)$$

where $m(t) = 0$ for $t < t_m$ and $m(t) = 1$ for $t \geq t_m$, and $k$ is a parameter representing the change caused by the intervention. Another example is a change in the level of a floodgate. This intervention itself acts as an independent stress on the system, and in such cases a new response function $\theta_m$ should be estimated. It is noted that it is possible that the hydrogeologic properties of the system are changed significantly by an intervention. In that case, the response of the system to all stresses is changed, and equation (5.1) becomes:

$$\hat{h}_i(t) = \int_{-\infty}^{t_m} R_i(\tau)\theta_{i_1}(t-\tau)d\tau + \int_{t_m+\Delta t}^{t} R_i(\tau)\theta_{i_2}(t-\tau)d\tau \qquad (5.6)$$

where $\theta_{i1}$ and $\theta_{i2}$ are the response functions before and after the intervention, respectively. A transition period of $\Delta t$, during which the system shifts from one state to another, should be omitted from the data. The length of this period depends on the response time of the system.

### 5.2.2 Response functions for different types of stresses

Under a wide variety of hydrogeologic settings, the response of an impulse of precipitation surplus may be simulated accurately with a scaled gamma distribution function (SG df), given by [*Von Asmuth et al., 2002*]:

$$\theta_p(t) = A \frac{a^n t^{n-1} \exp(-at)}{\Gamma(n)} \qquad (5.7)$$

where $A, a$ and $n$ are parameters that define the shape of $\theta_p$. The choice of this impulse response function was based on physical arguments, and it was shown that the SG df was (at least) as effective as an ARMA type transfer function of optimal order. Although the SG df is very flexible, it proves to be less effective for non distributed types of stress. For groundwater withdrawals (stress type $w$) and surface water level fluctuations (type $s$), we propose to use response functions inspired on analytic solutions of simple hydrogeologic schematizations. Notice that the crucial issue in selecting a parametric impulse response function is whether the range of shapes it can take is sufficient to approximate the true response of the system accurately. In this sense, our viewpoint is that of time series analysis, where the only assumptions regarding the transfer model are that the system is linear and that the 'model order' is adequate. Here, we assume that the functions chosen can capture the essential behavior of the stress type regardless of the exact hydrogeologic setting. A systematic comparison of the performance of different impulse response functions for different types of stresses, however, falls beyond the scope of this paper and will be dealt with in an upcoming paper.

For withdrawals (stress type $w$) we choose the well formula of Hantush [*Hantush,* 1956] as a blueprint. The Hantush formula assumes a fully penetrating well in an aquifer of infinite extent, with transmissivity $KH$ [$L^2T^{-1}$] and storage coefficient $S$ [-], covered by a storage-free aquitard with resistance $c$ [T]. While a standard pumping test yields a step response function, here we are looking for an impulse response function, which is the derivative of the step response function with respect to time:

$$\theta_w(t) = -\frac{1}{4\pi KHt}\exp(-\frac{r^2 S}{4KHt} - \frac{t}{cS}) \tag{5.8}$$

where $r$ denotes the distance between the observation and the pumping well. Since we intend to use this formula in other hydrogeologic settings as well, we convert this equation to the following parametric impulse response function:

$$\theta_w(t) = -\frac{\gamma}{t}\exp(-\frac{\alpha^2}{\beta^2 t} - \beta^2 t) \tag{5.9}$$

where $\alpha, \beta$ and $\gamma$ are just parameters that no longer have a transparent physical meaning (except for cases where Hantush' assumptions happen to be satisfied). For surface water fluctuations (stress type $s$) we choose the so-called polder function of Bruggeman as a blueprint [*Bruggeman,* 1999]. This function represents a sudden unit increase of the water level at the boundary of a one dimensional semi-confined aquifer of semi infinite extent. The derivative gives the impulse response function $\theta_s$, i.e. the response to a very short rise and fall of the surface water level:

$$\theta_s(t) = -\frac{1}{\sqrt{\dfrac{4\pi KHt^3}{x^2 S}}} \exp\left(-\frac{x^2 S}{4KHt} - \frac{t}{cS}\right) \tag{5.10}$$

where $x$ denotes the distance between the surface water feature and the observation well. We convert the physical parameters to the abstract parameters $\alpha', \beta'$ and a scaling factor $\gamma'$, so that $\theta_s$ becomes:

$$\theta_s(t) = -\frac{\gamma'}{\sqrt{\pi \dfrac{\beta'^2}{\alpha'^2} t^3}} \exp\left(-\frac{\alpha'^2}{\beta'^2 t} - \beta'^2 t\right) \tag{5.11}$$

For stress type $b$ (the barometric pressure) we chose the SG df, i.e. the same impulse response function as used for precipitation and evaporation, because pressure is a distributed type of stress, too. The adequacy of a time series model based on these impulse response functions may be checked with the aid of the methods described and illustrated in the coming sections. The practical usefulness of these impulse response functions has been demonstrated through application to thousands of wells in Europe, Australia, and North and South America.

### 5.2.3 Parameter estimation

The first step in the parameter estimation process is the evaluation of the model equations using an initial estimate of the parameters, which will result in a time series of model errors or residual series. When modeling hydrologic data, the value of a model residual at a certain time is often correlated with its value at earlier times, so residuals cannot simply be modeled as a set of independent Gaussian deviates. Here, the model residuals are modeled with a separate noise model, which is given by:

$$n(t) = \int_{-\infty}^{t} \phi(t-\tau) dW(\tau) \tag{5.12}$$

where $\phi$ is the noise impulse response function and $W$ is a continuous white noise (Wiener) process [L]. This is equivalent to an AR(1) model embedded in a Kalman-filter under the pure prediction scenario in discrete time, when the response function $\phi$ is an exponential [*Von Asmuth and Bierkens*, 2005]. In that case, the one-step-ahead prediction error or innovation series $v$ of the noise model may be obtained as:

$$v(t) = n(t) - n(t - \Delta t)\exp(-\alpha\Delta t) \tag{5.13}$$

The noise model is important for the parameter estimation process and for dealing with irregularly spaced data, but also for prediction, forecasting and stochastic

simulation purposes. For further details we refer to [*Von Asmuth and Bierkens,* 2005], as the parameter estimation process itself is not influenced by the fact that the transfer model contains multiple stresses.

### 5.2.4    General modeling procedure

The model identification procedure devised by Box and Jenkins [1970] is that, first, the model order is specified, second, the parameters are estimated, and third, the model results are checked. The adequacy of the model results may be checked with statistical criteria like the autocorrelation and crosscorrelation functions of the innovation series $\nu$. The autocorrelation function indicates whether the white noise assumption holds, which is a prerequisite of the algorithms used for the estimation of the model parameters and their covariance; it also indicates whether the order of the noise model is adequate. The crosscorrelation functions between the innovation series $\nu$ and the different input series indicate whether there are patterns left in the innovation series that could be explained by the input series. This gives an indication of the adequacy of the order of the transfer functions, but also of possible non-linearity in the relationships. If the model does not meet these so-called diagnostic checks, the model order is is updated. This procedure is repeated until the optimal model order is identified [*Box and Jenkins,* 1970].

As stated earlier, there is no need for identifying the order of the transfer model on statistical grounds in the PIRFICT approach, as the impulse response functions are chosen on physical grounds and span a whole range of ARMA model orders [*Von Asmuth et al.,* 2002]. In both cases, however, the modeler also has to identify the stresses that influence the groundwater dynamics, decide which stresses to use, and check the results. Stresses that are not incorporated in the model can lead to erroneous results when their dynamic behavior is correlated with one or more of the other (already considered) stresses. On the other hand, the model may have difficulty in



*figure 5.1: Proposed procedure for modeling time series of ground water heads in complex situations. The procedure for identifying the noise model is that of [Box and Jenkins, 1970]. When both the plausibility and diagnostic checks are met, the results can be accepted for further use.*

uniquely identifying all influences when there is a high number of stresses, a lack of pronounced dynamics of the stresses, or a low quality or scarcity of the data. The diagnostic checks introduced by Box and Jenkins are devised to assess whether the TFN model is adequate and optimal in a statistical sense, which remains important especially for the noise model, but does not exclude the possibility that the model

results are influenced by non-causal correlations. Here, we propose the use of plausibility checks to guide the modeler in assessing whether the results of the transfer model are physically realistic. The resulting modeling procedure is illustrated in figure 5.1.

The plausibility checks include the model residuals $n$, the evaporation factor $f$, the local drainage base $d$, and the moments of the impulse response functions and their standard deviations. The model residuals, uniting all factors that are not accounted for by the model, are an important aid in identifying possible unknown stresses that may be a source of model distortions. Non-random patterns of the residuals in space or time reveal the fact that there are still stresses missing in the model. The patterns themselves often give enough information to pinpoint the nature and location of the missing stresses. The evaporation factor f is important, as the seasonal cycle in the evaporation is often present in other natural or anthropogenic stresses such as groundwater withdrawals for agricultural or drinking water purposes as well. Regarding the drainage base, an estimate that is too low or too high may be caused by the influence of stresses that are not incorporated in the model or that are not well quantified. This can easily happen with stresses that do not show pronounced dynamics, such as more or less constant seepage or withdrawal rates. In addition, the moments of the impulse response functions of the different stresses provide relevant information. Moments can be used to characterize the functioning of the groundwater system and can be related to its geohydrologic properties [*Von Asmuth and Maas,* 2001*; Von Asmuth and Knotters,* 2004]. In contrast to physical parameters that are only defined in the context of a certain schematization, moments are related to common statistical terms and are more generally applicable. The $j^{th}$ moment of an impulse response function is defined as:

$$M_j = \int_{-\infty}^{\infty} t^j \theta(t)dt \qquad (5.14)$$

where $M_0$ represents the area under the impulse response function, $\mu = \dfrac{M_1}{M_0}$ is the mean of the impulse response function, and $\sigma^2 = \dfrac{M_2}{M_0} - \mu^2$ is the variance. Matching of moments is a common technique for solving differential equations, a.o. in transport modeling [e.g., *Yu et al.,* 1999*; Luo et al.,* 2006].

*table 5.1: Model results and parameter estimates for the groundwater head series observed in well 46DP0032*

| EVP | 92.0 % |
| --- | --- |
| **RMSE (m)** | 0.068 |
| **Drainage base (m)** | 11.53 |

| Input series | Stress type | Parameter | Value | $\sigma$ |
| --- | --- | --- | --- | --- |
| VENRAY | Precipitation | $A(M_{0,p})$ | 299.2 | 23 |
| | | $b$ | 0.003117 | 0.00033 |
| | | $n$ | 0.898 | 0.025 |
| EINDHOVEN | Evaporation | $f$ | 1.16 | 0.076 |
| P.S. BERGEN | Pumping well | $\alpha$ | 2.58 | 2.4 |
| | | $\beta$ | 0.0522 | 0.027 |
| | | $\gamma$ | 0.00585 | 0.03 |
| | | $M_{0,w}$ | -3.63E-05 | 5.68E-06 |
| SambeekBoven | River | $\alpha'$ | 0.02527 | 0.094 |
| | | $\beta'$ | 0.0368 | 0.13 |
| | | $\gamma'$ | 0.7656 | 0.057 |
| | | $M_{0,s}$ | 0.7279 | 0.0939 |

## 5.3    Example application

### 5.3.1    Single series

The functioning and results of a multiple input PIRFICT model on a single well located in the northern part of the province of Limburg (the Netherlands) are presented in this subsection. The well, with the national code 46DP0032, has two screens. We consider only the top one, which is located 13 meters below the surface. The well is located on the edge of the floodplain of the river Meuse, in an aquifer that consists of coarse, gravelly sands overlain by finer sands which originate from river deposits. In addition to the river level fluctuations, the head is influenced by precipitation and evaporation and by a pumping station near the town of Bergen where groundwater is withdrawn for drinking water production.

Time series data of all stresses are available. The precipitation and potential evaporation series originate from stations of the Royal Dutch Meteorologic Institute in Venray and Eindhoven, respectively. The river levels were monitored at a dam in Sambeek, downstreams of well 46DP0032; the pumping rates were obtained from the drinking water company of Limburg. The parameters of all three impulse response functions are optimized using the methods described in Von Asmuth and Bierkens [2005]. Here, we will consider the results of the transfer part of the model and of this

individual series only, after the first run of the model. The model results and parameter estimates are summarized in table 5.1. In the table, two parameters are given that define the goodness of fit: the explained variance percentage (EVP) and the Root Mean Squared Error (RMSE). In the definition of EVP, the residual variance $\sigma^2_{n(t)}$ is weighed according to the variance $\sigma^2_{h(t)}$ of the original signal in the following manner:

$$\text{EVP} = \frac{\sigma^2_{h(t)} - \sigma^2_{n(t)}}{\sigma^2_{h(t)}} * 100\% \tag{5.15}$$

The results of the time series model are shown in figure 5.2, where the measured heads $h$ (dots) and predicted heads $\sum_{i=1}^{N} h_i + d$ (solid) are plotted together in the upper graph, while $h_i$ is plotted for every stress in the graphs beneath.
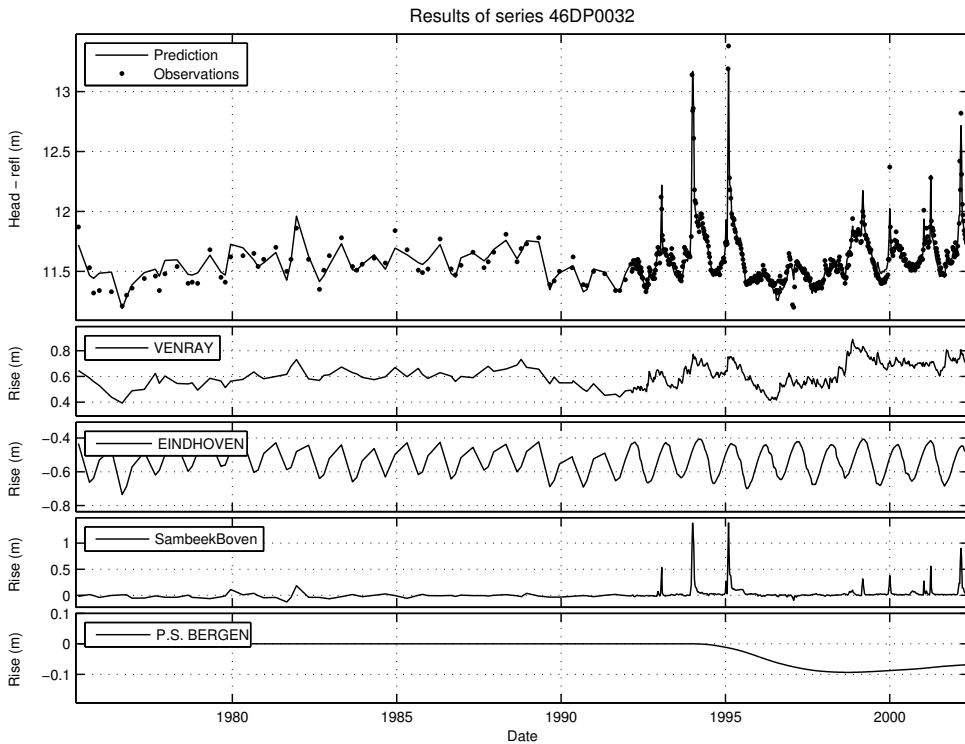
*figure 5.2: The ground water head fluctuations in well 46DP0032 (top) decomposed in four partial series due to (from top to bottom) rainfall, evaporation, river level fluctuations and a pumping station. Adding the partial series and the estimated drainage base ( $d$ , see table 5.1) results in the predicted series.*

Thus, the groundwater level series is decomposed into four partial series, which show the effects of the individual stresses. For example, it may be observed that the recent series of wet years in the Netherlands (1999-2002) has led to an overall increase in the groundwater levels. Furthermore, the effect of evaporation doesn't vary much from year to year, but shows a distinct seasonal pattern as it is mainly influenced by temperature and solar radiation. The river level has little effect as it is maintained by dams, except for high-water events when the river leaves its channel and enters the floodplain. In those cases, the head responds very quickly and shows distinct peaks. Finally, heads show a gradual decline since 1994 due to the groundwater withdrawal that started around that time. Note that these distinct differences in dynamic behavior allow the time series model to distinguish the effects of the individual stresses. Also note that the frequency of the groundwater level observations changes around 1992 from four times a year to once a week. This does not pose a problem for the model, as the impulse response equations are continuous in time and predictions are not fixed to a certain time discretization.

For each of the different stresses, an impulse response function is estimated (although not always independent from the other stresses, see equations (5.3) to (5.5)). The impulse response function forms the heart of the time series model and represents the response of heads to a unit impulse of the stress. The shape of the impulse response function depends on the position of the observation well and the properties of the system. As an example, the impulse response function of the level of the river Meuse is plotted in figure 5.3. From the graph, one can infer that the heads indeed respond quickly to a rise and fall in river level, as the impulse response function (solid line) peaks in less than half a day. Apart from its dynamic response, one is often also
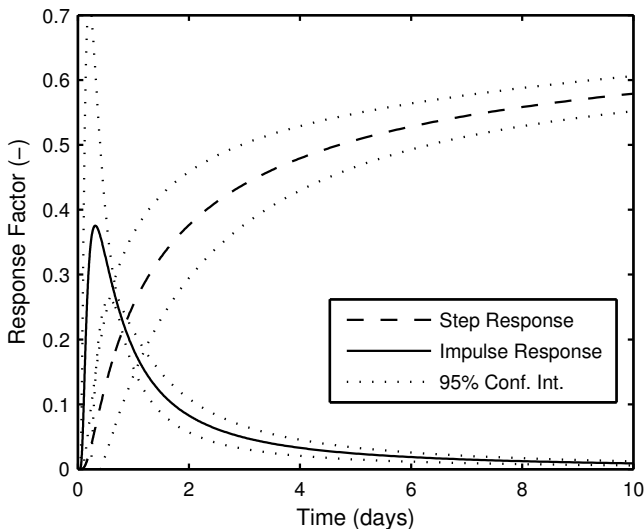


*figure 5.3:* *Estimated impulse and step response functions of observation well 46DP0032 for the level of the river Meuse (measured at dam Sambeek-Boven).*
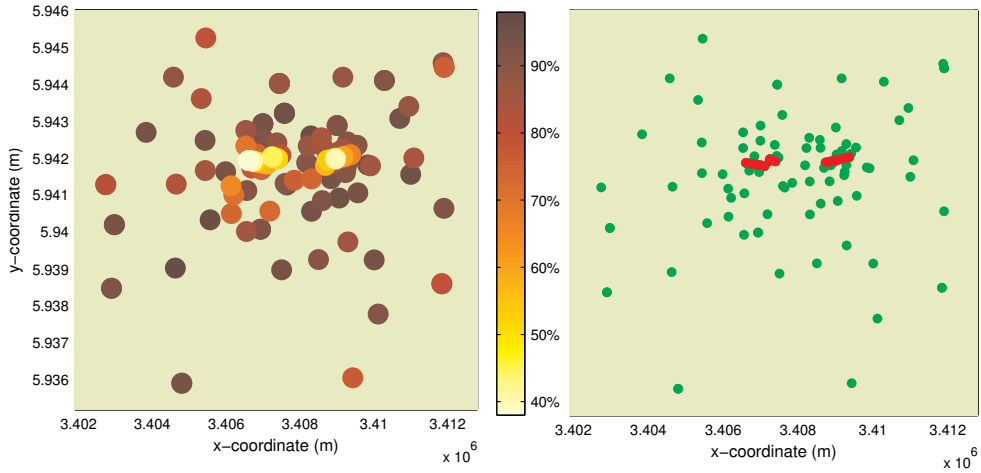
*figure 5.4: Spatial distribution of the explained variance percentage (EVP), in plan view (left figure). Model performance in general is lowest near the 15 pumping well screens, which are clustered in two rows (red dots in right figure).*

123

interested in the stationary influence of a certain stress. This is represented by the step response function (dashed line), which is the integral of the impulse response function with respect to time; the step response represents the response of the system to a sudden and then ongoing rise of the river level. The level that the step response function approaches towards infinity is the zeroth moment ($M_0$), also known as the gain in the time series analysis literature. Using (5.11), we can obtain $M_0$ for stresses of type $s$ as:

$$M_{0,s} = \int_0^\infty \theta_s(t)\mathrm{d}t = \gamma' \exp(-2\alpha') \tag{5.16}$$

In the case presented, $M_{0,s}$ for the river Meuse is 0.73 meter, which means that a river level rise of 1 meter will eventually lead to a groundwater level rise of 73 centimeters on this location. For as long as the assumption of linearity holds, the zeroth moment can be used for quick scenario calculations, as the multiplication of $M_{0,s}$ and a planned rise in the river level yields a prediction of the groundwater level rise. Note that the standard deviation of $M_{0,s}$ is reasonable, whereas the standard deviation of parameter $\alpha$ is large (a similar case is true for $M_{0,w}$ of the groundwater withdrawal). This is caused by the covariance between the parameters.
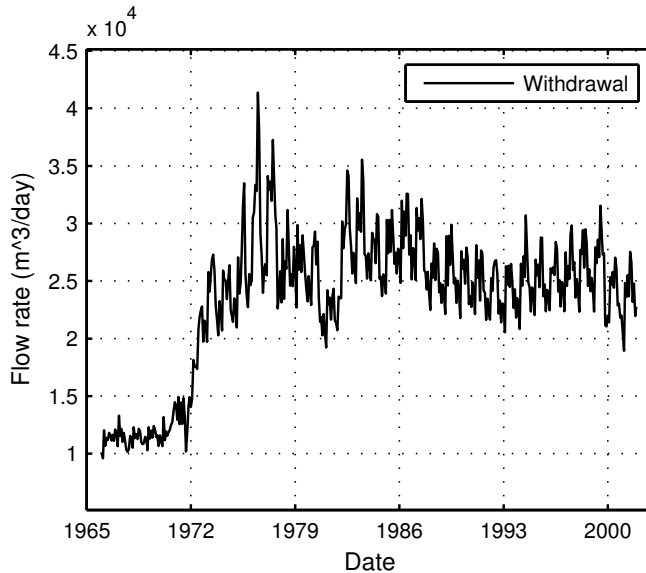
*figure 5.5: Groundwater withdrawal history at location 'Harlingerland'. The flow rate is obtained by summing the rates of the individual pumping wells.*

### 5.3.2 Multiple series

We will illustrate the procedure in which the results from multiple observation wells may be interpreted and checked in a second application. This application concerns a drinking water production area with multiple observation wells located in the north-western part of Germany, in an area called 'Harlingerland' near the Town of Esens. Groundwater is withdrawn by the drinking water company Oldenburgisch-Ostfriesischer Wasserverband at a current rate of 9 million cubic meters a year. The fifteen pumping wells are clustered in two rows, which are more or less placed in series (figure 5.4). The well screens are located at a depth of 25 to 40 meters from the soil surface. The withdrawal rate history (figure 5.5) shows a marked increase in the period from 1972 until 1976. Such a rate change is very important for a reliable estimate of the influence of a pumping well, as it will have caused a marked drawdown. Furthermore, the withdrawal rates show a distinct seasonal cycle, caused by factors like the increased watering of yards in dry periods. The withdrawal rate is obtained by summing the pumping rates of the individual wells. Data on the individual wells, which are controlled separately, is only available since 1992. Anyway, use of the individual pumping rates is not straightforward, as the number of 15 wells is too high to estimate their influence independently. Groundwater dynamics are recorded at 116 observation wells with a total of 139 screens in a circle with a radius of 5 kilometers around the pumping wells. The screens depths range from less than 1 to more than 90 meters from the surface. Part of the wells are monitored every two weeks, and the other part monthly. The period in which the wells were monitored differs from well to well, with the earliest measurements dating back to 1964, while other wells were installed as recently as 1999. In some of the wells, monitoring

*table 5.2: Range of model results and parameter estimates for all 139 groundwater head series.*

| Parameter | Minimum (+/- $2\sigma$) | Median | Maximum (+/- $2\sigma$) |
|:---:|:---:|:---:|:---:|
| EVP | 38.1 | 85.9 | 97.5 |
| $M_{0,p}$ | 59 (+/- 1.3e3) | 864 | 6580 (+/- 3.5e5) |
| $M_{0,e}$ | 172 (+/-122) | 790 | 7976 (+/-5913) |
| $M_{0,w}$ | 1.93e-8 (+/- 4.1e-5) | 4.50e-5 | 1.25 e-2 (+/- 0.20) |
| f | 0.09 (+/- 0.31) | 1.20 | 4.29 (+/- 87) |
| d | -3.21 | 1.26 | 7.01 |

stopped in 1975, while in some others the period between 1975 and 1998 is missing. The aquifer in the area consists mainly of sands; in some parts a clay layer separates the phreatic from the deeper aquifer.

The head fluctuations are modeled with precipitation, reference evapotranspiration and groundwater withdrawals; there are no rivers or other important fluctuating surface waters in the area. Results of the 139 time series models are summarized in table 5.2, where the minimum, median and maximum value of the parameters in all 139 models are given, along with their 95% confidence interval. The median value of EVP points out that the fit of most models is good, although there are clearly outliers in the results, as the extremes of most parameter estimates are not within a range that is physically plausible. This does not mean that the estimates are necessarily biased, as most extremes are accompanied by large standard deviations. For these cases, a confidence interval of +/- $2\sigma$ includes the value zero, but also a realistic value, indicating that the quality or quantity of the data is not sufficient to determine the value of that parameter.

The only estimate in table 5.2 that does seem biased is the minimum value for the evaporation factor $f$ , which is very low (upper 95% limit is 0.4). In this case, however, the zeroth moment of the evaporation $M_{0,e}$ which is defined as:

$$M_{0,e} = \int_0^\infty f\theta_p(t)\mathrm{d}t = fA \qquad (5.17)$$

does have a large confidence interval (+/- 941), as the influence of precipitation itself ($A$) cannot be determined from the data accurately (the probable cause is that the period in which data are available for this particular series only spans 16 months, with one observation per month).
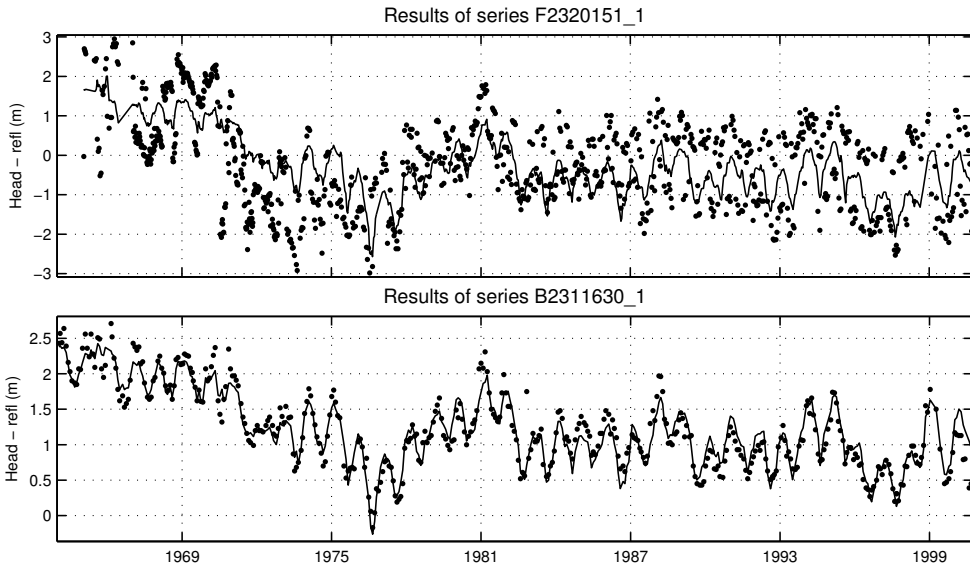
*figure 5.6: Head observations and predictions of wells F230151(near the pumping station) and B2311630 (at some distance), giving an indication of good and bad model fits.*

To determine the reason for the lower EVP of some models, its spatial distribution is plotted (figure 5.4). From the figure, it is clear that the model fits are lowest in the immediate vicinity of the pumping wells. This suggests that the influence of the individual wells is not modeled correctly, which is indeed the case, as the withdrawal rate of the total well field is incorporated in the model, not the rates of the individual wells. The individual model fits (figure 5.6) give the same indication. Near the well field, the head fluctuates wildly when the individual pumping wells are shut on and off. The predictions, however, only follow the general drawdown pattern. At some distance, the predictions fit the head fluctuations much better. The model fit is also lower in some of the shallower wells. The head series in those cases show distinct signs of (threshold) non-linearity [e.g., *Knotters and De Gooijer,* 1999]. The PIRFICT method can currently handle threshold non-linearity for precipitation and evaporation, but not yet for other stresses, so this was not considered further at this time.

Next, we focus on the estimated influence of the well field, which was an important objective for doing the time series analysis of this site. Using (5.9), we find that the zeroth moment of a well is given by [*Hantush,* 1956]:

$$M_{0,w} = \int_0^\infty \theta_w(t)\mathrm{d}t = -2\gamma\, \mathrm{K}_0(2\alpha) \tag{5.18}$$
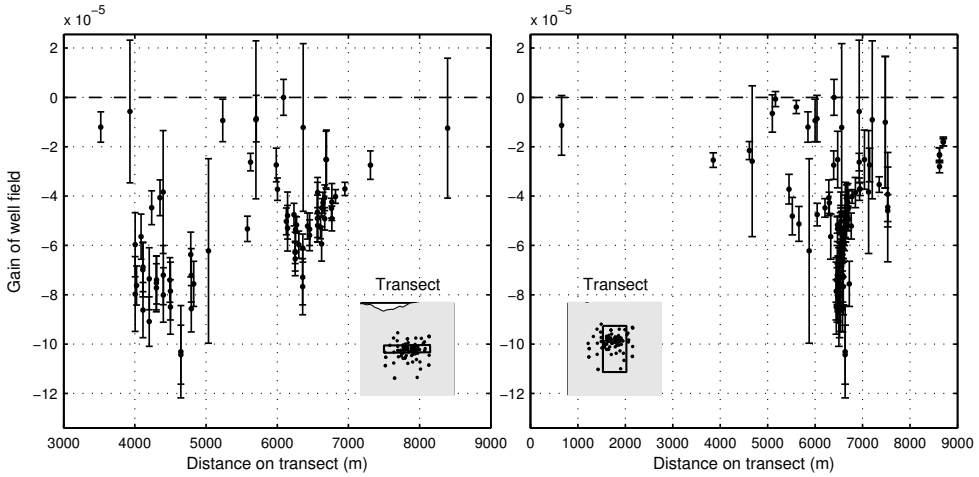
*figure* 5.7: *Estimated gains (dots) for the well field in the different observation wells, presented in west-east (left) and north-south (right) cross sections. The error bars indicate the 95% confidence interval of the estimates.*

in which $K_0$ is the modified Bessel function of the second kind and zeroth order [*Abramowitz and Stegun,* 1964]. First, we remove all wells with estimates of $M_{0,w}$ with large confidence intervals, as they are of little value and only blur our view on the other results; we choose a cut-off value of 8e-05. The 34 series that were disregarded all lack the period from 1972 to 1976 in which the withdrawal rate was increased significantly. In figure 5.7, the spatial distribution of the 105 remaining $M_{0,w}$ estimates and their confidence interval is plotted in two cross sections, one parallel and one perpendicular to the series of pumping wells. As expected, the $M_{0,w}$ estimates are highest near the pumping wells; the left row of wells, which have deeper screens, cause stronger drawdowns than the right row. In the perpendicular cross section, the pattern of $M_{0,w}$ approximates the shape of a drawdown cone. The $M_{0,w}$ estimates at shallow well screens are clearly lower than at deeper well screens; this may be caused by a resistance layer or by recharge from ditches or creeks. This is in line with the head series of these observation wells (not shown here), where head differences between higher and lower screens range up to 4 meters. Although the model fit is low in the vicinity of the pumping well screens (within a radius of about 200 to 300 meters), all in all, the $M_{0,w}$ estimates seem nevertheless reasonable. The estimates of other parameters near the wells, however, are clearly biased. In figure 5.8 cross sections are shown of the EVP and $M_{0,e}$ for the observation wells in the vicinity of the pumping wells. The low values of EVP near the pumping well screens correlate
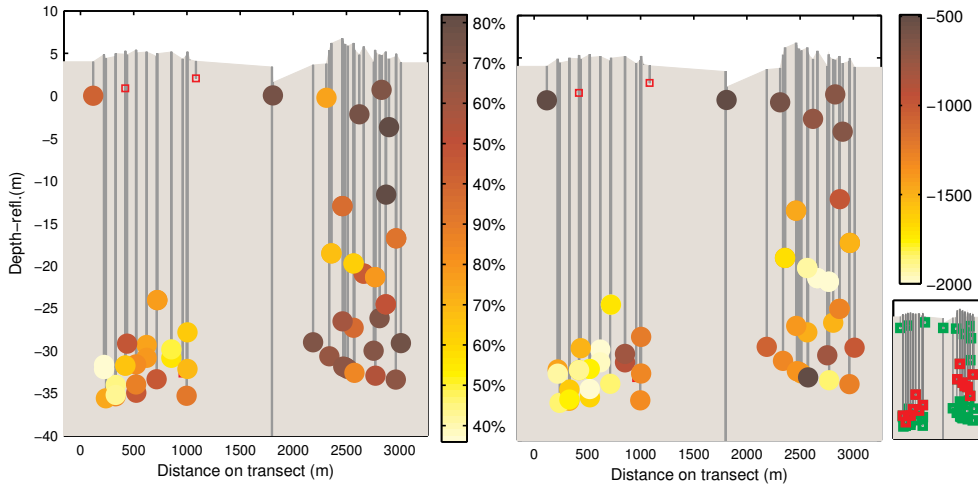
*figure 5.8: West-east cross sections near the pumping wells, showing the percentage of variance accounted for (left) and the estimated gain of the evapotranspiration (right) in the different observation wells. The values of both are lowest near the pumping well screens (the red markers in the figure bottom right).*

with the low values of $M_{0,e}$. As it is not logical for the influence of evapotranspiration to be that heterogeneous in deeper soil layers, this indicates that the influence of evapotranspiration is overestimated and partly correlates with the effect of the individual pumping wells.

## 5.4    Discussion and conclusions

In this paper, the PIRFICT model for time series analysis was extended to handle multiple inputs. For stresses other than areal recharge, analytic solutions of simple hydrogeologic schematizations were used as a guide to develop appropriate impulse response functions. Different stresses are not necessarily connected in parallel in the model, as is the standard procedure in ARMA models. A case with a single observation well was used to illustrate how the model can effectively decompose head series into partial series that each show the effect of an individual stress. In the example with multiple observation wells, it was shown that, next to its a priori use in defining the impulse response functions, physical knowledge is also valuable in checking the consistency and plausibility of the model results a posteriori. The parameter values should fall within a range that is physically plausible. The spatiotemporal patterns observed in the variables supply important and independent feedback on the results, as there is no spatial dependency imposed on the models. By focusing on the model residuals, missing stresses, processes or other sources of error may be readily identified. High error levels are a possible source of bias in the estimates, as other stresses may partly compensate missing stresses when their influence is correlated. In this case, the main source of error was the fact that the behavior of the individual pumping wells was not accounted for. In spite of this, the

overall drawdown pattern of the well field in total, which is often the factor of interest, was represented well. The spatial distribution showed that the estimates for the evaporation series were clearly biased. Such a problem may be corrected by incorporating data on the missing factors in the model (in this case the pumping at individual wells). If such data are not available (as in this case), the bias can be reduced by constraining the optimization problem to a realistic range, based on the values and patterns in the surrounding observation wells. Future research will include methods to impose a spatial coherence a priori in the model, so that the individual dynamics of a high number of pumping wells can be incorporated in the model, and the results of neighboring observation wells are linked and remain plausible.

The presented time series model may be applied to decompose series of head fluctuations into partial series, representing the influence of individual stresses. This enables the evaluation of individual effects of a stress, such as hydrologic interventions, pumping wells, climatic changes and surface water levels. Furthermore, the presented model may be used for forecasting, gap filling, scenario studies, trend analyzes, and optimization and control of hydrogeologic systems (using more or less standard time series analysis methods, see e.g., Hipel and McLeod [1994]. The PIRFICT approach is particularly suited for the batch processing of many series, it uses a small number of parameters, and it is not limited by irregular or high-frequency data. Although here, we specifically focus on groundwater heads, the approach presented can also be applied to (contaminant) transport problems, or for that matter, to hydrologic problems in general. In the field of transport, the scaled gamma distribution has proven to be well usable also (it matches the convection dispersion equation [see e.g., *Jury and Roth*, 1990; *Maas*, 1994]), whereas in other fields, care has to be taken to select appropriate impulse response functions and methods for the processes and stresses occurring there.

An attractive feature of time series models is that they are based on relatively few assumptions and the fits are generally high; the model lets the data more or less speak for itself. As such, time series models are a valuable tool for preprocessing groundwater level series before calibrating a groundwater model. Using time series models, missing stresses or series that are influenced by hydrologic interventions may be readily identified. Also, series may be identified that are not suitable for model calibration, for instance because they represent a hydrologic feature that is not incorporated in the model, such as perched water tables. One step further, transient groundwater models may be calibrated directly on the moments of the impulse response functions estimated with time series analysis [*Van de Vliet and Boekelman,* 1998*; Von Asmuth and Maas,* 2001], which requires the calibration of steady models only. Moments of the impulse response functions can also be modeled spatially with the analytic element method, creating a possibility for transient modeling with analytic elements [*Bakker et al.,* 2007].

**Abstract:**

For e.g., visual interpretation, mapping or empirical modeling purposes, the amount of information contained in a full spatiotemporal description of the groundwater table dynamics is simply too large. For such purposes, the data has to be compressed without loosing too much information. Methods have been developed to visualize the groundwater regime in overall graphs, or statistically characterize the dynamics with a limited set of parameters. More recently, methods have been sought to identify the properties that determine the dynamics of a groundwater system. In such approaches, it is believed that the spatial differences in the groundwater dynamics are determined by the system properties, while its temporal variation is driven by the dynamics of the input into the system. In this chapter, a method is presented that links the dynamics of the input to the spatially variable system properties, and results in a new set of parameters that characterize the groundwater dynamics (GD). While the dynamics of the input are characterized by its mean level and annual amplitude, the functioning of the groundwater system is characterized by its impulse response (IR) function. The IR function can for instance be estimated empirically using a time series model. Subsequently, the input and system characteristics are combined into a set of parameters that describe the output, or groundwater dynamics, using simple analytic expressions. It is shown that these so-called GD characteristics (the mean depth, convexity, annual amplitude and phase shift), can describe the groundwater dynamics in detail (for as far as the time series model can). In the example application, the GD characteristics are compared to other methods for characterizing the groundwater regime, using two example series of groundwater level observations. It is shown that the so-called MxGL statistics (Mean Highest, Lowest or Spring Groundwater Level) that are often used have some important drawbacks, as they filter out the low-frequency dynamics of a system and mix-up annual with higher frequencies. Consequently, it is concluded that the capability of MxGL statistics in characterizing the groundwater dynamics at different locations is less than that of GD characteristics.

# Chapter

6

# Characterizing groundwater dynamics based on response characteristics

**Adopted from:**

Von Asmuth, J. R., and M. Knotters (2004)
Characterising spatial differences in groundwater dynamics based on a system identification approach.
*Journal of Hydrology*, 296(1-4), 118-134.

## 6.1    Introduction

In deltas, wetlands, and other low-lying areas, the groundwater table can usually be found at shallow depths. In about 50% of the Dutch soils, for example, the mean highest groundwater level (MHGL) is less than 40 cm below the soil surface [*Van der Sluijs,* 1990]. In such areas, the groundwater table and its temporal variability can be of great economical as well as environmental importance. Of course, the most visible and direct economical impact of groundwater dynamics occurs through groundwater (related) inundation or droughts [*Baker et al.,* 1988*; Jarrett,* 1991*; Leeuwis-Tolboom and Peters,* 2002]. In urban areas, high groundwater levels can cause flooding of tunnels or basements, while for instance the wooden pile foundation of buildings can be damaged when water tables drop below the pile heads. In agricultural areas, high groundwater levels can cause crops to perish, agricultural fields to become inaccessible for machinery and harvesting, whereas moisture availability and growth is reduced when the capillary fringe drops below the root zone [*Feddes,* 1981*; Kroes et al.,* 2000]. Wetlands and other groundwater dependent sites are known to harbor many rare and species rich animal and plant communities, and can be very vulnerable to changes in the hydrologic circumstances [e.g., *Grootjans,* 1985, *Witte et al.,* 1992]. Apart from this direct impact, the hydrology of an area also greatly determines soil formation and other site conditions for agricultural crops or spontaneous vegetation [*Porporato et al.,* 2003]. Because of its economical impact, however, the groundwater regime has been adapted by man in order to optimize agricultural production and minimize water hazard. Moreover, the groundwater regime has often also been influenced by factors such as groundwater abstraction, land reclamation, urbanization and canalization of rivers and streams. Because of the strong spatial dependence of groundwater flow, such interventions did not only influence the groundwater levels in their immediate vicinity. Also in nearby and more remote areas, groundwater levels dropped, negatively influencing agricultural production during dry periods and threatening groundwater related habitats. In order to optimize and balance the interests of economical and ecological land use purposes, knowledge about the spatiotemporal dynamics of the groundwater table is very important.

The amount of information contained in a full description of the groundwater dynamics is simply too large for, e.g., visual interpretation, mapping or empirical modeling purposes. For such purposes, the data has to be compressed without loosing too much information. In models that empirically explain the influence of the groundwater regime on phenomena related to it, for example, the number of explanatory variables should be kept as low as possible, in order to avoid overfitting behavior or coincidental correlations. Consequently, methods have been developed to visualize the groundwater regime in overall graphs, or to statistically characterize the dynamics with a limited set of parameters. Examples of graphs that describe the groundwater regime at a certain location are groundwater hydrographs, regime curves or frequency of exceedence graphs. A groundwater hydrograph, of course, forms the most straightforward plot of the information. A regime curve displays the average annual course, and is used, e.g., to test whether the course of the groundwater in a certain year deviates from the normal pattern. Frequency of exceedence graphs give information about the extreme levels that occur at a certain location, and can be used to identify which natural habitats and ecosystems can occur

at which groundwater regimes [e.g., *Nieman,* 1973]. Examples of statistical characteristics that describe the groundwater regime are the Mean Highest, Mean Lowest and Mean Spring Groundwater Level (MHGL,MLGL,MSGL, together referred to as MxGL statistics [*Van Heesen,* 1970*; Van der Sluijs and De Gruijter,* 1985]). Such statistics are used, for instance, to analyze or model the relationship between the groundwater regime and factors such as crop growth [*Feddes et al.,* 1988*; Van Dam,* 2000], soil conditions or ecology [*Grootjans,* 1985*; Witte et al.,* 1992*; Van Ek et al.,* 2000]. Furthermore, they are used to generate and display spatial information about the groundwater regime [e.g., *Van Heesen,* 1970*; Boucneau et al.,* 1996*; Bierkens et al.,* 2000*; Hartung,* 2002] for urban and rural planning purposes or policy making. More recently, methods have been sought to identify the properties that determine the dynamics of a groundwater system. Such methods regard the problem from the viewpoint of system identification or time series analysis [*Box and Jenkins,* 1970*; Hipel and McLeod,* 1994*; Ljung,* 1999], or use the somewhat less sophisticated multiple regression methods. Here, the groundwater system is seen as a black box that transforms series of observations of the input or explanatory variables into a series of the output variable or groundwater level. In groundwater systems that are not disturbed by groundwater abstraction or other influences, the climatologic conditions can be considered the only input. This approach can be used to characterize or map the spatial variation in the groundwater dynamics, as it is believed that the spatial differences in the groundwater dynamics are determined by the spatial variation in the system properties, while its temporal variation is driven by the dynamics of the input into the system. The spatial variation of the input or climatologic conditions is taken to be low (at least on a small scale) and can therefore be neglected. Examples of studies where this approach is followed can be found in [*De Castro Ochoa and Munoz Reinoso,* 1997*; Lammerts et al.,* 2001*; Von Asmuth and Maas,* 2001*; Hartung,* 2002]. There are important drawbacks to the methods described above. Low-frequency dynamics in the meteorologic conditions pose a problem for all methods that are based on groundwater level observations only, as the groundwater level can vary considerably between periods. Van Heesen [*Van Heesen,* 1970], for example, already prescribed a minimal observation period of 8 years for the calculation of MxGL statistics, while Knotters and Van Walsum [*Knotters and Van Walsum,* 1997] showed that there is still considerable variation in the groundwater dynamics above the 8 year time scale. In order to filter out long-term fluctuations more effectively and thus obtain temporally invariable or climate-representative statistics, Knotters and Van Walsum have recommended the use of 30 year periods for the calculation of MxGL statistics. Since groundwater level series are generally not available over such long periods, they use a time series model to extend limited series of observations to 30 years. However, even with climate-representative MxGL statistics, a disadvantage is that much information is lost, as MxGL statistics focus only on the annual groundwater cycle and filter out the other frequencies (by selecting the extreme and spring levels every year, and then averaging over the observation period). Although the use of system characteristics seems promising for local studies in undisturbed groundwater systems, they can not take into account effects of spatial variable or non-stationary influences (e.g., global climate change). The spatial variability in climatologic conditions is often low on a small scale, but it can be significant for

research on larger scales or in mountainous terrain and also other influences such as groundwater abstraction are often non-stationary and spatially variable.

In this paper, a method is presented that links the response characteristics of groundwater systems to the dynamic behavior of the input, and combines them into parameters that describe the groundwater dynamics. By doing so, the limitations of using system characteristics can be circumvented, while the key characteristics of the output can be easily identified. Furthermore, knowledge is gained on the relationship between input dynamics, system properties and output dynamics. The method is based on the use of a continuous time Transfer Function Noise (TFN) model, which estimates the Impulse Response (IR) function of the system from the temporal correlation between time series of groundwater level and precipitation surplus. As stated earlier, the use of a time series model is necessary to correct for non-average meteorologic circumstances in the observation period. In an earlier paper [*Von Asmuth et al.,* 2002], it was shown that this so-called PIRFICT-model has practical advantages, but the same background and general applicability as Box & Jenkins TFN-models, which are widely used in the field of groundwater hydrology and beyond [*Tankersley et al.,* 1993*; Gehrels et al.,* 1994*; Hipel and McLeod,* 1994*; Van Geer and Zuur,* 1997]. For information about the parameter estimation process and model validation and application aspects, we refer to [*Von Asmuth et al.,* 2002].

The paper is organized as follows: first, the theory and background are given, and are summarized at the end of the section; in the example application, the GD characteristics are compared to MxGL statistics, regime curves and frequency of exceedence graphs, using data from two different groundwater systems; last, but not least, a discussion and conclusions are given.

## 6.2   Methods and theory

In this section the methods will be developed with which the meteorologic input, system properties and output dynamics are described, estimated and linked to each other.

### 6.2.1   Meteorologic characteristics

In order to characterize the variability of precipitation and evaporation, several methods could be used. First of all, because of the annual meteorologic cycle, a time series of precipitation surplus is periodic and a fast Fourier transform of the series is used to estimate and characterize its amplitude spectrum (figure 6.1a). Indeed, in the spectrum we see a clearly isolated peak at annual frequency, while the power in the rest of the signal appears to be rather randomly and uniformly distributed over both the lower and higher frequencies. Because of this, we can rely on a simple but effective method to estimate the average precipitation surplus intensity and its annual amplitude. A disadvantage of using Fourier analysis in order to obtain the annual amplitude is the fact that it requires the specification of a spectral window, which can be somewhat subjective when the annual frequency is not well isolated. To start with, the average level $\overline{p}$ of the precipitation surplus is, logically, obtained with:
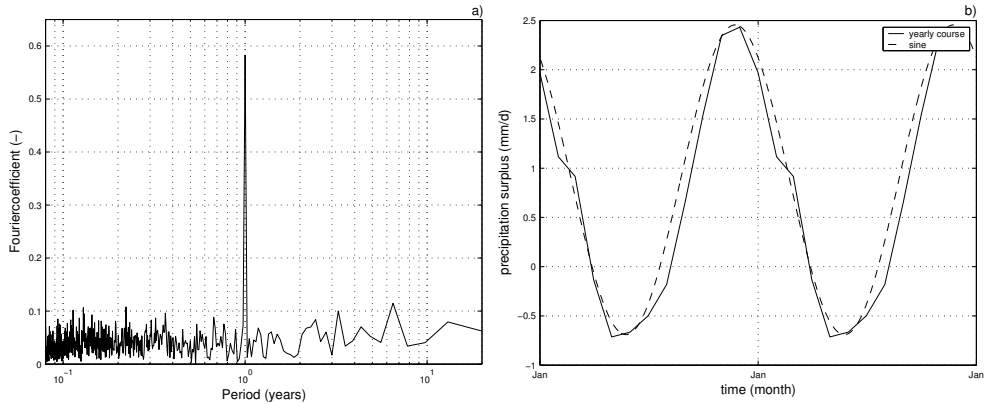
*figure 6.1: a) Amplitude spectrum of the precipitation surplus series at the town of De Bilt (period 1983-2001), showing an isolated peak at annual frequency. b) Average annual course of the precipitation surplus and a sine with the same frequency, phase and amplitude.*

$$\overline{p} = \frac{\int\limits_{t_{pb}}^{t_{pe}} p(\tau)d\tau}{t_{pe} - t_{pb}} \tag{6.1}$$

with $t_{pb}$ and $t_{pe}$ denoting the start and end of the period over which the meteorologic characteristics are calculated. Next, time is split into year $Y$ and Julian day $D$, and the precipitation surplus is averaged over $Y$, which effectively filters out its yearly course $\tilde{p}$:

$$\tilde{p}(D) = \frac{\sum\limits_{Y_{pb}}^{Y_{pe}} p(Y,D)}{Y_{pe} - Y_{pb}}, \quad 1 \le D \le 365 \tag{6.2}$$

Because the temperature largely determines the annual evaporation cycle and is more or less harmonic, so is the precipitation surplus, and the annual amplitude can be obtained by matching a sine to the yearly course (figure 6.1b).

### 6.2.2 Response characteristics

The functioning of linear systems can be completely characterized by their impulse response (IR) function [*Ziemer et al.,* 1998*; Von Asmuth and Maas,* 2001]. In this section, we will describe the methods with which the IR function is estimated, and use its moments to avoid restricting ourselves to the specific type of IR-function used in our model. An elegant way to obtain estimates of the response of hydrologic systems is the use of a transfer function noise (TFN) time series model [*Box and Jenkins,* 1970].

For groundwater systems, the response could also be obtained using a deterministic groundwater flow model such as MODFLOW. Here the use of an empirical time series model is preferred over the use of deterministic models, because the construction of time series models is less time-consuming and often also yields more accurate predictions of the groundwater level [*Hipel and McLeod,* 1994*; Knotters,* 2001]. In TFN models, a series of groundwater level observations is modeled as a linear [or non-linear, e.g., *Tong,* 1990*; Knotters and De Gooijer,* 1999] transformation of the time series that influence the groundwater level. The deviations between predicted and observed output are modeled with a separate so-called noise model. Where the transfer function in discrete Box-Jenkins models is seen as the response to a momentary impulse, from a continuous point of view it is actually the response to an input series with the shape of a block. As such, it is not an invariant system characteristic, because it is still dependent on the observation frequency. Because of this, and because Box-Jenkins TFN models have some important limitations as they only operate on a regular frequency, here we use a continuous time TFN model. For the simple case of a linear, undisturbed phreatic system that is influenced by precipitation surplus only, the following single input continuous TFN model can be used to model the relationship between groundwater dynamics and precipitation surplus [*Von Asmuth et al.,* 2002]:

$$
\begin{cases}
h(t) = \hat{h}(t) + d + n(t) \\[2mm]
\hat{h}(t) = \int_{-\infty}^{t} p(\tau)\theta(t-\tau)d\tau \\[2mm]
n(t) = \int_{-\infty}^{t} \phi(t-\tau)dW(\tau)
\end{cases}
\tag{6.3}
$$

where
$t$     : time [T]
$h$     : observed groundwater level [L]
$\hat{h}$     : predicted groundwater level [L] attributable to $p$ and relative to $d$
$p$     : precipitation surplus intensity [LT$^{-1}$]
$d$     : level of $h$ without precipitation, or local drainage level, relative to some reference level [L]
$n$     : residual series [L].
$\theta$     : transfer impulse response function [-]
$\phi$     : noise impulse response function [-]
$W(t)$ : continuous white noise (Wiener) process [L], with properties $\mathrm{E}\{dW(t)\} = 0$,

$\qquad \mathrm{E}[\{dW(t)\}^2] = dt$, $\mathrm{E}[dW(t_1)dW(t_2)] = 0, t_1 \neq t_2$.

TFN models are identified by choosing mathematical functions, or in the discrete case a set of so-called Auto-Regressive and/or Moving Average (ARMA) parameters, to represent the transfer and noise IR functions. Model identification can be done iteratively, using the correlation structures in the available data and model diagnostics, or based on insight into the physical behavior of the analyzed system. On the basis of physical argumentation [*Von Asmuth et al.,* 2002], we choose the following functions as IR-functions:

$$\theta(t) = A\frac{a^n t^{n-1}\exp(-at)}{\Gamma(n)} \tag{6.4}$$

$$\phi(t) = \sqrt{2\alpha\sigma_n^2}\exp(-\alpha t) \tag{6.5}$$

with $\sigma_n^2$ denoting the variance of the residuals. The physical meaning of (6.5) and its parameter set $\Psi = [A, a, n, \alpha]$ are explained in section 2.4.2. $\Psi$ is estimated from the available data using a Levenberg-Marquardt algorithm and an objective function based on the likelihood function of the noise model [*Von Asmuth and Bierkens,* 2005]. The local drainage level $d$ is obtained from the data in the following way:

$$d = \frac{\sum_{i=0}^{N}h(t_i)}{N} - \frac{\sum_{i=0}^{N}\hat{h}(t_i)}{N} - \frac{\sum_{i=0}^{N}n(t_i)}{N} \tag{6.6}$$

with $N$ being the number of groundwater level observations. $\theta$ is a scaled version of the gamma distribution function (SG df, [*Abramowitz and Stegun,* 1964]) and can, because of its flexible nature, adequately model the response of a broad range of groundwater systems, e.g., those with a simple exponential response [*Knotters and Van Walsum,* 1997; *Bierkens et al.,* 2000], a Gaussian response [*Chen et al.,* 2002] or more in general with a range of ARMA-type responses [*Von Asmuth et al.,* 2002]. For other, linear or approximately linear, systems, the SG df will probably capture the most important characteristics of the systems response well. Under the assumption of linearity, the deterministic part of the groundwater dynamics is completely determined by the IR function. As such, the characteristics of the IR function (i.e. its moments) could well be used to characterize the dynamics of groundwater systems for economical, ecological or other purposes. By using moments, the estimates of various types of TFN models and IR functions can be easily compared. The moments of the IR function have the added advantage that they can be obtained not only from time series models, but from analytic or distributed groundwater models as well [*Von Asmuth and Maas,* 2001]. The parameters $A, a$ and $n$ of the SG df are related to the first three moments in the following manner [*Abramowitz and Stegun,* 1964]:

$$M_i = \int_0^{\infty} t^i A\frac{a^n t^{n-1}\exp(-at)}{\Gamma(n)}\,dt, \; M_0 = A, \; M_1 = A\frac{n}{a}, \; M_2 = A\frac{n}{a^2}(1+n) \tag{6.7}$$

with $i$ being the moment order.

### 6.2.3    Combination into GD-characteristics

Often not the system properties, but the behavior of the groundwater level itself is of direct interest for water management or other purposes. In the following we will derive a set of GD characteristics from the moments of the estimated IR function and the dynamic characteristics of the precipitation surplus series. When a TFN model is calibrated on a limited set of observations, an estimate of the mean groundwater level $\bar{h}$ and annual amplitude $\tilde{h}$ can be obtained using the step and frequency response functions. The step response function is the response of a system to an input series with unit intensity and the shape of a step, and equals:

$$\Theta(t) = \int_0^t \theta(\tau)\,\mathrm{d}\tau \tag{6.8}$$

The step response function asymptotically reaches a constant level that is known as the gain [*Box and Jenkins,* 1970] and equals the zeroth moment $M_0$ of $\theta$. In hydrologic terms, it is the level to which the water table rises when the precipitation surplus intensity would be constant and of unit intensity. From this, $\bar{h}$ can be estimated in the following way, taking in mind that the rise starts from the local drainage base $d$ :

$$\bar{h} = d + \bar{p}M_0 \tag{6.9}$$

with $M_0$ equalling $A$ when $\theta$ equals a SG df (see (6.7)). Because most economical and ecological activity takes place at surface level ($s$), often the groundwater depth (s-$\bar{h}$) is of more interest than the groundwater level ($\bar{h}$). However, the use of a reference level $s$ or $d$ that varies in space introduces another degree of freedom, that is compensated by specifying the convexity or mean distance of the groundwater table from the local drainage base ($\bar{h} - d$).

The frequency response is the Fourier transform of the impulse response function, or the response of a system to an input series in the shape of a sine with unit amplitude and frequency $\xi$ [*Ziemer et al.,* 1998]:

$$\Xi(t) = \int_{-\infty}^t \sin(\xi t)\theta(t-\tau)\,\mathrm{d}\tau \tag{6.10}$$

When we solve equation (6.10) for a system with a SG df shaped IR function, we find (see appendix):

$$\Xi(t) = \frac{A}{(1+\frac{\xi^2}{a^2})^{\frac{n}{2}}} \sin\{\xi t - \frac{n}{\xi}\arctan(\frac{\xi}{a})\} \tag{6.11}$$

The term before the sine function is the amplitude response function and can be used to estimate the annual amplitude (i.e. frequency $\xi = \frac{2\pi}{365.24} = 0.0172$) of the groundwater level fluctuations:

$$\tilde{h} = \frac{\lambda A}{(1+\frac{0.0172^2}{a^2})^{\frac{n}{2}}} \tag{6.12}$$

with $\lambda$ being the average amplitude of the precipitation surplus over a certain period (figure 6.1b). The last term of (6.11) is the phase response function. From this a factor $\eta$ can be obtained that equals the phase shift between the annual groundwater and precipitation surplus cycle:

$$\eta = \frac{n}{0.0172}\arctan(\frac{0.0172}{a}) \tag{6.13}$$

### 6.2.4 Summary of method

In summary, the method proposed consists of the following steps (see figure 6.2). First, the key dynamic characteristics of the input to the system are calculated, for univariate time series models being the mean and annual amplitude of the precipitation surplus. Second, an estimate of the response of the groundwater system is made, that is characteristic of its functioning. For this purpose, a continuous time series model is used and calibrated on time series of groundwater level observations and precipitation surplus. From the IR function, moments can be derived and used as system characteristics. Third, both system and input characteristics are combined into characteristics of the
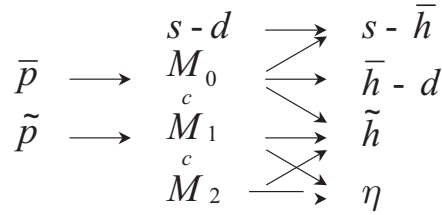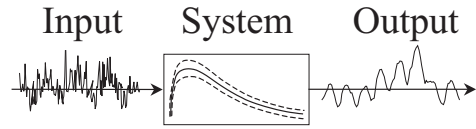
$\bar{p}$ = average precipitation surplus intensity [LT$^{-1}$]

$\tilde{p}$ = annual amplitude of $p(t)$ [LT$^{-1}$]

$s$ = surface level [L]

$d$ = local drainage level [L]

$M_n$ = $n^{th}$ order moment of the IR function [-]

$\bar{h}$ = average groundwater level [L]

$\tilde{h}$ = annual amplitude of $h(t)$ [L]

$\eta$ = time shift [T]

**figure 6.2:** *Relationship between input, system and output characteristics.*

output or groundwater table dynamics, using simple analytic expressions. The arrows in figure 6.2 illustrate the way in which the groundwater dynamics are determined by the properties of the system and the input, but the links can also be used vice versa. The average groundwater depth $s - \bar{h}$, for example, is determined by the distance of the soil surface from the local drainage level $s - d$, and by $M_0$ and $\bar{p}$, but vice versa, $\bar{p}$, or the average groundwater recharge, can also be estimated when $M_0$ and $\bar{h} - d$ are known. The deterministic part of the groundwater dynamics is completely determined by the four parameters that describe the output (for as far as the response of the real world systems fits the SG df response), since for linear systems, the transformation of input to output is completely determined by the IR function.

## 6.3 Example application

### 6.3.1 Set-up and data set

In the example application, the performance of the GD-characteristics in characterizing the groundwater dynamics is compared to that of MxGL statistics, regime curves and frequency of exceedence graphs. For this purpose, two series of groundwater level observations are selected that originated from rather different hydrologic systems, and also show a different type of dynamic behavior. Furthermore, both piezometers are undisturbed by groundwater abstraction, hydrologic interventions or other influences, thus allowing them to be modeled with precipitation surplus as the only input series. The first series originates from a dune reserve in the province of North-Holland, The Netherlands, near the town of Egmond. The dunes in that area form an approximately 1-5 kilometer wide ridge between the sea and the adjacent polder area, so the system is relatively large and isolated. The terrain is undulating and the soil consists of Holocenic aeolian sands. The groundwater level in this piezometer (19CNL5281) is well observed, with observations taken manually about the 14th and 28th of every month in the period from 5-12-1982 until 29-1-2001. As input for the time series model, observations of the precipitation and potential evaporation are used starting from 1-1-1973. The precipitation series is available on a

**table 6.1**: *Calibration results, estimated parameters and characteristics for piezometers 19CNL5281 and 32cl0034.*

| Piezometer | 19CNL5281 | | 32cl0034 | |
|---|---|---|---|---|
| $R^2$ | 0.919 | | 0.815 | |
| RMSE | 11.2 cm | | 8.4 cm | |
| $A$   $(+/- 2\sigma)$ | 1396 | $(+/- 336)$ | 172 | $(+/- 16)$ |
| $a$   $(+/- 2\sigma)$ | 0.0019 | $(+/- 0.0006)$ | 0.0151 | $(+/- 0.0040)$ |
| $n$   $(+/- 2\sigma)$ | 1.0687 | $(+/- 0.0442)$ | 0.6723 | $(+/- 0.0557)$ |
| [MLGL MGL MSGL MHGL] | [3.77   4.02   4.18   4.3] | | [0.75   0.96   1   1.26] | |
| $[\bar{h} - d \; \bar{h} \; \tilde{h} \; \eta]$ | [2.08   4.03   0.21 90.8] | | [0.24   0.96   0.14 33.3] | |

daily basis and is observed by the Provincial Water Company of North-Holland in the dunes near the town of Castricum, whereas the daily potential evaporation series originates from a station of the Royal Dutch Meteorologic Institute near the town of De Kooy. The second series originates from a piezometer (32cl0034) located on the main meteorologic field of the Royal Dutch Meteorologic Institute at the town of De Bilt in the centre of the Netherlands [see also *Bierkens et al.,* 1999]. The terrain is flat, lies at the edge of an ice-pushed ridge that is a remnant of the glaciers that covered the north of the Netherlands during the Saalien ice age, and borders a small river named the 'Biltse Grift'. The series is observed with a daily frequency during a 16-year period (5-1-1981 until 1-1-1997). However, the series is not totally complete but 8.75% of the observations are missing. The precipitation surplus is obtained from
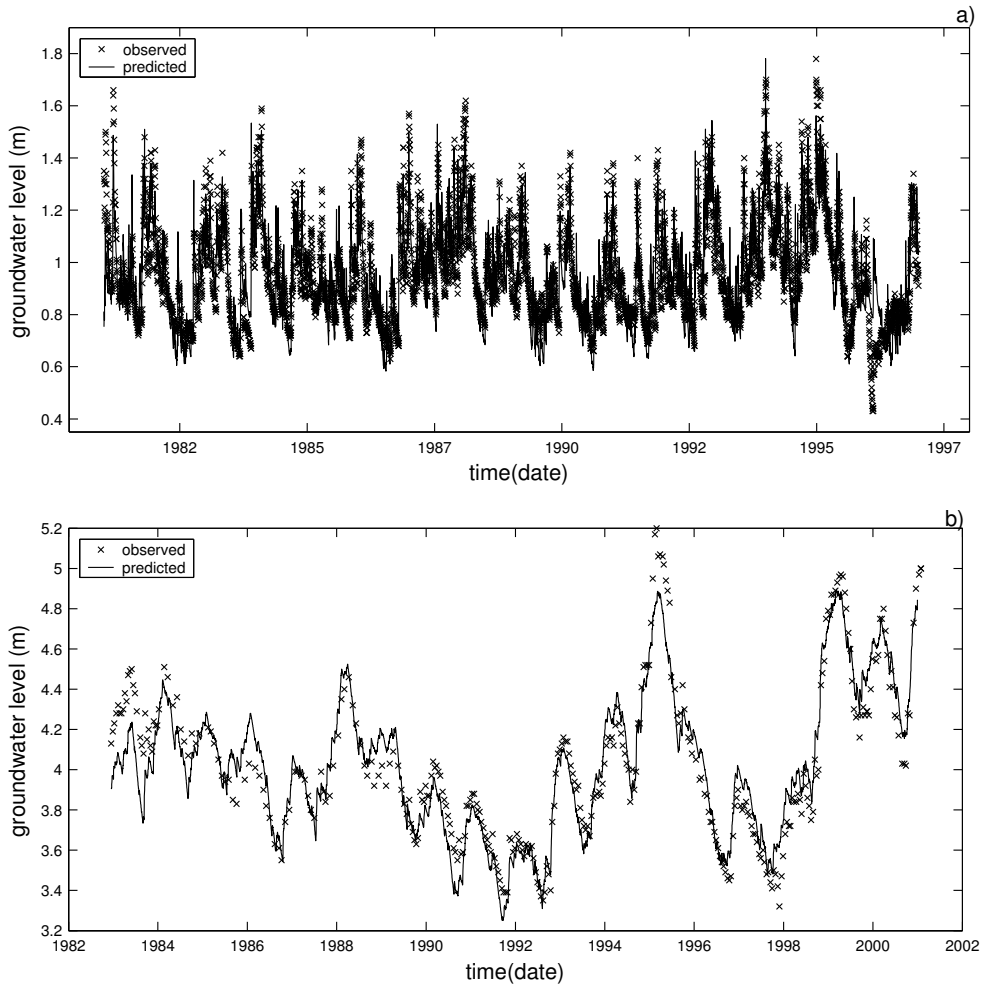


*figure 6.3:* Time plot of the available groundwater level observations for piezometers 19CNL5281 and 32cl0034, together with the predictions (transfer part, at daily frequency) from the time series model.

daily averaged observations of precipitation and potential evapotranspiration at the same meteorologic field, starting from 1-1-1971.

### 6.3.2    Comparison of MxGL statistics and GD-characteristics

In order to get an estimate of the response of these systems, a single input continuous time-TFN model (equation (6.3)) was calibrated on the data of both series, with a $R^2$ (coefficient of determination) of respectively 0.919 and 0.815. The results are summarized in table 6.1. Time plots of the available groundwater level observations for both series, together with output from the TFN models, are given in figure 6.3. Subsequently, MxGL statistics and GD characteristics were calculated using the piezometer data and the estimated IR functions. By definition, MxGL statistics are calculated from series with an interval of 14 days, so where necessary, the piezometer data was resampled to match this frequency. The MHGL and MLGL are calculated as the average of the HG3 or LG3 (i.e. the average of the three highest, respectively lowest values in a certain year) over a number of years. The MSGL, on the other hand, is the average-of-the-average of the three values neighboring the first of April (or October on the southern hemisphere) of every year. The MxGL statistics and GD characteristics have to be calculated over the correct period in order to make the results comparable. Note that the period used for calculating the meteorologic characteristics $\overline{p}$ and $\tilde{p}$ does not equal the period used for calculating the MxGL statistics. The reason for this is that the groundwater system has a memory, and the precipitation surplus $p$ outside the period of groundwater level observations also has a certain influence. Therefore, $p$ has to be weighted according to its influence on the groundwater level series, or in other words the following block response function:

$$\Theta(t) = \int_{t-(t_{he}-t_{hb})}^{t} \theta(\tau)\mathrm{d}\tau \qquad\qquad (6.14)$$

with $t_{hb}$ and $t_{he}$ denoting the start and end of the period from which groundwater level observations are available. Using (6.14), the weighted average and weighted amplitude of the precipitation surplus were calculated as:

$$\overline{p} = \frac{\displaystyle\int_{t_{pb}}^{t_{he}} p(\tau)\Theta(t_{he}-\tau)d\tau}{\displaystyle\frac{}{t_{he}-t_{pb}}\int_{0}^{} \Theta(t)d\tau}$$

$$\tilde{p}(D) = \frac{\displaystyle\sum_{Y_{pb}}^{Y_{he}} p(Y,D)\Theta(Y_{he}-Y)}{\displaystyle\frac{}{Y_{he}-Y_{pb}}\sum_{0}^{} \Theta(Y)}, \qquad 1 \le D \le 365$$

(6.15)

The application of a weighted average and amplitude implies that GD characteristics can be readily calculated for any period or climatologic scenario, or using any weighting (when for instance the recent climatologic circumstances are deemed more important than those of 30 years ago). However, the use of a 30-year period representing the present day climate (as suggested in [*Knotters and Van Walsum,* 1997] and also used in the meteorologic sciences [*Sluijter and Nellestijn,* 2002]), remains a logical choice for many applications. The MxGL statistics and GD characteristics of both piezometers are plotted in figure 6.4, together with a regime curve obtained analogous to equation (6.2). The figure clearly shows that MGL and $\bar{h}$ are almost equal, while MHGL-MLGL is greater than the annual amplitude $\tilde{h}$ for both piezometers. Remarkably, MHGL-MLGL for piezometer 19CNL5281 almost equals MHGL-MLGL for piezometer 32cl0034. Consequently, the MHGL and MLGL do not differentiate very well between the dynamics of these two, hydrologically different, systems. There is, however, a difference between their response amplitudes. From their amplitude spectra (figure 6.6), but also directly from the groundwater hydrograph (figure 6.3), it can be seen that the high frequency fluctuations of the small and fast meadow system are more pronounced than those of the bigger and slower dune system, while the opposite is true for the annual amplitude. Logically, the high frequency fluctuations influence the extremes (HG3 and LG3). Apparently, the different behavior in the high frequencies compensates the difference in the annual amplitudes, and results in comparable MHGL-MLGLs. A factor that does differentiate between both piezometers is the Mean Spring Groundwater Level (or more exactly MSGL-MGL) which is, in contrast to MHGL and MLGL, not a result of fluctuations in several frequencies, but is in fact an ordinate of the annual groundwater cycle. This factor is closely related to the factor $\eta$ (see table 6.1), which equals the phase shift between the annual groundwater and precipitation surplus cycle. Shifting the sinusoidal groundwater cycle in time will alter the value on the first of April and hence the MSGL. However, logically also the amplitude of the annual cycle is of influence to
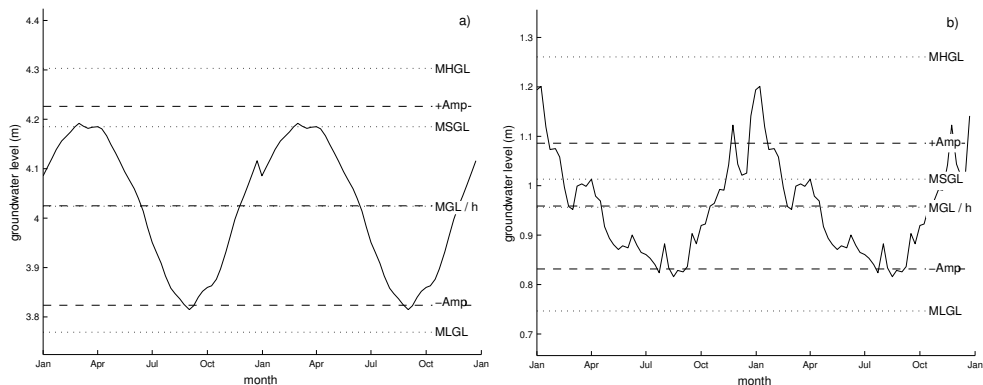


*figure 6.4:* *Comparison of MxGL statistics and GD characteristics for piezometers 19CNL5281 and 32cl0034, respectively. From the figure it can be seen that MHGL - MLGL is larger than two times the annual amplitude, and does not differentiate much between these systems. MGL and $\bar{h}$ are almost equal and their lines are hard to distinguish.*

**143**
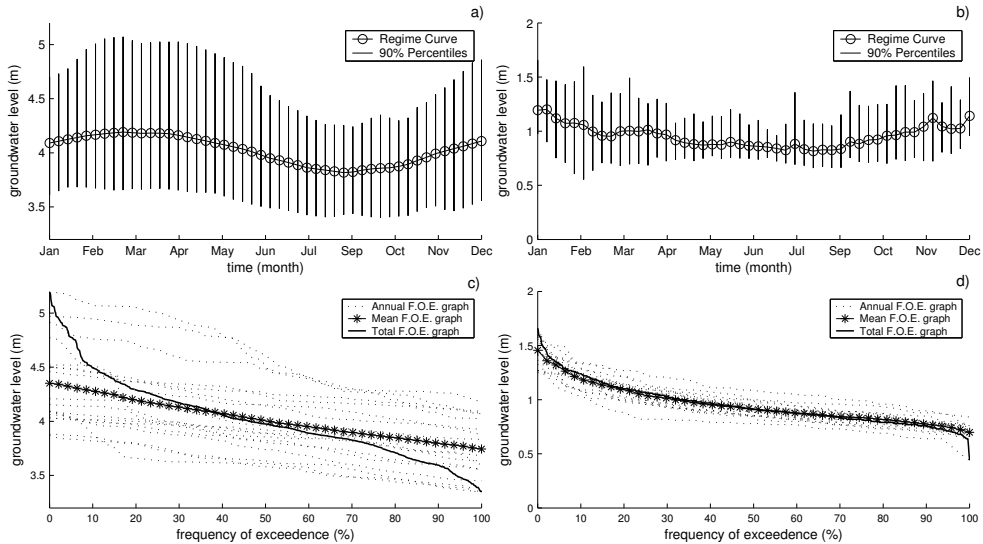
**figure 6.5:** *Plots of the regime curves* $\tilde{h}(D)$ *and the 90% confidence interval of* . $h(Y,D)$ . *of piezometers 19CNL5281 (a) and 32cl0034 (b), and respective annual, mean and total frequency of exceedence graphs (c and d).*

the MSGL.

### 6.3.3    GD-characteristics and fluctuations of non-annual frequency

A common method to illustrate and quantify the long-term variability of the groundwater level at a certain location is to plot both the mean (i.e. the regime curve $\tilde{h}(D)$) and two or more percentiles of the distribution function of $h(Y\,|\,D)$ (figure 6.5a and b). An alternative is the use of frequency of exceedence graphs (i.e. the cumulative relative distribution of $h(Y,D)$), which however do not contain information about the period of occurrence of the extremes (figure 6.5c and d). Here we have used a 90% confidence interval, and also plotted the MHGL and MLGL to be able to compare the annual, higher and lower frequency fluctuations. From the results it can be seen that both systems react rather differently to different frequencies in the input signal. For the slower system, it is clear that the low frequency fluctuations are large, and MHGL and MLGL are by far respectively lower and higher than the extremes that can occur in climatologically non-average years. For the fast system the high frequency fluctuations are more pronounced, and MHGL-MLGL exceeds the average annual cycle. The finding that different hydrologic systems react differently to fluctuations in different frequencies is in fact reported in many other publications [see also *Gehrels et al.,* 1994*; Leduc et al.,* 1997*; Molenat et al.,* 1999*; Chen et al.,* 2002]. The proportionality of fluctuations of different frequency within a signal can be examined in more detail by looking at its amplitude spectrum, obtained with a fast Fourier transform (figure 6.6a and b). From the amplitude spectrum, it can again be
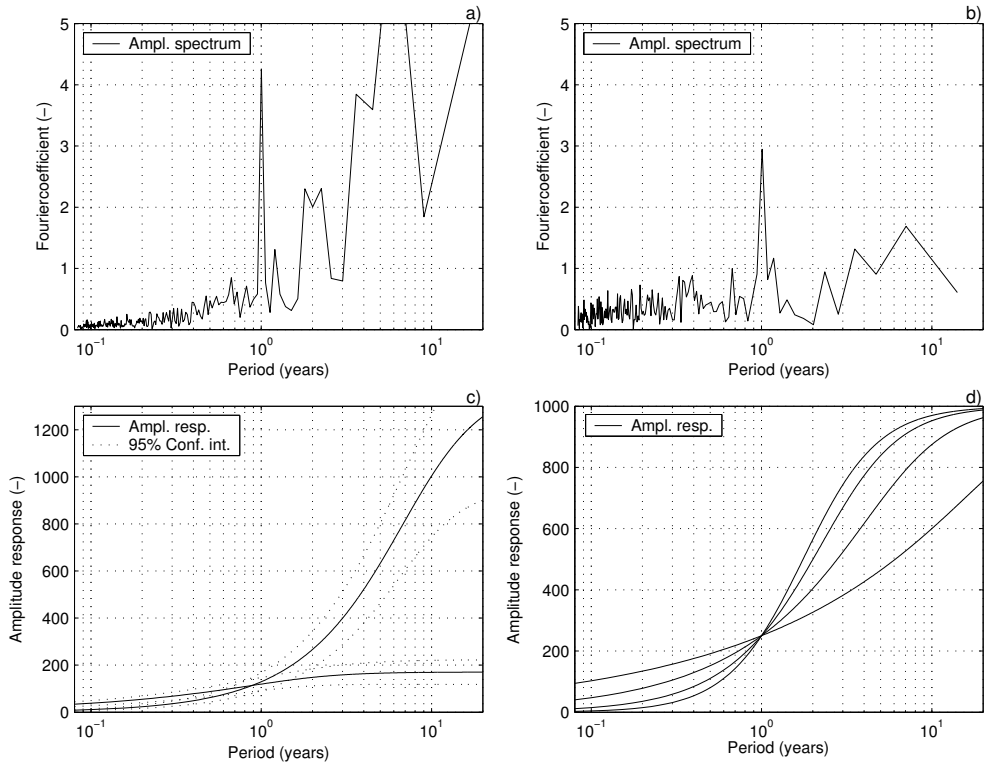
*figure 6.6: a) and b) Amplitude spectra of the groundwater level observations for piezometers 19CNL5281 and 32cl0034. c) Amplitude response of piezometers 19CNL5281 and 32cl0034 d) Amplitude response of a variety of synthetic systems with fixed annual Amplitude and $M_0$ but variable $\eta$ .*

seen that the annual amplitude of the slow system is higher than that of the faster system. Both systems, however, also show a very distinct overall pattern in their amplitude spectrum. Because in the frequency domain, the amplitude spectrum of the groundwater level series is the result of a multiplication of the amplitude response of the system and the amplitude spectrum of the precipitation surplus, it may be worthwhile to examine the amplitude responses of both systems. Using equation (6.11) and the parameters obtained with the time series model, we have plotted both amplitude responses in figure 6.6c, together with their 95% confidence interval, that was calculated with:

$$\sigma^2_{\Xi(t)} = (\frac{\partial \Xi(t)}{\partial A})^2 \sigma^2_A + (\frac{\partial \Xi(t)}{\partial a})^2 \sigma^2_a + (\frac{\partial \Xi(t)}{\partial n})^2 \sigma^2_n + 2(\frac{\partial \Xi(t)}{\partial A})(\frac{\partial \Xi(t)}{\partial a})\sigma^2_{A,a} + $$

$$+ 2(\frac{\partial \Xi(t)}{\partial A})(\frac{\partial \Xi(t)}{\partial n})\sigma^2_{A,n} + 2(\frac{\partial \Xi(t)}{\partial a})(\frac{\partial \Xi(t)}{\partial n})\sigma^2_{a,n}$$

(6.16)

with the variances $\sigma_a^2$ and covariances $\sigma_{A,a}^2$ being obtained from the Levenberg-Marquardt optimization routine. The amplitude responses prove to coincide well with the amplitude spectra of both systems. This is in fact not that surprising, as we also found that the amplitude spectrum of the precipitation surplus was rather flat, apart from the peak at annual frequency (figure 6.1a). From its amplitude response, the slow dune system appears to act as a low pass filter on the precipitation surplus fluctuations, which is corroborated by the fact that the high-frequency fluctuations of the water table are small. A similar result was also found in [*Chen et al.,* 2002] in a research aimed at predicting average annual groundwater levels. According to [*Gehrels et al.,* 1994], this phenomenon can be explained by the buffering effect of the unsaturated zone, but it can also occur with other types of inputs, e.g., infiltration of surface water, and because of a delay in the response of the saturated zone in the centric parts of a system [*De Zeeuw and Hellinga,* 1958]. The amplitude response of the fast piezometer again proves to be slightly lower than that of the slow system at annual frequency, but both are equal for a frequency close to a year. The fact that the amplitude responses of different systems almost cross at annual frequency indicates that the groundwater regime can still vary considerably when $s - \bar{h}$ and $\tilde{h}$ are equal. Therefore, we have to look at the non-annual frequencies in order to further differentiate between different systems. Except for the annual cycle, however, there are no other distinct signals at isolated frequencies, so it is difficult to get one well-defined characteristic for fluctuations at lower or higher frequencies. We can, however, given the fact that the amplitude spectrum of the precipitation input is rather flat, use our knowledge of how systems reacts to and amplify signals at different frequencies. From (6.12) it is clear that the amplitude response, when the frequency $\xi$ decreases, reaches an asymptotic level that equals $\lambda A$. The zeroth moment $M_0$, in other words, determines the magnitude of the groundwater level fluctuations towards the lower frequencies, so the low frequency fluctuations are covered by specifying $\bar{h} - d$. Vice versa, this also means that $M_0$ is hard to estimate from a time series of groundwater level observations of limited length, or in a period in which inter-annual fluctuations are small, and it could be worthwhile to use prior information on d, e.g., obtained from the level of surface waters in the vicinity. As an alternative for using $\bar{h} - d$, the distribution function of $h(Y,D)$ could also well be used, e.g., in constructing a ratio between the annual amplitude and the 95% percentile of $h(Y,D)$. Such a parameter or ratio has the advantage that its dimension is [L], and it therefore gives a more direct indication of the magnitude of long-term fluctuations. However, an analytic solution for this can not easily be found so it has to be solved numerically, which can be rather time consuming when for instance the GD characteristics have to be generated spatially.

The fourth and last GD characteristic is the phase shift between the annual groundwater and meteorologic cycle, which is related to the groundwater level dynamics in the sense that slow systems have a large phase shift, and are at the same time not very sensitive to high frequency fluctuations. In order to illustrate the relationship between phase shift and the amplitude spectrum of the groundwater

level fluctuations more clearly, we have plotted the amplitude response of several synthetic systems with fixed $s - \bar{h}$, $\bar{h} - d$ and $\tilde{h}$, but varying $\eta$ (figure 6.6d). The figure clearly shows that $\eta$ determines the proportionality between the magnitude of the higher and lower frequencies, while $M_0$ determines the asymptote of the amplitude response towards the lower frequencies. The phase shift was also identified as an important characteristic of groundwater systems in [*Larocque et al.,* 1998*; Manga,* 1999*; Lee and Lee,* 2000*; Chen et al.,* 2002]. In the latter publication, however, one phase shift is applied for the three basic frequencies that were identified in the precipitation surplus input, while the phase response is in fact a function of the frequency (see e.g., equation (6.11) or [*Ziemer et al.,* 1998]).

## 6.4    Discussion and conclusions

The methods in this paper were developed and illustrated using a single input time series model. The range of geohydrologic variety that the time series model covers is given in section 2.4.3, and validation results are given in [*Von Asmuth et al.,* 2002]. The methods of estimating GD-characteristics can be readily extended to include other influences such as groundwater abstraction or fluctuating surface water levels, as for linear systems, the stationary draw-down and annual amplitude attributable to a secondary factor can simply be added to that of precipitation surplus. Only in the case where other factors have a very different dynamic behavior, this has to be taken into account in the form of additional parameters to be able to better characterize the dynamics. A special case of another factor that influences the groundwater level is the noise that enters the system through the noise model. However, as the expected value of white noise is zero and its frequency spectrum is flat, both its stationary influence and its power at annual or any other specific frequency can be neglected. A limitation that arises from the use of a linear time series model is that it does not hold for non-linear systems. Future research will have to tackle this problem with the aid of non-linear time series models.

In the usefulness of the method presented, two aspects are of importance. The first is the direct link between the dynamics of the input, the properties of the groundwater system, and the dynamics of the groundwater table that is developed in this study. Using this link, one can correct for non-average meteorologic conditions in the observation period, in the same sense as MxGL statistics are made 'climate representative' [*Knotters and Van Walsum,* 1997] by using simulations over an historical 30 year period. Here, however, the key characteristics of the groundwater dynamics can be also readily obtained for any climatologic scenario, without having to simulate future meteorologic time series. This adds flexibility to the method necessary for dealing with non-stationary influences. Next to climate change, also factors such as groundwater abstraction can cause the groundwater dynamics to be non-stationary, so that the historical 30-year period does not necessarily reflect the current or future state. The system properties, in terms of the moments of the IR function, can be estimated using a time series model, but also with the aid of deterministic groundwater flow model such as MODFLOW [*Von Asmuth and Maas,* 2001]. The link described can thus be used vice versa also, to make estimates of the input or recharge

to an aquifer using groundwater level observations and the systems' response, as was proposed earlier in [*Chen et al.,* 2002].

A second important aspect is the way in which the groundwater dynamics at a certain location are characterized. In fact, the application of a time series model is not necessary to derive GD characteristics if long term observations of the groundwater level are available, as GD characteristics can also be inferred from the local drainage base and frequency analysis on the data itself (please note, however, that such information is also not error free). The four GD characteristics are the mean depth, the convexity, the annual amplitude and the phase shift between the annual groundwater and meteorologic cycle. By looking at their relationship with the impulse and amplitude response function, it is shown that GD characteristics each contain information about a separate aspect of the groundwater dynamics, and together can describe the dynamics in detail (at least, for as much as the time series model can). In the example application, it is shown that the methods for calculating MxGL statistics form a rather heuristic combination of averaging and selecting extremes, and that they do not differentiate optimally between different groundwater systems. As the extremes or highest (HG3) and lowest levels (LG3) in a year are the result of both the annual cycle and the fluctuations in higher frequencies, both frequencies interfere in the calculation of MxGL statistics. This is the reason that MHGL-MLGL exceeds the amplitude of the annual cycle, either calculated with the time series model or using a regime curve. Furthermore, the MxGL statistics are the results of averaging over a longer period (minimally 8 years), which effectively filters out lower frequency fluctuations and the corresponding inter-annual extremes. The extreme levels and their duration, however, often are of great interest because of their disproportionate economical damage or ecological influence. Thus, it is concluded that the capability of MxGL statistics in differentiating between groundwater regimes at different locations is less than that of GD characteristics.

Because GD characteristics provide a compact way of describing the groundwater dynamics at a certain location, they can be of use to get a better understanding on the relationship between groundwater dynamics and other factors, such as soil conditions or ecology. Next to that, they can be used to map the groundwater dynamics for planning purposes or policy making. In ecological sciences, for example, there is still little deterministic knowledge of how plant species are related and/or adapted to the groundwater dynamics at a certain location, and often use is made of expert knowledge [e.g., *Runhaar et al.,* 1987*; Ellenberg,* 1991]. Identifying the main characteristics of the groundwater dynamics can help to get a clearer picture of which part of the dynamics (e.g., the extremes, the average conditions or the low levels in summer) are important for the survival of endangered species.

**Menyanthes**
Data management
Visualisation
Statistical analysis
Time series analysis

*Menyanthes* was developed by:
Jos von Asmuth
Kees Maas
Martin Knotters

**KWR**
*Watercycle Research Institute*

**ALTERRA**
**WAGENINGEN**UR

**ŤU**Delft

*Menyanthes trifoliata*

**Abstract:**

The program Menyanthes combines a variety of functions for managing, editing, visualizing, analyzing and modeling hydrogeologic time series. Menyanthes was initially developed within the scope of the PhD research of the first author, whose primary aim was the integration of data and physically-based methods for modeling time series of groundwater heads. As such, time series analysis forms the heart of Menyanthes. Within Menyanthes, time series can be modeled using both the ARMA and PIRFICT methods. The PIRFICT method is a new method of time series analysis that has practical advantages and facilitates physical interpretation and implementation of knowledge on physical behavior. Analytic solutions to specific hydrogeologic problems may be used as response function, along with their physically-based parameters. A more general approach is possible using Skew-Gaussian distribution functions, which prove to fit the behavior of hydrogeologic (and other) systems well. Use of such functions within the PIRFICT method substantially simplifies the model identification procedure, as compared to the traditional Box-Jenkins procedure. PIRFICT models may be fitted to a large number of time series in batch. Spatial patterns that emerge in the results provide useful, additional, and independent information, which adds another dimension to time series analysis. Their interpretation is supported by the spatial visualization and analysis tools of Menyanthes. The PIRFICT method also facilitates the integration of time series and spatially-distributed models via, e.g., moment-generating differential equations. The PIRFICT method may prove to be of use for other types of time series as well, both within and outside the realm of environmental sciences.

# Chapter 7

# *Menyanthes*: software for groundwater head data

**Adopted from:**

Von Asmuth, J.R., C. Maas, M. Knotters, M.F.P. Bierkens, M. Bakker, T.N. Olsthoorn, D.G. Cirkel, I. Leunk, F. Schaars, and D.C. Von Asmuth (submitted)
Menyanthes: software for hydrogeologic time series analysis, interfacing data with physical insight.
*Environmental Modelling & Software*.

**Software availability**

| | |
|---|---|
| Developers: | KWR Watercycle Research Institute |
| | Delft University of Technology |
| | Alterra Wageningen UR |
| | Artesia |
| First available year: | 2003 |
| Software requirements: | Windows® XP, 2003, Vista, 7 |
| Program language: | Matlab, C |
| Availability and cost: | see www.menyanthes.nl |
| Primary contact: | Jos von Asmuth |

## 7.1 Introduction

### 7.1.1 Time series models: their strong points and limitations

Dynamic systems are ubiquitous in our environment, both natural and man-made, and form the subject of many different scientific disciplines. Consequently, the theory and techniques for analyzing and modeling collected time series data have a vast application area. One may discern different disciplines such as time series analysis, system identification, control engineering, signal processing and filtering, depending on the type of system addressed, their physical functioning, the available data and the problem to be solved. The theory and techniques used in
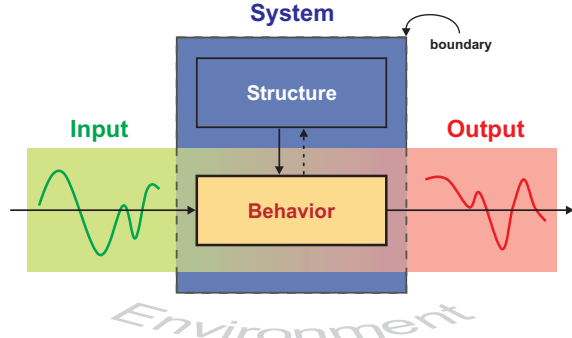


*figure 7.1:* *Schematic representation of an open physical system. The spatial structure of environmental systems is generally far more complex than their temporal behavior, and largely unknown. The structure does not have to be explicitly defined for modeling the behavior.*

these disciplines, however, are often remarkably similar. The ARIMA (AutoRegressive Integrated Moving Average) time series models are most widely used, although there has been a diversification of methods. The application of ARIMA models developed rapidly after the publication of the text book by Box and Jenkins [1970]. Although it took some time before the methods and principles were picked up in the hydrologic sciences, this was amply compensated by the publication of the voluminous book by Hipel and McLeod [1994]. In their book, Hipel and McLeod present many different applications, as well as several extensions to the ARIMA models as published by Box and Jenkins. Relatively simple data-based models like the ARIMA time series model are still widely applied in the environmental sciences, despite the rise of physically-based, spatially-distributed models. Advantages of time series models over physically-based, distributed models are their relative accuracy, ease of construction, and well-developed statistical background. The first two points follow directly from the fact that a time series model treats a physical system firstly as a 'whole'. The effective, overall behavior of a system at a certain location is modeled without requiring the explicit definition of its entire spatial structure (figure 7.1). In contrast, physically-based, distributed models start bottom-up from the so-called representative elementary volume (REV), and almost inherently invite the inclusion of all detail above that scale. As a consequence, such models generally require a great amount of data, while the larger part of the model still is the result of correlations, extrapolations and assumptions, which cannot be verified or validated [see e.g., *Konikow and Bredehoeft*, 1992*; Oreskes et al.,* 1994]. Such problems, which are more or less intrinsic to the REV scale approach, have lead some to proclaim that it will prove to be a dead end for complex, environmental systems, while others designed alternative approaches on

larger aggregation scales [e.g., *Reggiani et al.,* 1998*; Reggiani et al.,* 1999*; Beven,* 2002].

ARIMA models, however, also have several limitations. First of all, there is the problem of identifying the model order. A time series modeler has to identify the order of an ARIMA model, which contains at least 5 terms in case of a model with a deterministic component or Transfer Function-Noise (TFN) model. These terms define the number of parameters in different parts of the model, and the delay time [*Box and Jenkins,* 1970]. In addition, a modeler has to define a suitable degree of differencing to induce stationarity in the case of non-stationary series, and also has to choose the model frequency, which however interferes with the model order [*Von Asmuth et al.,* 2002]. Although Box and Jenkins devised a specific iterative procedure for identifying the model order, its results can be ambiguous [*Hipel and McLeod,* 1994]. The procedure itself is to some extent heuristic, subjective and knowledge and labor intensive [*De Gooijer et al.,* 1985]. The use of ARIMA models thereby requires an adequate understanding of the underlying theory. The available time series analysis literature is difficult for outsiders, as it has a strong statistical focus and uses a specific mathematical notation, jargon and viewpoint. The second problem is the handling of irregular or high-frequency data. In hydrological practice, where irregular data are common, the fact that ARIMA models can not deal with that poses a serious problem. High frequency data can also cause problems, as it causes the autocorrelation in the data to become high, autoregressive(AR)-parameters to become badly scaled, and the effect of moving average(MA)-parameters to be limited [*Von Asmuth et al.,* 2002]. Third, there is the problem of extrapolating results. Because of the data-based nature of time series models, there is a risk in extrapolating results to situations for which no data are available, e.g., for scenario calculations. Extrapolation to other points in space or situations where the system itself is changed is impossible without additional knowledge or techniques. Fourth, and in spite of their simple nature, ARIMA models are easily over-parameterized. This means that it will generally be difficult to separate the different effects and get correct, reliable and unique parameter estimates, when a system is affected by many inputs, when input(s) do not change much over time and/or are cross-correlated. The problem is compounded by the fact that the MA-part of ARIMA models is parameter-inefficient, especially when dealing with high-frequency data.

### 7.1.2  Recent developments and use of physical insight

A possible solution to the above mentioned limitations is to use physical knowledge on the functioning of a system and/or to combine the strong points of data- and physically-based methods. Relatively recent, attention is paid to the fact that the transfer function of time series models, which describes the response of systems through time, should have a physical basis, at least when applied to physical phenomena like groundwater level fluctuations (although this was already proposed by Van Geer [1987]). In the process of incorporating physical insight in the use of time series models, the following approaches can be distinguished. First, physical knowledge may be used for model identification, to *a posteriori* select the correct model order from a subset of predefined candidate models. The selection is based on

the question whether or not the results and computed transfer functions are plausible from a physical-hydrologic perspective [*Young and Beven,* 1994*; Price et al.,* 2000]. A simple example is precipitation, which should result in a rise and not a lowering of the water level. A second approach is the reversal of this process, in which case the model order is selected *a priori* based on a physical-hydrologic analysis of the system. This approach follows Knotters and Bierkens [2000], who selected the ARX(1) model (which is equivalent to an exponential response function) for modeling the relationship between precipitation surplus and water table depth, based on an analysis of the functioning of a simplified soil column. A third approach is to replace the statistically oriented, discrete-time ARIMA model structure by a model structure in which continuous-time response functions are used [*Von Asmuth et al.,* 2002*; Von Asmuth and Bierkens,* 2005]. This approach to time series analysis is named the Predefined Impulse Response Function In Continuous Time (PIRFICT) method. Analytic solutions to geohydrologic problems may be used directly as response functions in this method, i.e. functions that are more complex than the ARX(1) model proposed by Knotters and Bierkens. As a result, time series models are forced to show a certain, physically-plausible behavior in response to different excitations. Simple distribution functions like the scaled gamma distribution prove to fit the behavior of many hydrogeologic systems, regardless of their specific properties and structure. Other developments include a solution to the problem of modeling irregular or high-frequency data [see *Bierkens et al.*, 1999; *Berendrecht et al.*, 2003; *Yi and Lee*, 2004; *Von Asmuth and Bierkens*, 2005], non-linear time series modeling [*Knotters and De Gooijer*, 1999; *Berendrecht et al.*, 2004], and the combination of time series and spatially distributed models (see sections 2.2. and 2.3).

### 7.1.3    Scope and outline of paper

The methods related to the PIRFICT approach were presented and tested in several papers. It was shown that the PIRFICT method allows for the incorporation of physical behavior, reduces the number of parameters, simplifies model identification and avoids problems related to the data frequency in time series models [*Von Asmuth et al.,* 2002*; Von Asmuth and Bierkens,* 2005]. The PIRFICT method was extended and tested to include complex situations with multiple stresses of different types [*Von Asmuth et al.,* 2008]. The use of time series model results and/or impulse responses in distributed models was addressed in [*Von Asmuth and Maas,* 2001*; Bakker et al.,* 2007*; Bakker et al.,* 2008]. In the present paper, a synthesis is given of the methods and results, while specifically focusing on the interface between data and physically-based methods. In addition, the program *Menyanthes* is presented, in which the time series analysis methods that were developed are incorporated (and future developments will be). Finally, an example application is given where data from multiple locations are analyzed, illustrating the use of spatial patterns and physical insight to interpret results. The PIRFICT method is applied to groundwater head series, as in all the above mentioned papers, but it is not restricted to that. The case of groundwater heads may therefore serve as an example for linear or linearized systems in general, in or outside the environmental sciences.

## 7.2    Methods and theory

### 7.2.1    Differential equations, impulse responses and convolution

While many papers start from the time series analysis viewpoint, we choose to start from a physical viewpoint here to better serve the non-specialist community. The construction of a physically-based model for the dynamics of an arbitrary variable starts with the derivation of an appropriate mathematical expression, commonly a differential equation, for it. Differential equations are derived from two types of equations [e.g., *Bear,* 1972]:

- The continuity equation(s)
- The constitutive equation(s)

For groundwater, the continuity equation is found by applying the principle of mass conservation (while in other cases, also other conservation laws may apply, like those for energy or momentum). In simple words, this principle states that no mass can disappear without reason and it is valid for all types of masses. In case of water, it is known as the water balance equation. The constitutive equation specifies a property of the specific medium under consideration, which in case of groundwater is Darcy's law. This equation treats a representative elementary volume (REV) of a porous medium as a continuum, thereby disregarding the spatial variability below this scale (consisting of grains and pores, or molecules on an even smaller scale).

A standard technique for solving differential equations of linear dynamic systems exactly is to determine their solution for a Dirac delta function $\delta$ [*Dirac,* 1947] as input $p$. The delta function has the following properties:

$$\begin{cases} \delta(t) \ = 0, \qquad t \neq 0 \\ \int\limits_{-\infty}^{\infty} \delta(t)\mathrm{dt} = 1 \end{cases} \qquad (7.1)$$

and can be thought of as the limit of (for instance) a Gaussian distribution function with mean zero and a variance that also approaches zero. In words, the delta function is an instantaneous pulse or impulse of unit area. The impulse response function $\theta$ is defined as the effect of $\delta$ as input $p$ on the state $h$ of a system, i.e. the deviation in time from an otherwise steady state $d$. In terms of groundwater head and precipitation, $\theta$ can be thought of as the response to a very short shower of unit height, when the head is otherwise constant or equals the local drainage base. In mathematical terms, $\theta$ is defined by the following conditions:

$$
\begin{cases}
\theta(t) = h(t) - d \\
h(t) = d, & t < 0 \\
p(t) = \delta(t)
\end{cases}
\tag{7.2}
$$

If $\theta$ and $d$ are known, $h$ can be obtained for an input $p$ that varies arbitrarily in time through convolution (Duhamel's principle [*Duhamel*, 1833]):

$$
h(t) - d = \int_{-\infty}^{t} \theta(t - \tau) p(\tau) \mathrm{d}\tau \equiv \int_{0}^{\infty} \theta(t) p(t - \tau) \mathrm{d}\tau \equiv (\theta * p)(t)
\tag{7.3}
$$

where the last form is just a compact way of symbolizing a convolution product. Equation (2.35) implies that the dynamics of an arbitrary, linear, dynamic system at a certain location are completely governed by the dynamics of the input and the impulse response function. The impulse response function $\theta$ completely characterizes the dynamically relevant, physical properties of a system, and is as such an integral property thereof. $\theta$, however, is a function of the location in the system, and of the specific excitation under consideration.

### 7.2.2 Methods of constructing response functions

The solution to (2.35) only fits the dynamics of a system exactly if $\theta$ (and $p$) are also exactly known (and the system is linear and time-invariant). Of course, $\theta$ is never known exactly in practice, but there are several methods to obtain an approximation for it. The first, bottom-up and nowadays mainstream approach, is to construct a spatially-distributed model to solve the governing differential equation. Distributed models may be transformed into sets of impulse responses, as outlined in [*Sahuquillo*, 1983]. Although this procedure is not common, its advantages over the direct use of the distributed model are its computational efficiency and the fact that the differential equation can be solved continuously in time by doing so. A second approach starts top-down and approximates $\theta$ as an ARMA transfer function. In that case, (2.35) is approximated by discrete convolution:

$$
h_t - d = \sum_{i=-\infty}^{t} \Theta_{t-i} p_i \equiv \sum_{i=0}^{\infty} \Theta_i p_{t-i} \equiv \Theta(B) p_t \equiv (\Theta * p)_t
\tag{7.4}
$$

where $t$ is a discrete time index ($t \in \mathbb{N}$), $\Theta$ is the transfer or block response function and $B$ is the backward-shift operator defined by:

$$
B^b p_t = p_{t-b}
\tag{7.5}
$$

$\Theta$ can be written symbolically as [see section 2.2; *Box and Jenkins*, 1970]:

$$\Theta(B) = \delta^{-1}(B)\omega(B) \tag{7.6}$$

where $\delta(B)$ is the autoregressive operator (not to be confused with the Dirac delta function), and $\omega(B)$ the moving average operator. The actual computation of $h_t$ is done recursively as:

$$h_t - d - \delta_1(h_{t-1} - d) - \ldots - \delta_{nr}(h_{t-nr} - d) = \omega_0 p_t + \omega_1 p_{t-1} + \ldots + \omega_{ns} p_{t-ns} \tag{7.7}$$

where $nr$ and $ns$ define the number of parameters or order of the transfer function. Although ARMA models are used in a variety of sciences and are usually perceived as 'black box' models, ARX(1) models may be seen as representing the physics of a simple soil column [*Knotters and Bierkens,* 2000] and an ARX($nr$) model as representing a (linear) distributed groundwater model with $nr$ cells (see section 2.3.1). The recursive computation of (7.7) is computationally less demanding than the direct evaluation of (7.3) or (7.4) and as ARMA models have a long history, this was probably an important advantage. With modern technology, however, this is not an issue anymore, and consequently as a third approach, $\Theta$ and/or $\theta$ can also be approximated in other ways.

### 7.2.3    The PIRFICT method and use of distribution functions

Von Asmuth et al. [2002] presented a new method of time series analysis, called the PIRFICT method.  A difference with ARIMA-type and kindred time series models, whether discrete or continuous time [e.g., *Brockwell*, 2001; *Young and Garnier*, 2006], is that it is formulated as a (continuous time) convolution integral (7.3), instead of as a temporal difference or differential equation. In this formulation, $\theta$ may in principle be an arbitrary continuous mathematical function having a known integral. The PIRFICT method has several practical advantages [*Von Asmuth et al.,* 2002], and facilitates the integration of data and physically-based methods in two ways. First, the physical behavior of a system may be implemented directly in the model as response function $\theta$, for example as analytic solution to the differential equation in a specific situation. Elementary physical principles may be used to reduce the degrees of freedom in the model, e.g., by limiting the sign of the effect to either positive or negative or by sharing
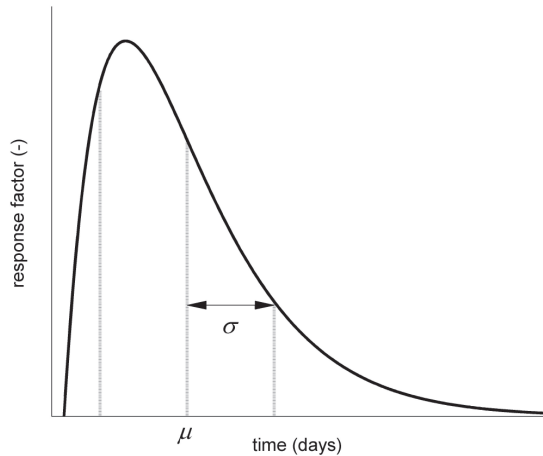
*figure* 7.2: *Skewed impulse response function with mean* $\mu$ *and variance* $\sigma^2$.

parameters of response functions across several inputs [*Von Asmuth et al.,* 2008]. Continuous response functions also tend to contain fewer parameters than their ARMA counterparts, because the number of parameters is not linked to the model frequency, as in ARMA models [*Von Asmuth et al.,* 2002]. Second, the results of time series models and physically-based, distributed models can be matched using moments of the impulse response functions (which in principle is also possible for ARMA-models, by correcting the moments of the transfer functions obtained for the model frequency). The use of moments to characterize response functions, as is common practice with statistical distributions, has been proposed independently by Nash [1959], Jury and Roth [1990] and Maas [1994]. A more recent notion is that moment-generating differential equations may be derived directly from partial differential equations [*Maas,* 1995; *Harvey and Gorelick,* 1995; *Govindaraju and Das,* 2007], which provides a direct link between time series and distributed models [*Von Asmuth and Maas,* 2001; *Bakker et al.,* 2008].

Although system-specific, physically-based analytic solutions may be used for $\theta$, a more general approach is also possible. A first step in this direction was taken by several authors, who independently and for different problems, observed that impulse response functions typically take the shape of a skewed distribution function (figure 7.2). A pioneer in this respect was Nash [1958], who proposed the use of the gamma distribution for modeling the transformation of rainfall into run-off. The gamma distribution is given by:

$$\theta(t) = \frac{a^n t^{n-1} \exp(-at)}{\Gamma(n)}$$
(7.8)

where $a$ and $n$ are parameters. The physical analog of the gamma distribution is a cascade of linear reservoirs, also known as the Nash cascade. The use of distribution functions for modeling the transfer of solutes through soil was suggested by Jury [1982] and Maas [1994] (discussed are the lognormal, gamma and Pearson type III distributions). Jury and Roth [1990] showed that their shapes have a striking similarity with that of the impulse response solution to the physically-based convection-dispersion equation. When using distribution functions as response function, the temporal behavior of a system is defined, but the physical or spatial structure is not. In that sense, they are not 'physically-based'. Jury and Roth, however, consider this a relative issue and point out that it is a curious phenomenon that physically-based laws like the convection-dispersion equation are commonly accepted merely because their solution can be manipulated into matching the shape of the outflow concentrations. Similar statements are made by Beven in the context of his equifinality thesis [e.g., *Beven,* 1993; *Beven,* 2006]. For groundwater heads, the use of distribution functions is perhaps less intuitive. Von Asmuth et al. [2002] and numerous studies afterwards showed that a scaled gamma distribution (the gamma distribution multiplied by a constant) may be used effectively to model the head response to recharge, and that it yields results that are at least as good as those of ARMA models (in terms of calibration and validation error). In fact, skewed distribution functions prove to be applicable as response functions also outside the earth sciences, e.g., in psychology

[*McGill,* 1965] or pharmaceutics [*Sun,* 1998]. Maas [1994] gives a partial explanation for the wide area of application. He shows that transferring a signal through $n$ linear, independent systems is mathematically equivalent to adding up $n$ random, independent variables. Consequently, the central limit theorem of statistics also applies to the transfer of signals through linear systems and like probability distribution functions, impulse response functions will approach Gaussian distributions when $n$ approaches infinity. Maas also shows that before reaching this limit, impulse responses tend to have a skew-Gaussian shape. Bakker et al. [2008] point out that the scaled gamma distribution as well as Hantush' well function [*Hantush,* 1956*; Veling and Maas,* 2010] and Bruggeman's polder function [*Bruggeman,* 1999] are all special cases of the following function:

$$\theta(t) = At^{\upsilon} \exp(-at - \frac{b}{t}) \qquad (7.9)$$

where $A, \upsilon, b$ are parameters. Veling [2010] named (7.9) the generalized moving Gaussian distribution and gave its mathematical properties. In [*Von Asmuth,* in print], it is pointed out that the central limit theorem and skew-Gaussian convolutional limit of Maas [1994] are not directly applicable to the case of groundwater heads. From (7.9) , however, it follows that a variety of systems with convection-dispersion or diffusion-type processes do tend to the same, skew-Gaussian and relatively simple behavior.

## 7.3    Software architecture

### 7.3.1    Why Menyanthes?

The methods discussed in section 2.3 were developed within the scope of a PhD research project that initially encompassed three fields of modeling: time series, groundwater and eco-hydrologic modeling [*Von Asmuth and Maas,* 2001]. The three 'leaves' in this research led to naming the developed computer program after *Menyanthes trifoliata*, the scientific name of the Marsh trefoil or Bogbean. As the natural habitat of *Menyanthes trifoliata* is on the verge of ground and surface water, as it is threatened by anthropogenic influences and also is one of the wonderful flowers in such areas, it serves as the natural ambassador for this research.

### 7.3.2    Information systems hierarchy and data management

For several reasons, Matlab was chosen as programming environment for *Menyanthes* (see www.mathworks.com). Matlab is a versatile tool for mathematical computing and analysis, which incorporates a high-level, matrix-oriented programming language that allows scientists and engineers to easily express their own algorithms. In addition, Matlab applications may be compiled into stand-alone applications, which greatly reduces the effort for both developing the methods and implementing them in user-friendly programs. Furthermore, Matlab is strong in data management and (2D and 3D) visualization, which, as discussed below, are important facets of *Menyanthes*. The design of *Menyanthes* reflects the goal to integrate empirical and physically-based

methods. We use the data–information–knowledge–wisdom hierarchy (DIKW) of Ackoff [1989] as a metaphor, and distinguish a corresponding hierarchy in information systems (figure 7.3). This concept was adopted from [*Rowley,* 2007], and adapted to reflect that knowledge on how a system functions generally crystallizes into models in the earth and technical sciences. In that view expert and decision support systems are higher up in the hierarchy, and contain a reflection of the wisdom that is necessary to make the right decisions. The functionality of *Menyanthes* starts bottom-up
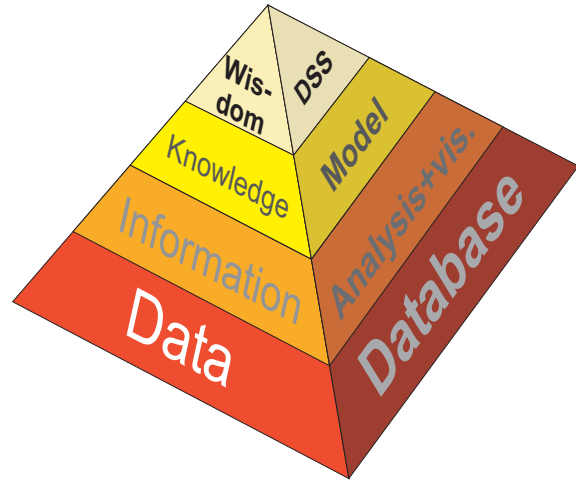


*figure 7.3: Knowledge and information systems hierarchy, after [Rowley, 2007].*

with a database, in order to keep a close link with the original observations. The PIRFICT method can operate on time series data as-is, so there is no need for the interpolation, gap filling and /or time discretization steps that accompany ARMA models. *Menyanthes* incorporates various tools and functions for processing, visualizing and checking time series data. The basic philosophy with regards to data management is that the original data are never deleted or modified. Changes and corrections are stored 'on top' and can always be reevaluated and undone. A relatively new development is that groundwater monitoring is increasingly being performed with the aid of pressure sensors. Pressure sensors, however, show specific problems and types of errors [*Von Asmuth,* 2010*; Sorensen and Butcher,* 2011] and *Menyanthes* contains tools for specifically dealing with these problems.

### 7.3.3    Data analysis and visualization

Analysis and visualization tools are the primary means of transforming crude data into useful information, at the second level of the DIKW hierarchy. In case of environmental systems, data sets are essentially 4D ( $x, y, z, t$ ) and the amount of information contained in them is simply too large for direct use in, e.g., visual interpretation, mapping, or empirical modeling purposes. For such purposes, the data has to be compressed without losing too much significant information. Specific methods have been developed for groundwater heads to visualize the dynamics in overall graphs such as frequency of exceedence graph**s** or regime curves (figure *7.4a*) , and for statistically characterizing the dynamics with a limited set of parameters (for which time series models can be of use, [see *Knotters and Van Walsum,* 1997*; Von Asmuth and Knotters,* 2004]). In *Menyanthes*, parts or characteristics of (groups of) groundwater level time series may be presented in various ways and in 2D or 3D. Visualization extracts information but also aids validation of the data. For example, outliers are typically easily detected in a simple plot of a set of groundwater level
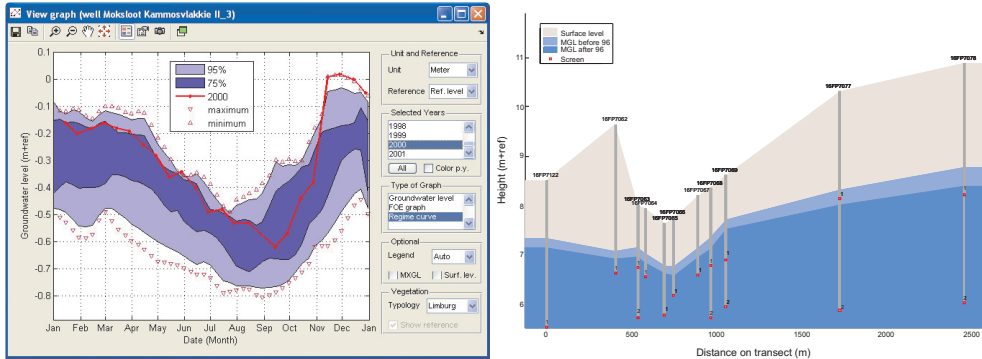
*figure 7.4:* Examples of visualization of series of groundwater level observations and their characteristics (a) regime curves (b) in a 2D cross-section.

series. On the other hand, visualization also aids 'visual system identification', as spatial patterns in the groundwater heads reveal much of the basic structure and functioning of a system (e.g., figure *7.4*b). As environmental systems are usually heterogeneous and have layers with variable resistance, monitored locations at different depths can be more or less isolated from each other and display different behavior. The depth of well screens and monitored locations is therefore important to understand the observed behavior, and 2D or preferably 3D visualization is needed to interpret data and modeling results correctly.

### 7.3.4 Main screen and modeling tools

The time series modeling tools and methods discussed in the previous section are at the third level of the DIKW hierarchy, and form the heart of Menyanthes. The set-up of the main screen (figure 7.5) follows the schematization of an open system as depicted in figure 7.1, which adds to the intuitiveness of the user interface. The upper right quadrant is devoted to groundwater level series (the output), the lower right quadrant to explanatory series (the input) and the lower left quadrant for model properties and results (the system). The fourth and final quadrant in the upper left-hand corner contains a map, in which spatial selections can be performed and data are presented in top or cross-section view (figure *7.4*b). A separate screen is used for 3D visualization. In principle, the same visualization tools are used for the data themselves and for checking and interpreting time series model results. Each time series is modeled independently of the other series. Spatial patterns that emerge from the results of multiple time series therefore yield valuable and independent information on the spatial structure and functioning of a system, and for checking the plausibility of results [*Von Asmuth et al.,* 2008]. Menyanthes offers the possibility to use both PIRFICT and ARMA(1, *ns* ) type response functions for time series analysis. This allows both methods to be compared directly, as was published in [*Von Asmuth et al.,* 2002]. It can also serve to asses the validity of the continuous response function used. In addition, Menyanthes also contains methods of non-linear time series modeling, following the threshold or TARSO approach [*Knotters and De Gooijer,* 1999]. These methods are useful for modeling systems with periodic run-off or periodically uncovering drainage means.
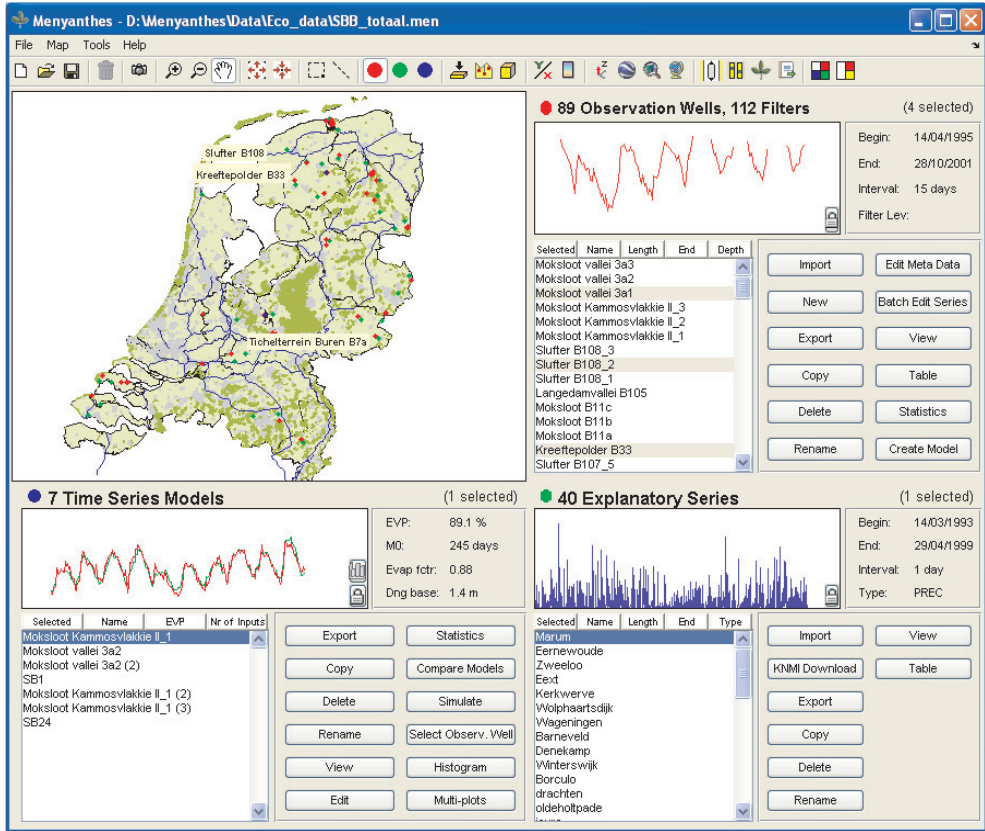
*figure 7.5: Main screen of Menyanthes, consisting of four quadrants. The set-up and coloring (dot markers) reflect the system schematization and colors in figure 7.1.*

### 7.3.5    Simultaneous time series analysis

One of the foremost advantages of the PIRFICT method is that it greatly reduces the amount of effort needed for model identification. As stated in the introduction, the iterative Box-Jenkins style model identification procedure for defining the number of parameters in an ARMA model can be very knowledge and labor intensive [*De Gooijer et al.,* 1985], whereas functions like the scaled gamma distribution are very flexible and may be used effectively to model the response of a system regardless of the hydrogeologic setting of the observation wells [*Von Asmuth et al.,* 2002]. The PIRFICT method allows for the use of a standard response function for a specific type of input, and model identification consequently reduces to the problem of identifying which inputs or stresses influence the dynamics of a system [*Von Asmuth et al.,* 2008]. In case of groundwater heads, or any environmental variable for that matter, stresses tend to have a spatial effect and influence all locations within a certain region. Assessment of the region of influence of a certain stress (for instance groundwater extraction) is in fact one of the possible applications of time series models. Using the PIRFICT method, stress identification therefore generally applies to all observation

wells in an area, making it a batch process. In figure 7.6, the user interface of *Menyanthes* is shown which supports the process of simultaneous model identification for multiple time series. In the box marked with a one, there is a listbox with locations or groundwater level series. All settings made in the user interface apply to the series selected in this listbox. In the box marked with a two, there is a listbox with stresses or explanatory series. *Menyanthes* automatically selects the nearest meteorologic stations, as precipitation and evaporation always influence the groundwater level. The explanatory series shown in listbox two are not necessarily used for *all* locations in listbox one, but in contrast, it is a list of explanatory series used for *any* of them. Properties of explanatory series that are set, are set for all locations selected in listbox one for which they are used. An example is the type of impulse response function used for a certain stress [see *Von Asmuth et al.,* 2008], which can be selected in the drop-down menu that is marked with a three. When appropriate settings, explanatory series and response functions are chosen, the model parameters can be optimized by pressing the 'estimate' button, after which the models are optimized one-by-one. Details on the optimization procedure and noise model, which are also based on a continuous-time formulation, are presented in [*Von Asmuth and Bierkens,* 2005]. The model identification procedure of Box and Jenkins includes diagnostic checking of results, using statistical methods and criteria. Correlations that are valid from a statistical viewpoint can, of course, be non-causal or physically non-plausible. The example application will demonstrate how physical insight may also be of use for checking the plausibility and adequacy of time series analysis results.
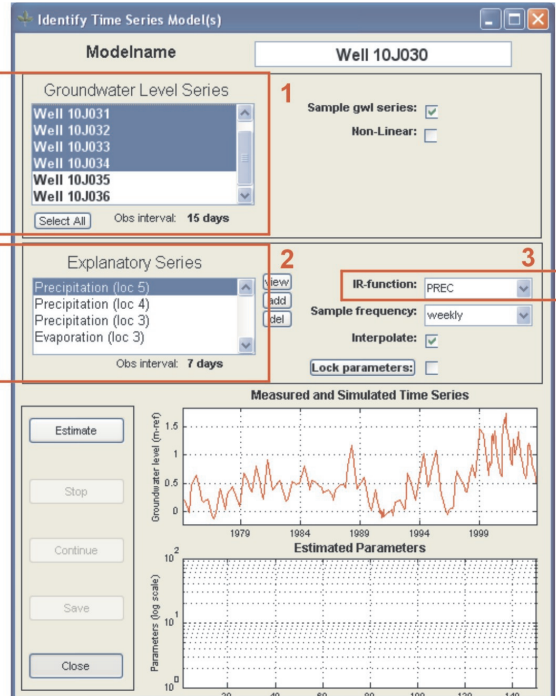


figure 7.6: *User interface for 'identifying' multiple time series models simultaneously.*

### 7.3.6    Menyanthes versus alternative software

The combination of time series modeling tools and the focus on geohydrologic systems and their corresponding data make *Menyanthes* unique in its sort, and hard to compare to other software. There is, of course, software that overlaps and is comparable to part of the functionality of *Menyanthes*. Hydrogeologic modeling software such as Visual Modflow (see www.swstechnology.com), Groundwater

Modeling System and Groundwater Vistas (see www.scisoftware.com) solve the differential equation for groundwater flow, using the finite difference method. Software like MLAEM (see www.scisoftware.com) or TTim (see ttim.googlecode.com) use the analytic element method for this. See, e.g., [*Reggiani and Schellekens,* 2005; *Todini,* 2007] for more information on models that (also) deal with surface water hydrology. *Menyanthes* differs from such models in the sense that it only models the temporal and not the spatial dynamics, and it models groundwater heads instead of flow. For application in practice, an important property of *Menyanthes* is that it does not require any geohydrologic schematization or parameterization. There are of course also other time series modeling packages like Matlab's System Identification Toolbox [*Ljung,* 2011], the Captain Toolbox [*Taylor et al.,* 2007], or R [*Cryer and Chan,* 2008]. Compared to these, the user interface of *Menyanthes* is more 'friendly' to non-specialist hydrolgeologists, as its primary focus is on the groundwater system itself, and not on the statistical aspects and properties of the models and data. Also, *Menyanthes* uniquely contains the PIRFICT method, which enables the fast and easy analysis of large numbers of time series simultaneously.

## 7.4    Example application

### 7.4.1    Time series decomposition

Standard time series models are linear, which may pose limitations to their usability. Linearity, however, also has a number of advantages and many applications are based on it. Whereas equation (2.35) only applies to the single input case, it can be readily extended to multiple inputs using superposition, which is valid for linear systems. In case of $n$ inputs, their combined effect is obtained simply by summing the individual effects:

$$h(t) - d = \sum_{i=1}^{n} (\theta_i * p_i)(t) \tag{7.10}$$

As stated previously, the $\theta_i$ 's in (7.10) are not necessarily independent. Parameters can be shared across several inputs in order to reduce their number. Also, for different types of inputs, different response functions can be used [see *Von Asmuth et al.,* 2008]. A consequence of (7.10) is that time series models can decompose time series into partial series showing the effects of individual inputs (figure 7.7). This property enables intervention analysis as described by [*Hipel et al.,* 1975; *Hipel and McLeod,* 1994, Ch. 19]. Time series are decomposed in an 'explained' and an 'unexplained' part (even for the single input case), allowing for trend detection or the quantification of anthropogenic influences when there is only one natural factor such as precipitation surplus [e.g., *Rolf,* 1989; *Gehrels,* 1999; *Yihdego and Webb,* 2011]. In the case of multiple inputs, the primary interest is usually the impact of one of these inputs, such as pumping [e.g., *Von Asmuth et al.,* 2008; *Harp and Vesselinov,* 2011] or hydrologic measures. More in general, time series analysis can reveal much of the

functioning of a system and of the interactions between its parts [e.g., *Lee and Lee,* 2000*; Fabbri et al.,* 2011]. An archetypical application of time series models lies in the area of 'control engineering', where the problem is how to manipulate the inputs into a system to obtain the desired output. For hydrogeologic purposes, this means that time series models can be used to optimize (or identify problems with) the water management in an area [e.g., *Bidwell et al.,* 1991*; Lee et al.,* 2005].

### 7.4.2    Impact of hydrologic measures in a Dutch dune area

As an example, an application of *Menyanthes* is presented in which the primary aim is to assess the impact of hydrologic measures. Other application examples include the characterization of groundwater dynamics [*Von Asmuth and Knotters,* 2004], the assessment of the influence of pumping [*Von Asmuth et al.,* 2008], the development of maps of the risk of extreme levels [*Manzione et al.,* 2010], the evaluation of effects of climatic changes on stagnant surface water [*Lehsten et al.,* 2011] and trend detection and impact of land use changes [*Yihdego and Webb,* 2011]. The area under consideration here is part of the Amsterdam Water Supply dunes, which is the area that has supplied the city of Amsterdam with water since the nineteenth century. The area contains facilities for storing and infiltrating surface water (figure 7.8), and pumping wells for groundwater extraction. It is perhaps the most intensively monitored hydrogeologic system world wide. The focus is on the southern part, which is relatively undisturbed by the water supply activities. Two main canals were dug in
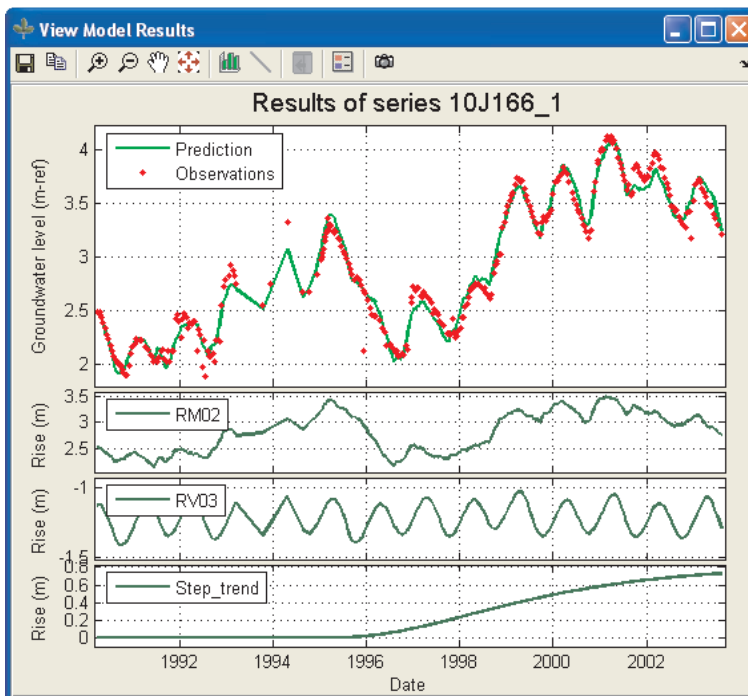


*figure 7.7: Groundwater level series decomposed into partial series showing the effects of individual inputs. The figure shows results from one of the 438 locations in the example application.*
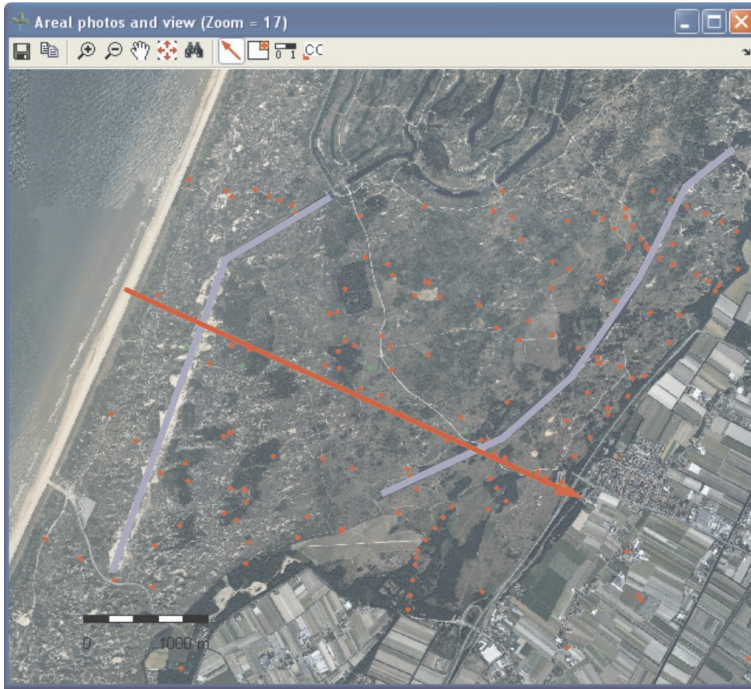
*figure 7.8: Location of the 438 selected piezometer screens (red dots) and meteorologic stations (green dots). The Van Limburg Stirumkanaal (left) and Oosterkanaal (right) are indicated by a blue line, whereas a red arrow shows the location of the transects of figure 7.10.*

this part of the dune area, and drain the system. The canal in the West near the sea is called the 'Van Limburg Stirumkanaal', whereas the canal in the East near the polder area is appropriately called the 'Oosterkanaal' or eastern canal. In order to increase the natural value and restore the ecological functioning of the area, to raise the groundwater levels and increase the area of (threatened) wet dune slack ecosystems, hydrologic measures were taken in the period from 1995 to 1996. As part of these measures, the Van Limburg Stirumkanaal was backfilled with sand and the level of the Oosterkanaal was raised by 50 centimeters. In addition, the few pumping wells that were present in the area were closed (pumping station Noordwijk and artesian wells bordering the Oosterkanaal). Finally, also the level in the storage canals was raised. In springtime following the completion of these measures in 1997, considerable crop damage was recorded in the flower bulb fields neighboring the dune area. The damage was estimated to be around three million dollars, and as the cause allegedly was oxygen shortage or 'drowning' of the bulbs, a link was made with the measures taken by the water supply company, which was blamed for the damage [e.g., *Olsthoorn,* 2000]. While previous attempts to quantify the effects of these measures using ARIMA models failed, a new attempt was made when *Menyanthes* became available. A summary of the results of this investigation are presented in the following, while restricting ourselves to the effects of the measures on the groundwater heads at the monitored locations.

In the area where the measures were taken, 680 locations (piezometer screens) are monitored by the water supply company. Precipitation data was collected at five different locations in a transect going from the sea to the polder area (see figure 7.8, one of which is 'RM3' in figure 7.7). Reference evaporation data was obtained from the meteorologic station 'Leiduin' of the water supply company ('RV03' in figure 7.7). Apart from these and the dates of the hydrologic measures, no other data were used. However, not every available time series proved to be usable for the analysis. From these 680 screens, series were selected that had measurements both some time before and after the measures, span a period that is not too short and contain a sufficient number of measurements in total. Based on these criteria, 438 screens of 194 observation wells were selected for the analysis (figure 7.8). As the area is relatively undisturbed by the current water supply activities, only the effects of natural stresses (precipitation and evaporation) and those of the measures were taken into account. Each measure thereby has a different effect in different parts of the system, but all measures were taken in a relatively short period (considering the long memory of the dune system). Their effect is therefore strongly correlated, and it is not possible to separate and estimate the individual effects directly from the data. It is, however, possible to estimate the combined effects of the measures, by modeling them as a shifted, single Heaviside step function $H(t)$ or step trend as input series. This function may be defined by:

$$H(t) = \int_{-\infty}^{t} \delta(\tau - T) \mathrm{d}\tau \qquad (7.11)$$

and it changes instantaneously from zero to one at time $T$, for which the 28th of February 1995 is chosen here (the date that the Van Limburg Stirumkanaal was backfilled). In words, a step function represents a sudden and lasting change in an
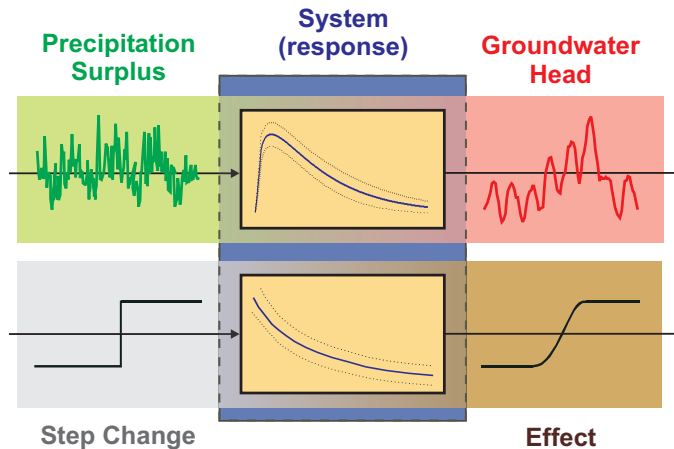
*figure 7.9: A step change in the input is transformed by the system into a sigmoid-shaped effect. When the change is not linked to one of the other input series, it has its own independent response function.*

input series. The effect of a step function on the output, which is called the step response function, is generally not instantaneous but it has a gradual, sigmoid-like shape (figure 7.9). When the step change is not linked to one of the other input series (as, e.g., a sudden change in evapotranspiration by forest clearing would), it should be given its own independent impulse response function, for which a scaled gamma distribution is used here. An example of results for a single series was already shown in figure 7.7. It may be seen from the lower graph in this figure that the response of the dune system is indeed slow, and the estimated effect of the measures has not reached its stationary value after 8 years, at the end of the time series. Four output screens of *Menyanthes* are shown in figure 7.10. They contain results from multiple series in 2D cross-sections. In *Menyanthes*, the model performance is characterized in several ways, one of which is the explained variance percentage (EVP). The EVP is defined as:

$$\text{EVP} = \frac{\sigma_h^2 - \sigma_n^2}{\sigma_h^2} * 100\% \tag{7.12}$$

where $\sigma_h^2$ and $\sigma_n^2$ denote the variance of the groundwater level observations and error or residual series respectively. The EVP for each series is plotted as a colored dot at the location of its piezometer screen in figure 7.10a. The estimated effect of the measures, based on a selection of screens that are placed higher than -15m+ref, is plotted in figure 7.10b. As the effects are still non-stationary and increasing, the estimated final effect can only be found through extrapolation and is therefore inexact. The estimated effect on 31-8-2003, at the end of the time series, is used here, along with its confidence interval (+/- $2\sigma$ ). Next to that, the values of the local drainage base and the estimated evaporation factor are shown in figures 7.10c and d, respectively. In all figures 7.10a) to 7.10d), spatial patterns emerge from the individual results. The EVP's, for instance, are clearly highest in the shallower screens (and range even up to 98.4%), while the lowest values can be found on the Western side near the sea. From this simple pattern, several conclusions may be drawn. First, there must be a (semi)confining layer at a depth of about -10 to -20 meters + ref (which corresponds with soil profile data), as otherwise, the lower screens would not behave differently than the higher screens . For this reason, only the shallower screens are depicted in figure 7.10b). Second, low EVP's indicate that there are factors missing in the models, and the spatial pattern clearly points to the tide as being one of them. A logical candidate for a second missing factor is the pumping that is going on in the Northern part of the area. The influence of pumping, like that of the tide, spreads much farther in a confined layer, explaining why these factors are not necessary to model the dynamics in the shallower, phreatic screens adequately. From the pattern that is visible in the effects in figure 7.10b), the individual effects of the backfilling of the Van Limburg Stirumkanaal and the rise of the level of the Oosterkanaal are visible. The effect of backfilling a canal is logically highest near the canal itself (and ranges up to 1m74, on 31-8-2003). Similarly, the effects of a rise in canal stage is highest at the canal. The effects gradually fade in the direction of the sea and the polder area, and towards the middle of the dune area. The local drainage base shows a gradient going
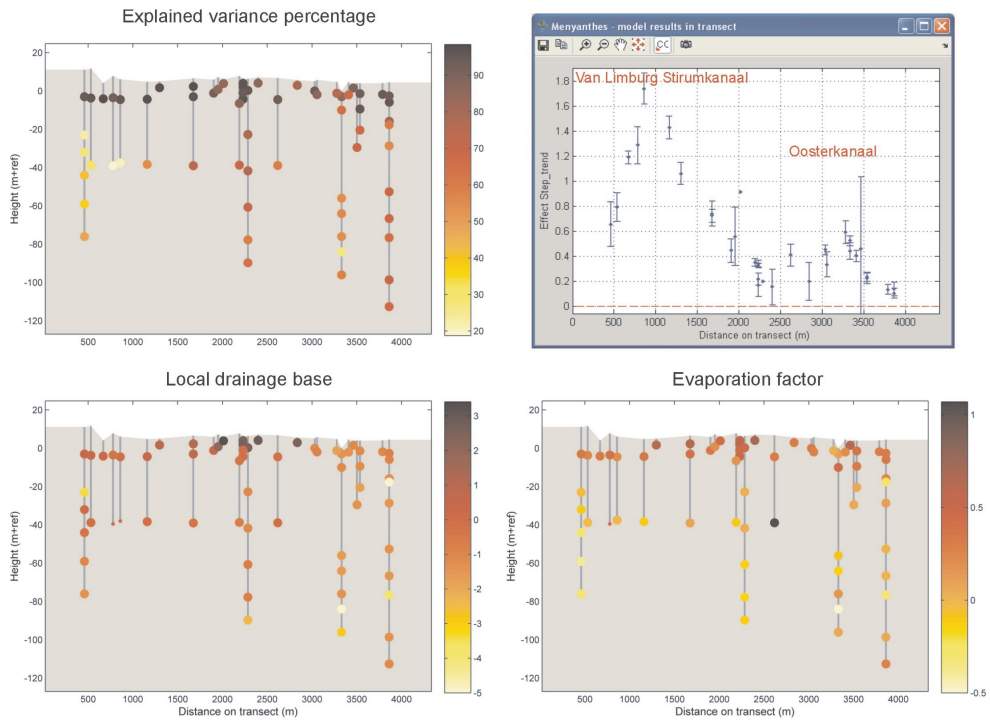
*figure 7.10: Model results along the transect displayed in figure 7.8, showing a) the explained variance percentages as colored dots at the locations of the piezometer screens. b) the estimated effect in screens higher than -15m+ref, on 31-8-2003. c) the local drainage base, d) the estimated evaporation factor.*

from the sea towards the (actively drained) polder area. Next to that, the phreatic screens in the center of the dune system clearly show a higher drainage base. This may be due to the fact that the precipitation surplus creates a 'lens' of fresh water that floats on top of the salt water beneath. The estimated evaporation factor is clearly lower beneath the confining layer, and even reaches negative values. As this is physically implausible, it suggests that the evaporation factor estimates below the confining layer are biased, probably due to the fact that the effects of pumping are not taken into account and the model performance is low. As water usage often shows a seasonal pattern, the effects of pumping are easily correlated with those of evaporation. In all figures, also several 'outliers' are visible. These outliers in part concern the same locations, and the results at these locations should be checked for plausibility. The results of the phreatic screens, however, where the model performance is high, are much more coherent. Additional detail on the modeling procedure, the model output, its interpretation and plausibility is given in [*Von Asmuth et al.*, 2008].

Note that the filling in of a canal or drainage means can also be seen as a change in system properties, making the system time-variant. In principle, time-variant systems

can not be modeled with one response function. In such cases, time series should be split up into periods in which the system is invariant, and be modeled separately. From a physical-hydrologic viewpoint, however, the issue of time-variance proves to be ambiguous, as the filling in of a canal can also be seen as a change in its flux from some value to zero, leaving the system unchanged. The approach presented above follows the latter viewpoint, and the generally very good model performance and plausible results add to the confidence in it.

## 7.5    Discussion and conclusions

Among the limited number of user-friendly tools for time series analysis that exist, *Menyanthes* has a special position. *Menyanthes* first integrates a variety of functions, starting from the (data)base of the DIKW pyramid, above which are visualization and analysis tools, and ending with the time series modeling tools on top. The combination of these functions strengthens each one of them. For example, the visualization tools and close link with the data are useful during the time series modeling process. *Vice versa*, time series models may also be used for improved characterization and visualization of a system's dynamics. They can even be of use for data management, e.g., for error or outlier detection. Second, *Menyanthes* facilitates the use of physical principles and physical interpretation. In this respect, the PIRFICT method of time series analysis has a parallel with the data-based mechanistic methodology [*Young and Beven,* 1994*; Young,* 1998]. A difference is that physical laws or physically inspired behavior are implemented *a priori* in the model, in addition to the use of physical insight to check or select results *a posteriori*. Furthermore, temporal moment matching allows for a direct link between time series and spatially-distributed, physically-based models. Scrutiny should, however, be practiced in the use of physical insight. The choice of a predefined continuous response function within the PIRFICT method is an assumption, whose adequacy may be checked in *Menyanthes* by comparing results with those of ARMA models. Also, the assessment of the physical plausibility of model results should be made carefully. Time series (and also physically-based) model results may be influenced by non-causal cross-correlations. On the other hand, unexpected model results are not necessarily implausible, but may lead to new insights.

The possible applications of the PIRIFCT method are essentially the same (and as broad) as those of ARMA time series models. An attractive feature of time series models in general is that they are based on relatively few assumptions and the fits are generally high; the model lets the data more or less speak for itself. As such, time series models may be a valuable tool for preprocessing groundwater level series before calibrating a distributed model. Using time series models, missing stresses or series that are influenced by hydrologic interventions may be identified readily. Also, series may be identified that are not suitable for use in model calibration, for instance because they represent a hydrologic feature that is not incorporated in the model, such as perched water tables. A specific feature of the PIRIFCT method is its increased efficiency and ease of use. The finding that skew-Gaussian distribution functions prove to fit the behavior of hydrogeologic (and other) systems well in general, allows for all time series in an area to be modeled in batch. Batch modeling is not 'only' a

practical advantage, but the simultaneous visualization and analysis of results from multiple locations literally adds another dimension to the technique and its results (i.e. the spatial dimension). Whereas the spatial structure that is implemented in physically-based models is based on inferences and assumptions that cannot be truly verified (as was argued in the introduction), there is no spatial dependency imposed on the time series models. Spatial patterns that emerge in time series analysis results are therefore obtained independently, and may yield valuable information on both the model results themselves and the functioning and properties of systems. Furthermore, effects of individual stresses that cannot be separated through time series analysis of individual series alone, may become separable on the basis of the observed spatial patterns. The spatial visualization and analysis tools of *Menyanthes* greatly enhance the use and interpretation of these patterns. As *Menyanthes* facilitates a shift in focus from data and statistical aspects of time series towards the physical properties of systems and spatial variation thereof, the term 'system identification' as coined by Zadeh [1956] and adopted by, e.g., Ljung [1999], becomes more appropriate than the term 'time series analysis' that has its origin in sciences like econometrics. To date, *Menyanthes* is used on three continents, in six countries, by 76 different organizations. A few hundred ecologists and hydrogeologists have followed a course on hydrogeologic time series analysis and use of *Menyanthes* [*Von Asmuth et al.,* 2006], and several tens of thousands of groundwater level series have been analyzed with it. The program and its methods therefore prove to fulfill a need that is hard to fulfill otherwise. As stated, the fact that standard time series models are linear is a strong point, but also a limitation. Methods of non-linear time series analysis following the threshold concept have been added to *Menyanthes*, and prove to be of value. The authors welcome initiatives to test the methods on a wider range of systems, and extend its functionality with other (non-linear) processes.

# Chapter 8

# Summary and conclusions

## 8.1    Objectives, contributions and background

As perhaps is not uncommon in research, the initial objectives differ from the final results presented in this thesis. Initially, the primary objective was to improve current methods with which the relationship between groundwater dynamics (and impact of hydrologic measures) and the species composition of groundwater dependent ecosystems is modeled. In the alternative proposed originally by [*Maas,* 1995], a key notion is that the spatial differences in groundwater level dynamics are mainly determined by the spatially variable properties of groundwater systems, while the variation through time or temporal dynamics are mainly driven by the spatially less variable meteorologic dynamics. Consequently, it was hypothesized that spatial differences in vegetation are also primarily determined by system properties, or in other words, that they can be modeled more accurately by 'filtering out' the temporal, meteorologic dynamics. To be more specific, Maas proposed the use of time series models for inferring the so-called impulse response function from series of groundwater level observations, which in turn can be characterized using statistical moments (see section 2.3.4). Moments are scalars and constants in linear, time-invariant systems. Together with the spatially less variable driving forces, they completely characterize the (deterministic part of the) dynamics at a certain point in space. As an additional advantage, moments can also be simulated spatially using a standard distributed groundwater model. The scope of this research thus initially encompassed three fields of modeling, i.e. time series, groundwater and eco-hydrologic modeling. These fields of modeling together constitute the method of impulse response moments, from which moments can be derived and between which they can be mutually exchanged. Although it was envisaged that the link between time series models and groundwater models would be a subject in its own right, at the time the existing ARMA time series models were thought to be well developed and to be directly usable without modification. ARMA models, however, turned out to have several important limitations (see section 2.2.3). What followed was a process of improving and adapting time series analysis methods to the needs set by the initial objectives of this research (but not limited to that). As such, most of the actual contributions of this thesis are in the field of time series analysis.

Appreciating the contents of this thesis requires some knowledge of its background, i.e. the methods and theory on (hydrogeologic) time series analysis and system identification. In chapter 2, a basic introduction is therefore given, and different views on the subject are presented. First, some thought is given to the systems approach in general, as to date, most hydrologist are less familiar with system identification methods than with physically-based, spatially-distributed models. In short, the system viewpoint can be characterized by the fact that in essence it is top-down. The system viewpoint treats a groundwater system firstly as a 'whole' and not bottom-up, as an aggregate of cells, layers and/or elements. Having said this, also from the systems perspective, there are different ways in which time series or observed dynamics can be perceived and modeled. Time series analysis is a method that originates from the statistical sciences. In principle, it does not require any knowledge of the physical functioning of the system under consideration and basically, it can be seen as a variant of simple, linear regression. From a physical point of view, on the other hand,

the central concept is the so-called impulse response function. Impulse response functions can be inferred from a data set through time series analysis, but also using 'purely' physically-based methods, either analytic or numeric. A link between time series and distributed models can, as stated previously, be established using moments of impulse response functions. Moments can be generated directly and spatially using moment-generating differential equations, implemented in a standard groundwater model. Convolution of an impulse response function with a time series that describes the dynamics of the driving force yields a simulation of the groundwater level dynamics at a certain location (see section 2.3.2). It can serve as an alternative method for solving partial differential equations or groundwater models, next to current methods as the analytic element, finite element or finite difference method. Examination of the impulse responses of several elementary groundwater systems yields insight in its shape, as a function of elementary physical variation.

In this thesis, a 'mix' between statistical and physically-based time series modeling methods is developed and presented. In this approach, which was named the PIRFICT method, the time series analysis problem is formulated in a continuous time domain, having a convolution integral as its basis. It allows for the use of (continuous) distribution functions that have a statistical origin, as well as physically-based analytic response functions. Distribution functions of skew-Gaussian nature, however, like the scaled gamma distribution, prove to fit the behavior of a wide range of systems quite well. The PIRFICT method has several advantages over ARMA models and perhaps most importantly, it allows for the application of time series analysis methods on a large scale and in a standardized, less knowledge and labor intensive fashion. As for application in practice, an important step is that a user friendly computer program named *Menyanthes* was developed, which contains most of the methods presented. Use of the methods and tools developed allows researchers and (eco)hydrologists to infer the effects, impulse responses and moments of different driving forces effectively from a given data set, and separate them. Simply put, this thesis provides the methods and tools that are better put up to the original tasks set for it. In the next sections, a summary and conclusions of the different steps taken and peer reviewed papers published are given.

**175**

## 8.2    Time series modeling using continuous response functions
As discussed in section 2.2.3, the traditional methods for time series analysis (the discrete-time AutoRegressive-Moving Average (ARMA) -type transfer function noise model of [*Box and Jenkins,* 1970]) have several limitations. In chapter 3, an alternative method is presented that is formulated in continuous time and has a convolution integral at its basis. In this so-called PIRFICT model, in principle any analytic function can be used as impulse response function, as long as it has a known integral or step response function. A scaled gamma distribution, however, whose physical analog is a cascade of linear reservoirs, proves to fit the groundwater level response to precipitation and evaporation closely in general. This is shown in the example application of chapter 3 (and in many later results). Furthermore, the results of the PIRFICT model prove to be equivalent to those of ARMA models, which can be concluded from the fact that the transfer functions of both models have the same

general shape, when the models are applied to the same data. ARMA transfer functions, however, show an irregular pattern in the part that is influenced by MA parameters, whereas the shape of the gamma distribution is smooth. On the other hand, the performance of both models in terms of calibration and validation RMSE (Root Mean Squared Error) also proved to be highly comparable. When comparing results from 15 piezometer series, the only difference found was that the calibration error of the PIRFICT model was slightly higher, whereas its validation error was actually slightly lower. Although the differences are small, they can be explained by the different model structure. As the shape of the MA part of an ARMA transfer function is free, it is also free to fit coincidental cross-correlations between input and output. Such 'overfitting' behavior will result in a lower calibration RMSE, a higher validation RMSE, and in a partly coincidental or random pattern of the transfer function, as was found in the example application. Because the continuous response functions of the PIRFICT method are smooth and in general have fewer parameters, it is less sensitive to overfitting. However, the choice of a specific response function in the PIRFICT method also imposes restrictions on its shape. These restrictions can also negatively influence the model results, when the function chosen does not adequately match the 'true' response. All in all, the PIRFICT method has the following attractive features:

- The parameters in the PIRFICT approach are in principle constants and insensitive to the data frequency. This is important when aiming to assess time invariant response characteristics of systems, as in the original objective of this thesis. The example application of chapter 3 shows that, as expected, the results of ARMA models do depend on model order and data frequency (which, however, may also be 'repaired' given the continuous time viewpoint).
- Because the equations are continuous in time, the model can simulate and operate on data at any frequency. The frequency of the input and output do not have to be equal and can even be irregular. In comparison to the combined Kalman filter and ARX model approach [*Bierkens et al.,* 1999], the PIRFICT method is more general. In principle any analytic response function can be used, not only the exponential function of the ARX model. As discussed in section 2.3.5, this is important in many situations, e.g., for taking into account the relative position in a system, for modeling the effects of the unsaturated zone and excitations other than precipitation and evaporation.
- Use of the PIRFICT approach strongly simplifies the model identification procedure. The continuous time formulation firstly enhances physical insight in the model structure and its results. The physical functioning of a specific system, including its physically-based parameters, can in principle be directly implemented in the model, in the shape of an analytic solution. The order of ARMA models, on the other hand, contains at least six terms (when including the frequency). Its definition is a complex and partly subjective matter. The model order has to be identified more or less manually, or simply put, the six terms or numbers have to be 'manually optimized' by the modeler. In contrast, response functions like the scaled gamma distribution have a wide applicability; they are very flexible and comprise a range of ARMA model

orders. When using such functions, model identification generally reduces to identifying which factors influence a certain system. After that, the same model can be applied to all piezometers in an area, regardless of their geohydrologic setting and location. This feature, together with the previous point, makes the PIRFICT method especially suited for standardized and automated analysis of large numbers of time series.

## 8.3    A continuous noise model for autocorrelated data of irregular frequency

Groundwater head series, and most environmental time series for that matter, share the common feature that the observations (ánd the errors of any model applied to them) are autocorrelated. In short, autocorrelation means that the value of a variable at a certain time step is correlated with the values at previous time steps. Whenever model errors are autocorrelated, they cannot be modeled as simple, independent Gaussian deviates, as such a practice would yield inefficient parameter estimates and covariance estimates that in general are too low (see also section 2.2.2). In such cases, the model error should be modeled in a way that accounts for the autocorrelation, as is the case when a separate, so-called noise model is used.

In chapter 4, a continuous time noise model is presented that can handle irregularly spaced data and whose basis is the so-called Ornstein-Uhlenbeck process. Because of the continuous time formulation, it combines well with the PIRFICT model. It is, however, not restricted to that, and can be applied for modeling irregularly spaced residuals of various deterministic models. For gaining insight in its functioning, a comparison is made between the AR(1) model, 'conventional' or embedded in a Kalman filter, and the Ornstein-Uhlenbeck based (OUB) noise model. In this respect, the Kalman filter is treated in the pure prediction scenario or its 'degenerate' form [*Ahsan and O'Connor,* 1994], meaning that the variance of measurement noise is regarded to be zero or negligible compared to the other sources of noise. It is shown that the equations for handling the predictions in the time update of the 'degenerate' Kalman filter, those of the forecasting mode in the AR(1) model, and those yielding the innovations in the continuous noise model are mathematically equivalent. The OUB model, however, is exact and computationally more efficient, as it is not evaluated recursively. Also, because of its simplicity, the OUB model can be easily implemented. A limitation of the OUB model, however, is that it is limited to processes showing exponential decay.

As for the optimization process, a simplification that is made is that a weighted least squares criterion was derived from the likelihood function used otherwise. This criterion was named the Sum of Weighted Squared Innovations (SWSI) criterion. Parameter estimates based on the SWSI criterion converge to maximum likelihood estimates for larger sample sizes. An advantage of the SWSI criterion is that it can be optimized using standard nonlinear least squares regression methods, e.g., the method of Levenberg-Marquardt, which are very computationally efficient. Next to the OUB model, also the SWSI criterion can be used in combination with other models, e.g., for the innovations of a Kalman filter when exponential decay is assumed.

Last but not least, this chapter introduces the innovation variance function as a tool for diagnostically checking the validity of the response function of a noise model. This check can be performed by comparing the estimated innovation variances for different time steps with the theoretical innovation variance function, accompanied by its confidence interval. The innovation variance function provides additional information on the validity of a model, especially in light of its use on data of irregular frequency.

## 8.4    Modeling groundwater head series subjected to multiple stresses

In ARMA transfer function noise models, the standard procedure for modeling time series that are subjected to multiple stresses is simply to connect them in parallel in the model. In other words, each individual stress is assigned its own, independent transfer function, and their effect is summed to give the total (deterministic) prediction. Although the same procedure could be followed in the PIRFICT approach, the subject deserves more attention than that, because:

- The scaled gamma distribution (see section 2.4.1) generally fits the response to precipitation and evaporation well, but not necessarily that to other types of stresses. In chapter 5, impulse response functions for other stresses are derived from analytic solutions to the appropriate elementary geohydrologic problems.
- The response to different types of stresses can be a function of the same, physical parameters. In such cases, the number of parameters and degrees of freedom in the model can be reduced by taking this into account. An example is evaporation, whose response can be modeled as the reverse of that of precipitation, apart from a reduction factor.
- In time series literature, the focus lies on analyzing individual series, using statistically based methods and criteria. In case of groundwater heads, however, a physical-hydrologic perspective, in combination with the simultaneous analysis of multiple series, can be very helpful. It can even be a necessity, when results from single series are ambiguous.

The PIRFICT approach all in all calls for a different perspective on modeling procedures and interpretation of results. As it is more complex than the single input case, in the multiple input case also the interpretation of results is more complex. Physical knowledge can, as stated, be very valuable in checking the consistency and plausibility of results, and forms a welcome addition to statistical diagnostic checks. First and logically, the parameter values should fall within a range that is physically plausible. Next to that, whenever multiple locations are monitored and analyzed, spatial patterns supply important and independent feedback on the results. These patterns are independent, as there is no spatial dependency imposed *a priori* on the individual models. Via the model residuals and their patterns, missing stresses, measures or other sources of error may be readily identified. High levels of error can cause parameter estimates to be biased, as other stresses may partly compensate missing

stresses when their influence is correlated. A stress whose estimated effect is easily biased is evaporation, as it displays a seasonal pattern that easily correlates with other (seasonal) factors. An example is given in chapter 5, where the spatial distribution in the evaporation factor clearly showed its bias. In this case, the main source of error was the fact that the behavior of the individual pumping wells was not accounted for, whereas the total pumping rate indeed displayed a seasonal pattern. Such a problem may be solved by incorporating data on missing factors in the model, in this case the pumping of the individual wells. If such data are not available (as in this case), the bias can be reduced by constraining the optimization problem to a realistic range, based on the values and patterns in the surrounding observation wells.

Next to these methodological issues, chapter 5 serves as an example of the fact that a time series model can effectively decompose head series into partial series, which each show the effect of an individual stress. This is perhaps one of the foremost 'practical values' of a time series model. Under certain conditions, by doing so, the effect of individual hydrologic measures or factors such as groundwater pumping can be identified and quantified. Furthermore, in the model also the effect of changes in the level of the stresses is contained, by means of the individual response functions. This allows its results to be used for scenario calculations, optimization of groundwater heads and/or their management.

## 8.5 Characterizing groundwater dynamics based on response properties

As discussed in section 1.1, the initial aim of this thesis was to improve current methods for modeling the relationship between the groundwater level and groundwater dependent vegetation. Maas [1995] proposed the use of moments of the impulse response function as system characteristics, in stead of statistics describing the groundwater level dynamics themselves. Descriptive statistics are influenced by climatic variation and only capture certain aspects, whereas moments are constants and completely characterize the dynamics. However, it is not the moments but the groundwater level itself that determines growth conditions in wet ecosystems (together with other abiotic and biotic factors, and more or less direct). Moments may not ne directly or uniquely usable for eco-hydrologic modeling, because:

- Moments are constants, whereas ecosystems are not. While it may be an attractive feature in relatively stable ecosystems, this limits the direct use of moments in modeling the dynamics of ecosystems, e.g., in response to climatic dynamics.
- In densely populated areas like the Netherlands, (ground)water levels are often managed and/or influenced by anthropogenic activities. In such cases, not one but several sets of moments are of importance, and the activities themselves are spatially variable also.
- Wet ecosystems by definition have high water tables, resulting in predominant non-linear behavior. Also in case of non-linearity, not one but several sets of moments are of importance.

In order to describe the problem more accurately, different methods for characterizing the groundwater level dynamics are discussed in chapter 6. Also, the relationship between groundwater level characteristics and impulse response moments is investigated, based on the scaled gamma distribution and PIRFICT approach to time series analysis. In order to provide alternatives, a frequency domain analysis of the dynamics and response of systems was made. From this, a method was developed that combines the temporal dynamics of the input (characterized by its mean level and annual amplitude) with the spatially variable impulse response moments. This method results in a new set of parameters that characterizes the output or groundwater dynamics (GD), using simple analytic expressions. These so-called GD characteristics are the mean depth, convexity, annual amplitude and phase shift of the groundwater level, and it is shown that they characterize the groundwater dynamics in detail (for as far as the time series model does). The link between input, moments and GD characteristics can also be used vice versa, for instance to make estimates of the recharge to an aquifer, analogous to what was proposed earlier in [*Chen et al.,* 2002].

In the example application in chapter 6, the GD characteristics are compared to other methods of characterizing the groundwater regime, using two series of groundwater level observations. It is shown that the so-called and often used MxGL statistics (i.e. the Mean Highest, Lowest and Spring Groundwater Level) have some important drawbacks and do not differentiate well between different groundwater regimes. In MxGL statistics, different aspects of the fluctuations are mixed up, as the three highest (HG3) and lowest levels (LG3) in a year are the result of both annual and higher frequencies. By averaging the xG3 over a longer period (minimally 8 years), low-frequency dynamics and inter-annual extremes are effectively filtered out, and are consequently disregarded. By averaging the three extreme values in a year, and by imposing an observation period of 14 days, the MxGL statistics also do not reflect the annual extremes directly. Both the annual and inter annual extreme levels, however, are of great interest because of their disproportionate economic damage or ecologic influence. Consequently, it is concluded that the capability of MxGL statistics in characterizing the groundwater dynamics at different locations is less than that of GD characteristics. The methods by which GD-characteristics are obtained can be readily extended to include excitations such as groundwater pumping or surface water levels. For linear systems, the stationary level and annual amplitude of the effect of all stresses can simply be summed. Only when stresses have a very different dynamic behavior, such an approach does not adequately characterize the dynamics and additional characteristics might be needed. A special case of another factor that influences the groundwater level is the model error or noise that enters the system. However, as the expected value of white noise is zero and its frequency spectrum is flat, both its stationary influence and its amplitude at annual or any other specific frequency can be neglected.

All in all, GD characteristics provide a compact way of describing the groundwater dynamics at a certain location. They can be used not only for eco-hydrologic modeling, but whenever the amount of information contained in a full spatiotemporal description of the groundwater level dynamics is simply too large, e.g., visual

interpretation, mapping or empirical modeling purposes. As an additional advantage, GD characteristics can be readily obtained for any input (e.g., climatic) scenario using the analytic expressions given. This adds flexibility to the method, which is necessary for dealing with situations where groundwater level fluctuations are non-stationary. Non-stationarity is caused by climatic changes, but also factors like groundwater pumping. In such cases, the characteristics of an historical 8- or 30-year record of measurements do not adequately reflect the current or future state.

## 8.6 *Menyanthes*: software for groundwater head data

In two ways, this chapter provides a synthesis of the contents and results of this thesis. First, it contains a short overview of the methods developed and results obtained, in relation to alternative methods. While many papers approach time series analysis methods from a statistical viewpoint, this chapter starts from a physical viewpoint to better serve those that are not initiated in time series analysis. It furthermore focuses specifically on the interface between data and physically-based methods, as that is the main characteristic of the PIRFICT method and their integration reflects the original objectives of this thesis. Second, the program *Menyanthes* is presented, which was developed within the scope of this research, and in which most of the methods developed and functionality needed for performing the research were incorporated. In that sense, *Menyanthes* also provides a synthesis of the methods and results, as they are joined in the program.

Among the limited number of user-friendly tools for time series analysis that exist, *Menyanthes* takes a special position. *Menyanthes* first integrates a variety of functions, starting from the (data)base of the DIKW pyramid of [*Ackoff,* 1989], above which are visualization and analysis tools, and ending with the time series modeling tools on top. The combination of these functions strengthens each one of them. For example, the visualization tools and close link with the data may be valuable during the time series modeling process. *Vice versa*, time series models may also be used for improved visualization and characterization of the dynamics of hydrologic systems. They can even be of use for data management, e.g., for error or outlier detection. Second, *Menyanthes* has a strong physical-hydrologic focus, and facilitates the use of physical principles and physical interpretation. In this respect, the PIRFICT method of time series analysis has a parallel with the data-based mechanistic modeling (DBM) methodology [*Young and Beven,* 1994*; Young,* 1998]. A difference is that physical principles or physically inspired behavior can be implemented *a priori* in a PIRFICT model, in addition to the use of physical insight to check or select results *a posteriori* as in the DBM approach. Scrutiny should, however, be practiced in the use of physical insight. The choice of a predefined continuous response function within the PIRFICT method is in fact an assumption, whose adequacy may be checked in *Menyanthes* by comparing results with those of ARMA models. Also, the assessment of the physical plausibility of model results should be made carefully. Time series (and also physically-based) model results may be influenced by non-causal cross-correlations. On the other hand, unexpected model results are not necessarily implausible, but may lead to new insights.

The possible applications of the PIRIFCT method are essentially the same (and as broad) as those of ARMA time series models. An attractive feature of time series models is that they are based on relatively few assumptions and the fits are generally high. As such, time series models may be a valuable tool for preprocessing groundwater level series before calibrating distributed models. Using time series models, missing stresses or series that are influenced by hydrologic interventions may be readily identified. Also, series may be identified that are not suitable for use in model calibration, for instance because they represent a hydrologic feature that is not incorporated in the model, such as perched water tables. One step further, temporal moment matching allows for a direct link between time series and spatially-distributed, physically-based models [*Von Asmuth and Maas*, 2001; *Bakker et al.*, 2007; *Bakker et al.*, 2008]. A specific feature of the PIRIFCT method is its increased efficiency and ease of use. The finding that skew-Gaussian distribution functions prove to fit the behavior of hydrogeologic (and other) systems well in general, allows for all time series in an area to be modeled in batch. Batch modeling is not 'only' a practical advantage, but the simultaneous visualization and analysis of results from multiple locations literally adds another dimension to the technique and its results (i.e. the spatial dimension). Whereas the spatial structure that is implemented in physically-based models is based on inferences and assumptions that cannot be truly verified (as was argued in the introduction), there is no spatial dependency imposed on the time series models. Spatial patterns that emerge in time series analysis results are therefore independent and free of assumptions (except for the fact that the representativeness of input series may vary in space), and may yield valuable information on both the model results themselves and the functioning and properties of systems. Furthermore, effects of individual stresses that cannot be separated through time series analysis of individual series alone, may become separable on the basis of the observed spatial patterns. The spatial visualization and analysis tools of *Menyanthes* are important in this respect, and greatly enhance the use and interpretation of these patterns. As such, the methods and functionalities of *Menyanthes* facilitate a shift in focus from data and statistical aspects of time series towards the physical properties of systems and spatial variation thereof. The term 'system identification', as coined by [*Zadeh,* 1956] and adopted by [e.g., *Ljung,* 1999], thus becomes more appropriate than the term 'time series analysis', which has its origin in sciences like econometrics.

To date, *Menyanthes* is used on three continents, in six countries, by 76 different organizations. A few hundred ecologists and hydrogeologists have followed a course on hydrogeologic time series analysis and use of *Menyanthes* [*Von Asmuth et al.,* 2006]*,* and several tens of thousands of groundwater level series have been analyzed with it. The program and its methods therefore prove to fulfill a need that is hard to fulfill otherwise. As stated, the fact that standard time series models are linear is a strong point, but also a limitation. Methods of non-linear time series analysis following the threshold concept [*Knotters and De Gooijer*, 1999] have been added to *Menyanthes*, and prove to be of value. Initiatives to test the methods on a wider range of systems, and extend its functionality with other (non-linear) processes are welcomed.

# Hoofdstuk 9

# Samenvatting en conclusies

## 9.1 Doelstelling, bijdrage en achtergrond

De oorspronkelijke doelstelling van dit promotieonderzoek verschilt, zoals wellicht niet ongebruikelijk bij onderzoek, van de uiteindelijke resultaten die hier gepresenteerd worden. Doelstelling van het onderzoek was oorspronkelijk het verbeteren van de gangbare methoden waarmee de relatie tussen grondwaterdynamiek (en ingrepen daarin) en grondwaterafhankelijke ecosystemen gemodelleerd wordt. In het oorspronkelijk door [*Maas,* 1995] voorgestelde alternatief, staat de notie centraal dat ruimtelijke verschillen in grondwaterdynamiek hoofdzakelijk worden bepaald door de ruimtelijk variërende eigenschappen van het grondwatersysteem. De variatie in de tijd ofwel temporele dynamiek daarentegen, wordt hoofdzakelijk gedreven door de ruimtelijk minder variabele meteorologische dynamiek. Hieruit kan vervolgens de hypothese afgeleid worden dat ook de ruimtelijke verschillen in vegetatie hoofdzakelijk afhangen van systeemeigenschappen, of dat deze, in andere woorden, wellicht nauwkeuriger gemodelleerd kunnen worden door de meteorologische dynamiek 'weg te filteren'. Om preciezer te zijn, Maas stelde voor om met behulp van tijdreeksmodellen de zogenaamde impulsresponsfunctie af te leiden uit tijdreeksen van de grondwaterstand, en die op zijn beurt te karakteriseren met statistische momenten (zie paragraaf 2.3.4). Momenten zijn scalair en constant in lineaire, stabiele systemen, en ze bepalen samen met de ruimtelijk minder variabele, drijvende krachten volledig (het deterministisch deel van) de dynamiek op een gegeven plaats in de ruimte. Een bijkomend voordeel is dat momenten ook ruimtelijk gemodelleerd kunnen worden met behulp van een standaard grondwater model. Het onderwerp van dit proefschrift omvatte dus oorspronkelijk drie velden van (eco)hydrologische modellering, te weten tijdreeks-, grondwater- en ecohydrologische modellen. Uit elk type model kunnen momenten afgeleid worden. Tezamen vormen ze de impulsrespons-momenten-methode, waarbinnen momenten onderling uitgewisseld kunnen worden. Alhoewel voorzien was dat de link tussen tijdreeks- en grondwatermodellen niet eenvoudig te leggen is en een aparte studie zou kunnen vergen, was destijds de inschatting dat de bestaande ARMA-tijdreeksmodellen uitontwikkeld en zonder meer bruikbaar zouden zijn. Bij het begin van het promotieonderzoek bleek echter al gauw dat ARMA-modellen verschillende belangrijke beperkingen kennen (zie paragraaf 2.2.3). Wat hierop volgde was een proces van het aanbrengen van wijzigingen en verbeteringen in de tijdreeksanalyse methodiek, om ze beter bruikbaar te maken voor de oorspronkelijke doelstellingen van dit onderzoek (maar niet beperkt daartoe). Het grootste deel van de eigenlijke, uiteindelijke bijdragen van dit proefschrift ligt dan ook op het vlak van de tijdreeksanalyse.

Het kunnen plaatsen van de inhoud van dit proefschrift vergt enige kennis van haar achtergronden, te weten de theorie en methodiek van tijdreeksanalyse (van grondwaterstanden) en systeemidentificatie. Hoofdstuk 2 bevat daarom een basale introductie op het onderwerp, vanuit verschillende gezichtspunten. Allereerst wordt enige aandacht besteed aan de systeembenadering in het algemeen, omdat de meeste (eco)hydrologen op het moment minder vertrouwd zijn met systeem identificatie en haar methoden dan met op de fysica gebaseerde, ruimtelijk gedistribueerde modellen. De systeembenadering kan kortweg gekarakteriseerd

worden door te stellen dat haar gezichtspunt in essentie 'top-down' is. Een grondwatersysteem wordt daarbij in eerste instantie als 'geheel' benaderd, en niet 'bottom-up' als een verzameling van cellen, lagen en/of andere elementen. Hier zij echter opgemerkt dat er ook vanuit systeemperspectief verschillende manieren zijn waarop tijdreeksen en/of de waargenomen dynamiek beschouwd en gemodelleerd kunnen worden. Tijdreeksanalyse is een methode die haar oorsprong vindt in de statistiek. Tijdreeksanalyse kan opgevat worden als variant van eenvoudige, lineaire regressie, en voor het toepassen ervan is in beginsel geen kennis nodig over het fysieke functioneren van het beschouwde systeem. Vanuit fysisch oogpunt, aan de andere kant, neemt de zogenaamde impulsresponsfunctie, of kortweg impulsrespons, een centrale plaats in. Ze kan uit metingen afgeleid worden via tijdreeksanalyse, maar ook via 'puur' fysisch gebaseerde, analytische dan wel numerieke, methoden. De link tussen tijdreeks- en grondwatermodellen kan, zoals gezegd, gelegd worden via de momenten van de impulsrespons, die ruimtelijk gesimuleerd kunnen worden via zogenaamde moment-genererende differentiaalvergelijkingen, geïmplementeerd in een willekeurig grondwatermodel. Convolutie van de impulsrespons met een tijdreeks die de dynamiek van de drijvende kracht beschrijft, geeft een simulatie van de grondwaterspiegeldynamiek op een bepaalde plek. Het kan dienen als alternatieve methode voor het oplossen van een grondwatermodel of partiële differentiaal vergelijking, naast gangbare methoden als de analytische elementen, eindige elementen of eindige differentie methode (zie paragraaf 2.3.2). Beschouwing van de impulsrespons van een aantal elementaire grondwatersystemen geeft inzicht in de vorm ervan, als functie van elementaire fysische variatie.

Als onderdeel van dit proefschrift is een nieuwe methode van tijdreeksanalyse ontwikkeld, die als zodanig een 'mix' vormt tussen statistische en fysische gebaseerde methoden van tijdreeksmodellering. In deze benadering, die PIRFICT-methode gedoopt is, is het tijdreeksanalyseprobleem geformuleerd in continue tijd, en het heeft een convolutieintegraal als basis. De formulering staat het gebruik van (continue) verdelingsfuncties uit de statistiek toe als responsfunctie, maar ook fysisch-gebaseerde, analytische functies. Verdelingsfuncties die de vorm aannemen van een scheve normale verdeling, zoals de geschaalde gamma verdeling, blijken echter in het algemeen goed in staat het gedrag van een reeks van systemen te beschrijven. De PIRFICT-methode heeft een aantal voordelen ten opzichte van de traditionele ARMA-modellen en -methode. Het heeft als wellicht belangrijkste eigenschap dat het op een gestandaardiseerde, en minder kennis- en arbeidsintensieve manier, toegepast kan worden op een groot aantal meetreeksen tegelijk. Een belangrijke stap in de toepassing van de ontwikkelde methoden in de praktijk is de ontwikkeling van een gebruiksvriendelijk computerprogramma geweest, genaamd *Menyanthes*. Het gebruik van dit programma en de daarin opgenomen methoden stelt onderzoekers en (eco)hydrologen in de praktijk in staat om de effecten, impulsresponsfuncties en momenten van verschillende invloedsfactoren af te leiden uit de data, en te scheiden. Kort gezegd levert dit promotieonderzoek de methoden en hulpmiddelen die nodig zijn voor het uitvoeren van de taken die het oorspronkelijk ten doel gesteld waren. De nu volgende paragrafen bevatten de samenvatting en conclusies van de verschillende stappen uit het onderzoek, en de wetenschappelijke artikelen die het heeft opgeleverd.

## 9.2 Tijdreeksmodellering met behulp van continue responsfuncties

De gangbare methoden voor tijdreeksanalyse (de discrete ARMA (AutoRegressive-Moving Average) transferruis-modellen van [*Box and Jenkins,* 1970] hebben, zoals besproken in paragraaf 2.2.3, verschillende beperkingen. In hoofdstuk 3 wordt een alternatieve methode gepresenteerd die geformuleerd is in een continu tijdsdomein. De basis van deze zogenaamde PIRFICT-methode is een convolutieintegraal. Binnen de PIRFICT-methodiek kan in principe elke analytische functie als impulsrespons gebruikt worden, zolang deze een bekende integraal of staprespons heeft. Een geschaalde gammaverdeling, met als fysisch analoog een cascade van lineaire reservoirs, blijkt echter in staat om de respons van de grondwaterstand op neerslag en verdamping in het algemeen nauwgezet te kunnen beschrijven. Dit blijkt uit de beschreven toepassing in hoofdstuk 3 (en uit de vele latere resultaten). De resultaten van het PIRFICT-model blijken equivalent te zijn met die van ARMA-modellen, wat geconcludeerd kan worden uit het feit dat de transferfuncties van beide modellen globaal dezelfde vorm aannemen, wanneer de modellen toegepast worden op dezelfde data. De transferfuncties van ARMA-modellen vertonen echter een onregelmatig patroon in het deel dat beïnvloed wordt door MA-parameters, terwijl de vorm van de gammaverdeling glad is. Ook de prestaties van beide modellen, in termen van de kalibratie- en validatiefout ofwel Root Mean Squared Error (RMSE), blijken in hoge mate vergelijkbaar. Bij het vergelijken van de resultaten van 15 stijghoogtereeksen kwam als enige verschil naar voren dat de kalibratiefout van het PIRFICT-model iets groter en de validatiefout juist iets kleiner was. De verschillen, hoewel klein, kunnen verklaard worden door het verschil in modelstructuur. Omdat de vorm van de ARMA-transferfunctie vrij is in het MA-deel, kan het deels ook vrijelijk afhangen van toevallige kruiscorrelaties tussen in- en uitvoer. Dergelijk 'overfittingsgedrag' zal resulteren in een kleinere kalibratie-RMSE, een grotere validatie-RMSE, en in een deels toevallig of random patroon van de transferfunctie, zoals gevonden is in de voorbeeldtoepassing. Omdat de continue responsfuncties van het PIRFICT-model glad zijn, en in het algemeen minder parameters hebben, is het model ook minder gevoelig voor overfitting. Hier zij echter gesteld dat de keuze van een specifieke responsfunctie ook beperkingen stelt aan de vorm ervan. Deze beperkingen kunnen de modelresultaten ook negatief beïnvloeden, wanneer de gekozen functie niet goed overeenkomt met de 'werkelijke' respons. De PIRFICT-methode heeft al met al de volgende, attractieve eigenschappen:

- De modelparameters zijn in de PIRFICT-benadering in principe constanten en onafhankelijk van de datafrequentie. Dit is belangrijk wanneer beoogd wordt de niet-tijdsafhankelijke responskarakteristieken van systemen vast te stellen, zoals in de originele doelstellingen van dit proefschrift. De resultaten van ARMA-modellen blijken in hoofdstuk 3, zoals verwacht, wel afhankelijk van de modelorde en datafrequentie (wat vanuit de continue tijdsgedachte echter wel 'te repareren' zou zijn).
- Het PIRFICT-model kan, vanwege het feit dat de modelvergelijkingen continu zijn in de tijd, voorspellingen doen en overweg met data met een willekeurige frequentie. De frequentie van modelinput en output zijn daarbij niet

gekoppeld, en kunnen bovendien onregelmatig zijn. De PIRFICT-methode is in vergelijking met de gecombineerde Kalman-filter- en ARX-modelbenadering [*Bierkens et al.,* 1999] bovendien meer generiek. In principe kan elke analytische responsfunctie gebruikt worden binnen het model, niet alleen de exponentiële functie van het ARX-model. Dit is, zoals aan de orde is gekomen in paragraaf 2.3.5, in veel situaties van belang, bijvoorbeeld wanneer rekening gehouden moet worden met de relatieve positie binnen een systeem, met de effecten van de onverzadigde zone, en met andere factoren dan neerslag en verdamping.

- Het gebruik van de PIRFICT-benadering vereenvoudigt in hoge mate het probleem van de modelidentificatie. De continue formulering bevordert allereerst het fysisch-hydrologische inzicht in de modelstructuur en – resultaten. Het fysische functioneren van een specifiek systeem, inclusief de fysisch gebaseerde parameters, kan in principe direct geïmplementeerd worden in het model, in de vorm van een analytische oplossing ervan. De modelorde van ARMA-modellen omvat, ter vergelijking, tenminste zes getallen (wanneer ook de frequentie in beschouwing wordt genomen). Het definiëren ervan is een complexe en ten dele subjectieve aangelegenheid. De modelorde dient min of meer handmatig geïdentificeerd te worden door de modelleur, of kort gezegd, de zes getallen dienen 'handmatig geoptimaliseerd' te worden. Responsfuncties als de geschaalde gammaverdeling zijn echter breed toepasbaar, zeer flexibel van aard zijn, en ze omvatten een reeks van ARMA-modelordes. Bij gebruik daarvan komt modelidentificatie doorgaans neer op het identificeren van de factoren die een systeem beïnvloeden. Hierna kan hetzelfde model toegepast worden op alle peilbuizen uit een gebied, ongeacht de geohydrologische situatie en locatie. Deze eigenschap maakt, samen met het punt hierboven, dat de PIRFICT-methode in bijzonder geschikt is voor gestandaardiseerde en geautomatiseerde analyse van grote aantallen tijdreeksen.

## 9.3 Een continu ruismodel voor reeksen met autocorrelatie en een onregelmatige frequentie

Reeksen van de grondwaterstand (en daarmee ook de fouten van modellen die erop toegepast worden) delen de eigenschap dat de verschillende waarnemingen in de reeks autocorrelatie vertonen. Autocorrelatie betekent kortweg dat de waarde van een waarneming op een bepaald tijdstip samenhang vertoont met de waarde op vorige tijdstippen. Modelfouten die behept zijn met autocorrelatie kunnen niet eenvoudigweg opgevat worden als onafhankelijke, normaal verdeelde afwijkingen. Een dergelijke praktijk zou inefficiënte parameter schattingen opleveren, en schattingen van de (co)variantie die in het algemeen te laag uitvallen (zie ook paragraaf 2.2.2). In dit soort gevallen dienen de modelfouten gemodelleerd te worden op een manier die recht doet aan de autocorrelatie, bijvoorbeeld door het gebruik van een apart, zogenaamd ruismodel.

In hoofdstuk 4 wordt een ruismodel gepresenteerd dat geformuleerd is in continue tijd, en dat overweg kan met data met een onregelmatige frequentie. De basis ervan

is het zogenaamde Ornstein-Uhlenbeckproces, en het wordt kortweg aangeduid met OUB (Ornstein-Uhlenbeck based)-model. Het OUB-model past vanwege de formulering goed bij het eerder beschreven PIRFICT-model, maar de bruikbaarheid ervan is daartoe niet beperkt. Het kan gebruikt worden voor modelfouten met een onregelmatige frequentie van allerlei deterministische modellen. Om inzicht te verkrijgen in haar functioneren wordt het OUB-model vergeleken met het AR(1)-model, in conventionele zin en ingebed in een Kalman-filter. Het Kalman-filter wordt daarbij behandeld in haar 'gedegenereerde' vorm [*Ahsan and O'Connor,* 1994], ofwel in een scenario waarbij het slechts voor het voorspellen van meetwaarden gebruikt wordt. Dit betekent *in concreto* dat de variantie van de fout in de metingen zelf geacht wordt nul te zijn, dan wel verwaarloosbaar klein in vergelijking met andere foutenbronnen. Hoofdstuk 4 laat zien dat a} de vergelijkingen voor het doen van de voorspellingen in de zogenaamde tijdsupdate van het 'gedegenereerde' Kalman-filter, b) die van de voorspellingsmodus van het AR(1)-model, en c) de vergelijkingen waarmee de innovaties in het OUB-model berekend worden, wiskundig equivalent zijn. De voorspellingen van het OUB-model zijn echter exact. Ze zijn bovendien rekenkundig efficiënter, omdat ze direct en niet recursief geëvalueerd worden. Vanwege haar eenvoud is het OUB-model bovendien eenvoudig implementeerbaar. Het model is echter beperkt tot processen die een min of meer exponentiële demping vertonen.

Voor wat betreft het optimalisatieproces, ook dit is vereenvoudigd door een gewogen kleinste kwadraten criterium af te leiden uit de zogenaamde Likelihood-functie die normaliter gebruikt wordt. Dit criterium is het 'Sum of Weighted Squared Innovations (SWSI) criterion' gedoopt. Parameterschattingen op basis van het SWSI-criterium convergeren naar zogenaamde 'Maximum Likelihood'-schattingen voor grotere aantallen metingen. Een voordeel van het SWSI-criterium is dat het geoptimaliseerd kan worden met standaard niet-lineaire regressiemethoden, zoals de methode van Levenberg-Marquardt, die rekenkundig erg efficiënt zijn. Naast het OUB-model kan ook het SWSI-criterium gebruikt worden in combinatie met andere modellen, bijvoorbeeld voor de innovaties van een Kalman-filter, wanneer daarbij exponentieel verval wordt aangenomen.

In dit hoofdstuk wordt tenslotte ook de zogenaamde innovatievariantie-functie geïntroduceerd, als hulpmiddel voor het diagnostisch toetsen van de validiteit van de responsfunctie van een ruismodel. Deze toets kan uitgevoerd worden door de geschatte innovatievarianties voor verschillende tijdstappen te vergelijken met de theoretische innovatievariantie-functie en haar betrouwbaarheidsinterval. De innovatievariantie-functie levert additionele informatie op over de validiteit van een model, in het bijzonder met oog op haar toepassing op tijdreeksen met een onregelmatige frequentie.

## 9.4 Het modelleren van stijghoogtereeksen die beïnvloedt worden door meerdere factoren

Bij ARMA-transferruismodellen is de standaard procedure voor het modelleren van tijdreeksen die worden beïnvloed door meerdere factoren, eenvoudigweg het parallel

verbinden ervan in het model. Anders gezegd, elke individuele factor krijgt hierbij zijn eigen, onafhankelijke transferfunctie toegewezen. De effecten van alle factoren worden gesommeerd, wat aldus de totale (deterministische) voorspelling oplevert. Alhoewel in de PIRFICT-benadering dezelfde procedure gevolgd zou kunnen worden, verdient het onderwerp meer aandacht dan dat, omdat:

- De geschaalde gamma-verdeling (zie paragraaf 2.4.1) is in het algemeen goed bruikbaar als respons op neerslag en verdamping, maar niet noodzakelijkerwijs voor andere typen factoren. In hoofdstuk 5 worden impulsresponsfuncties voor andere factoren afgeleid uit de analytische oplossing van het toepasselijke, elementair geohydrologische probleem.
- De respons op verschillende typen factoren kan een functie zijn van dezelfde fysische parameters. In zulke gevallen kan het aantal parameters en vrijheidsgraden in het model gereduceerd worden door hiermee rekening te houden. Een eenvoudig voorbeeld is de respons op verdamping, die gemodelleerd kan worden als omgekeerde van de respons op neerslag, afgezien van een reductiefactor.
- In de tijdreeksanalyse literatuur ligt de focus op het analyseren van individuele reeksen, op grond van statistische methoden en criteria. In geval van grondwaterstanden kan echter een fysisch-hydrologisch perspectief, in combinatie met de analyse van meerdere reeksen tegelijk, erg behulpzaam zijn. Het is zelfs noodzakelijk, wanneer de resultaten van individuele reeksen geen klaar antwoord bieden.

De PIRFICT-benadering werpt al met al een ander licht op de modelleerprocedures en de interpretatie van resultaten. De interpretatie van resultaten is bovendien ingewikkelder in geval van meerdere invloedsfactoren dan bij een enkele factor, omdat ook het model ingewikkelder is. Fysisch-hydrologische kennis kan, zoals gezegd, erg behulpzaam zijn bij het toetsen van de consistentie en de plausibiliteit van resultaten, en vormt een welkome aanvulling op de statistische, diagnostische toetsen. Allereerst dienen de geschatte parameterwaarden logischerwijze binnen een fysisch plausibel bereik te liggen. Wanneer er bovendien metingen zijn verricht op meerdere locaties, kunnen ruimtelijke patronen belangrijke en onafhankelijke 'feedback' geven op de resultaten. Deze patronen zijn onafhankelijk, omdat er geen ruimtelijke afhankelijkheid vooraf is opgelegd aan de individuele modellen. Via de modelfouten of residuen en de patronen daarin kunnen missende invloeden, maatregelen of andere foutenbronnen eenvoudig opgespoord worden. Bij een slechte fit kunnen de parameterschattingen afwijkingen vertonen, omdat de invloeden die wel meegenomen zijn deels kunnen compenseren voor de ontbrekende factoren, wanneer deze correlaties vertonen. Een factor wiens geschatte effect vaker zulke afwijkingen vertoont is verdamping. Verdamping vertoont een seizoensafhankelijk patroon dat snel correleert met andere seizoensafhankelijke factoren. Een voorbeeld hiervan is te vinden in hoofdstuk 5, waar de ruimtelijke verdeling van de geschatte verdampingsfactor duidelijk wijst op een dergelijke afwijking. In dit specifieke geval was de belangrijkste foutenbron in het model het feit dat er geen rekening werd gehouden met het gedrag van de individuele onttrekkingsputten, terwijl de totale

189

onttrekkingshoeveelheid een seizoenspatroon laat zien. Een dergelijk probleem kan verholpen worden door de ontbrekende informatie mee te nemen in het model, in dit geval de onttrekkingen van de individuele putten. Wanneer die informatie niet voorhanden is (zoals in dit geval), kunnen de afwijkingen beperkt worden door het optimalisatieprobleem in te perken tot een realistisch bereik, gebaseerd op de waarden en patronen in de omliggende waarnemingsputten.

Naast deze methodologische kwesties dient hoofdstuk 5 ook als voorbeeld van het feit dat een tijdreeksmodel in staat is om stijghoogtereeksen effectief te ontbinden in deelreeksen, die de invloed tonen van de afzonderlijke factoren. Dit is wellicht een van de belangrijkste praktische toepassingen van een tijdreeksmodel. Op die manier kan, onder bepaalde voorwaarden, het effect van afzonderlijke hydrologische maatregelen of factoren als grondwateronttrekkingen in beeld gebracht en gekwantificeerd worden. Het model levert bovendien een schatting op van het effect van veranderingen in de invloedsfactoren, in de vorm van de individuele responsfuncties. De resultaten kunnen hierdoor ook gebruikt worden voor scenarioberekeningen, het optimaliseren van de grondwaterstanden en/of het beheer daarvan.

## 9.5 Het karakteriseren van de grondwaterstandsdynamiek op basis van responseigenschappen

Zoals besproken is in paragraaf 1.1 had dit proefschrift oorspronkelijk ten doel het verbeteren van de gangbare methoden waarmee de relatie tussen grondwaterstand en vegetatie in ecohydrologische modellen gemodelleerd wordt. Maas [*Maas,* 1995] stelde het gebruik van momenten van de impulsresponsfunctie voor, als alternatief voor het gebruik van statistieken die de grondwaterspiegeldynamiek zelf beschrijven. Zulke beschrijvende statistieken worden beïnvloedt door meteorologische variatie, én ze kenmerken bovendien slechts een beperkt aantal facetten van de gemeten dynamiek. Het zijn, welbeschouwd, echter niet de momenten, maar het is de grondwaterstand zelf die de groeiomstandigheden in natte ecosystemen bepaalt (samen met andere abiotische en biotische factoren, en meer of minder direct). Momenten zijn daarbij niet altijd, of niet zonder meer, bruikbaar voor ecohydrologische modellering, om de volgende redenen:

- Momenten zijn constant, terwijl ecosystemen dat niet zijn. Alhoewel dit wellicht een aantrekkelijke eigenschap is voor relatief stabiele ecosystemen, compliceert het het gebruik van momenten voor het modeleren van de ecosysteemdynamiek zelf, e.g., in respons op klimaatdynamiek en -verandering.
- In dichtbevolkte gebieden zoals Nederland wordt de (grond)waterstand gewoonlijk actief beheerd of anderszins beïnvloedt door menselijke activiteiten. In zulke gevallen zijn niet één, maar verschillende responsfuncties van belang, en dit soort activiteiten zijn bovendien ook zelf variabel in de ruimte.
- Natte ecosystemen kennen per definitie hoge (grond)waterstanden, en gedragen zich daardoor vaak sterk niet-lineair. De

grondwaterspiegeldynamiek kan ook bij niet-lineariteit niet volledig gekarakteriseerd worden door middel van slechts één verzameling momenten.

Om de hierboven beschreven problematiek te verhelderen worden in hoofdstuk 6 verschillende methoden voor het karakteriseren van de grondwaterspiegeldynamiek behandeld. In het hoofdstuk wordt bovendien een relatie gelegd tussen zulke grondwaterstandskarakteristieken en momenten van impulsresponsfuncties. Als uitgangspunt is hierbij de geschaalde gamma-verdeling en PIRFICT-benadering van tijdreeksanalyse gekozen. Om tot alternatieve oplossingen te komen is een analyse gemaakt van zowel de grondwaterspiegeldynamiek als de impulsresponsfunctie in het frequentiedomein. Hieruit is een methode ontwikkeld die de temporele dynamiek van de invloedsfactoren (gekarakteriseerd door haar gemiddelde en jaarlijkse amplitude) combineert met de ruimtelijk variabele impulsresponsmomenten. Deze methode resulteert in een nieuwe set van parameters die de uitvoer of grondwaterspiegel-dynamiek zelf karakteriseren, en maakt gebruik van eenvoudige analytische formules. Deze zogenaamde GD-karakteristieken zijn de gemiddelde diepte, opbolling, jaarlijkse amplitude en faseverschuiving van de grondwaterstand. Het hoofdstuk laat bovendien zien dat deze vier karakteristieken samen de grondwaterspiegeldynamiek in detail karakteriseren (in dezelfde mate als het tijdreeksmodel). De beschreven relatie tussen invloed, momenten en GD-karakteristiek kan ook in omgekeerde zin gebruikt worden, bijvoorbeeld om de aanvulling van een aquifer in te schatten, analoog aan wat eerder voorgesteld is door [*Chen et al.,* 2002].

In de voorbeeldtoepassing in hoofdstuk 6 worden de GD-karakteristieken vergeleken met andere methoden voor het karakteriseren van het grondwaterregime, voor twee reeksen met sterk afwijkende eigenschappen. De vergelijking laat zien dat de veelgebruikte, zogenoemde GxG-waarden (d.w.z. de Gemiddeld Hoogste, Laagste en VoorjaarsGrondwaterstand) belangrijke nadelen hebben, en niet goed onderscheid maken tussen de verschillende regimes. Bij de berekening van GxG-waarden worden verschillende facetten van de fluctuaties door elkaar gehaald, aangezien de drie hoogste (HG3) en laagste (LG3) standen in een jaar een resultante zijn van zowel de jaarlijkse als hogere frequenties in het signaal. Door het middelen van de xG3-waarden over langere periodes (minimaal 8 jaar) worden bovendien de lagere frequenties en langjarige extremen effectief weggefilterd, en dus buiten beschouwing gelaten. Doordat de drie extreme waarden in een jaar gemiddeld worden tot xG3, en doordat een waarnemingsinterval van 14 dagen voorgeschreven is, reflecteren de GxG-waarden bovendien de jaarlijkse optredende extremen niet goed. Zowel de jaarlijkse als langjarige extremen zijn echter van groot belang vanwege de disproportionele ecologische invloed en economische schade die er uit ontstaan. Al met al kan geconcludeerd worden dat GxG-waarden minder goed in staat zijn om de grondwaterspiegeldynamiek op verschillende locaties te karakteriseren dan GD-karakteristieken. De methodiek waarmee GD-karakteristieken worden verkregen kan eenvoudig worden uitgebreid naar situaties waar meerdere factoren invloed hebben, zoals grondwateronttrekkingen of oppervlaktewaterstanden. Voor lineaire systemen kan het stationaire niveau en de jaarlijkse amplitude van het effect van alle invloedsfactoren gesommeerd worden. Bij factoren met een sterk afwijkend dynamisch gedrag voldoet deze benadering echter niet volledig, en zijn wellicht

aanvullende karakteristieken nodig. Een bijzondere factor die van invloed is op de grondwaterstand is de modelfout, c.q. de ruis die het systeem beïnvloedt. Omdat de verwachtingswaarde van witte ruis nul is, en het frequentiespectrum vlak, kan echter zowel de stationaire invloed als de specifiek jaarlijkse amplitude ervan verwaarloosd worden.

GD-karakteristieken vormen al met al een compacte methode voor het beschrijven van de grondwaterspiegeldynamiek op een bepaalde locatie. Ze kunnen niet alleen toegepast worden in het kader van ecohydrologische modellering, maar in alle gevallen waar de hoeveelheid informatie in een volledige beschrijving van grondwaterspiegeldynamiek in ruimte en tijd eenvoudig te groot is. Dit is bijvoorbeeld het geval bij karteringsdoeleinden, visuele interpretatie van de gegevens, of empirische modellen. Een bijkomend voordeel is dat de effecten van veranderingen in de invloedsfactoren op de GD-karakteristieken, bijvoorbeeld verschillende klimaatscenario's, eenvoudig berekend kunnen worden met behulp van de gegeven analytische formules. Deze eigenschappen geven de methode de flexibiliteit die nodig is om om te kunnen gaan met situaties waarbij de grondwaterspiegeldynamiek niet-stationair is. Klimaatverandering kan een oorzaak zijn van dergelijke niet-stationariteit, maar ook menselijke ingrepen zoals grondwateronttrekkingen. In dit soort gevallen zullen karakteristieken van de metingen van de afgelopen 8 of 30 jaar de huidige of toekomstige toestand niet goed weergeven.

## 9.6    *Menyanthes*: software voor grondwaterstandsgegevens

Dit hoofdstuk vormt in twee opzichten een synthese van de inhoud van de rest van dit proefschrift. Het hoofdstuk geeft allereerst een kort overzicht van de ontwikkelde methoden en behaalde resultaten, in relatie tot alternatieve methoden. Terwijl veel publicaties de tijdreeksanalysemethodiek primair vanuit een statistisch oogpunt benaderen, begint de tekst hier vanuit de fysica, in de hoop het materiaal voor een breder publiek toegankelijk te maken. Het hoofdstuk richt zich bovendien specifiek op de combinatie van data- en fysisch-gebaseerde methoden. Die combinatie is niet alleen karakteristiek voor de PIRFICT-methode, de integratie ervan is ook een van de oorspronkelijke doelstellingen van het onderzoek. In dit hoofdstuk wordt, ten tweede, het programma *Menyanthes* gepresenteerd. *Menyanthes* is ontwikkeld in het kader van dit promotieonderzoek, en de meeste functionaliteiten die voor het onderzoek nodig waren of er uit voort komen zijn er in opgenomen. In dat opzicht vormt *Menyanthes* ook een synthese van de inhoud van het proefschrift, omdat de ontwikkelde methoden en behaalde resultaten in het programma samenkomen.

Het programma *Menyanthes* neemt tussen het (toch al beperkte) aantal gebruiksvriendelijke programma's voor tijdreeksanalyse dat verkrijgbaar is, een speciale positie in. *Menyanthes* integreert allereerst verschillende functies. De functionaliteit begint bij de (data)basis van de DIKW piramide van [*Ackoff,* 1989], wordt aangevuld met visualisatie en analysefuncties daarbovenop, en eindigt met tijdreeksmodellerings-functies. De kracht van deze combinatie van functies is groter dan die van de som der losse onderdelen. De visualisatiefuncties en directe link met de data kunnen bijvoorbeeld waardevol zijn gedurende het tijdreeksmodelleringsproces.

Tijdreeksmodellen kunnen *vice versa* ingezet worden om de visualisatie en karakterisatie van de dynamiek van hydrologische systemen te verbeteren. Ze kunnen zelfs een bruikbaar hulpmiddel zijn voor databeheer, bijvoorbeeld voor het detecteren van fouten en *outliers*. *Menyanthes* heeft, ten tweede, een sterk fysisch-hydrologische inslag en vergemakkelijkt de toepassing van fysische principes en fysisch-hydrologische interpretatie. In dat opzicht heeft de PIRFICT-methode een parallel met de *data-based mechanistic modeling* (DBM) methode voor tijdreeksanalyse [*Young and Beven,* 1994*; Young,* 1998]. Een verschil met die methodiek is dat fysische principes of gedrag dat is geïnspireerd op de fysica *a priori* in een PIRFICT-model kunnen worden opgenomen, naast het gebruik van fysisch inzicht *a posteriori* bij het controleren en selecteren van resultaten zoals in de DBM-methode. Bij het gebruik van fysisch inzicht dient echter wel enige terughoudendheid betracht te worden. De keuze van een bepaalde continue responsfuncties in de PIRFICT-methode komt in principe neer op het doen van een aanname. De toepasselijkheid van die functie kan binnen *Menyanthes* gecontroleerd worden door de resultaten ervan te vergelijken met die van ARMA modellen. Ook de vraag of modelresultaten al dan niet fysisch-plausibel zijn dient met zelf-kritisch vermogen beantwoord te worden. De resultaten van tijdreeks- (en fysisch gebaseerde) modellen zijn, aan de ene kant, mogelijk beïnvloedt door niet-causale kruiscorrelaties. Onverwachte resultaten zijn, aan de andere kant, niet noodzakelijkerwijs onwaarschijnlijk, maar kunnen ook tot nieuwe inzichten leiden.

De mogelijke toepassingen van de PIRFICT-methode zijn in essentie gelijk aan (en even breed als) die van ARMA-tijdreeksmodellen. Een aantrekkelijke eigenschap van tijdreeksmodellen is dat er betrekkelijk weinig aannamen aan ten grondslag liggen, en dat de nauwkeurigheid ervan in het algemeen groot is. Tijdreeksmodellen zijn als zodanig een waardevol hulpmiddel voor het voorbewerken van grondwaterstands-reeksen, voorafgaand aan het kalibreren van een ruimtelijk model. Ontbrekende invloedsfactoren en reeksen die beïnvloedt worden door hydrologische ingrepen kunnen met behulp van tijdreeksmodellen eenvoudig geïdentificeerd worden. Daarnaast kan tijdreeksanalyse behulpzaam zijn om reeksen te identificeren die ongeschikt zijn voor modelkalibratie, bijvoorbeeld omdat ze een hydrologisch object of fenomeen vertegenwoordigen dat niet in het model opgenomen is, zoals een schijnspiegelsysteem. Nog een stap verder gaat het zogenaamde 'matchen' van de momenten van tijdreeksmodellen en ruimtelijk gedistribueerde, fysisch gebaseerde modellen [*Von Asmuth and Maas*, 2001; *Bakker et al.*, 2007; *Bakker et al.*, 2008]. Een eigenschap die specifiek is voor de PIRFICT-methode is de grote efficiëntie en het gebruiksgemak. De constatering dat scheef-normale verdelingen het gedrag van grondwater (en andere) systemen in het algemeen goed kunnen beschrijven, maakt het mogelijk om alle tijdreeksen in een gebied groepsgewijs te modelleren. Dit is niet 'alleen' een praktisch voordeel. De gelijktijdige visualisatie en analyse van de resultaten op meerdere locaties voegt letterlijk een dimensie toe aan de techniek van tijdreeksanalyse (zijnde de ruimtelijke dimensie). Terwijl de ruimtelijke structuur die is opgenomen in fysisch gebaseerde modellen gebaseerd is op afleidingen en aannamen die niet werkelijk geverifieerd kunnen worden (zoals betoogd in de inleiding van hoofdstuk 9), is er geen ruimtelijke afhankelijkheid opgelegd aan de tijdreeksmodellen. De ruimtelijke patronen die verschijnen in tijdreeksanalyseresultaten zijn daarom onafhankelijk en vrij van aannamen (afgezien van het feit dat de representativiteit van

193

invoerreeksen kan varieren in de ruimte), en leveren waardevolle informatie over zowel de modelresultaten zelf als de eigenschappen en het functioneren van systemen. De effecten van individuele invloedsfactoren die niet gescheiden kunnen worden op basis van tijdreeksanalyse van de afzonderlijke reeksen alleen, kunnen daarbij wellicht wel geschat en gescheiden worden op basis van de waargenomen ruimtelijke patronen. De ruimtelijke visualisatie- en analysefuncties van *Menyanthes* vervullen in dat opzicht een belangrijke rol, en ondersteunen het gebruik en de interpretatie van dergelijke patronen. De methoden en functionaliteiten van *Menyanthes* brengen aldus een verschuiving van aandachtspunt teweeg van data- en statistische aspecten van tijdreeksen naar het gedrag en de fysische eigenschappen van systemen, en de ruimtelijke variatie daarvan. De term 'system identification', zoals gesuggereerd door [*Zadeh,* 1956] en overgenomen door [e.g., *Ljung,* 1999], wint hiermee aan toepasselijkheid, en is te prefereren boven de term 'tijdreeksanalyse', die zijn oorsprong vindt in wetenschappen als de statistiek en econometrie.

*Menyanthes* wordt op het moment van schrijven gebruikt op drie continenten, in zes landen, door 76 verschillende organisaties. Een paar honderd ecologen en hydrologen hebben een cursus gevolgd omtrent de analyse van geohydrologische tijdreeksen en het gebruik van *Menyanthes* daarbij [*Von Asmuth et al.,* 2006]. Enige tienduizenden grondwaterstandsreeksen zijn inmiddels met het programma geanalyseerd. Het programma en de daarin opgenomen methoden blijken aldus een behoefte te vervullen die niet eenvoudig op een andere manier te vervullen is. Zoals gezegd is het feit dat traditionele tijdreeksmodellen lineair zijn een kracht, maar ook een beperking. Niet-lineaire methoden voor tijdreeksanalyse die uitgaan van de drempelbenadering [*Knotters and De Gooijer*, 1999] zijn opgenomen in *Menyanthes*, en blijken van waarde te zijn. Initiatieven die gericht zijn op het testen van de methodiek op een breder bereik van systemen, of op het verbreden van de functionaliteit met andere (niet-lineaire) processen worden door de auteurs verwelkomd.

# Appendix A

# Derivations

# Derivations

## A.1 Auto- and crosscorrelation functions of the residuals and innovations

The autocorrelation of the innovations equals (with t''< t'< t):

$$E\{v(t)v(t')\} = E\{\int_{t'}^{t}\sqrt{\frac{2\alpha\sigma_n^2}{\beta}}\,\mathrm{e}^{-\alpha(t-\tau)}\,\mathrm{dW}(\tau)\int_{t''}^{t'}\sqrt{\frac{2\alpha\sigma_n^2}{\beta}}\,\mathrm{e}^{-\alpha(t-\tau)}\,\mathrm{dW}(\tau)\}=0 \tag{A.1}$$

due to the properties of the Wiener process ( $E\{\mathrm{dW}(t)\mathrm{dW}(t')\}=0$ if $t \neq t'$ ). Using (11), the autocorrelation of the residuals can be written as (for t'< t):

$$E\{\tilde{n}(t)\tilde{n}(t')\} = E(\mathrm{e}^{-\alpha(t-t')})\tilde{n}(t')\tilde{n}(t')+\tilde{n}(t')\int_{t-t'}^{t}\sqrt{\frac{2\alpha\sigma_n^2}{\beta}}\,\mathrm{e}^{-\alpha(t-\tau)}\}\mathrm{dW}(\tau)\} \tag{A.2}$$

which gives:

$$E\{\tilde{n}(t)\tilde{n}(t')\} = \mathrm{e}^{-a(t-t')}\,\sigma_n^2 \tag{A.3}$$

Using (12), the crosscorrelation between residuals and innovations equals, for any t'< t:

$$E\{v(t)\tilde{n}(t')\} = E\{\tilde{n}(t)\tilde{n}(t') - \mathrm{e}^{-\alpha\Delta t}\,\tilde{n}(t-\Delta t)\tilde{n}(t')\} \tag{A.4}$$

which gives, using (A.3):

$$E\{v(t)\tilde{n}(t')\} = \mathrm{e}^{-a(t-t')}\,\sigma_n^2 - \mathrm{e}^{-\alpha\Delta t}\,\mathrm{e}^{-a(t-\Delta t-t')}\,\sigma_n^2 = 0 \tag{A.5}$$

## A.2 Relation between residual variance and individual innovations

Starting with (4.18), we can write the innovation variance as an expected value and get:

$$\sigma_n^2(\Psi) = (\frac{1}{1-\mathrm{e}^{-2\alpha\Delta t_i}})\,\mathrm{E}\{v^2(t_i,\Psi)\} \tag{A.6}$$

using every $v(t_i)$ individually, $N$ single sample estimates of $\hat{\sigma}_n^2(t_i, \Psi)$ can be obtained with:

$$\hat{\sigma}_{n,t_i}^2(\Psi) = (\frac{1}{1 - e^{-2\alpha\Delta t_i}})v^2(t_i, \Psi) \tag{A.7}$$

next, we can get a more accurate estimate if we average the $N$ estimates of $\hat{\sigma}_n^2(t_i, \Psi)$ :

$$\hat{\sigma}_n^2(\Psi) = \frac{\sum_{i=1}^{N}(\frac{1}{1 - e^{-2\alpha\Delta t_i}})v^2(t_i, \Psi)}{N} \tag{A.8}$$

## A.3    SWSI (Sum of Weighted Squared Innovations) criterion

Given the following log likelihood function:

$$\Lambda\{\Psi \mid O\} = -0.5N \ln(2\pi) - 0.5\sum_{i=1}^{N} \ln\{\sigma_v^2(\Delta t_i, \Psi)\} - 0.5\sum_{i=1}^{N} \frac{v^2(t_i, \Psi)}{\sigma_v^2(\Delta t_i, \Psi)} \tag{A.9}$$

we can replace $\sigma_v^2(\Delta t_i, \beta)$ by $\sigma_n^2(\beta)$ using (4.18) and get:

$$\Lambda\{\Psi \mid O\} = -0.5N \ln(2\pi) - 0.5\sum_{i=1}^{N} \ln\{(1 - e^{-2\alpha\Delta t_i})\sigma_n^2\} - 0.5\sum_{i=1}^{N} \frac{v^2(t_i, \Psi)}{(1 - e^{-2\alpha\Delta t_i})\sigma_n^2} \tag{A.10}$$

next, we get by placing $N / \sigma_n^2(\Psi)$ in the last term outside the summation sign:

$$\Lambda\{\Psi \mid O\} = -0.5N \ln(2\pi) - 0.5\sum_{i=1}^{N} \ln\{(1 - e^{-2\alpha\Delta t_i})\sigma_n^2\} - ...$$

$$0.5\frac{N}{\sigma_n^2}\frac{\sum_{i=1}^{N}(\frac{1}{1 - e^{-2\alpha\Delta t_i}})v^2(t_i, \Psi)}{N} \tag{A.11}$$

and replace in both terms $\sigma_n^2(\Psi)$ by (A.8) to get :

$$\Lambda\{\Psi \mid O\} = -0.5N \ln(2\pi) - 0.5\sum_{i=1}^{N} \ln\{(1 - e^{-2\alpha\Delta t_i})\frac{\sum_{j=1}^{N}\frac{1}{1 - e^{-2\alpha\Delta t_j}}v^2(t_j, \Psi)}{N}\} - 0.5N \tag{A.12}$$

In the last term several items have thus been eliminated. Because the sum of several logarithms is the logarithm of their product, (A.12) equals:

$$\Lambda\{\Psi\,|\,O\} = -0.5N\ln(2\pi) - 0.5\ln[\prod_{i=1}^{N}\{(1-e^{-2\alpha\Delta t_i})\frac{\sum_{j=1}^{N}\frac{1}{1-e^{-2\alpha\Delta t_j}}v^2(t_j,\Psi)}{N}\}]-0.5N \quad (A.13)$$

As it is constant for any set of the model parameters $\Psi$, the summation term can be placed outside the product sign, giving:

$$\Lambda\{\Psi\,|\,O\} = -0.5N\ln(2\pi) - 0.5\ln[\{\frac{\sum_{j=1}^{N}\frac{1}{1-e^{-2\alpha\Delta t_j}}v^2(t_j,\Psi)}{N}\}^N\prod_{i=1}^{N}\{(1-e^{-2\alpha\Delta t_i})\}]-...(A.14)$$

$$0.5N$$

which equals, by taking an $N^{\text{th}}$ power root to the power of $N$, and placing $N$ outside the logarithm

$$\Lambda\{\Psi\,|\,O\} = -0.5N\ln(2\pi) - 0.5N\ln[\frac{\sum_{j=1}^{N}\frac{1}{1-e^{-2\alpha\Delta t_j}}v^2(t_j,\Psi)}{N}\sqrt[N]{\prod_{i=1}^{N}(1-e^{-2\alpha\Delta t_i})}]-...(A.15)$$

$$0.5N$$

The $N^{\text{th}}$ power root of the product term is the geometrical mean which is also constant given $n(t_i)$ and can be placed inside the summation term:

$$\Lambda\{\Psi\,|\,O\} = -0.5N\ln(2\pi) - 0.5N\ln[\frac{\sum_{j=1}^{N}\frac{\sqrt[N]{\prod_{i=1}^{N}(1-e^{-2\alpha\Delta t_i})}}{1-e^{-2\alpha\Delta t_j}}v^2(t_j,\Psi)}{N}]-0.5N \quad (A.16)$$

As the first and last terms of the likelihood function are now constant, (A.16) can be maximized by minimizing a sum of weighted squared innovations:

$$S^2\{\Psi\,|\,O\} = \sum_{j=1}^{N}\frac{\sqrt[N]{\prod_{i=1}^{N}(1-e^{-2\alpha\Delta t_i})}}{1-e^{-2\alpha\Delta t_j}}v^2(t_j,\Psi) \quad (A.17)$$

## A.4 Amplitude and phase response of a scaled gamma distribution function

(derivation thanks to Kees Maas)

The amplitude and phase response of a scaled gamma distribution function (SG df, see section 2.4.1) can be found by convoluting it with a harmonic signal $f(t)$ with time $t$ and frequency $\xi$:

$$f(t) = e^{i\xi t} \tag{A.18}$$

which gives:

$$g(t) = \int_0^\infty e^{i\xi(t-\tau)} A \frac{a^n \tau^{n-1}}{\Gamma(n)} e^{-a\tau} \, d\tau = A \frac{e^{i\xi t}}{\Gamma(n)} \int_0^\infty (a\tau)^{n-1} e^{-a\tau(1+i\frac{\xi}{a})} \, da\tau \tag{A.19}$$

The gamma function $\Gamma(n)$ equals [*Abramowitz and Stegun,* 1964]:

$$\Gamma(n) = k^n \int_0^\infty t^{n-1} e^{-nt} \, dt \qquad (\Re n > 0, \Re k > 0) \tag{A.20}$$

Using (A.20), (A.19) can be written as:

$$g(t) = A \frac{e^{i\xi t}}{\Gamma(n)} \frac{\Gamma(n)}{(1+i\frac{\xi}{a})^n} = \frac{A}{(1+i\frac{\xi}{a})^n} e^{i\xi t} \tag{A.21}$$

Next, by replacing i in the first part of (A.21), we have:

$$g(t) = \frac{A}{(1+\frac{\xi^2}{a^2})^{\frac{n}{2}}} e^{i\xi(t-\frac{n}{\xi}\mathrm{atan}\frac{\xi}{a})} \tag{A.22}$$

Ergo, the amplitude response of a SG-type IR function to a harmonic signal (A.18) is:

$$\frac{A}{(1+\frac{\xi^2}{a^2})^{\frac{n}{2}}} \tag{A.23}$$

And its phase response is:

$$\frac{n}{\xi}\operatorname{atan}\frac{\xi}{a} \qquad\qquad (A.24)$$

## A.5    Impulse response of a system of linear reservoirs
(derivation thanks to Kees Maas)

In matrix notation, the vector of impulse reponses of a system of $n$ linear reservoirs is (see section 2.3.1):

$$\boldsymbol{\theta} = \frac{1}{c_1 A_0}\exp(-\mathbf{A}t)\begin{bmatrix}1\\0\end{bmatrix} \qquad\qquad (A.25)$$

where the exponential is a so-called matrix exponential. One of the possible definitions of a matrix function $f(\mathbf{A})$ is:

$$f(\mathbf{A}) = \mathbf{E}f(\boldsymbol{\Lambda})\mathbf{E}^{-1} \qquad\qquad (A.26)$$

where $\mathbf{E}$ is the matrix of eigenvectors of $\mathbf{A}$, and $\boldsymbol{\Lambda}$ the matrix of eigenvalues. If all reservoirs are identical, $\mathbf{A}$ is symmetric and $\mathbf{E}^{-1} = \mathbf{E}^{\mathrm{T}}$. Combining (A.26) and (A.25) in that case yields:

$$\begin{bmatrix}\theta_1\\\theta_2\\.\\\theta_n\end{bmatrix} = \frac{1}{c_1 A_0}\begin{bmatrix}e_{11}&e_{12}&.&e_{1n}\\e_{21}&e_{22}&.&e_{2n}\\.&.&.&.\\e_{n1}&e_{n2}&.&e_{nn}\end{bmatrix}\begin{bmatrix}e^{-\lambda_1 t}&&&\\&e^{-\lambda_2 t}&&\\&&.&\\&&&e^{-\lambda_n t}\end{bmatrix}\begin{bmatrix}e_{11}&e_{21}&.&e_{n1}\\e_{12}&e_{22}&.&e_{n2}\\.&.&.&.\\e_{1n}&e_{n2}&.&e_{nn}\end{bmatrix}\begin{bmatrix}1\\0\\.\\0\end{bmatrix} \qquad (A.27)$$

Because the last columnvector mostly contains zeros, this equals:

$$\begin{bmatrix}\theta_1\\\theta_2\\.\\\theta_n\end{bmatrix} = \frac{1}{c_1 A_0}\begin{bmatrix}e_{11}&e_{12}&.&e_{1n}\\e_{21}&e_{22}&.&e_{2n}\\.&.&.&.\\e_{n1}&e_{n2}&.&e_{nn}\end{bmatrix}\begin{bmatrix}e^{-\lambda_1 t}&&&\\&e^{-\lambda_2 t}&&\\&&.&\\&&&e^{-\lambda_n t}\end{bmatrix}\begin{bmatrix}e_{11}\\e_{12}\\.\\e_{1n}\end{bmatrix} \qquad (A.28)$$

or

$$\begin{bmatrix} \theta_1 \\ \theta_2 \\ . \\ \theta_n \end{bmatrix} = \frac{1}{c_1 A_0} \begin{bmatrix} e_{11} & e_{12} & . & e_{1n} \\ e_{21} & e_{22} & . & e_{2n} \\ . & . & . & . \\ e_{n1} & e_{n2} & . & e_{nn} \end{bmatrix} \begin{bmatrix} e_{11}\,\mathrm{e}^{-\lambda_1 t} \\ e_{12}\,\mathrm{e}^{-\lambda_2 t} \\ . \\ e_{1n}\,\mathrm{e}^{-\lambda_n t} \end{bmatrix}$$ 

(A.29)

so that:

$$\theta_i = \frac{1}{c_1 A_0} \begin{bmatrix} e_{i1} & e_{i2} & . & e_{in} \end{bmatrix} \begin{bmatrix} e_{11}\,\mathrm{e}^{-\lambda_1 t} \\ e_{12}\,\mathrm{e}^{-\lambda_2 t} \\ . \\ e_{1n}\,\mathrm{e}^{-\lambda_n t} \end{bmatrix}$$ 

(A.30)

or:

$$\theta_i = \frac{1}{c_1 A_0} \sum_{j=1}^{n} e_{ij} e_{1j}\,\mathrm{e}^{-\lambda_j t}$$ 

(A.31)

In case of non-identical reservoirs, we first use a similarity transform in order to get a symmetrical matrix, as:

$$\mathbf{A}^* = \mathbf{J}\mathbf{A}\mathbf{J}^{-1}$$ 

(A.32)

where $\mathbf{J}$ is:

$$\mathbf{J} = \begin{bmatrix} \sqrt{A_1} & & & \\ & \sqrt{A_2} & & \\ & & . & \\ & & & \sqrt{A_n} \end{bmatrix}$$ 

(A.33)

and $\mathbf{A}$ and $\mathbf{A}^*$ are similar. Because $\mathbf{A}^*$ is symmetric, (A.32) can be written as:

$$\mathbf{A}^* = \mathbf{J}\mathbf{A}\mathbf{J}^{-1} = \mathbf{J}\mathbf{E}\mathbf{\Lambda}\mathbf{E}^{-1}\mathbf{J}^{-1} = (\mathbf{J}\mathbf{E})\mathbf{\Lambda}(\mathbf{J}\mathbf{E})^{-1} = (\mathbf{J}\mathbf{E})\mathbf{\Lambda}(\mathbf{J}\mathbf{E})^{\mathrm{T}}$$ 

(A.34)

so that:

$$\mathbf{A} = \mathbf{J}^{-1}(\mathbf{J}\mathbf{E})\mathbf{\Lambda}(\mathbf{J}\mathbf{E})^{T}\mathbf{J}$$ 

(A.35)

and:

$$f\{\mathbf{A}\} = \mathbf{J}^{-1}(\mathbf{JE})f\{\mathbf{\Lambda}\}(\mathbf{JE})^{T}\mathbf{J} \qquad (A.36)$$

In our reservoir case, this yields:

$$\mathbf{\theta} = \frac{1}{c_1 A_0}\mathbf{J}^{-1}(\mathbf{JE})\exp\{-\mathbf{\Lambda}t\}(\mathbf{JE})^{T}\mathbf{J}\begin{bmatrix} 1 \\ 0 \\ . \\ 0 \end{bmatrix} \qquad (A.37)$$

which may be written as:

$$\theta_i = \frac{1}{c_1 A_0}\sum_{j=1}^{n}\alpha_{ij}\,e^{-\lambda_j t} \qquad (A.38)$$

In words, the solution is a weighted sum of exponentials, in which the weights $\alpha$ depend on the eigenvectors of $\mathbf{A}$.

GROUNDWATER LEVEL FLUCTUATIONS — J.C. Gehrels — 1999

SCHWARTZ / ZHANG — F U N D A M E N T A L S   O F   GROUND WATER

Ziemer / Tranter / Fannin — SIGNALS & SYSTEMS — Continuous and Discrete — Fourth Edition — PHIPE — PRENTICE HALL

LJUNG — SYSTEM IDENTIFICATION — THEORY FOR THE USER — second edition — PH PTR

ON CONVOLUTIONAL PROCESSES AND DISPERSIVE GROUNDWATER FLOW   C. MAAS

Box and Jenkins — TIME SERIES ANALYSIS forecasting and control — Revised Edition — Holden Day

Gardiner — Stochastic Methods — 4th Edition

VERRUIJT — THEORY OF GROUNDWATER FLOW — MACMILLAN

William A. Jury · Kurt Roth — Transfer Functions and Solute Movement through Soil

46 — DEVELOPMENTS IN WATER SCIENCE — BRUGGEMAN — ANALYTICAL SOLUTIONS OF GEOHYDROLOGICAL PROBLEMS

BEAR — DYNAMICS OF FLUIDS IN POROUS MEDIA — ELSEVIER

DEVELOPMENTS IN WATER SCIENCE — HIPEL / McLEOD — TIME SERIES MODELLING OF WATER RESOURCES

**Appendix** B

# References

# References

Abramowitz, M., and I. A. Stegun (1964), *Handbook of Mathematical Functions,* Dover Publications Inc., New York.

Ackoff, R. (1989), From Data to Wisdom, *Journal Of Applied Systems Analysis*, 16(1), 3-9.

Ahsan, M., and K. M. O'Connor (1994), A reappraisal of the Kalman filtering technique, as applied in river flow forecasting, *Journal of Hydrology*, 161, 197-226.

Akaike, H. (1970), Statistical predictor identification, *Ann. Inst. Statist. Math.*, 22, 203-217.

Akaike, H. (1974), A new look at the statistical model identification, *IEEE Trans. Autom. Control*, 19, 716-723.

Akaike, H. (1979), A Bayesian extension of the AIC procedure, *Biometrika*, 66, 237-242.

Amorocho, J., and W. E. Hart (1964), A critique of current methods in hydrologic systems investigation, *Trans. Am. Geophys. Un.* 45, 307-321.

Backlund, A. (2000), The definition of system, *Kybernetes*, 29(4), 444-451.

Baker, V. R., R. C. Kochel, and P. C. Patton (1988), *Flood geomorphology,* John Wiley, New York.

Bakker, M., K. Maas, F. Schaars, and J. R. Von Asmuth (2007), Analytic modeling of groundwater dynamics with an approximate impulse response function for areal recharge, *Advances in Water Resources*, 30(3, doi:10.1016/j.advwatres.2006.04.008), 493-504.

Bakker, M., K. Maas, and J. R. Von Asmuth (2008), Calibration of transient groundwater models using time series analysis and moment matching, *Water Resources Research*, 44(W04420), doi:10.1029/2007WR006239.

Barndorff-Nielsen, O. E. (1999), Superposition of Ornstein-Uhlenbeck type processes, *Theory probab. appl.*, 45(2), 175-194.

Barndorff-Nielsen, O. E., and N. Shephard (2001), Non-Gaussian Ornstein-Uhlenbeck-based models and some of their uses in financial economics, *Journal of the Royal Statistical Society (B)*, 63, 1-42.

Bartholomeus, R. P., J. P. M. Witte, P. M. van Bodegom, and R. Aerts (2008), The need of data harmonization to derive robust empirical relationships between soil conditions and vegetation, *Journal of Vegetation Science*, 19(doi: 10.3170/2008-8-18450), 799-808.

Bear, J. (1972), *Dynamics of fluids in porous media,* American Elsevier Publishing Company, New York.

Beck, M. B. (1987), Water quality modelling: a review of the analysis of uncertainty, *Water Resources Research*, 23(8), 1393–1442.

Berendrecht, W. L. (2004), *State space modeling of groundwater fluctuations,* Phd thesis, Delft University of Technology, Delft.

Berendrecht, W. L., A. W. Heemink, F. C. Van Geer, and J. C. Gehrels (2003), Decoupling of modeling and measuring interval in groundwater time series analysis based on response characteristics, *Journal of Hydrology*, 278, 1-16.

Berendrecht, W. L., A. W. Heemink, F. C. Van Geer, and J. C. Gehrels (2006), A non-linear state space approach to model groundwater fluctuations, *Advances in Water Resources*, 29, 959–973.

Bergstrom, A. R. (1990), *Continuous time econometric modeling,* Oxford University Press, New York.

Besbes, M., and G. de Marsily (1984), From infiltration to recharge: Use of a parametric transfer function, *Journal of Hydrology*, 74, 271-293.

Beven, K. J. (1993), Prophecy, reality and uncertainty in distributed hydrological modelling, *Advances in Water Resources*, 16, 41-51.

Beven, K. (2002), Towards an alternative blueprint for a physically based digitally simulated hydrologic response modelling system, *Hydrological Processes*, 16, 189-206, DOI: 10.1002/hyp.343.

Beven, K. J. (2006), A manifesto for the equifinality thesis, *Journal of Hydrology*, 320, 18-36.

Beven, K. J., and A. M. Binley (1992), The future of distributed models: model calibration and uncertainty prediction, *Hydrological Processes*, 6, 279-298.

Bidwell, V. J., P. F. Callander, and C. R. Moore (1991), An application of time-series analysis to groundwater investigation and management in Central Canterbury, New Zealand, *Journal of Hydrology*, 30(1), 16–36.

Bidwell, V. J. (2005), Realistic forecasting of groundwater level, based on the eigenstructure of aquifer dynamics, *Mathematics and Computers in Simulation*, 69, 12–20.

Bierkens, M. F. P., and D. J. J. Walvoort (1998), Simple stochastic models for groundwater level fluctuations. Part 2: Combined soil and groundwater model with stochastic input (in Dutch), *Stromingen*, 4(3), 5-20.

Bierkens, M. F. P., and W. A. Bron (2000), VIDENTE: a graphical user interface and decision support system for stochastic modelling of water table fluctuations at a single location: includes documentation of the programs KALMAX, KLATFN, SSD and EMERALD and introductions to stochastic modelling. Alterra report 118, Wageningen.

Bierkens, M. F. P., M. Knotters, and T. Hoogland (2000), Space-time modeling of water table depth using a regionalized time series model and the Kalman filter, *Water Resources Research*, 37(5), 1277-1290.

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.

Boucneau, G., M. Van Meirvenne, J. Desmet, and G. Hofman (1996), A methodology to evaluate the reliability of the Belgian soil map for predicting the actual water table characteristics, *Geoderma*, 69(3-4), 193-207.

Box, G. E. P., and G. M. Jenkins (1970), *Time Series Analysis: Forecasting and Control,* Holden-Day, San Fransisco.

Brockwell, P. J. (2001), Continuous-time ARMA processes, in *Handbook of Statistics, Vol. 19, Stochastic Processes: Theory and Methods*, edited by Shanbhag, D.N., and Rao, C.R., North-Holland, Amsterdam, 249-276.

Brubacher, S. R., and G. T. Wilson (1976), Interpolating time series with applications to the estimation of holiday effects on electricity demand, *Journal of the Royal Statistical Society, Part C, Applied Statistics*, 25(2), 107-116.

Bruggeman, G. A. (1999), Analytical Solutions of Geohydrological Problems, Elsevier, Amsterdam.

Bryson, A. E., and L. J. Henrikson (1965), Estimation using sampled data containing correlated noise, *J. Spacecr. Rockets*, 5, 662-665.

Chatfield, C. (1989), *The analysis of time series, an introduction,* Chapman and Hall, London.

Chen, Z., S. E. Grasby, and K. G. Osadetz (2002), Predicting average annual groundwater levels from climatic variables: an empirical model, *Journal of Hydrology*, 260, 102-117.

Crank, J. (1957), *The mathematics of diffusion,* Clarendon Press, Oxford.

Cryer, J. D., and K.-S. Chan (2008), *Time Series Analysis With Applications in R,* Springer, New York.

De Castro Ochoa, F., and J. C. Munoz Reinoso (1997), Model of long-term water-table

dynamics at Donana National Park, *Water Resources*, 31(10), 2586-2596.

De Haan, M. W. A. (1992), Characteristics of frequency of exceedence graphs of several groundwater dependent vegetation types (in Dutch), rapport nr. SWE 92.030, Kiwa, Nieuwegein.

De Gooijer, J. G., B. Abraham, A. Gould, and L. Robinson (1985), Methods for determining the order of an autoregressive-moving average process: A survey, *International Statistical Review*, 53(3), 301-329.

De Keizer, O. (2003), Estimation of the actual evapotranspiration in the PIRFICT time series model, estimation of the crop coefficient on the basis of groundwater level dynamics, Wageningen University, Wageningen.

De Zeeuw, J. W., and F. Hellinga (1958), Precipitation and Run-off (in Dutch), Landbouwkundig Tijdschrift, 70, 405-422.

Dirac, P. A. M. (1947), *The Principles of Quantum Mechanics,* Clarendon Press, Oxford.

Dooge, J. C. I. (1959), A General Theory of the Unit Hydrograph, *J. Geophys. Res.*, 64(2), 241–256, doi:10.1029/JZ064i002p00241.

Dooge, J. C. I. (1973), *Linear theory of hydrologic systems.,* Technical Bulletin 1468, Agricultural Research Service, U.S. Department of Agriculture, Washington D.C.

Duhamel, J. M. C. (1833), Memoire sur la methode generale relative au mouvement de la chaleur dans les corps solides plonges dans les milieux dont la temperature varie avec le temps, *Journal d' Ecole polytechnique de Paris*, 14(22), 20.

Dumm, L. D. (1954), Drain spacing formula, *Agr. Eng.*, 35, 726-730.

Dupuit, J. (1863), Etudes theoriques et pratiques sur le mouvement des eaux dans les canaux decouverts et a travers les terrains permeables, Dunod, Paris.

Ellenberg, H. (1991), Zeigerwerte der Gefässpflanzen (ohne Rubus), *Scripta Geobotanica*, 18, 9-166.

Fabbri, P., C. Gaetan, and P. Zangheri (2011), Transfer function-noise modelling of an aquifer system in NE Italy, *Hydrological Processes*, 25, 194-206, DOI: 10.1002/hyp.7832.

Feddes, R. A. (1981), Water use models for assessing root zone modification., in *Modifying the root environment to reduce crop stress.*, edited by Arkin, G.F., and Taylor, H.M., *ASAE Monograph*, St. Joseph, Michigan, 347-390.

Feddes, R. A., P. Kabat, P. J. T. Van Bakel, J. J. B. Bronswijk, and J. Halbertsma (1988), Modelling soil water dynamics in the unsaturated zone: state of the art., *Journal of Hydrology*, 100, 69-111.

Fernandez, C., J. Osiewalski, and M. F. J. Steel (1995), Modeling and inference with v-spherical distributions, *Journal of the American Statistical Association*, 90, 1331-1340.

Forchheimer, P. (1901), Wasserbewegung durch Boden, *Z. Ver. Deutsch. Ing.*, 45, 1782-1788.

Freeze, R. A., and R. L. Harlan (1969), Blueprint for a physically-based,digitally-simulated hydrologic response model, *Journal of Hydrology*, 9, 237-258.

Gardiner, C. W. (1994), *Handbook of stochastic methods,* Springer-Verlag, New York.

Gehrels, J. C., F. C. Van Geer, and J. J. De Vries (1994), Decomposition of groundwater level fluctuations using transfer modelling in an area with shallow to deep unsaturated zones, *Journal of Hydrology*, 157, 105-138.

Gehrels, J. C. (1999), Groundwater Level Fluctuations, separation of natural from anthropogenis influences and determination of groundwater recharge in the Veluwe area, the Netherlands, Doctoral Thesis Vrije Universiteit Amsterdam, Amsterdam.

Gelhar , L. W., A. Mantaglou, C. Welty, and K. R. Rehfeldt (1985), *A review of field scale physical solute transport process in saturated and unsaturated porous media,* Topical rep. EA-4190, EPRI  Electrical Power Research Institute , Palo Alto, California.

Govindaraju, R. S., and B. S. Das (2007), *Moment analysis for subsurface hydrologic applications,* Water Science and Technololgy Library, vol. 61. Springer, Dordrecht.

Grootjans, A. P. (1985), *Changes of groundwater regime in wet meadows,* PhD thesis, University of Groningen, Groningen.

Hadamard, J. (1902), Sur les problèmes aux dérivées partielles et leur signification physique, *Princeton University Bulletin*, 49-52.

Hantush, M. S. (1956), Analysis of Data From Pumping Tests In Leaky Aquifers, *Transactions, American Geophysical Union*, 37(6), 702-714.

Harp, D. R., and V. V. Vesselinov (2011), Identification of Pumping Influences in Long-Term Water Level Fluctuations, *Ground Water*, 49, 403-414. doi: 10.1111/j.1745-6584.2010.00725.x.

Hartung, A. (2002), Bestimmung langfristiger hydrologischer Charakteristika von Fluss und Grundwasser zur Beschreibung von Standortbedingungen in Auenwäldern, PhD Thesis, Technical University Dresden, Dresden.

Harvey, A. C. (1989), Forecasting, structural time series models and the Kalman filter, Cambridge University Press, Cambridge.

Harvey, A. C., and R. G. Pierse (1984), Estimating missing observations in economic time series, *Journal of the American Statistical Association*, 79(385), 125-131.

Harvey, C. F., and S. M. Gorelick (1995), Temporal moment-generating equations: Modeling transport and mass-transfer in heterogeneous aquifers, *Water Resour. Res.*, 31(8), 1895-1911.

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.

Jakeman, A. J., and G. M. Hornberger (1993), How much complexity is warranted in a rainfall-runoff model?, *Water Resources Research*, 29, 2637-2649 .

Jarrett, R. D. (1991), Paleohydrology and its value in estimating floods and droughts, in *National Water Summary 1988-89--Hydrologic Events and Floods and Droughts*, edited by Paulson, R.W., Chase, E.B., Roberts, R.S., and Moody, D.W., U.S. Geological Survey Water-Supply Paper 2375, 105-116.

Jones, R. H. (1980), Maximum likelihood fitting of ARMA models to time series with missing observations, *Technometrics*, 22, 389-395.

Journel, A. G., and C. Huijbregts (1978), *Mining Geostatistics,* Academic Press, London.

Jury, W. A. (1982), Simulation of solute transport using a transfer function model, *Water Resources Research*, 18, 363-368.

Jury, W. A., and K. Roth (1990), Transfer functions and solute movement through soil: theory and applications, Birkhäuser Verlag, Basel, Switzerland.

Jury, W. A., and G. Sposito (1985), Field calibration and validation of solute transport models for the unsaturated zone, *Soil Science Society of America Journal*, 49, 1331-1341.

Kalman, R. E. (1960), A new approach to linear filtering and prediction problems, *J. bas. engng. trans. ASMSE D*, 82, 35-45.

Kalnay, E. (2002), Atmospheric Modeling, Data Assimilation, and Predictability, Cambridge University Press, Cambridge.

Khabie-Zeitoune, E. (1982), Prediction in continuous time, in *Time Series Analysis: Theory and Practice 2`*, edited by Anderson, O.D., North-Holland, Amsterdam, 7–24.

Klemes, V. (1978), Physically based stochastic hydrologic analysis, *Advances in Hydroscience*, 11, 285-356.

Klir, G. J. (2001), *Facets of Systems Science,* ISFR International Series on Systems Science and Engineering, Volume 7, Plenum Press, New York.

Knotters, M. (2001), *Regionalised time series models for water table depths,* PhD Thesis, Wageningen University, Wageningen.

Knotters, M., and M. F. P. Bierkens (2000), Physical basis of time series models for water table depths, *Water Resources Research*, 36(1), 181-188.

Knotters, M., and J. G. De Gooijer (1999), Tarso modelling of water table depths, *Water Resources Research*, 35(3), 695-705.

Knotters, M., and P. E. V. Van Walsum (1997), Estimating fluctuation quantities from time series of water-table depths using models with a stochastic component, *Journal of Hydrology*, 197, 25-46.

Konikow, L. F., and J. D. Bredehoeft (1992), Groundwater models cannot be validated, *Advances in Water Resources*, 15, 75-83.

Koutsoyiannis, D. (2001), Coupling stochastic models of different timescales, *Water Resources Research*, 37(2), 379-391.

Kraijenhoff van de Leur, D. A. (1958), A study of non-steady groundwater flow with special reference to a reservoir coefficient, *De Ingenieur* , 70(B), 87-94.

Kroes, J. G., J. G. Wesseling, and J. C. Van Dam (2000), Integrated modelling of the soil-water-atmosphere-plant system using the model SWAP 2.0 an overview of theory and an application., *Hydrological Processes*, 14(11/12), 1993-2002.

Kruithof, A. J. (2001), The impulse response function of the water table, separation of the influence of the unsaturated and saturated zone (in Dutch), Master Thesis, TU Delft, BTO no. 2001.163(s) Kiwa Water Research, Delft, Nieuwegein.

Kuhn, T. S. (1962), *The structure of scientifiic revolutions,* University of Chicago Press, Chicago, Illinois.

Lammerts, E. J., C. Maas, and A. P. Grootjans (2001), Groundwater variables and vegetation in dune slacks, *Ecological Engineering*, 17, 33-47.

Lankester, J., and C. Maas (1996), Research on characterising site conditions for vegetation, on the basis of the impulse response (in Dutch), *Stromingen*, 2(3), 5-17.

Larocque, M., A. Mangin, M. Razack, and O. Banton (1998), Contribution of correlation and spectral analysis to the regional study of a large karst aquifer (Charente, France), *Journal of Hydrology*, 205, 217-231.

Leduc, C., J. Bromley, and P. Schroeter (1997), Water table fluctuation and recharge in semi-arid climate: some results of the HAPEX-Sahel hydrodynamic survey (Niger), *Journal of Hydrology*, 188-189, 123-138.

Lee, J.-Y., and K.-K. Lee (2000), Use of hydrologic time series data for identification of recharge mechanism in a fractured bedrock aquifer system, *Journal of Hydrology*, 229, 190-201.

Lee, K.-K., Y. Hyun, and S.-J. Kim (2005), Time series modeling for evaluation of groundwater discharge rates into an urban subway system, *Geosciences Journal*, 9(1), 15-22.

Leeuwis-Tolboom, J. A. M., and J. H. Peters (2002), Position paper groundwater nuisance, definition of the position of the drinking water sector (In Dutch), reportnr. 2002/2/4277, VEWIN, Rijswijk.

Lehsten, D., J. R. Von Asmuth, and M. Kleyer (2011), Simulation of Water Level Fluctuations in Kettle Holes Using a Time Series Model, *Wetlands*, 31, 511-520, DOI 10.1007/s13157-011-0174-7.

Li , W., W. Nowak, and O. A. Cirpka (2005), Geostatistical inverse modeling of transient

pumping tests using temporal moments of drawdown, *Water Resources Research*, 41:W08403. doi:10.1029/2004WR003874.

Little, R. J., and D. B. Rubin (1987), *Statistical Analysis with Missing Data,* Wiley, New York.

Ljung, L. (1999), *System Identification, Theory for the user,* Prentice Hall, Upper Saddle River.

Ljung, L. (2011), *System Identification Toolbox ,User's Guide,* The MathWorks, Inc., Natick, MA.

Luo, J., O. A. Cirpka, and P. K. Kitanidis (2006), Temporal-moment matching for truncated break-through curves for step or step-pulse injection, *Advances in Water Resources*, 29(9, doi:10.1016/j.advwatres.2005.10.005), 1306-1313.

Maas, C. (1994), *On Convolutional Processes and Dispersive Groundwater Flow,* PhD Thesis, Delft University of Technology, Delft.

Maas, C. (1995), Groundwater dynamics, a research proposal aimed at improving the duration line method (in Dutch), rapp. nr. SWI 95.121, KIWA NV, Nieuwegein.

Manga, M. (1999), On the timescales characterizing groundwater discharge at springs, *Journal of Hydrology*, 219, 56-69.

Manzione, R., M. Knotters, G. Heuvelink, J. Von Asmuth, and G. Camara (2010), Transfer function-noise modeling and spatial interpolation to evaluate the risk of extreme (shallow) water-table levels in the Brazilian Cerrados, *Hydrogeology Journal*, 18(8), 1927-1937 .

Marchal, J. H. (1975), On the concept of a system, *Philosophy of Science*, 42(4, doi:10.1086/288663), 448-468.

Marquardt, D. W. (1963), An algorithm for least-squares estimation of nonlinear parameters, *Journal of the Society of Industrial and Applied Mathematics*, 11(2), 431-441.

McGill, W. J. G. J. (1965), The General-Gamma Distribution and Reaction Times, *Journal Of Mathematical Psychology*, 2, 1-18.

McLaughlin, D. (1995), *Recent developments in hydrologic data assimilation,* Reviews of Geophysics, U.S. Natl. Report to Intl. Geophysical Union, 1991-1994.

Molenat, J., P. Davy, C. Gascuel-Odoux, and P. Durand (1999), Study of three subsurface hydrologic systems based on spectral and cross-spectral analysis of time series, *Journal of Hydrology*, 222, 152-164.

Mélard, G. (1984), Algorithm AS197. A fast algorithm for the exact likelihood of autoregressive-moving average models, *Journal of the Royal Statistical Society, Part C, Applied Statistics* (33), 104-114.

Nash, J. E. (1958), Determining runoff from rainfall, *Proc. Inst. Civ. Eng.*, 10, 163-184.

Nash, J. E. (1959), Systematic Determination of Unit Hydrograph Parameters, *Journal of Geophysical Research*, 64(1), 111–115, doi:10.1029/JZ064i001p00111.

Nieman, E. (1973), Grundwasser und Vegetationsgefüge, *Nova Acta Leopoldina Suppl.*, 6, 1-172.

Olsthoorn, T. N. (2000), Hydrologic restoration measures in the dunes, cause of millions of damage to flower bulbs? (in Dutch), *H2O*, 33(25/26), 23-24.

Olsthoorn, T. N. (2008), Do a Bit More with Convolution, *Ground Water*, 46(1), 13–22.

Oreskes, N., K. Schrader-Frechette, and K. Belitz (1994), Verification,validation and confirmation of numerical models in the earth sciences, *Science*, 263, 641–646.

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.

Penman, H. L. (1948), Natural evaporation from open water, bare soil and grass, *Proc. Roy. Soc. London, series A*, 193(1032), 120-146.

Polderman, J. W., and J. C. Willems (1998), *Introduction to Mathematical Systems Theory: A*

*Behavioral Approach,* Springer-Verlag, New York.

Popper, K. (1959), *The logic of scientific discovery,* Harper and Row, New York.

Porporato, A., P. D'Odorico, F. Laio, and I. Rodriguez-Iturbe (2003), Hydrologic controls on soil carbon and nitrogen cycles. I Modeling scheme, *Adv. Water Res.*, 26(1), 45-58.

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, 63-78.

Pulido-Velazquez, M. A., A. Sahuquillo-Herraiz, J. C. Ochoa-Rivera, and D. Pulido-Velazquez (2005), Modeling of stream-aquifer interaction: the embedded multireservoir model, *Journal of Hydrology*, 313 , 166-181.

Quimpo, R. G. (1971), Structural relation between parametric and stochastic hydrology models, *IAHS Publ.*, 100, 151-157.

Reggiani, P., S. M. Hassanizadeh, M. Sivapalan, and W. G. Gray (1999), A unifying framework for watershed thermodynamics: constitutive relationships, *Advances in Water Resources*, 23(1), 15-39.

Reggiani, P., M. Sivapalan, and S. M. Hassanizadeh (1998), A unifying framework for watershed thermodynamics: balance equations for mass, momentum, energy and entropy and the second law of thermodynamics, *Advances in Water Resources*, 22(4), 367-398.

Reggiani, P., and J. Schellekens (2005), Rainfall-runoff modeling: distributed models, in *Encyclopedia of Hydrological Sciences*, edited by Anderson, M.G., Wiley, Chichester, UK.

Richards, L. A. (1931), Capillary conduction of liquids through porous mediums, *Physics*, 318–333.

Rifai, M. N. E., W. J. Kaufman, and D. K. Todd (1956), *Dispersion Phenomena in Laminar Porous Media,* University of California, Berkely.

Rolf, H. L. M. (1989), Lowering of the water table in the Netherlands. Analysis of the period 1950-1986 (in Dutch), DGV-TNO, Delft.

Rolf, H. L. M., and J. J. Lebbink (1998), Hydrologic effects caused by the reduction of the dune water abstraction in the North-Holland Dune reserve (in Dutch), Provinciaal Waterleidingbedrijf Noord-Holland, Heemskerk.

Rowley, J. (2007), The wisdom hierarchy: representations of the DIKW hierarchy, *Journal of Information Science*, 33(2), 163-180, DOI: 10.1177/0165551506070706.

Runhaar, J., C. L. G. Groen, R. Van der Meijden, and R. A. M. Stevers (1987), A new division of the Dutch flora into ecological groups (in Dutch), *Gorteria*, 13, 276-359.

Sahuquillo, A. (1983), an eigenvalue numerical technique for solving unsteady linear groundwater models continuously in time, *Water Resources Research*, 19(1), 87-93, doi:10.1029/WR019i001.

Sánchez-Vila, X., and J. Carrera (2004), On the striking similarity between the moments of breakthrough curves for a heterogeneous medium and a homogeneous medium with a matrix diffusion term, *Journal of Hydrology*, 294, 164-175.

Savenije, H. G. (2001), Equifinality, a blessing in disguise, *Hydrological Processes*, 15, DOI: 10.1002/hyp.494, 2835-2838.

Schoups, G., and J. A. Vrugt (2010), A formal likelihood function for parameter and predictive inference of hydrologic models with correlated, heteroscedastic, and non-Gaussian errors, *Water Resources Research*, 46, W10531, doi:10.1029/2009WR008933.

Schweppe, F. C. (1973), *Uncertain dynamic systems,* Prentice-Hall , Eaglewood Cliffs, N.J.

Sherman, L. K. (1932), Streamflow from rainfall by unit-graph method, *Engineering News Record*, 108, 501-505.

Shibata, R. (1976), Selection of the order of an autoregressive model by Akaike's information criterion, *Biometrika*, 63, 117-126.

Skyttner, L. (2005), *General Systems Theory, Problems, Perspectives, Practice,* World Scientific Publishing Co. Ptc. Ltd., Singapore.

Sloan, W. T. (2000), A physics-based function for modeling transient groundwater discharge at the watershed scale, *Water Resources Research*, 36(1), 225-241.

Sluijter, R., and J. Nellestijn (2002), Climate atlas of the Netherlands, the normal period 1971-2000 (in Dutch), uitgeverij Elmar, Rijswijk.

Snedecor, G. W., and W. G. Cochran (1967), *Statistical Methods,* The Iowa State University Press, Ames.

Sorensen, J. P. R., and A. S. Butcher (2011), Water Level Monitoring Pressure Transducers-A Need for Industry-Wide Standards, *Ground Water Monitoring & Remediation*, doi: 10.1111/j1745–6592.2011.01346.x, 1-7.

Sun, Y.-N. J. W. J. (1998), Transit Compartments versus Gamma Distribution Function To Model Signal Transduction Processes in Pharmacodynamics, *Journal of Pharmaceutical Sciences*, 87(6), 732-737.

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.

Taylor, C. J., D. J. Pedregal, P. C. Young, and W. Tych (2007), Environmental time series analysis and forecasting with the Captain toolbox, *Environmental Modelling & Software*, 22, 797-814.

Te Stroet, C. B. M. (1995), Calibration of Stochastic Groundwater Models. Estimation of system noise statistics and model parameters, PhD thesis, Delft University of Technology, Delft.

Todini, E. (2007), Hydrological catchment modelling: past, present and future, *Hydrology and Earth System Sciences*, 11, 468-482.

Tong, H. (1990), Non-linear Time Series: A Dynamical System Approach, Clarendon, Oxford.

Tüxen, R. (1954), Pflanzengesellschaften und Grundwasser-Ganglinien, *Angewandte Pflanzensoziologie*, 8, 64-97.

Uhlenbeck, G. E., and L. S. Ornstein (1930), On the Theory of the Brownian Motion, *Physical Review*, 36, 823-841.

Van Dam, J. C. (2000), Field-scale water flow and solute transport, SWAP model concepts, parameter estimation, and case studies, PhD-thesis, Wageningen University, Wageningen.

Van de Vliet, R. N., and R. Boekelman (1998), Areal coverage of the impulse response with the aid of time series analysis and the method of moments [In Dutch]., *Stromingen*, 4(1), 45-54.

Van de Vliet, R. N., H. L. M. Rolf, and J. J. Lebbink (2000), Effects of the reduction and stopping of the water abstraction Zuid-Kennemerland. Time series analysis of observed groundwater level series for urgent areas (in Dutch), PWN / IWACO, Rotterdam.

Van der Sluijs, P. (1990), Groundwater level classes, in *Soils of the Netherlands (in Dutch)* , edited by Locher, W.P., and De Bakker, H., Malmberg, Den Bosch, 167-180.

Van der Sluijs, P., and  J. J. De Gruijter (1985), Water table classes: a method to describe seasonal fluctuation and duration of water tables on Dutch soil maps, *Agricultural Water Management*, 10, 109-125.

Van Ek, R., J. P. M. Witte, H. Runhaar, and F. Klijn (2000), Ecological effects of water management in the Netherlands: the model DEMNAT, *Ecological Engineering*, 16, 127-

141.

Van Geer, F. C. (1987), Application of Kalman filtering in the analysis and design of groundwater monitoring networks, PhD Thesis, Delft University of Technology, Delft.

Van Geer, F. C., P. K. Baggelaar, and P. R. Defize (1988), Application of time series analysis to time series of groundwater head (in Dutch), *H₂0*, 21(161), 438-442.

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.

Veling, E. J. M. (2010), Approximations of impulse response curves based on the generalized moving Gaussian distribution function, *Advances in Water Resources*, 33, 546-561.

Veling, E. J. M., and C. Maas (2010), Hantush Well Function revisited, *doi:10.1016/j.jhydrol.2010.08.033, Journal of Hydrology*, 393, 381-388.

Von Asmuth, J. R. (2000), Water table dynamics from an ecologic perspective, research proposal (in Dutch), Kiwa N.V., Nieuwegein.

Von Asmuth, J. R. (2010), *On the quality, frequency and validation of pressure sensor series (in Dutch),* Rapportnr. KWR 2010.001, KWR Watercycle Research Institute, Nieuwegein.

Von Asmuth, J. R., and M. F. P. Bierkens (2005), Modeling irregularly spaced residual series as a continuous stochastic process, *Water Resources Research*, 41(12), W12404, doi:10.1029/2004WR003726.

Von Asmuth, J. R., M. F. P. Bierkens, and K. Maas (2002), Transfer function noise modeling in continuous time using predefined impulse response functions, *Water Resources Research*, 38(12), 23_1-23_12.

Von Asmuth, J. R., and M. Knotters (2004), Characterising spatial differences in groundwater dynamics based on a system identification approach, *Journal of Hydrology*, 296(1-4), 118-134.

Von Asmuth, J. R., M. Knotters, and K. Maas (2006), *Time series analysis for (eco)hydrologists, background documentation and course manual (in Dutch),* Kiwa Water Research/Alterra, Nieuwegein/Wageningen.

Von Asmuth, J. R., and K. Maas (2001), The method of impulse response moments: a new method integrating time series-, groundwater- and eco-hydrological modelling, in *Impact of Human Activity on Groundwater Dynamics*, edited by Gehrels, J.C., Peters, N.E., Hoehn, E., Jensen, K., Leibundgut, C., Griffioen, J., Webb, B., and Zaadnoordijk, W.J., IAHS Press, Centre for Ecology and Hydrology, Wallingford, 51-58.

Von Bertalanffy, L. (1950), An Outline of General System Theory, *British Journal for the Philosophy of Science*, 1 (2), 134-165.

Whitehead, P. G., and P. C. Young (1979), Water quality in river systems: Monte-Carlo analysis, *Water Resources Research*, 15, 451-459.

Wiener, N. (1949), Extrapolation, Interpolation and Smoothing of Stationary Time Series, MIT, 56-103, MA.

Willems, J. C. (1991), Paradigms and Puzzles in the Theory of Dynamical Systems, *IEEE transactions on automatic control*, 36(3), 259-294.

Witte, J. P. M., F. Klijn, F. A. M. Claessen, C. L. G. Groen, and R. Van der Meijden (1992), A model to predict and assess the impacts of hydrologic changes on terrestrial ecosystems in The Netherlands, and its use in a climate scenario, *Wetland Ecology and Management*, 2(1/2), 69-83.

Yi, M. J., and K. K. Lee (2004), TFN modeling of irregularly observed groundwater heads, *Journal of Hydrology*, 288, 272-287.

Yihdego, Y., and J. A. Webb (2011), Modeling of bore hydrographs to determine the impact of climate and land-use change in a temperate subhumid region of southeastern Australia, *Hydrogeology Journal*, 19, 877-887.

Young, P. C., and K. J. Beven (1994), Data-based mechanistic modelling and the rainfall-flow non-linearity, *Environmetrics*, 5, 335-365.

Young, P., S. Parkinson, and M. Lees (1996), Simplicity out of complexity in environmental modelling: Occam's razor revisited, *Journal of Applied Statistics*, 23, 165-210.

Young, P. C. (1998), Data-based mechanistic modelling of environmental, ecological, economic and engineering systems, *Environmental Modelling & Software*, 13, 105-122.

Young, P. C., and H. Garnier (2006), Identification and estimation of continuous-time, data-based mechanistic (DBM) models for environmental systems, *Environmental Modelling & Software*, 21(8), 1055-1072, DOI : 10.1016/j.envsoft.2005.05.007.

Yu, C., A. W. Warrick, and M. H. Conklin (1999), A moment method for analyzing breakthrough curves of step inputs, *Water Resources Research*, 35(11), 3567-3572.

Zadeh, L. A. (1956), On the identification problem, *IRE Transactions on Circuit Theory*, 3, 277-281.

Ziemer, R. E., W. H. Tranter, and D. R. Fannin (1998), *Signals and systems: continuous and discrete,* Prentice-Hall, Upper Saddle River.

Zwamborn, M. H. (1995), Modeling of the unsaturated zone for the benefit of groundwater models (in Dutch), Kiwa N.V., Nieuwegein.

# Appendix C

# Publications and curriculum vitae
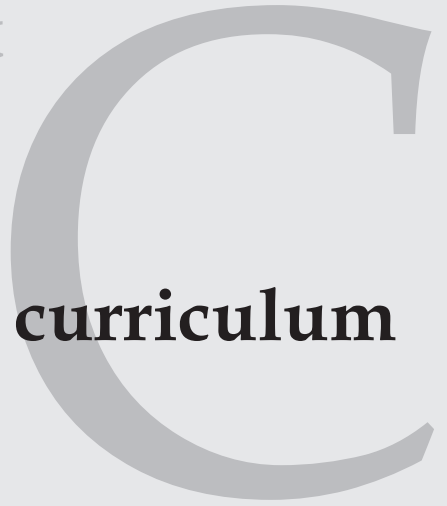
# Publications

**Peer reviewed publications (11)**

Von Asmuth, J. R., C. Maas, M. Knotters, M. F. P. Bierkens, M. Bakker, T. N. Olsthoorn, D. G. Cirkel, I. Leunk, F. Schaars, and D. C. Von Asmuth (submitted), Menyanthes: software for hydrogeologic time series analysis, interfacing data with physical insight, *Environmental Modelling & Software*.

Post, V. E. A. and J. R. Von Asmuth (in prep.), Hydraulic head measurements: New technologies, classic pitfalls. *Hydrogeology Journal*.

Lehsten, D., J. R. Von Asmuth, and M. Kleyer (2011), Simulation of Water Level Fluctuations in Kettle Holes Using a Time Series Model, *Wetlands*, 31, 511-520, DOI 10.1007/s13157-011-0174-7.

Manzione, R., M. Knotters, G. Heuvelink, J. Von Asmuth, and G. Camara (2010), Transfer function-noise modeling and spatial interpolation to evaluate the risk of extreme (shallow) water-table levels in the Brazilian Cerrados, *Hydrogeology Journal*, 18(8), 1927-1937 .

Von Asmuth, J. R., K. Maas, M. Bakker, and J. Petersen (2008), Modeling time series of groundwater head fluctuations subjected to multiple stresses, *Ground Water*, 46, doi: 10.1111/j.1745-6584.2007.00382.x(1), 30-40.

Bakker, M., K. Maas, and J. R. Von Asmuth (2008a), Calibration of transient groundwater models using time series analysis and moment matching, *Water Resources Research*, 44(W04420), doi:10.1029/2007WR006239.

Bakker, M., K. Maas, F. Schaars, and J. R. Von Asmuth (2007), Analytic modeling of groundwater dynamics with an approximate impulse response function for areal recharge, *Advances in Water Resources*, 30(3, doi:10.1016/j.advwatres.2006.04.008), 493-504.

Von Asmuth, J. R., and M. F. P. Bierkens (2005), Modeling irregularly spaced residual series as a continuous stochastic process, *Water Resources Research*, 41(12), W12404, doi:10.1029/2004WR003726.

Von Asmuth, J. R., and M. Knotters (2004), Characterising spatial differences in groundwater dynamics based on a system identification approach, *Journal of Hydrology*, 296(1-4), 118-134.

Witte, J. P. M., and J. R. Von Asmuth (2003), Do we really need phytosociological classes to calibrate Ellenberg's indicator values?, *Journal of Vegetation Science*, 14, 615-618.

Von Asmuth, J. R., M. F. P. Bierkens, and K. Maas (2002c), Transfer function noise modeling in continuous time using predefined impulse response functions, *Water Resources Research*, 38(12), 23_1-23_12.

**Book chapters and conference proceedings (6)**

Manzione, R. L., M. Knotters, G. M. B. Heuvelink , J. R. Von Asmuth, and G. Câmara (2009), Predictive risk mapping of water table depths in a Brazilian Cerrado area, in *Quality aspects in spatial data mining*, edited by Stein, A., Shi, W., and Bijker, W., CRC Press, Taylor & Francis Group, Boca Raton, Florida,

Bakker, M., C. Maas, and J. R. Von Asmuth (2008b), Transient calibration of flow to ditches with entry resistance using measured moments of response functions, in *Calibration and Reliability in Groundwater Modelling: Credibility of Modeling* Proceedings of ModelCARE 2007 Conference, IAHS Publ. 320,

Maas, C., J. R. Von Asmuth, M. Bakker, E. Veling, and T. N. Olsthoorn (2007), *Groundwater Modeling: In Search of a Different Paradigm, in: Excursions Into Ecohydrology, Featuring the 2007 Darcy Lecture,* Boussinesq Center, TU-Delft, UNESCO-IHE, Delft.

Putters, B., C. Maas, D. G. Cirkel, F. Vaessen, M. Juhàsz, and J. R. von Asmuth (2006), Prevention and prediction of nitrate concentrations in limestone drinking water wells in South Limburg (NL), *Proceedings of the 10th International Specialized Conference on Diffuse Pollution and Sustainable Basin Management*.

Von Asmuth, J. R., A. P. Grootjans, E. J. Lammerts, and A. J. M. Jansen (2002a), Eco-hydrological modeling using response characteristics of hydrological systems, in *Water Resources and Environment Research*, edited by Schmitz, G.H., Proceedings ICWRER Conference, Dresden University of Technology, Dresden, 335-340.

Von Asmuth, J. R., and K. Maas (2001), The method of impulse response moments: a new method integrating time series-, groundwater- and eco-hydrological modelling, in *Impact of Human Activity on Groundwater Dynamics*, edited by Gehrels, J.C., Peters, N.E., Hoehn, E., Jensen, K., Leibundgut, C., Griffioen, J., Webb, B., and Zaadnoordijk, W.J., IAHS Press, Centre for Ecology and Hydrology, Wallingford, 51-58.

**Publications in professional journals (11)**

De Meij, T., and J. R. Von Asmuth (2011), Correctie van eigen luchtdrukmetingen is noodzakelijk, *H₂O*, 4, 29-32.

Maas, C., J. R. Von Asmuth, and J. Runhaar (2008), Kanttekeningen bij 'Oorzaak en gevolg van numerieke verdroging', *H₂O*, 9, 22-24.

Von Asmuth, J. R. (2006), Kiwa Water Research Builds a Tool for Monitoring Groundwater Levels , *The MathWorks News & Notes*, 14-16.

Cirkel, D. G., and J. R. Von Asmuth (2005b), Trends in grondwater! Schatgraven in grondwaterstandsreeksen, *Trendsinwater.nl*, 16, 2.

Bakker, D., C. J. S. Aggenbach, J. R. von Asmuth , and J. P. M. Witte (2004), De dynamiek van vennen in schijnspiegelsystemen, *Stromingen*, 10(4), 5-12 .

Von Asmuth, J. R., C. Maas, and D. G. Cirkel (2004), Tijdreeksanalyse van grondwaterstanden nu binnen ieders bereik, *H₂O*, 24, 31-33.

Von Asmuth, J. R., W. L. Berendrecht, C. Maas, and F. C. Van Geer (2003), Tijdreeksanalyse van grondwaterpeilen: de invloed van de calibratiemethode?, *Stromingen*, 9(2), 61-63.

Peerboom, J. M. P. M., C. J. S. Aggenbach, and J. R. Von Asmuth (2003), Hydrologie van de Beegderheide: ontstaanswijze en functioneren van de vennen en hydrologische ontwikkelingen tussen 1997-2002, *Natuurhistorisch Maandblad*, 92, 145-152.

Von Asmuth, J. R., M. F. P. Bierkens, and C. Maas (2002b), Soms is weten beter dan meten (tenzij je verkeerd zit natuurlijk). Het discrete Box-Jenkins- versus het continue PIRFICT-tijdreeksmodel, in praktijk, *Stromingen*, 8(1), 5-14.

Von Asmuth, J. R., C. Maas, and M. F. P. Bierkens (2001), Waarom doen alsof de neerslag eens per maand valt? Het discrete Box-Jenkins- versus het continue PIRFICT-tijdreeksmodel, in theorie, *Stromingen*, 7(4), 33-44.

Von Asmuth, J. R. (2000), Boekbespreking van: A.J.M Jansen (2000), Hydrology and restoration of wet heathland and fen meadow communities, Proefschrift Rijksuniversiteit Groningen, *Stromingen*, 6(3), 57-60.

**Reports and other publications (31)**

Von Asmuth, J. R., J. G. Streefkerk, and C. Maas (2011), *Natte natuur in het Drents-Friese Wold. Overzicht gegevens, hydrologische situatie en effecten van herstelmaatregelen.*, Rapportnr.

KWR 2011.106, KWR Watercycle Research Institute, Nieuwegein.

Von Asmuth, J. R., A. P. Grootjans, and S. Van der Schaaf (2011), Over de dynamiek van peilen en fluxen in vennen en veentjes. Eindrapport deel 2, OBN-onderzoek 'Herstel van biodiversiteit en landschapsecologische relaties in het natte zandlandschap', Rapport nr. 2011/OBN147-2-NZ, Bosschap, bedrijfschap voor bos en natuur, Driebergen.

Von Asmuth, J. R. (2010), *Over de kwaliteit, frequentie en validatie van druksensorreeksen,* Rapportnr. KWR 2010.001, KWR Watercycle Research Institute, Nieuwegein.

Von Asmuth, J. R., S. Van der Schaaf, A. P. Grootjans, and C. Maas (2010), *Weerstand en wegzijging in natte natuurgebieden, schatting via analyse van gemeten (grond)waterpeilen,* Delft University of Technology, Delft.

Leunk, I., K. J. Raat, and J. R. Von Asmuth (2010), *Snel en nauwkeurig detecteren van putverstopping met tijdreeksanalyse (Menyanthes),* rapport BTO 2010.036(s), KWR Watercycle Research Institute, Nieuwegein.

Maas, C., J. R. Von Asmuth, R. Agtersloot, F. Schaars, and P. Maas (2009), *Het Grensmaasproject en de Vlaamse VHR-gebieden, een signalerings- en alarmeringsprotocol ten aanzien van de grondwaterstanden,* KWR 08.075, KWR Watercycle Research Institute, Nieuwegein.

Von Asmuth, J. R., C. Maas, and M. Knotters (2009), *Handleiding Menyanthes, versie 1.9,* KWR Watercycle Research Institute, Nieuwegein.

Knotters, M., S. P. J. Van Delft, H. E. Keizer-Vlek, J. R. Von Asmuth, P. C. Jansen, F. P. Sival, and C. E. Van 't Klooster (2008), *Evaluatie monitoring Deurnese Peel en Mariapeel. Kwantificering van effecten van maatregelen en advies over het monitoringplan,* Alterra-Document2, Alterra, Wageningen.

Maas, C., and J. R. Von Asmuth (2008), *Advies Monitoring Grensmaas,* KWR 08.020, KWR Watercycle Research Institute, Nieuwegein.

Cirkel, D. G., C. Maas, and J. R. Von Asmuth (2007), Temporele variaties van nitraat in onttrokken grondwater. Ontwikkeling en toepassing van een tijdreeksmodel voor de modellering en voorspelling van nitraatconcentraties in onttrokken grondwater, Rapport KWR 07.003, Kiwa Water Research, Nieuwegein.

Von Asmuth, J. R., M. Knotters, and C. Maas (2006), *Tijdreeksanalyse voor (eco)hydrologen, achtergronddocumentatie en cursushandleiding,* Kiwa Water Research/Alterra, Nieuwegein/Wageningen.

Maas, C., D. G. Cirkel, and J. R. Von Asmuth (2005), *Tijdreeksanalyse Voornes Duin, deel 1: Afzonderlijke peilbuizen,* rapport KWR 05.055, Kiwa Water Research, Nieuwegein.

Cirkel, D. G., and J. R. Von Asmuth (2005a), *Hydrologische evaluatie anti-verdrogingsmaatregelen Terwisscha,* rapport nr. KWR 05.29, Kiwa Water Research, Nieuwegein.

Cirkel, D. G., and J. R. Von Asmuth (2004), *Effecten van de vernattingsmaatregelen in de Amsterdamse waterleidingduinen ,* rapport nr. KWR 03.092, Kiwa Water Research, Nieuwegein.

Cirkel, D. G., C. Maas, and J. R. Von Asmuth (2004), What Els? Evaluatie van het voorlopige meetnet voor Extreem Lage Stijghoogten van de provincie Noord-Brabant, KWR.04.073 Kiwa Water Research, Nieuwegein.

Maas, C., and J. R. Von Asmuth (2004), Tijdreeksanalyse Mander, Een onderzoek naar de hydrologische wisselwerking tussen de stuwwal van Ootmarsum en de Slenk van Reutum, Kiwa Water Research, Nieuwegein.

Maas, C., and J. R. Von Asmuth (2003), Tijdreeksanalyse grondwaterstanden 1975-2002,

Effectiviteit van het stand-still-beleid van de Provincie Limburg, Kiwa Water Research, Nieuwegein.

Jansen, A. J. M., J. R. Von Asmuth, J. Bunnik, and A. C. Zuidhoff (2001), *Vijf NB-wet-terreinen op het landgoed Twickel (Overijssel),* rapportnr. KOA 01.049, Kiwa N.V., Nieuwegein.

Von Asmuth, J. R. (2000b), Onderzoeksvoorstel grondwaterspiegeldynamica in ecologisch perspectief, Kiwa N.V., Nieuwegein.

Koppejan, H., P. J. M. Melman, J. R. Von Asmuth, and D. De Jong (1999), *Standaardvoorschrift kwelderkartering in Nederland,* rapport MDGAE-98-02, Meetkundige dienst, Delft.

Von Asmuth, J. R., B. Van Gennip, and A. M. De Meulmeester (1999), *P.Q.-onderzoek Rottumerplaat en -oog,* rapp.nr. MDGAE 9911, RWS-Meetkundige Dienst, Delft.

Van Gennip, B., J. R. Von Asmuth, J. Cools, and M. Bakker (1998), *De buitendijkse gebieden langs het Haringvliet en Hollandsch Diep,* RWS - Meetkundige Dienst, Delft.

Von Asmuth, J. R., M. Tinga, and J. A. M. Janssen (1998), *Project Stroomlijnen Werkproces Vegetatiekartering, eindrapport fase 1.,* Intern rapport, RWS-Meetkundige Dienst, Delft.

Von Asmuth, J. R., A. G. Knotters, and L. J. Schulpen (1997a), *Arc Info Gebruikers Cursus. Cursus met beknopte cursushandleiding,* RWS-Meetkundige Dienst, Delft.

Von Asmuth, J. R., B. Gennip, and P. J. M. Melman (1997b), *P.Q.-onderzoek 'Groene Strand' Terschelling,* rapp.nr. MDGAT 9753, RWS-Meetkundige Dienst, Delft.

Zonneveld, L. M. L., J. R. Von Asmuth, and J. A. M. Van Dongen (1997), *Vegetatiekartering 'de Grie' Terschelling 1993,* rapp.nr. MDGAT 9536, , RWS-Meetkundige Dienst, Delft.

Reitsma, J. M., J. R. Von Asmuth, and E. R. Stenfert-Steehouwer (1996), *De schorren van de Westerschelde 1990 / 1993,* rapp.nr. MDGAT 9623, RWS-Meetkundige Dienst, Delft.

Von Asmuth, J. R., and M. Tolman (1996), *Vegetatiekartering Schiermonnikoog 1992, rapportage en ecologische interpretatie ,* rapp.nr. MDGAT 9603, RWS-Meetkundige Dienst, Delft.

Von Asmuth, J. R. (1995a), Monitoring in de Millingerwaard, aanzet tot een vijf-jaarlijkse vegetatiekartering en thematische / geometrische analyse m.b.v. GIS, Afstudeerrapport, Landbouwuniversiteit Wageningen, Wageningen.

Von Asmuth, J. R. (1995b), Project Monitoring Rottum, Rottumeroog en -plaat als testcase voor het vergelijken van vegetatiekaarten, rapp.nr. MDGAT 9548, RWS-Meetkundige Dienst, Delft.

Von Asmuth, J. R. (1994), *Disturbance, regrowth and replanting in the tropical rainforests of Rabi, Gabon,* Afstudeerrapport, Landbouwuniversiteit Wageningen, Wageningen.

# Curriculum vitae

Jos von Asmuth werd geboren op 9 juni 1970 te Eindhoven. Hij haalde zijn gymnasium diploma aan het Eindhovens Protestants Lyceum[1] in 1988. Nadat hij bij de Vrije Hogeschool[2] te Driebergen het zogenaamde propaedeutische jaar gevolgd had, ging hij biologie studeren aan de Landbouwuniversiteit te Wageningen[3]. Hij studeerde in 1995 af met het predicaat Cum Laude in twee oriëntaties en met drie afstudeeronderwerpen. De hoofdrichtingen van zijn studie waren vegetatiekunde, landschapsecologie en geografische informatie systemen. Aansluitend aan zijn laatste stage werkte hij een viertal jaren bij de Meetkundige Dienst[4] van Rijkswaterstaat, op het gebied van ecologische kartering, monitoring en analyse. In die periode groeide hij uit van projectmedewerker, via projectmanager tot innovatiecoördinator.

Om zich inhoudelijk verder te verdiepen begon hij, na een korte positie als interim secretaris van het ProGIS-programma van het ministerie van Landbouw, Natuurbeheer en Visserij[5], in 1999 aan een promotieonderzoek aan de Technische Universiteit te Delft. Dit onderzoek, dat in nauwe samenwerking met Kiwa N.V.[6] is uitgevoerd, omvatte oorspronkelijk de integratie van drie typen modellen, te weten tijdreeks-, grondwater- en ecohydrologische modellen. Met dit doel voor ogen is de nadruk van het promotieonderzoek uiteindelijk vooral bij het eerste onderdeel komen te liggen, te weten tijdreeksanalyse en haar toepassing op, en relatie met, de fysica van grondwater. Jos heeft als onderdeel van het onderzoek een nieuwe methode voor tijdreeksanalyse ontwikkeld, de zogenaamde PIRFICT-methode. Hij heeft deze en andere methoden voor het analyseren van (grond)waterspiegelfluctuaties bovendien vormgegeven in het gebruiksvriendelijke programma *Menyanthes.*

Nog tijdens het promotieonderzoek hebben Jos en Kiwa gezamenlijk besloten om zowel methodiek als programma in praktijk te brengen, en het gebruik ervan waar nodig te faciliteren. Jos is vervolgens in 2003 in dienst van Kiwa getreden, als ecohydroloog en hoofdauteur van *Menyanthes*. Het programma wordt ten tijde van het verschijnen van dit proefschrift gebruikt op drie continenten, in zes landen, door 76 verschillende organisaties. Naar schatting enige honderden (eco)hydrologen hebben inmiddels een één- of tweedaagse cursus gevolgd die door hem, Martin Knotters en Kees Maas ontwikkeld en gegeven is. Enige tienduizenden grondwaterstandsreeksen zijn inmiddels met behulp van *Menyanthes* geanalyseerd.

Jos is lid van het deskundigenteam 'Nat Zandlandschap' van het ministerie van EL&I, van de redactie van het hydrologische tijdschrift Stromingen, van de webredactie van de Nederlandse Hydrologische Vereniging en van de DINO-gebruikersraad van TNO. Hij kreeg in 2005 de 'Best Poster Award' op het symposium van het Water Research Centre Delft, voor zijn poster over tijdreeksanalyse met drempel-niet-lineariteit.

---

[1] Thans het Christiaan Huygens College
[2] Thans het Bernard Lievegoed College for Liberal Arts
[3] Thans Wageningen University & Research centre
[4] Thans de Data-ICT-Dienst van Rijkswaterstaat
[5] Thans het ministerie van Economische Zaken, Landbouw en Innvovatie (EL&I)
[6] Thans KWR Watercycle Research Institute

NW1/4 SECT. 1
T51N, R4W
JUNE 22
JOHN GLINDINNING

Gentianaceae                    #57

menyanthes trifoliata
(L)

NW1/4 sec 1, T51N R4W
WIS.
in marsh low behind beach ridge

John Glindinning    June 22

# Groundwater System Identification

## through Time Series Analysis

Groundwater, water in the ground. Although it is invisible, it is a vital resource for all terrestrial life (whether direct or indirect). Many processes interact with it. Rain recharges it, as it infiltrates the soil. Plant roots take it up, and their leaves evaporate it. It discharges to rivers and streams, and is abstracted with pumping wells. It is controlled with ditches and drainage means. Such processes and activities leave their traces in the groundwater level fluctuations. Careful analysis of these fluctuations may reveal much of the functioning of systems, and of the effects of individual factors. This is shown by many, but practiced by too few, as traditional time series analysis theory and software are complex.

In this thesis, a new method of time series analysis is presented. Its continuous time formulation fits existing physical-hydrologic theory and methods well. It is shown that groundwater level responses generally take the shape of simple distribution functions. This notion, combined with the program *Menyanthes* that was developed, enable the quick and easy analysis of large numbers of time series. The spatial patterns that emerge in the results of multiple models literally add another dimension to the technique. As time series models are usually accurate also, they may be valuable to every (eco) hydrologist.