

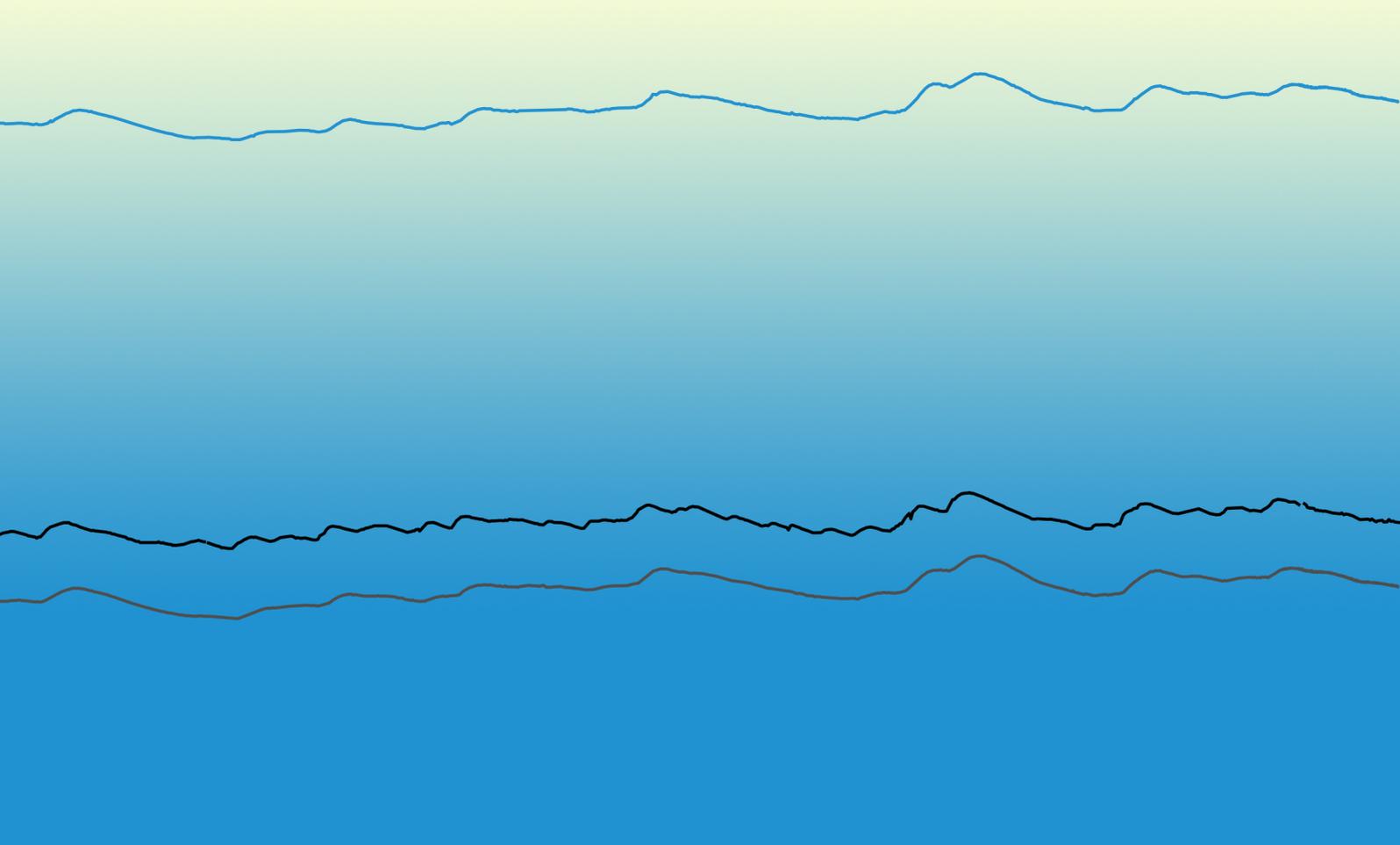


Technische Universiteit Delft

Non-Linear Time Series Analysis of Deep Ground- water Levels

An application to the Veluwe

R.A. Collenteur



NON-LINEAR TIME SERIES ANALYSIS OF DEEP GROUNDWATER LEVELS

AN APPLICATION TO THE VELUWE

by

R.A. Collenteur

in partial fulfillment of the requirements for the degree of

Master of Science

in Civil Engineering

at the Delft University of Technology,

to be defended publicly on Thursday the 24th, 2016 at 14:30 AM.

Student number:	4241665	
Thesis committee:	Dr. ir. M. Bakker	TU Delft
	Dr. ir. W. J. Zaadnoordijk	KWR, TU Delft
	Dr. ir. M. Hrachowitz	TU Delft
	Prof. dr. ir. R.M. Hanssen	TU Delft

An electronic version of this thesis is available at <http://repository.tudelft.nl/>.

ABSTRACT

The objective of this study is to improve the simulation of deep groundwater levels by time series models with pre-defined impulse response functions. This is attempted by adding a conceptual non-linear root zone model to simulate the recharge series to the model and by testing the use of a separate response function for the percolation zone. Three root zone models are developed based on two different recharge mechanisms: preferential flow, percolation, and a combination of the two. The performance of these models is compared to a linear model that is commonly used in time series models to simulate the recharge. The approach is applied to groundwater level measurements in the Veluwe, a largely forested area in the Netherlands characterized by thick unsaturated zones. The effect of groundwater extractions and land reclamations is added to the model to further improve the simulation of the groundwater levels. The models are tested on three observation wells with increasing thickness of the unsaturated zone, varying from 7 m to 29 m to 49 m.

The results show that model performance is improved by the implementation of a non-linear root zone model, particularly in simulating the peaks and lows in the groundwater levels. The recharge fluxes simulated by the non-linear models show different patterns that are physically more realistic than those simulated by the linear model. It is shown that different recharge series result in simulated groundwater levels that are very similar. This is a clear example of equifinality and it is recommended to introduce new sources of information to validate the modelled processes (e.g., water content measurements of the percolation zone or actual evaporation data).

For the shallow well, the models with a single response function are selected as the best. The largest improvements for the deeper groundwater levels are obtained by the addition of a separate response for the percolation zone. For example, the average deviation from the observed groundwater levels decreased 0.18 m to 0.08 m for the deepest observation well by applying the separate response function. The models with an additional response function were better at simulating the estimated time to peak, the time it takes a recharge pulse to cause a peak in the groundwater levels. The time to peak is introduced in this research as a qualitative indicator to validate the modelled processes. The simulated responses indicate that the groundwater levels respond very quickly to water that leaves the root zone, even though the percolation zone is tens of metres thick.

For each of the observation wells it is investigated if adding the effect of groundwater extractions or land reclamations of Flevoland to the models improves the simulation of the groundwater levels. For the shallow well it is concluded that either the effect of land reclamations or groundwater extractions needs to be taken into account. Since these two stresses are correlated, it is concluded that only one of these should be taken into account when no further information is available to constrain the models. For the medium deep well, the additional stresses did not significantly improve model performance and it is concluded that they do not have to be taken into account for this well. For the deep well, model performance is improved by both stresses. The largest improvements are observed when the effect of groundwater extractions is considered in the model. For implementation of these stresses, the entire simulation period should be used for calibration, or constraints have to be implemented to obtain realistic results.

PREFACE

This master thesis is the result of one of the most pleasurable periods of my time at university. It is also the final step in pursuing a MSc. degree in Water Management at the faculty of Civil Engineering at the TU Delft. It goes without saying that all this would not have been possible without the help of many colleagues, friends and family.

I would like to thank my graduation committee for their active involvement in my research. I am grateful to Mark Bakker for being my daily supervisor and for sharing his enthusiasm for Python with me. I express my gratitude Willem Jan Zaadnoordijk who always took the time to explain or discuss methodological issues with me. Markus Hrachowitz, I greatly appreciate all your help with the conceptualization of the unsaturated zone. I thank Prof. Hanssen for being part of my graduation committee.

Joanne van der Spek, I am thankful to you for teaching me many of the techniques used in time series analysis. Tanja Euser I thank for the help on the estimation of the size of the root zone reservoir. I am grateful to Huub Savenije for his suggestion to try local precipitation data, significantly changing the results of this study. I will never make the mistake not to use local precipitation data again.

Graduating for over 16 months has one big advantage, sharing the graduation room with many students for a longer period of time. I greatly appreciate all my fellow students for the many 'professional' discussions, python jokes, bouldering sessions, Thursdays drinks after the colloquium and the many social moments that we shared during our studies.

My ambition for this thesis project was shared with my ambition to train for and compete at the world championships in freestyle kayaking in the summer of 2015 in Ottawa, Canada. As a consequence, this graduation project has taken a little more time than average (I would suggest the distribution of the time to graduate is now skewed slightly more to the right). I am grateful to my parents for supporting me throughout my entire studies, even when kayaking seemed to prolong that activity forever!

*R.A. Collenteur
Delft, 14th of March, 2016*

CONTENTS

List of Figures	ix
List of Tables	xv
1 Introduction	1
2 Study area 'the Veluwe'	5
2.1 Topography & Geology	5
2.2 Selection of the observation wells.	5
2.3 Meteorological data	10
3 Methodology	13
3.1 Modelling Strategy	14
3.2 The unsaturated zone	14
3.3 The transfer function.	18
3.4 The noise model	20
3.5 A separate response for the percolation zone	21
3.6 Model Calibration	21
3.7 Parameter Sensitivity.	23
3.8 Model Performance	26
3.9 Software description	29
4 Application of the four recharge models to three observation wells	31
4.1 Shallow observation well B27D00010.	31
4.2 Medium deep observation well B27C00490	39
4.3 Deep observation well B33A01130	43
4.4 Separate percolation and saturated zone responses.	47
4.5 Discussion	55

5 Additional explaining stresses	61
5.1 Anticipated influences on the groundwater levels	61
5.2 Shallow observation well B27D00010	62
5.3 Medium deep observation well B27C00490	69
5.4 Deep observation well B33A01130	71
5.5 Discussion	75
6 Conclusions	79
6.1 Introduction	79
6.2 Does a conceptual root zone model improve the model performance?	79
6.3 Does the use of a separate response function for the percolation zone improve model performance?	80
6.4 Can model performance be improved by adding stresses?	80
7 Recommendations	83
7.1 The use of local precipitation data	83
7.2 Validating the recharge simulation	83
7.3 Is the percolation zone linear?	84
7.4 Improving process consistency	84
7.5 The need for open source flexible models	85
Bibliography	87
A Derivations	91
A.1 Objective function revisited	91
A.2 Solving the soil model	92
A.3 Time to peak and the impulse response function	93
B Tables	95
B.1 Chapter 4	95
B.2 Chapter 5	98
C Figures	103
C.1 Chapter 4	103
C.2 Chapter 5	116

LIST OF FIGURES

2.1	Overview map of the study area and the location of the observation wells	6
2.2	Observed groundwater levels	7
2.3	Standardized observed groundwater levels	7
2.4	Photos of the environment around the observation wells	8
2.5	Meteorological data for the study area	11
3.1	General Modelling Strategy	14
3.2	Schematic figure of the unsaturated zone	15
3.3	The conceptual root zone models	16
3.4	Example of impulse, step and block response of a unit input	19
3.5	Example of the model residuals and the innovations	20
3.6	The impulse response function for different values of A , a and n	24
3.7	Effect of the β parameter on the recharge coefficient C_r	25
3.8	Graphical representation of the concepts of time to peak T_{peak} and the shift in the time to peak ΔT_{peak}	29
4.1	Simulated groundwater levels using four recharge models and a 30-year calibration period for well B27D00010.	32
4.2	Residual series for all four models calibrated on a 30-year period for well B27D00010.	33
4.3	The recharge time series for the linear, preferential, percolation and combination model for observation well B27D00010 in the period 1997-2004.	34
4.4	Average yearly recharge fluxes and the 2.5% and the 97.5% percentiles (the whiskers) to provide information on the uncertainty in the recharge calculation.	36
4.5	Simulated groundwater levels using four recharge models and a 20-year calibration period for well B27D00010.	37

4.6	Residual series for the four recharge models for observation well B27D00010 calibrated on a 20 year period	38
4.7	Time steps between the observations for observation well B27D00010.	38
4.8	Observed and simulated groundwater levels for observation well B27C00490 using four recharge models. The vertical dashed line denotes the end of the calibration period (1974-2004).	40
4.9	Residual series for all four recharge models calibrated on a 30-year period for observation well B49C00490.	40
4.10	Observed and simulated groundwater levels for observation well B27C00490 using four recharge models. The vertical dashed line denotes the end of the calibration period (1974-1995).	42
4.11	Residual series for all four recharge models calibrated on a 20-year period for observation well B27C00490.	43
4.12	Observed and simulated groundwater levels for observation well B33A01130 using four recharge models. The vertical dashed line denotes the end of the calibration period (1974-2004).	45
4.13	Residual series for all four recharge models calibrated on a 30-year period for observation well B33A01130.	45
4.14	Observed and simulated groundwater levels for observation well B33A01130 using four recharge models. The vertical dashed line denotes the end of the calibration period (1974-1995).	46
4.15	Residual series for all four recharge models calibrated on a 20-year period for observation well B33A01130.	47
4.16	Observed and simulated groundwater levels of the four recharge models with a double-response function for observation well B27D00010. The vertical dashed line denotes the end of the calibration period (1974-2004).	48
4.17	Observed and simulated groundwater levels of the four recharge models with a double-response function for observation well B27D00010. The vertical dashed line denotes the end of the calibration period (1974-1995).	48
4.18	Step (left) and block response (right) for the percolation zone for the models calibrated on a $\sqrt{30}$ -year period.	49
4.19	Observed and simulated groundwater levels for observation well B27C00490 using four recharge models with a double-response function. The vertical dashed line denotes the end of the calibration period (1974-2004).	50
4.20	Observed and simulated groundwater levels for observation well B49C00490 using four recharge models with a double-response function. The vertical dashed line denotes the end of the calibration period (1974-1995).	50

4.21	Observed and simulated groundwater levels for observation well B33A01130 using four recharge models with a double-response function. The vertical dashed line denotes the end of the calibration period (1974-2004).	54
4.22	Observed and simulated groundwater levels for observation well B33A01130 using four recharge models with a double-response function. The vertical dashed line denotes the end of the calibration period (1974-1995).	54
5.1	Observed and simulated groundwater levels and the model residuals for observation well B27D00010 simulated for the period 1960-2004.	63
5.2	Simulated groundwater levels, residuals, and the drawdown for all four recharge models with the effect of the reclamation for well B27D00010. The models are calibrated on the period 1960-2004.	65
5.3	Discharges from the Vitens Pumping wells in Haere and Epe	66
5.4	Simulated groundwater levels, residuals, and the drawdown for all four recharge models with the effect of extractions for well B27D00010. The models are calibrated on the period 1960-2004.	68
5.5	Simulated declines due to the reclamation of Zuidelijk Flevoland in 1967 for the models calibrated on a 30-year period.	70
5.6	Simulated decline of the groundwater levels due to the reclamation of Zuidelijk Flevoland in 1967 for the model calibrated on a 30-year period.	72
5.7	Discharge from the Vitens Pumping well at the Amersfoortseweg	73
5.8	Simulated groundwater levels, residuals, and the drawdown for all four recharge models with the effect of extractions for well B33A01130. The models are calibrated on the period 1974-2004.	74
C.1	The recharge time series for the linear, preferential, percolation and combination model for observation well B27D00010 in the period 1997-2004. The calibration period is 1974-1995.	104
C.2	Average yearly recharge fluxes and the 2.5% and the 97.5% percentiles (the whiskers) to provide information on the uncertainty in the recharge calculation.	105
C.3	The recharge time series for the linear, preferential, percolation and combination model for observation well B27C00490 in the period 1997-2004. The calibration period is 1974-2004.	106
C.4	Average yearly recharge fluxes and the 2.5% and the 97.5% percentiles (the whiskers) to provide information on the uncertainty in the recharge calculation.	107
C.5	The recharge time series for the linear, preferential, percolation and combination model for observation well B27C00490 in the period 1997-2004. The calibration period is 1974-1995.	108

C.6	Average yearly recharge fluxes and the 2.5% and the 97.5% percentiles (the whiskers) to provide information on the uncertainty in the recharge calculation.	109
C.7	The recharge time series for the linear, preferential, percolation and combination model for observation well B33A01130 in the period 1997-2004. The calibration period is 1974-1995.	110
C.8	Average yearly recharge fluxes and the 2.5% and the 97.5% percentiles (the whiskers) to provide information on the uncertainty in the recharge calculation.	111
C.9	The recharge time series for the linear, preferential, percolation and combination model for observation well B33A01130 in the period 1997-2004. The calibration period is 1974-1995.	112
C.10	Average yearly recharge fluxes and the 2.5% and the 97.5% percentiles (the whiskers) to provide information on the uncertainty in the recharge calculation.	113
C.11	Residual series for the models with a double response function calibrated on the period 1974-2004 for well B27D00010.	114
C.12	Residual series for the models with a double response function calibrated on the period 1974-1995 for well B27D00010.	114
C.13	Residual series for the models with a double response function calibrated on the period 1974-2004 for well B27C00490.	114
C.14	Residual series for the models with a double response function calibrated on the period 1974-1995 for well B27C00490.	115
C.15	Residual series for the models with a double response function calibrated on the period 1974-2004 for well B33A01130.	115
C.16	Residual series for the models with a double response function calibrated on the period 1974-1995 for well B33A01130.	115
C.17	Simulated groundwater levels, residuals, and the drawdown for all four recharge models with the effect of the reclamation for well B27D00010. The models are calibrated on the period 1960-1995.	117
C.18	Simulated groundwater levels, residuals, and the drawdown for all four recharge models with the effect of extractions for well B27D00010. The models are calibrated on the period 1960-1995.	118
C.19	Simulated groundwater levels, residuals, and the drawdown for all four recharge models with the effect of the reclamation for well B27C00490. The models are calibrated on the period 1974-1995.	119
C.20	Simulated groundwater levels, residuals, and the drawdown for all four recharge models with the effect of the reclamation for well B33A01130. The models are calibrated on the period 1974-1995.	120

C.21 Simulated groundwater levels, residuals, and the drawdown for all four recharge models with the effect of extractions for well B33A01130. The models are calibrated on the period 1974-1995. 121

LIST OF TABLES

2.1	Vegetation types and lithology around the observation well	6
2.2	Overview of the characteristics of the studied observation wells	7
3.1	Overview of the parameters, types and constraints	23
3.2	Overview of the time to peak (T_{peak}) and the ΔT_{peak} is the expected shift in peak compared to the T_{peak} of borehole B27D0001.	28
4.1	Overview of the parameters values for the four recharge models for observation well B27D00010	33
4.2	Overview of the parameters values for the four recharge models calibrated on two different periods for observation well B27C00490. *To obtain reasonable results the parameter α is constrained to 1000 days.	41
4.3	Overview of the parameters values for the four recharge models calibrated on two different periods for observation well B33A01130.	44
4.4	Overview the impulse response parameters values for all recharge models with a double-response function calibrated on a 30-year period for three observation wells.	52
4.5	Diagnostic statistics for the four recharge models with a single response function, calibrated on the period 1974-1994 and validated from 1995-2004.	53
4.6	Summary table of the diagnostic statistics provided in table B.2 and B.1. The values constitute the number of times a model has the highest performance for a statistic. . .	56
4.7	Overview of the average annual recharge [m/year] calculated for each model calibrated on a 30 and a 20-year period using a single response function. The length of the calibration period is added in subscript to the model name.	57
4.8	Time to peak (T_{peak} in days) for all models with a single and a double response function calibrated on two different periods. The length of the calibration period is added in subscript to the model name.	58
4.9	Diagnostic statistics for the four recharge models using a single (1 IRF) or a double (2 IRF) response function, calibrated on a 20-year period.	59

B.1	Diagnostic statistics for the four recharge models using a single (1 IRF) or a double (2 IRF) response function, calibrated on a 30-year period. The best diagnostic statistics per model are highlighted in grey.	96
B.2	Diagnostic statistics for the four recharge models using a single (1 IRF) or a double (2 IRF) response function, calibrated on a 20-year period. The best diagnostic statistics per model are highlighted in grey.	97
B.3	Diagnostic statistics for the four recharge models using a single response function calibrated on a 30-year period, for well B27D00010.	98
B.4	Diagnostic statistics for the four recharge models using a single response function calibrated on a 20-year period, for well B27D00010.	99
B.5	Diagnostic statistics for the four recharge models using a double response function calibrated on a 30-year period, for well B27C00490.	99
B.6	Diagnostic statistics for the four recharge models using a double response function calibrated on a 20-year period, for well B27C00490.	100
B.7	Diagnostic statistics for the four recharge models using a double response function calibrated on a 20-year period, for well B33A01130.	100
B.8	Diagnostic statistics for the four recharge models using a double response function, calibrated on a 30-year period, for well B33A01130.	101

1

INTRODUCTION

For a long time, groundwater modelling has been dominated by the application of numerical models (e.g. MODFLOW (Harbaugh, 2005)) to investigate and understand groundwater systems. This approach is based on the physics of the system and described by (partial) differential equations. A fundamentally different approach is that of system identification (Freeze and Harlan, 1969). In system identification, it is tried to explain time series of observed data (e.g., groundwater levels) by measured input data (e.g. evaporation and precipitation) through a mathematical model. This approach does not require the explicit description of the flow domain and its boundary conditions. The downside of system identification is the absence of spatially distributed results. The time to develop a model, however, is reduced and allows for a relatively easy and economical analysis of groundwater levels.

Time series analysis became widespread with the publication of a book by Box and Jenkins (1970) and more specifically in water-related applications through the book of Hipel and McLeod (1994). Application of the methods of time series analysis to groundwater modelling started in the early '80s (e.g., Besbes and De Marsily (1984)). Initially, most time series models were based on autoregressive models as described by Box and Jenkins (e.g., Gehrels et al. (1994), Bierkens et al. (1999), Van Geer et al. (1991)). These models were later extended with Kalman filters to deal with irregular time steps (Yi and Lee, 2004) and to quantify uncertainties (van Geer and Zuur, 1997).

Another class of times series models was developed with some basis in the physics of groundwater systems. Maas (1994) showed how the groundwater response relates to the well-known Gamma or Pearson III distributions, and Knotters and Bierkens (2000) showed the physical basis of autoregressive models. von Asmuth et al. (2002) presented a method to effectively deal with the irregular time intervals of groundwater level observations through application of predefined impulse response functions, and developed a new objective function to use in the parameter estimation process (von Asmuth and Bierkens, 2005). This work was implemented in the software package Menyanthes (von Asmuth et al., 2012), making the tools of time series analysis available to a wide public. The techniques for time series analysis used in Menyanthes have been applied successfully in many case studies. There is evidence, however, that not all groundwater systems can be modelled efficiently with the existing methods.

One of the fundamental assumptions underlying many existing methods, is the linear response of the groundwater level to recharge, independent of the system state. Moreover, this recharge is

calculated as a linear function of precipitation and (potential) evapotranspiration. This linearity assumption is not valid for all systems. Different sources of non-linear behaviour have been observed by Knotters and Bierkens (2000) and Besbes and De Marsily (1984). The first source of non-linearity stems from a change in the boundary conditions of the system, causing a different system response. This type of behaviour can be accounted for through the implementation of threshold non-linearities, applying different response functions depending on the system state (Knotters and Bierkens, 2000). A second source of non-linear behaviour, the focus of this study, stems from the generation of recharge. A common definition of recharge, that is also applied in this study, is as follows (Fitts, 2013):

"Water in the unsaturated zone that moves downward and flows into the saturated zone"

It is clear from this definition that the recharge flux is dependent on the unsaturated zone, making the relationship between precipitation, evaporation, and recharge dependent on the system state. When the unsaturated zone is thin and the distance to the groundwater table is small, the non-linear effect of the unsaturated zone is negligible, the retardation and dispersion effect of the unsaturated zone can be modelled using a Pearson III distribution as the impulse response function. The linearity assumption appears to be invalid when the unsaturated zone is thick (Gehrels et al., 1994). For these cases, it has been suggested to apply an unsaturated zone model to calculate the recharge time series (Gehrels et al., 1994; Yi and Lee, 2004). Berendrecht et al. (2006) used a conceptual unsaturated zone model in combination with an autoregressive time series model to simulate deep groundwater levels in the Veluwe, improving the simulating of peaks. More recently, Peterson and Western (2014) applied an unsaturated zone model based on the VIC model (Wood et al., 1992) in combination with predefined impulse response functions to a site in Australia with deep ground water levels.

Along with these recent advances in groundwater time series modelling, hydrologists studying rainfall-runoff processes face similar challenges in dealing with the unsaturated zone. Rainfall-runoff models are usually calibrated on observed river discharges, with few possibilities to validate the internal processes of a model. As groundwater is an important source of water for the river discharge, it has been suggested to use groundwater levels to check the internal processes and further constrain model parameters (Seibert, 2000; Fenicia et al., 2008). The use of groundwater levels could improve parameter estimation in hydrological models, and De Weerd et al. (2014) even tried to use time series analysis of groundwater levels to constrain parameters. All these studies were all performed with a focus on simulating river discharges, and linear time series models to simulate groundwater levels.

The objective of this study is to improve the simulation of deep groundwater levels by time series models with pre-defined impulse response functions. This is attempted by adding a conceptual non-linear root zone model to simulate the recharge series to the model and by testing the use of a separate response function for the percolation zone. Different recharge mechanisms are conceptualized based on the FLEX models commonly applied in rainfall-runoff simulations (Fenicia et al., 2006). For the calculation of the groundwater levels, a time series model with pre-defined impulse response functions is applied. The focus of this study is on the simulation of groundwater levels. The conceptualisation of the unsaturated zone however, is envisioned to allow rainfall-runoff modellers to apply the results and methods of this study in future research. The model will be applied to an area in the Netherlands with thick unsaturated zones and deep groundwater levels, known as 'De Veluwe'. The main research question is:

"Can the simulation of deep groundwater levels be improved by adding a conceptual root zone model and an alternative impulse response function?"

The following three subquestions are considered:

1. Does a conceptual root zone model improve the model performance?
2. Does the use of a separate response function for the percolation zone improve model performance?
3. Can model performance be improved by adding stresses?

This report is structured as follows. The study area of this research, its meteorological conditions and the characteristics of the observation wells are discussed in chapter 2. In the third chapter, the methodology is presented. This involves a review of the concepts of time series analysis, the conceptualization of the unsaturated zone, and the methods used for calibration and validation of the model. The results of the different recharge models and impulse response functions are presented in chapter 4. In the fifth chapter, it is tried to improve model performance by adding different trends. The conclusions of this research and the answer to the research questions are presented in chapter 6. Recommendations for future research are given in the final chapter.

2

STUDY AREA 'THE VELUWE'

2.1. TOPOGRAPHY & GEOLOGY

The study area for this research is the Veluwe, a hilly area in the east of the Netherlands that forms one of the main groundwater bodies in the country (see Figure 2.1 later in this chapter). Covering over 1000 km², it serves as an important recharge area and source of drinking water. The groundwater table is located at a depth of 2 m below the surface at the borders to up to 40 m in the center of the Veluwe. The system drains to the river IJssel on the East, Lake Veluwe on the Northwest and the Rhine to the South. The groundwater system is characterized by long travel times ranging from decades to centuries. The largest part of the groundwater system is unconfined, except at the borders and at places where boulder clay layers cap the aquifer.

The subsurface of the Veluwe area consists of fluvial and marine sands and clays of Pleistocene age. Glacial conditions formed the ice-pushed ridges we now call the Veluwe during the Saalien. These ridges primarily contain reworked pre-Saalien deposits (sand and clay) along with glacial (boulderclay) and glaciofluvial deposits (sand, clay and gravel) (Berendsen, 2008). The boulderclay has been deposited in dipping layers in the sandy sediments, forming barriers for groundwater flow and recharge alike (Goes, 2000; Bakker and Meer, 2003). Aeolian sands were deposited on top of part of the Veluwe during the second ice-age, the Weichselien. Most of the study area consists of sand with little heterogeneity and a high hydraulic conductivity, despite its relatively complex geological history.

2.2. SELECTION OF THE OBSERVATION WELLS

The groundwater levels in the Veluwe have been monitored for a long period, at various locations and depths, due to its importance for drinking water supply. Three observation wells have been selected for this study, based on two selection criteria. The first criterion is a monitoring period of at least 30 years, ensuring a long enough time series for the slow-responding system. In the preliminary phase of this study, it was found that the response times of the system reached up to 10 years. The somewhat arbitrary value of two times the response time is used as a minimal calibration period. It is stressed that no formal investigation is performed in this study on how long the minimum calibration period should be. The second criterion is to cover a range of different thicknesses

Name	Vegetation	Lithology
B27D0001	Deciduous trees, grasses	intermediate fine to coarse sand with occasional gravel
B27C0049	Deciduous trees, shrubs (heath fields and coniferous trees <200 m)	intermediate fine to coarse sand with occasional gravel
B33A0113	Deciduous trees (shrubs nearby <50 m)	coarse to very coarse sand

Table 2.1: Overview of the vegetation types around the wells that were observed during the field visit and the lithology found in the Dinoloket database.

of the unsaturated zone, to be able to investigate the effect of the thickness of the unsaturated zone on time to peak of the groundwater response.

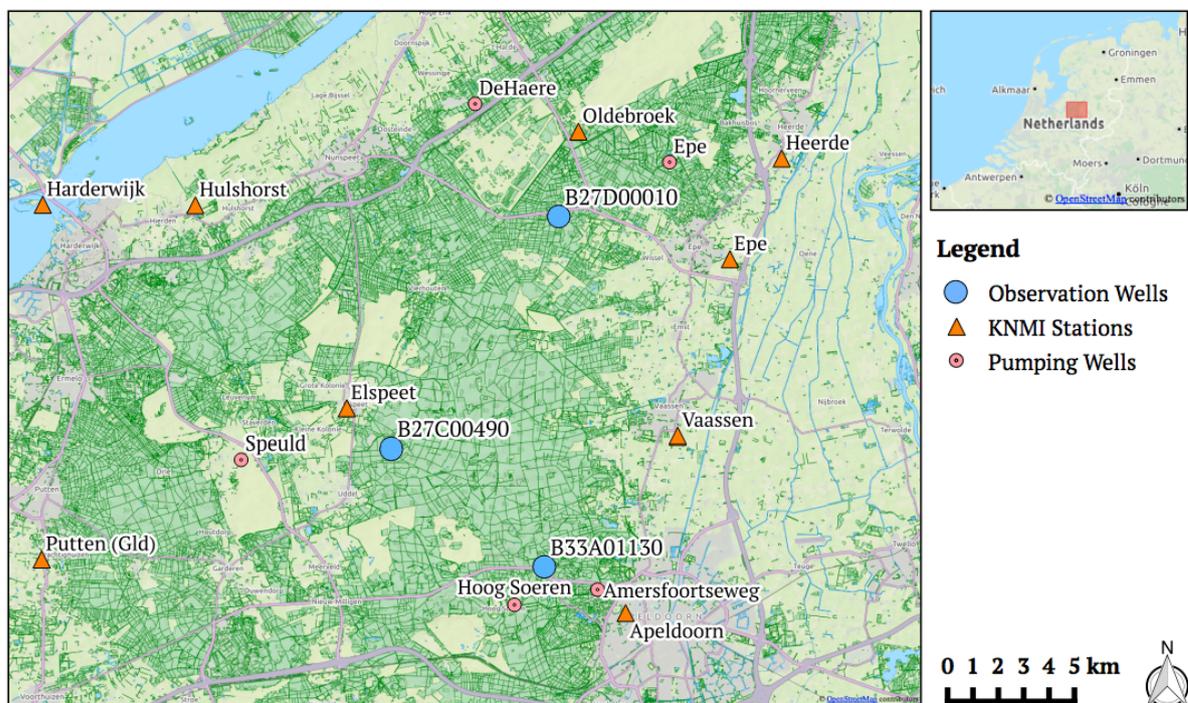


Figure 2.1: Location of the observation in the veluwe and the weather station used in this study.

A field visit to the selected observation wells was conducted on the 24th of December 2015 with the primary goal to observe the vegetation present at these locations. Information on the sub-surface in the Netherlands is centrally stored and publicly available through Dinoloket (<https://www.dinoloket.nl>). Groundwater levels and lithological data for all observation wells has been downloaded from this website. Tables 2.1 and 2.2 give an overview of the characteristics of the selected observation wells. The observed groundwater levels for the observation wells are shown in Figure 2.2. The thickness of the unsaturated zone is estimated by subtracting the maximum observed groundwater level from the ground surface level, and hence constitutes a minimum thickness.

Name	Code	Surface level	Filter level (bottom/top)	level	Mean level [m]	Variation [m]	Minimum thickness unsaturated zone [m]
Shallow	B27D00010	27.56	13.66/14.66		19.3	2.5	7.0
Medium	B27C00490	47.23	-6.60/-4.60		16.8	2.9	29.0
Deep	B33A01130	78.37	-2.80/1.20		28.1	3.0	49.0

Table 2.2: Overview of the characteristics of the observation wells. All levels are reported with reference to the Dutch National Datum (+NAP).

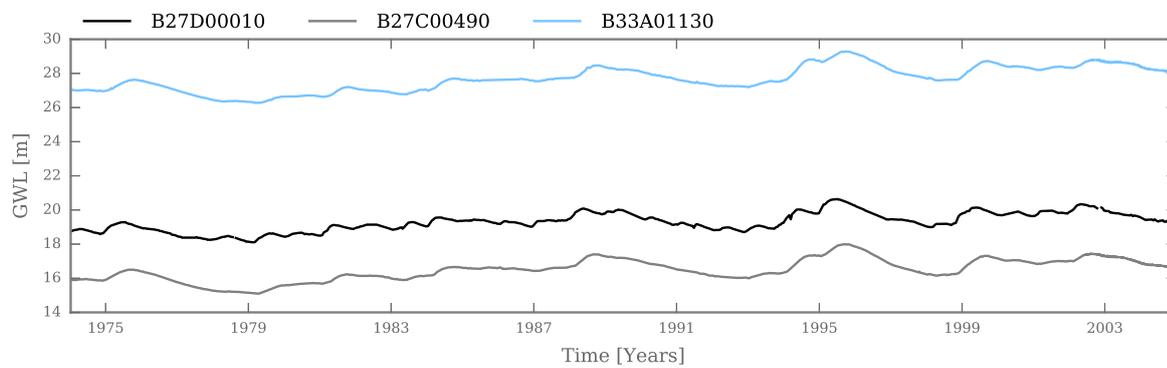


Figure 2.2: Observed groundwater levels for the selected observation wells. The groundwater levels are measured with reference to +NAP.

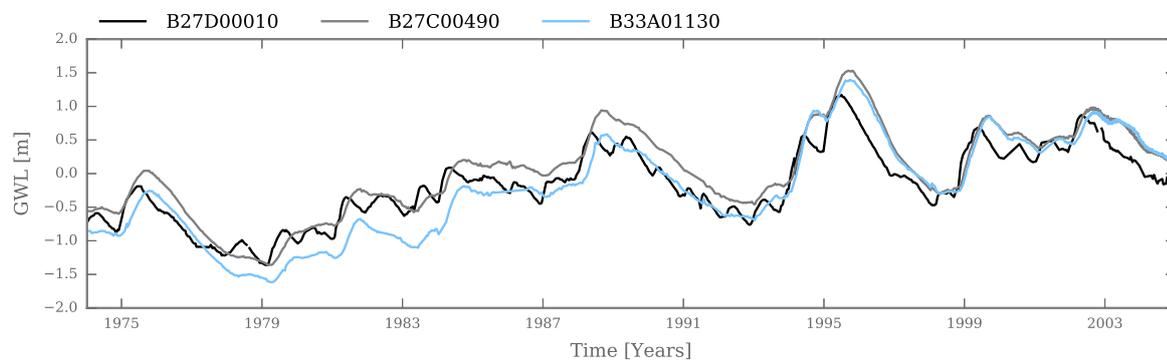


Figure 2.3: Standardized groundwater levels where the mean is subtracted from the observed groundwater level.



Figure 2.4: Photos a and b were taken near the shallow observation well. The actual observation well can be seen in the right bottom corner of photo b. Photos c and d were taken near the medium deep observation well and the actual observation well can be seen in the bottom left corner of photo d. Photos e and f were taken at the reported location of the deep observation well. No observation well was found at this location.

2.2.1. B27D00010: THE SHALLOW WELL

Observation well B27D00010 is located in the northern part of the Veluwe and has the thinnest unsaturated zone of the three observation wells with a minimum thickness of 7 m (see Table 2.2). In the remainder of this report there is referred to B27D00010 as the "shallow observation well". The environment around the observation well is shown in Figures 2.4a and 2.4b. The vegetation consists primarily of deciduous trees (primarily beech), but the surrounding area is also characterized by some fields of grass. Nearby (<200 m) are extensive plots of coniferous trees and open fields with heath. The lithology that was investigated based on samples taken during drilling, consists of intermediate fine to very coarse sands.

2.2.2. B27C00490: THE MEDIUM DEEP WELL

Observation well B27C00490 is located in the centre of the Veluwe close to the village of Elspeet. The minimum thickness of the unsaturated zone is estimated at 29 m (see Table 2.2) and there is referred to this observation well as the "medium deep well" in the remainder of this report. The environment around the observation well is shown in Figures 2.4c and 2.4d. The vegetation consists of deciduous trees (beech and birch), shrubs and an incidental coniferous tree. The lithology consists primarily of intermediate fine to very coarse grained sand. At 13.0-14.0 m below the ground surface a sandy clay layer is present. The exact thickness of this layer can not be given due to the coarse sampling interval during drilling.

2.2.3. B33A01130: THE DEEP WELL

Observation well B33A01130 is located east of the centre of the Veluwe close to the city of Apeldoorn. The minimum thickness of the unsaturated zone is estimated at 49 metres (see Table 2.2) and in the remainder of this report there is referred to this observation well as the "deep well". Figures 2.4e and 2.4f show the environment where the observation well was located. The observation well could not be found at the indicated location during the field visit. The responsible organization (Vitens) confirmed the removal of the observation well after it exploded, possibly due to a lightning strike (H. Willemsz, personal communication, 20-01-2016). The vegetation consists of deciduous trees, with some shrub vegetation nearby (<50 m). There is quite some relief in the area, most notably in the Northern direction where the elevation drops around 10 m in less than 50 m. The lithology of the subsurface is characterized by coarse to very coarse sand.

2.2.4. COMPARISON OF THE OBSERVATION WELLS

The vegetation at all locations consists primarily of deciduous trees, with local deviations in the forms of grasses, heath fields, shrubs, and coniferous trees. The lithology is similar for all observation wells, mainly consisting of sand with a variety of grain sizes. The minimum thickness of the unsaturated zone differs between observation wells as a result of the second selection criterion and ranges from 7 metres to 49 m. The trend in the groundwater level is similar for all observation wells, as is shown in Figure 2.3 where the mean groundwater level is subtracted from the observed levels. The dispersive effect of the unsaturated zone can be observed by comparing the fluctuations at shallow observation well and the deep observation well, where the groundwater levels of the latter are smoother and show virtually no yearly variations.

Looking at Figure 2.3 in more detail and comparing the groundwater levels shows that three periods can be identified where the level deviate from one another. In the first period, 1975-1985, the levels of the shallow and medium deep wells follow each other closely, while the levels of the deep well are consistently lower. In the second period, 1985-1994, the levels for the medium deep well are consistently higher, while the levels of the other two wells follow each other closely. In the third period, 1994-2005, the medium and deep well follow each other closely and the levels of shallow well are consistently lower. These local and temporal deviations may provide an indication of processes that need to be accounted for in the modelling process. It is interesting to see if these deviations are also found in the simulation results when not accounting for any additional processes apart from precipitation and evaporation.

2.3. METEOROLOGICAL DATA

Meteorological data is collected by the Royal Netherlands Meteorological Institute (KNMI) at multiple locations in the Netherlands. Precipitation data from the weather stations nearest to the observations wells is used: station Oldebroek 3 km from the shallow observation well, station Elspeet 2 km from the medium deep well, and station Apeldoorn 4 km from the deep observation well. In Figure 2.5 the total annual precipitation for each weather station for the period 1958-2004 is shown. The yearly precipitation differs for each location from 863 mm/year in Oldebroek to 915 mm/year in Elspeet and, 914 mm/year in Apeldoorn. The precipitation does not only differ in the amount but also in the rainfall pattern.

The evaporation data is taken from the KNMI weather station de Bilt, as no long sufficiently long time series were available at weather station nearby. The weather station in de Bilt is located 51 km from the deep observation well, 47 km from the the medium deep observation well and 57 km from the shallow observations well. The average annual evaporation is 543 mm/year calculated over the period 1958-2004. In Figure 2.5 the total annual evaporation in de Bilt are shown. The KNMI reports the Makkink reference evaporation (ET_{ref}), the evaporation of grass when sufficiently watered. The potential evaporation for other crops or vegetation types is calculated as follows:

$$ET_{pot} = k_c \cdot ET_{ref} \quad (2.1)$$

where k_c is the crop factor that is vegetation specific. Deciduous trees was the main vegetation type that was observed around all three observation wells during the field visit. Different values for the crop factor k_c for deciduous forest are reported. The average of three reported values for deciduous forests $k_c = 1.0$ is applied in this study (Droogers, 2009). The crop factor is assumed to be constant over time in this study. The potential evaporation is used as the input data for all models. The actual evaporation is simulated by the root zone model and depends on the amount of intercepted water and soil moisture that is available for evaporation and transpiration.

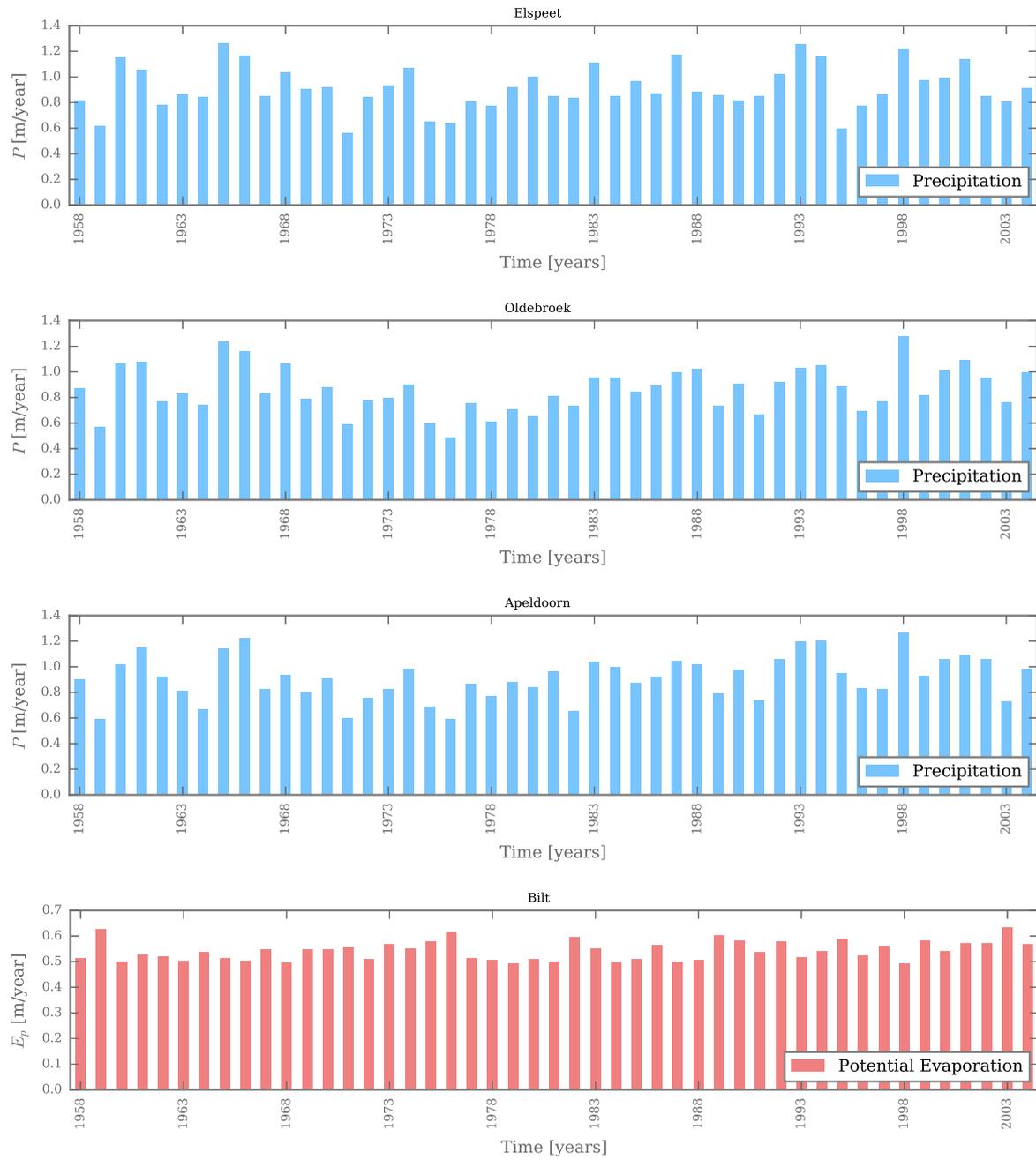


Figure 2.5: Precipitation time series from KNMI stations Elspeet, Oldebroek and Apeldoorn and reference evaporation calculated for KNMI weather station de Bilt.

3

METHODOLOGY

This chapter is structured as follows. The general modelling strategy is presented in the first section. The conceptualization of the unsaturated zone is discussed in the second section. In the third section the use of transfer functions in time series analysis is discussed, followed by a discussion on the noise model in the fourth section. In the fifth section the model calibration and the objective function that is used are discussed. The parameter sensitivities and ranges for plausible parameter values are discussed in the sixth section. Different measures that are used to test the model performance are presented in the seventh section. The final section of this chapter the software used in this study is introduced.

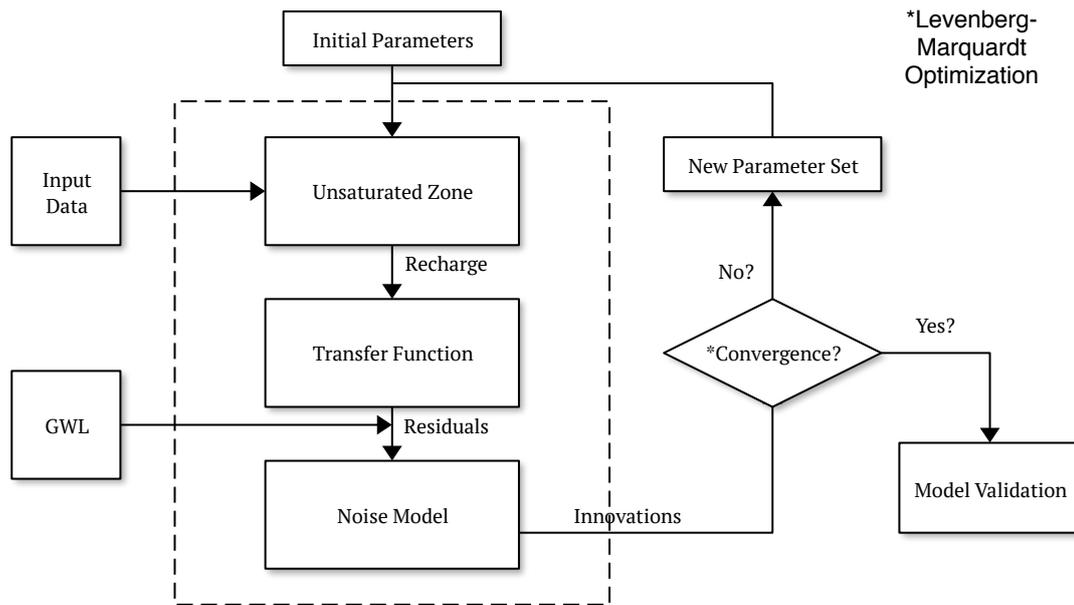


Figure 3.1: Schematization of the general modelling strategy that is used in this study.

3.1. MODELLING STRATEGY

Different conceptual models for the unsaturated zone are investigated to generate a recharge time series that serves as an input for the transfer function. The transfer function translates the recharge input to a time series of groundwater levels. The residuals (observed heads - modelled heads) are often correlated in time, hence in the final step a noise model is applied to obtain uncorrelated residuals called the innovations. The parameters of all three model components (unsaturated zone, transfer function, and noise model) are estimated simultaneously using the least squares method to minimize the innovations. This entire process of modelling groundwater levels is summarized in Figure 3.1.

3.2. THE UNSATURATED ZONE

3.2.1. THE LINEAR RECHARGE MODEL

One of the goals of this study is to test different conceptualizations of the unsaturated zone to generate recharge time series. These models are compared to a simple method for generating the recharge series that is commonly applied.

$$N(t) = P(t) - f \cdot ET_{pot} \quad (3.1)$$

Where $N(t)$ is the recharge, $P(t)$ is the precipitation, ET_{pot} the potential evaporation for that specific site (see Section 2.3), and f a parameter that is calibrated. This model is referred to in the remainder of this report as the linear recharge model.

3.2.2. GENERAL MODEL STRUCTURE

There is a wide range of models that can be applied to extend the time series model with an unsaturated zone model. The goal of the unsaturated zone model is to transform the precipitation and potential evaporation forcings to a recharge input for the transfer function. Experience in rainfall-runoff modelling is used to develop a conceptual model of the unsaturated zone. Rainfall-runoff modelling is concerned with simulating river discharges, but it also considers the non-linearities that result from the unsaturated zone. A minimal number of parameters is desired to prevent equifinality (Beven, 1993, 2006; Peterson and Western, 2014). The Flux Exchange (FLEX) hydrological model developed by Fenicia et al. (2006) is chosen as the starting point to conceptualize the unsaturated zone.

The FLEX model is aimed at simulating river discharges, and uses precipitation and potential evaporation as input data. The basic configuration consists of four reservoirs to account for the dominant processes in a catchment: an interception reservoir, an unsaturated reservoir, a slow reacting, and a fast reacting reservoir. Only the interception reservoir (IR) and the unsaturated soil reservoir (Sr) are used since we are interested in the recharge to the groundwater. The unsaturated reservoir is used to model the dynamics of the root zone, where the direction of flow can be both upward and downward. The flux coming out of the root zone as recharge travels through the percolation zone, where flow is directed downwards only. Typically in FLEX models, a lag function is applied to the recharge flux, representing the retardation and dispersion in the percolation zone. In this study, the retardation and dispersion effects are simulated by applying a scaled gamma function to simulate the groundwater response. The different zones as they are conceptualized in this study are presented in Figure 3.2.

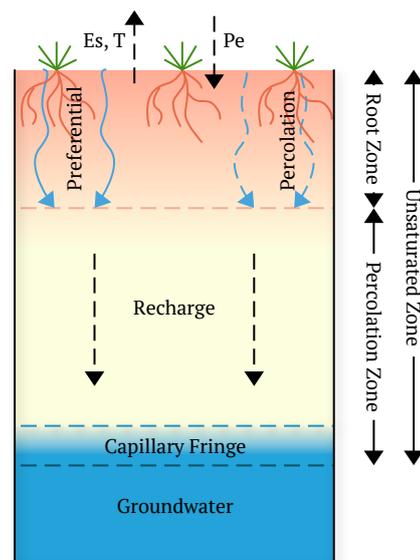


Figure 3.2: Schematic figure of the unsaturated zone

Different conceptualizations of the root zone are possible depending on which recharge mechanisms are perceived to be dominant. In this study area, two recharge mechanisms have been identified based on the analysis of the stable isotope ^{18}O , water percolating through the root zone, and water by-passing the root zone through preferential flow paths (Gehrels et al., 1998). The prevailing mechanism depends on local vegetation and soil characteristics. Three models were investigated (see Figure 3.3), one for each mechanism and one that combines the two mechanisms in one model. The implementation of these three recharge mechanisms in a mathematical model is discussed in the next two subsections.

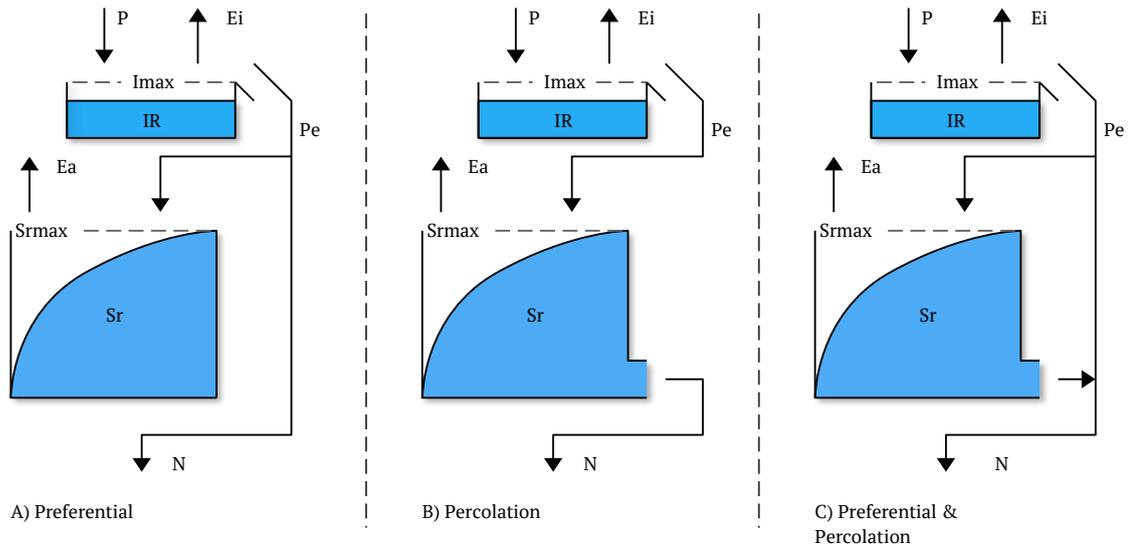


Figure 3.3: Three conceptual models, implementing different recharge mechanisms

3.2.3. INTERCEPTION RESERVOIR

A process that is often forgotten in conceptual hydrology is the interception of precipitation, causing threshold behaviour in the system (Savenije, 2004). Part of the precipitation is intercepted by plants and relief, from which it can evaporate. This evaporation flux from the interception reservoir is denoted by $E_i(t)$ [L/T] and cannot be larger than the amount of the intercepted water $I(t)$ [L] or the potential evaporation $E_p(t)$ [L/T].

$$E_i(t) = \min(I(t), E_p(t)) \quad (3.2)$$

The $E_i(t)$ is subtracted from $E_p(t)$ to get the updated potential evaporation $E_{pu}(t)$. Only if the interception capacity I_{max} [L/T] is exceeded, the excess water will continue down the system as effective precipitation $P_e(t)$ [L/T]. Values for I_{max} are dependent on local conditions in vegetation and relief, and typically lie in the order of 1-5 mm.

3.2.4. ROOT ZONE RESERVOIR

where $F(t)$ is the infiltration, $G(t)$ the groundwater entering through capillary rise, $TF_{in}(t)$ and $TF_{out}(t)$ the through flow, $T(t)$ the transpiration by plants, $E_s(t)$ the evaporation from the soil and finally the recharge to the groundwater, $N(t)$. The horizontal flow through the root zone is assumed to be negligible ($TF_{in} = TF_{out} = 0$), a valid assumption when lateral gradients are low as is the case in this study area. The upward flux $G(t)$ from the groundwater to the root zone is assumed to be absent ($G = 0$), motivated by the existence of percolation zones that are thicker than the water can reach through capillary rise (Wösten et al., 2001). Finally, since there is no individual data on transpiration and soil evaporation these two fluxes are truncated to an actual evaporation flux $E_a(t)$ [L/T]. The water balance for the root zone now becomes:

ROOT ZONE MODEL 1: PREFERENTIAL FLOW

In this model, recharge to the groundwater table occurs solely through preferential flow through the root zone. The hypothesis is that the soil retains water up to a point where these flow paths are saturated enough to drain water. Any excess water then 'shoots' along these pathways to the percolation zone, effectively bypassing the root zone reservoir. All water that does not add to the recharge will infiltrate into the root zone, where it can only leave via transpiration and soil evaporation (Eq. 3.4).

The transpiration and soil evaporation are modelled as one combined flux, the actual evaporation $E_a(t)$. The magnitude of the actual evaporation flux $E_a(t)$ depends on the soil moisture in the root zone relative to the storage capacity of the root zone S_{rmax} . This flux generally reaches its maximum amount when 50% of S_r is reached, limited by the updated potential evapotranspiration E_{pu} , so that $E_a(t)$ can be written as:

$$E_a(t) = E_{pu}(t) \min\left(1, \frac{S_r(t)}{0.5S_{rmax}}\right) \quad (3.3)$$

The recharge coming out of the root zone is directly related to the effective rainfall and the ratio of the storage S_r and the maximum root zone storage S_{rmax} to a power of β (eq. 3.5). The resulting equations for the root zone storage S_r and recharge $N(t)$ are:

$$\frac{dS_r(t)}{dt} = P_e(t) \left(1 - \left(\frac{S_r(t)}{S_{rmax}}\right)^\beta\right) - E_a(t) \quad (3.4)$$

$$N(t) = P_e(t) \left(\frac{S_r(t)}{S_{rmax}}\right)^\beta \quad (3.5)$$

ROOT ZONE MODEL 2: PERCOLATION

In this model, all effective precipitation infiltrates into the root zone, and percolates through the root zone to form recharge or evaporates from the root zone. In case the storage capacity of the root zone reservoir S_{rmax} is exceeded, excess precipitation will cause overland flow ($Q_r(t)$) and leave the system. The flux leaving the root zone as recharge is dependent on the percolation capacity K_p [L/T] (i.e. how easy the root zone percolates) and the ratio between the storage S_r and the maximum storage capacity S_{rmax} to a power γ . The resulting equations for the root zone storage S_r and recharge $N(t)$ are then:

$$\frac{dS_r(t)}{dt} = P_e(t) - K_p \left(\frac{S_r(t)}{S_{rmax}}\right)^\gamma - E_a(t) - Q_r(t) \quad (3.6)$$

The actual evaporation $E_a(t)$ is calculated with equation 3.3.

$$N(t) = K_p \left(\frac{S_r(t)}{S_{rmax}}\right)^\gamma \quad (3.7)$$

ROOT ZONE MODEL 3: COMBINATION

The previous two models have either preferential flow or percolation as a recharge mechanism. This model is a combination of the two. The resulting equations for the root zone storage S_r and recharge $N(t)$ are then:

$$N(t) = P_e(t) \left(\frac{S(t)}{S_{rmax}} \right)^\beta + K_p \left(\frac{S_r(t)}{S_{rmax}} \right)^\gamma \quad (3.8)$$

The final equation for the root zone reservoir is then as follows:

$$\frac{dS_r(t)}{dt} = P_e(t) \left(1 - \left(\frac{S_r(t)}{S_{rmax}} \right)^\beta \right) - K_p \left(\frac{S_r(t)}{S_{rmax}} \right)^\gamma - E_a(t) \quad (3.9)$$

The actual evaporation $E_a(t)$ is calculated with equation 3.3.

3.3. THE TRANSFER FUNCTION

A physical system that experiences an excitation, has a response to that excitation. In ground-water systems the groundwater head responds to (among others) the recharge $N(t)$. The response is dependent on the characteristics of the system. The transformation of input data (e.g. recharge) to output data (e.g. groundwater head) is modelled using a transfer function model in this study. The groundwater head is modelled as:

$$h(t) = h^*(t) + d + r(t) \quad (3.10)$$

where $h(t)$ is the observed head, h^* is the part of the groundwater head that can be explained by the input data, d is the base elevation for the model (assumed to be constant over time), and r is the residual error. The relation between input and output is modelled using an impulse response function, describing how the system responds to an excitation. This relation is assumed to be linear, meaning that twice the excitation gives twice the head response, regardless of the magnitude of the input or the value of the head. The principles of convolution can be applied to linear systems, and the responses of different excitations can be superimposed (Olsthoorn, 2008). The convolution integral for the part of the head explained by excitation $N(t)$ is:

$$h^*(t) = \int_{-\infty}^t N(\tau) \theta(t-\tau) d\tau \quad (3.11)$$

where h^* is the head [L], N is the input (e.g. recharge) [LT^{-1}], and θ is the impulse response function. The impulse response function is approximated by a scaled gamma distribution function, as proposed by Nash (1958), and shown to be appropriate in many groundwater cases (von Asmuth et al., 2002). In this study, the terms a^n and $\Gamma(n)$ are omitted from the gamma distribution as they only serve as scaling parameters (Peterson and Western, 2014). The adapted form of the scaled gamma distribution is:

$$\theta(t) = At^{n-1} e^{-\frac{t}{a}} \quad (3.12)$$

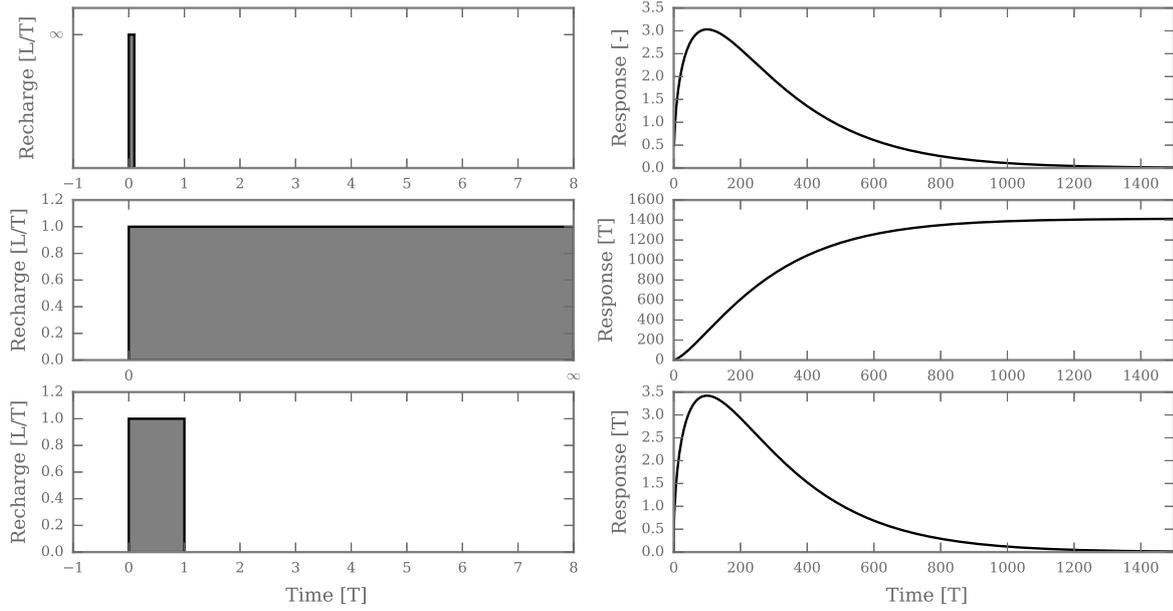


Figure 3.4: Example of impulse, step and block response of a unit input $N(t)$ to the groundwater heads ($A=10$; $a=5$; $n=1.5$).

where A , a and n are calibration parameters. The functions (3.11) and (3.12) are both in continuous time while precipitation and evaporation are measured in time intervals and can be seen as a block input. The block response is derived from the step response θ_s of the system, a sudden and infinite increase in the input (e.g. $N(t)$ in equation (3.13)).

$$N(t) = \begin{cases} 0 & \text{if } t < 0 \\ 1 & \text{if } t \geq 0 \end{cases} \quad (3.13)$$

$$\theta_s(t) = \int_0^t \theta(t-\tau) d\tau \quad (3.14)$$

$$\theta_s(t) = -At^n \left(\frac{t}{a} \right)^{-n} \Gamma \left(n, \frac{t}{a} \right) \quad (3.15)$$

The step response is equivalent to the integral of the impulse response function θ . The block response θ_b for a unit input (e.g. $N(t)$ in Equation (3.16)) in one time period Δt (e.g. 1 day), is obtained:

$$N(t) = \begin{cases} 0 & \text{if } 0 > t > \Delta t \\ 1 & \text{if } 0 \leq t \leq \Delta t \end{cases} \quad (3.16)$$

$$\theta_b(t) = \theta_s(t) - \theta_s(t - \Delta t) \quad (3.17)$$

Equation (3.17) gives the block response function that can be applied in the model. An example impulse, step and block response function are shown in Figure 3.4.

3.4. THE NOISE MODEL

The model described in the previous sections can be supplemented with other deterministic model components if needed (e.g. to account for groundwater pumping or step non-linearities). Despite a modellers' best effort to explain as much of the fluctuations in the groundwater levels as possible with the deterministic model, there will always be a residual series left. The source of this residual can be subdivided in three categories. The first is errors in the input data, caused, for example, by a rain gauge that is located at some distance from the actual study area. The second source of error is the system noise, caused by the model structure and the parametrisation of the model. The third source of error is the measurement error of the dependent variable, the groundwater levels in this case. In practice it is difficult to identify these individual errors from the residual series.

The residuals that result from the different sources of errors are often found to be correlated in time which means that part of the residual error can be explained by the error at the previous (von Asmuth and Bierkens, 2005). For the prediction of confidence intervals, the residuals should be uncorrelated with a mean of zero and be homoscedastic (constant variance). A noise model is applied to obtain uncorrelated errors. An additional advantage is that the noise model can be used to provide information on time steps where there are no measurements (von Asmuth and Bierkens, 2005). An exponential decay model of the error is applied as the noise model:

$$r(t) = r(t - \Delta t)e^{-\frac{\Delta t}{\alpha}} + v(t) \quad (3.18)$$

$$v(t) = r(t) - r(t - \Delta t)e^{-\frac{\Delta t}{\alpha}} \quad (3.19)$$

where r is the residual error, α [T] is a calibration parameter, and v is the innovation. Hence, the error $r(t)$, is a result of the error at time $t - \Delta t$ that decays exponentially with time, plus a random part, the innovation $v(t)$. These innovations form an uncorrelated time series that can be used for model calibration if the noise model works appropriate. It needs to be checked whether the innovation series should have a mean of zero, a fixed variance and, zero auto-correlation. Figure 3.5 gives an example of some residuals that are autocorrelated in time, and the innovations that result after the application of an exponential decay noise model.

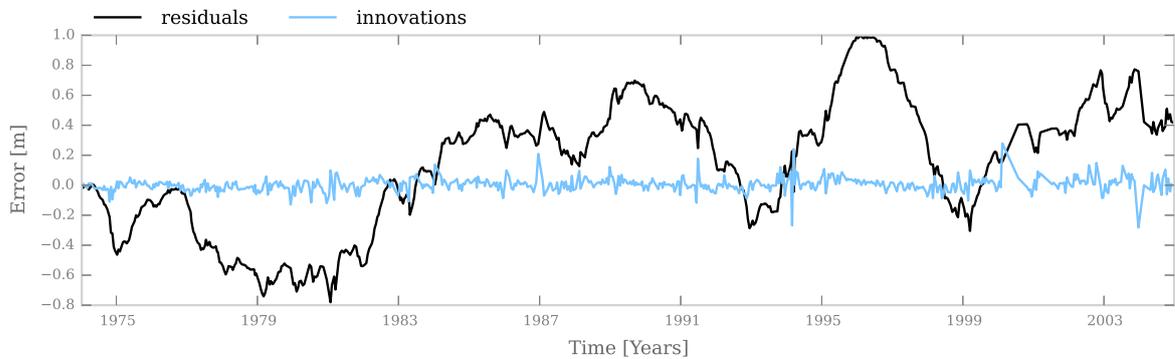


Figure 3.5: Example of the model residuals $r(t)$ (autocorrelated in time) and the innovations $v(t)$.

3.5. A SEPERATE RESPONSE FOR THE PERCOLATION ZONE

In the previous section an impulse response function was introduced that is commonly used in time series analysis. While many different forms of this function have been applied, see for example Peterson and Western (2014), Bakker et al. (2008) and Besbes and De Marsily (1984), they all resemble some form of the Gamma distribution with three parameters. This form of the impulse response function has been shown to relate to the physics of the system (von Asmuth et al., 2002; Besbes and De Marsily, 1984; Maas, 1994), and has proven useful in many applications. However, despite these successful applications other forms of impulse response functions remain an option. In this section we explore different possible impulse response functions that are tested in this study, with the main goal of testing their performance for boreholes with thick unsaturated zones.

The unsaturated zone model employed in this study simulates the dynamics that occur in the root zone, and is therefore better described as a root zone model. The recharge that drains from the root zone flows through the percolation zone, where the flux is dispersed and retarded. Ignoring all non-linear behaviour in the percolation zone, the retardation and dispersion can be described by an impulse response function that is represented by the probability density function of the normal distribution. This impulse response function θ_p can be described by two variables, μ and σ , and has the following form:

$$\theta_p(t \geq 0) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(t-\mu)^2}{2\sigma^2}} \quad (3.20)$$

Equation (3.20) is used for the first convolution on the recharge that results from the root zone model. The recharge flux that results from this convolution represents the recharge at the point where the unsaturated zone connects to the saturated zone. The response of the saturated zone can be described by an instantaneous increase in the groundwater level, followed by an exponential decay. The following function is used for the impulse response function:

$$\theta_g(t) = Ae^{-\frac{t}{a}} \quad (3.21)$$

Equation (3.21) is the same as (3.12), when parameter $n = 1$. Equations (3.20) and (3.21) are both impulse response functions. This model structure includes one additional parameter to be calibrated compared to model introduced in the previous section. There is referred to the combination of Equations (3.20) and (3.21) as 'the double response function' in the remainder of this report.

3.6. MODEL CALIBRATION

The parameters discussed in the previous section can not be measured and need to be estimated from calibration. Traditionally this was done by trial-and-error where the parameters were adjusted manually to improve the model fit. The model fit is quantified with an objective function (discussed in the following section), a measure of how well the model performs. The objective function is minimized by an iterative approach until it converges to a (mathematical) optimum. The methods used for the optimization are discussed in the second section.

3.6.1. OBJECTIVE FUNCTION

A common approach for parameter estimation is to minimize an objective function based on the model residuals. These residuals are often assumed to be independent, identically distributed according to a normal distribution with mean zero, and a constant variance σ^2 (Schoups and Vrugt, 2010). These simplifying assumptions then lead to a standard least squares method for parameter optimization, like root mean squared error or a Nash-Sutcliffe objective function. However, time series of groundwater levels often violate these assumptions and the standard least square method can not be applied straightforwardly.

von Asmuth and Bierkens (2005) developed an objective function, dubbed the sum of weighted squared innovations (SWSI), that deals with non-equidistant observations. This criterion assumes a Gaussian distribution of the innovations and parameter estimates obtained with the SWSI criterion approximate maximum likelihood estimates for large sample sizes (von Asmuth and Bierkens, 2005). Peterson and Western (2014) modified the objective function to increase numerical stability (derivation given in the Appendix A.1). The modified objective function is given by:

$$S^2(t, \Omega) = \sum_{j=1}^N \left(\frac{e^{\frac{1}{N} \sum_{i=1}^N \ln(1 - e^{-\frac{-2\Delta t_i}{\alpha}})}}}{1 - e^{-\frac{-2\Delta t_j}{\alpha}}} v^2(\Omega, t_j) \right) \quad (3.22)$$

where Ω is the parameter set, α the noise model parameter and v are the innovations as calculated in equation (3.19). Equation (3.22) is the objective function that is minimized to find the best model fit.

3.6.2. LEAST-SQUARES MINIMIZATION

Many algorithms are available to find a parameter set that minimizes equation 3.22. In the initial attempts of calibrating the model, it was tried to apply a global optimization scheme as proposed by Peterson and Western (2014), the "covariance matrix adaption evolutionary strategy" (Hansen et al., 2003). In testing this algorithm on both benchmark data and observed groundwater levels, the reproducibility of the parameter values proved to be a problem, and different optimal parameter sets were found when calibrating the model multiple times. The existence of multiple optimal parameter sets became evident and it proved to be difficult to improve the reproducibility using of global optimization scheme. It was therefore decided to switch to a local search method and combined with improved the initial estimates of the model parameters.

The gradient-based Levenberg-Marquardt optimization method was applied to minimize Equation 3.22. Since this is a local search method, selection of the initial parameter set is important to find an optimal solution. This apparent drawback can also be viewed as an advantage as it requires the modeller to better understand the system and to be able to select reasonable initial estimates. The least squares fitting procedure is implemented using LMFIT, a Python package developed by Newville et al. (2014). A great advantage of this implementation is that it allows to provide parameter boundaries for optimization and automatically calculates many statistics commonly used in non-linear least squares fitting (e.g., confidence intervals, covariances).

3.7. PARAMETER SENSITIVITY

In the previous sections, the different model components have been described. Each model component takes a number of parameters to be estimated. For the models with a single response function A , a , and n need to be estimated. The models with a double-response function A , a , μ , and σ need to be estimated. The models with a linear recharge model need the factor f to be estimated. The non-linear root zone models all take I_{max} and S_{rmax} as parameters to be estimated, and depending on the model K_p , β , γ , or a combination of these. All models need the parameters α to be estimated for the noise model. Finally, the base elevation d needs to be estimated for all models.

In this section, the different model parameters are explored. A sensitivity analysis is performed for each parameter, studying the impact of the change in parameter value on the model outcome. The covariance matrix is automatically scaled by LMFIT such that the parameters have a similar sensitivity in the optimization process. Some parameters are allowed to vary freely in the estimation process, while others are bounded or fixed to an estimated value. Table 3.1 provides an overview of all the parameter ranges in the models.

Parameter	Unit	Type	Range	Equation(s)
Response functions				
A	[-]	Lower Bound	0.0 - ∞	3.12
n	[-]	Lower Bound	0.0 - ∞	3.12
a	[Days]	Lower Bound	0.0 - ∞	3.12
μ	[Days]	Bounded	0.0 - ∞	3.20
σ	[Days]	Bounded	0.0 - ∞	3.20
Linear				
f	[-]	Bounded	0.0-1.5	3.1
Noise				
α	[Days]	Bounded	0.0 - 1000	3.19, 3.22
Base elevation				
d	[Meter]	Free	-	3.10
Root zone				
β	[-]	Lower Bound	0.0 - ∞	3.4, 3.5, 3.8, 3.9
γ	[-]	Lower Bound	0.0 - ∞	3.6, 3.7, 3.8, 3.9
K_p	[Meter/Day]	Bounded	0.001-0.1	3.6, 3.7, 3.8, 3.9
I_{max}	[Meter]	Fixed	0.0015	
S_{rmax}	[Meter]	Fixed	0.226 - 0.305*	3.4, 3.5, 3.6, 3.7, 3.8, 3.9

Table 3.1: Parameter ranges and constraints. *The value for S_{rmax} is separately estimated and fixed for each observation well.

3.7.1. TFN PARAMETERS: A , a , n

Parameters A , a , and n describe the shape of the impulse response function. All of these parameters need to be estimated and have no direct physical meaning. A is a scaling parameter and a and n are shape parameters. They all have a lower bound of 0 and no upper bound. For the double-response models, n is fixed to $n = 1$ and the impulse response reduces to an exponential decay function. The impulse response functions for different values of these parameters are shown in Figure 3.6.

3.7.2. A SECOND IMPULSE RESPONSE FUNCTION: μ AND σ

The parameters μ and σ are estimated for the models with a double-response function. The parameters determine the shape of the Gaussian distribution function as the mean μ and the standard deviation σ and model the retardation and dispersion of the unsaturated zone. The parameter μ is indirectly related to the time to peak of the system, T_{peak} . The parameter σ determines the dispersion that occurs in the unsaturated zone. Both parameters have a lower bound of zero and no upper bound.

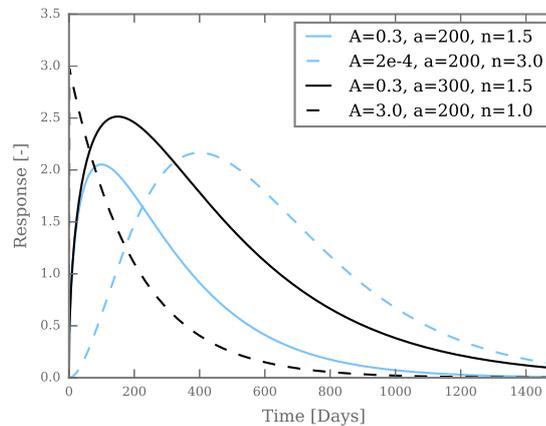


Figure 3.6: The impulse response function for different values of A , a and n .

3.7.3. NOISE MODEL: α

In principle, the α parameter used in the noise model has a lower bound of 0 and no upper bound. It frequently occurred that the noise model is used to explain all the head variations during optimization. For the simulations where this occurred, the upper boundary of 1000 days is implemented. This somewhat arbitrary upper bound means that the error at $t_0 + \Delta t$ can for 98.61% be explained by the error at t_0 , where $\Delta t = 14$ days.

3.7.4. THE EVAPORATION FACTOR: f

The parameter f is used in the linear model to simulate the recharge series. The potential evaporation is multiplied by this factor, implicitly accounting for various processes that might occur. While the crop factor K_c is already applied to calculate the potential evaporation (see Equation 2.1), factor f can compensate for an error in the estimated value for K_c . When the factor f is larger than unity, it can also account for a decrease in the recharge to the groundwater table. This can for ex-

ample be caused by increased drainage. A upper boundary of 1.5 is used, at which the mean annual evaporation is approximately equal to the mean annual precipitation.

3.7.5. PARAMETERS β AND γ

The β parameter determines the non-linear behaviour of the fraction of the effective precipitation that causes recharge through preferential flow. This fraction is also called the recharge coefficient C_r [-] and recalling from equation (3.5) is defined as:

$$C_r = \left(\frac{S_r(t)}{S_{rmax}} \right)^\beta \quad (3.23)$$

β has a lower bound of zero and no upper bound. The effect of a change in the values of β on the value of C_r decreases when the value of β increases. This decreasing effect can be observed in Figure 3.7. The parameter β can be interpreted as a parameter determining how saturated the root zone must be for before a fraction of the preferential flowpaths is activated. The γ parameter behaves in a similar fashion as β , and determines how much water is drained from the root zone reservoir.

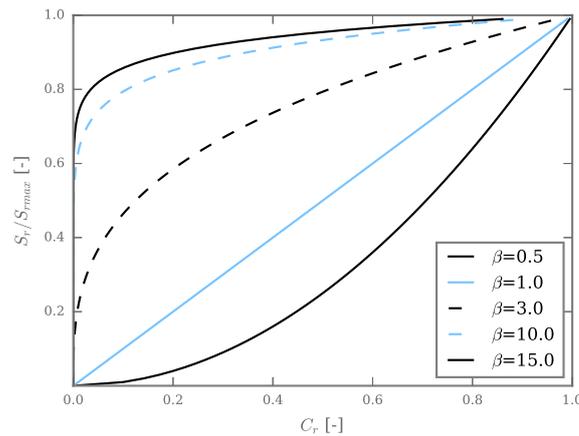


Figure 3.7: Effect of the β parameter on the recharge coefficient C_r .

3.7.6. THE INTERCEPTION CAPACITY I_{max}

The parameter I_{max} determines the size of the interception reservoir, and the amount of water that can be intercepted. Typical values for the interception capacity range from 0.6 to 6.0 millimetres in temperate climate zones, depending on the vegetation cover and soil types (Breuer et al., 2003). A value of 1.5 mm is used, based on the average of four vegetation types: herbs, deciduous and coniferous trees, and shrubs. In the sensitivity analysis that was performed for this parameter, it was found that the model outcomes are insensitive to this parameter. A 100% increase or decrease of this parameter from its default value of 1.5 mm, only causes minor changes in the objective function and the other diagnostic tests that are used. Therefore, the value of this parameter is fixed at 1.5 mm in the calibration process.

3.7.7. THE ROOT ZONE STORAGE CAPACITY: S_{rmax}

The root zone storage capacity may be estimated by considering many spatially distributed factors such as climate, geology, and vegetation types. Gao et al. (2014) tested the hypothesis that ecosystems size their root zone storages to cope with soil moisture deficits that occur with a certain return period. They showed that the storage capacity is controlled by climatic factors and can be estimated reasonably well using the mass curve technique, a technique that is used to size reservoirs. It is based on the relationship between the cumulative water supply (i.e., precipitation) and the cumulative demand (i.e. actual evaporation) over a long period of time. A maximum soil moisture deficit is calculated for each year, assuming an infinite soil moisture reservoir. The estimate for S_{rmax} is calculated as follows:

1. Run the interception reservoir to calculate the effective precipitation P_e .
2. Estimate the average daily actual evaporation $E_{a,avg}$ from a simple water balance of the root zone, where the change in storage is assumed to be zero: $\frac{ds}{dt} = P_e - E_a - N = 0$.
3. Estimate the actual evaporation by applying the seasonal pattern of potential evaporation E_p (this can also be seen as rescaling E_p to $E_{a,avg}$).
4. Calculate the soil moisture deficit for each time step with a maximum of 0, so only accounting for negative storages (or water shortage to evaporate).
5. Calculate the maximum soil moisture deficit for each year and fit a Gumbel distribution.
6. Determine the storage deficit with return period of 20 years, this is the estimate of S_{rmax} .

In the first step, a value of 1.5 mm for I_{max} is assumed. The final estimated value for S_{rmax} is relatively insensitive to changes in this parameter value. The water balance in step two does not involve runoff, as it is assumed to be zero in this study (discharge only happens through groundwater recharge). The recharge N is estimated at 360 mm/year, as calculated by Gehrels et al. (1998) for the Veluwe. In this study, a return period of twenty years is used to estimate S_{rmax} , but the estimate changes only marginally when longer return periods are used.

3.7.8. PERCOLATION CAPACITY K_p

This parameter constitutes the maximum percolation recharge flux. If this value is very small, little water is drained from the root zone and the flux is small but lasts for a long time. A large value means that the root zone drains quickly, and the recharge flux exhibits larger variations. The minimum is arbitrarily chosen at 1e-4 m/d (0.1 mm/d), which represents a percolation flux that is practically absent. A maximum value of 0.1 m/d is selected, approximately twice the maximum daily recharge that was found in the preliminary phase of this study. These larger values change the recharge pattern to one that is more similar to preferential flow.

3.8. MODEL PERFORMANCE

Several diagnostic statistics are provided in this section to investigate the performance of the different models. Each statistic focusses on different aspects of how well the model simulates the

observed groundwater levels. These statistics can also be used to compare the performance of different models. The Bayesian and Akaike informations criteria (BIC, AIC) are used to account for the number of parameters. Model performance is assessed through a validation period, for which some of the statistics can be calculated separately. Finally, a more qualitative test based on the time to peak of the system is proposed.

3.8.1. ROOT MEAN SQUARED ERROR

The root mean squared error (RMSE) is calculated as follows:

$$RMSE = \sqrt{\sum \left(\frac{(h_o - h_m)^2}{N} \right)} \quad (3.24)$$

where h_o is the observed head, h_m the modelled head and N the number of observations. This diagnostic statistic gives information on the goodness-of-fit of the deterministic part of the model, considering the residuals and not the innovations.

3.8.2. EXPLAINED VARIANCE PERCENTAGE

The explained variance percentage (EVP) is calculated as follows:

$$EVP = \frac{\sigma_o^2 - \sigma_r^2}{\sigma_o^2} 100\% \quad (3.25)$$

where σ_o^2 and σ_r^2 are the variances of the observed groundwater levels and the residuals series respectively. This diagnostic statistics gives information on the percentage of the variance in the observed groundwater heads that is explained by the model.

3.8.3. AKAIKE AND BAYESIAN INFORMATION CRITERIA

Akaike (1974) and Schwarz (1978) introduced the Akaike and Bayesian Information Criteria (AIC, BIC) to compare model performance while considering the number of free model parameters. The criteria are suitable for comparing models using the same data sets, but cannot be used for hypothesis testing. The Bayesian variant has a higher punishment for the number of parameters that is estimated. The AIC and BIC for model i are calculated as:

$$\chi^2 = \sum_j^N v_j^2 \quad (3.26)$$

$$AIC_i = N \ln \left(\frac{\chi^2}{N} \right) + 2k_i \quad (3.27)$$

$$BIC_i = N \ln \left(\frac{\chi^2}{N} \right) + k_i \ln(N) \quad (3.28)$$

where $\frac{\chi^2}{N}$ is the likelihood of the model, k_i is the number of free model parameters, and N is the number of observations. The AIC punishes for the number of model parameters through the factor $2k_i$ and the BIC through the factor $k_i \log N$. Models can be compared through the differences between models in the values of the AIC and the BIC.

3.8.4. TIME TO PEAK

In this subsection a qualitative performance indicator is proposed, the time to peak (T_{peak}). where the previous performance indicators are all numeric, this indicator is proposed as an indicator of the process consistency. The response of the groundwater to an impulse of recharge, is characterized by a fast rise of the heads, followed by a slow decline. The moment of rise to decline is the time to peak (T_{peak} , see Figure 3.8). The rising limb (and hence the time to peak) of this response is a result of the percolation zone, where dispersion and damping of the recharge signal from the unsaturated zone occurs. The time to peak is expected to increase with the thickness of the unsaturated zone, if all variables are kept constant (e.g. lithology).

The observed groundwater levels for all the three wells show a distinct peak in 1995 (see Figure 2.3). The time of the peak in 1995 is slightly different for each well, as can be observed in the second column of Table 3.2. As the measurement interval is approximately once every two weeks, there is an error of ± 14 days. Calculating the time to peak is impossible when the recharge event that caused the peak cannot be isolated. Based on several studies performed on the effect of the thickness of the unsaturated zone in the study area, a maximum of 10 days per meter unsaturated zone is proposed as a rule of thumb (www.grondwaterformules.nl). The resulting maximum values for the time to peak for each observation well can be found in Table 3.2.

It is assumed that the same wet period and subsequent recharge caused the 1995 peak at all observation wells. From the observed groundwater levels we can calculate the difference in the time to peak (ΔT_{peak}) based on the differences of the timing of the observed peaks in 1995. The ΔT_{peak} is calculated for all wells with reference to the most shallow well. As can be expected based on the thickness of the unsaturated zone, the ΔT_{peak} increases from the shallow to the deepest well (see Table 3.2). Based on the thickness of the unsaturated zone the differences in the ΔT_{peak} of B27C0049 and B33A0113 is rather small. Possibly this can be explained with the coarser sand that is found at B33A0113, showing that the thickness of the unsaturated zone as the sole predictor of the T_{peak} is not always sufficient.

Borehole	Groundwater peak in 1995	Thickness un-saturated zone (m)	ΔT_{peak} [Days]	Maximum time to peak [days]
B27D0001	03-07-1995	7.0	0	70
B27C0049	14-09-1995	29.0	73	290
B33A0113	28-09-1995	49.0	87	490

Table 3.2: Overview of the time to peak (T_{peak}) and the ΔT_{peak} is the expected shift in peak compared to the T_{peak} of borehole B27D0001.

The data provided in Table 3.2 is only based on the analysis of one peak, and therefore only serves as an indication. Examination of other peaks showed that the time shift is not constant, but is in the same order of magnitude. Wetting and drying of the percolation zone may cause a

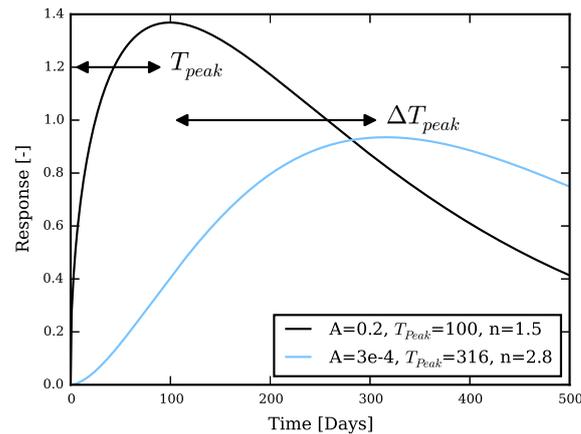


Figure 3.8: Graphical representation of the concepts of time to peak T_{peak} and the shift in the time to peak ΔT_{peak}

time to peak that is dependent on the system state of the percolation zone. Despite these nonlinearities and uncertainty in the exact values, the analysis above provides three qualitative tests to check process consistency of the models. First, the time to peak of well B27D00010 is about 80 days shorter than the other two wells. Second, a maximum time to peak is calculated for each well, based on the thickness of the unsaturated zone.

3.9. SOFTWARE DESCRIPTION

The models are implemented in a Python based software package dubbed 'GroundWater Time Series Analysis', or GW TSA in short. This object-oriented program can be imported as a module in a Python script or IPython notebook through standard import statements. The more computationally demanding unsaturated zone model is written in Cython, which can be ported to the programming language C and compiled and imported into the Python script. Using the compiled C-version of the unsaturated zone model instead of the same function in Python reduced computation times up to a factor of 25, depending on the chosen root zone model. The optimization process is built on LMFIT, (Newville et al., 2014), a wrapper around the scipy.optimize package (<http://www.scipy.org>).

```

1
2 from GWISA import * # Import the entire GWISA Toolbox
3
4 bores = [ 'Test_Data/B27D0001001_1.csv' ] # Observed groundwater levels
5 forcing = 'Test_Data/KNMI_Bilt.txt' # Climatic data
6 calibration = [ '01-01-1974', '31-12-1994' ] # Calibration period
7 validation = [ '01-01-1995', '31-12-2004' ] # Validation period
8
9 ml = Model(bores, forcing, calibration, validation) # Create a model object and prepare for
simulation
10 X0 = Parameters()
11 # (Name, Value, Vary, Min, Max, Expr)
12 X0.add_many(( 'A', 3, True, 1e-6, None, None),
13             ( 'a', 200., True, 0.0, None, None),
14             ( 'n', 1.0, True, None, None, None),
15             ( 'alpha', 100.0, True, 0.0, 1000.0, None),
16             ( 'Smax', 0.305, False, None, None, None),
17             ( 'Imax', 1.5e-3, False, None, None, None),
18             ( 'Beta', 2.0, True, 0.0, None, None))

```

```
19  
20 ml.solve(X0, RM='preferential', method='leastsq')  
21 ml.bores_list[0].plot_results()  
22 ml.bores_list[0].plot_diagnostics()
```

An example script to simulate groundwater levels using GW TSA is shown above. An instance of the Model class is created for each borehole and the data is prepared automatically (in line 9). The initial parameters and their boundaries are provided for use in the optimization method of LMFIT. The model is solved by calling the 'solve' function, and the parameter estimation process starts. Upon completion, GW TSA reports if the optimization was successful and a report text-file is created, containing information on the optimization. Calling the 'plot_results' function on each borehole instance returns a figure containing the basic information of the model, its parameters, and the diagnostic statistics. If more information is required, the function 'plot_diagnostics' produces a figure with an autocorrelation graph, a frequency of exceedance graph, and the covariance matrix. A more detailed description of the software can be found in the manual on the github repository (<https://github.com/raoulcollenteur/GWTSA>).

4

APPLICATION OF THE FOUR RECHARGE MODELS TO THREE OBSERVATION WELLS

In this chapter the results of the four recharge models in combination with different forms of the impulse response function are presented. The first three sections describe the results of the four recharge models for each observation well, in increasing order of the thickness of the unsaturated zone. Subsequently, the results of the alternative impulse response functions in combination with all four recharge models are presented. The chapter concludes with a discussion of the results and an outlook on how the model performance can be improved in the next chapter.

4.1. SHALLOW OBSERVATION WELL B27D00010

The observation well B27D00010 has the most shallow groundwater levels considered in this study, with an estimated thickness of the unsaturated zone of 7 meters. The seasonal pattern that is present in the evaporation and precipitation data is also observed in the groundwater time series. The maximum root zone storage S_{rmax} for this location is estimated at 0.30 m using the method presented in Section 3.7.7. The recharge is simulated for the entire period for which the input variables $P(t)$ and $E(t)$ are available: 1958-2015. Two periods are used for calibration of the time series models: a 30-year (1974-2004) and a 20-year (1974-1994) period. The models calibrated on a 20-year period have a 10-year period for model validation (1995-2004).

4.1.1. 30 YEAR CALIBRATION PERIOD

The observed and simulated groundwater levels for all four recharge models are shown in Figure 4.1. All four models are able to simulate the observed groundwater levels well, as is confirmed by high explained variances of over 99%. The peaks in the groundwater level in 1983, 1995 and 2003 are underestimated by all four models. Moreover, the lows in 1979, 1987 and 1993 are overestimated by all models. The preferential and combination model simulate similar groundwater levels, as can be observed in Figure 4.1 by the overlap in the simulated groundwater levels. The parameter K_p sits at the lower boundary ($1e-4$) and the value of γ is high (see Table 4.1), causing a minimal percolation flux in the combination model.

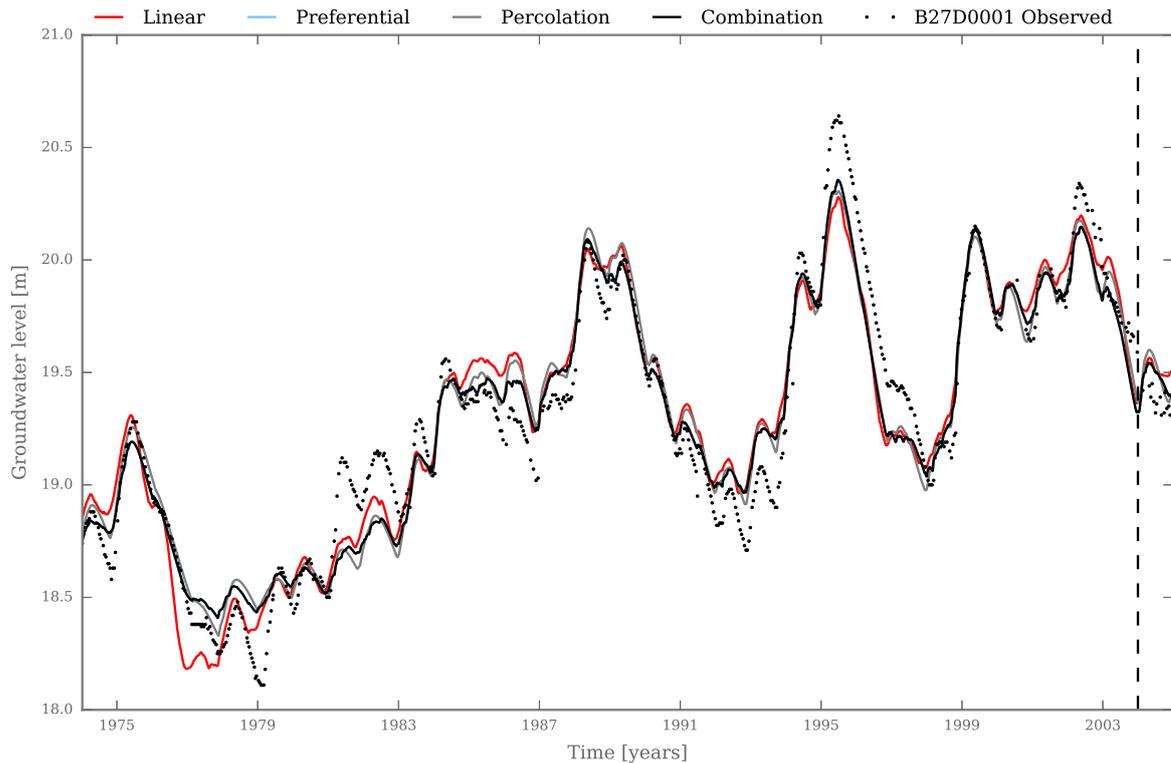


Figure 4.1: Observed and simulated groundwater levels for observation well B27D00010 using four recharge models. The vertical dashed line denotes the end of the calibration period (1974-2004).

The residual series for the model are shown in Figure 4.2. The residuals of all models follow a similar pattern. There are periods of consistent over- and underestimation present in the residual series. The most notable difference occurs around 1977, where the linear model underestimates the groundwater levels up to 40 centimetres. Berendrecht et al. (2006) observed a similar difference between a linear time series model and one that included a non-linear root zone model. Berendrecht et al. (2006) attributed this difference to a reduction in transpiration rates due to a lack of soil moisture needed to transpire. The year 1976 was extremely dry and these results suggest that under such conditions evaporation reduction because of low soil moisture is important.

The daily recharge simulated by the four recharge models for the period 1997-2005 are shown in Figure 4.3. The winter period November-March is shaded in blue for easier interpretation. The differences in the simulated recharges series are surprising, in particular because the simulated groundwater levels are largely similar. The recharge calculated by the preferential and the combination model is very similar, explained by the minimal percolation flux in the combination model. The linear model simulates recharge throughout the year, while the non-linear models show a recharge that occurs primarily in the winter periods. The most troubling feature of the linear recharge is the simulation of negative values, indicating the unlikely event of groundwater uptake.

A Monte Carlo analysis is performed to assess the uncertainty in the calculated recharge series. One thousand recharge series are calculated with 1000 different parameter sets. These parameter sets are drawn from a multivariate Gaussian distribution that is based on the covariance matrix returned by LMFIT. The parameter boundaries imposed on the parameter estimation are also respected when drawing these parameter sets. For each simulation, the total yearly recharge is calculated and uncertainties in the yearly recharge are considered in this analysis. Since the pa-

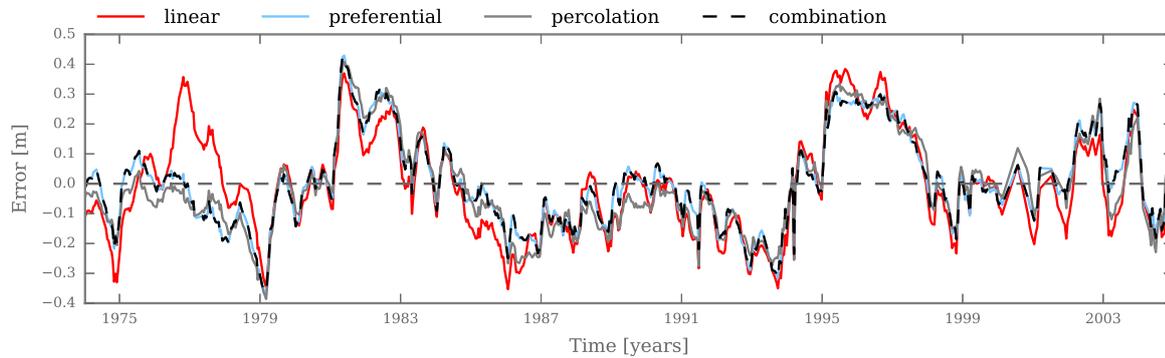


Figure 4.2: Residual series for all four recharge models calibrated on a 30-year period for observation well B27D00010.

Observation Well B27D00010

1974-2004	A	a	n	f	α	β	γ	K_p	d
Linear	4.83e-2	363.20	1.77	0.75	117.69	-	-	-	17.31
Pref.	3.60e-1	440.69	1.41	-	107.04	2.48	-	-	17.66
Perc.	1.95e-1	438.32	1.50	-	101.63	-	25.86	1.72e-2	17.73
Comb.	3.60e-1	437.02	1.41	-	107.00	2.28	57.95	1.00e-4	17.63
1974-1994									
Linear	3.05e-2	348.55	1.84	1.03	74.61	-	-	-	18.11
Pref.	5.37e-1	454.21	1.30	-	82.68	6.07	-	-	17.97
Perc.	9.54e-1	566.49	1.21	-	94.77	-	12.21	3.86e-3	17.60
Comb.	5.12e-1	453.77	1.31	-	81.34	6.97	128.52	1.00e-4	17.97

Table 4.1: Overview the parameters values for the four recharge models for observation well B27D00010, for two different calibration periods.

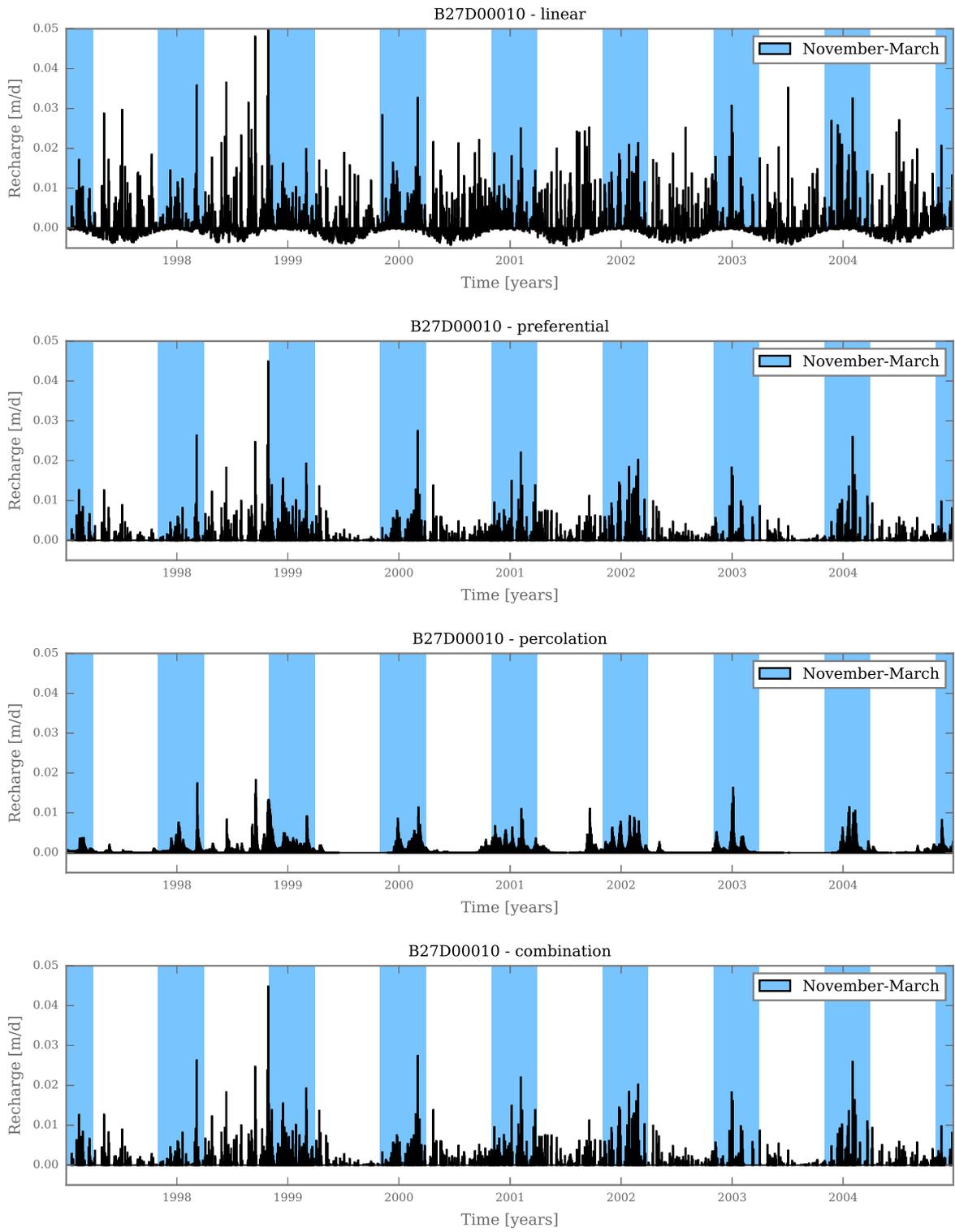


Figure 4.3: The recharge time series for the linear, preferential, percolation and combination model for observation well B27D00010 in the period 1997-2004.

parameters are constrained and the recharge models are non-linear the probability distribution of the yearly recharge for each year is not (necessarily) Gaussian.

The results of the analysis explained in the previous paragraph are shown in Figure 4.4. The blue bars indicate the average annual recharge and whiskers denote the 2.5% and the 97.5% percentiles as an indication of the uncertainties. The preferential model has the smallest uncertainty bounds, followed by the linear model, the percolation model, and finally the combination model. The difference between the 2.5% and the 97.5% percentiles is approximately 0.15 m/year for the preferential model and 0.3 m/year for the linear model. The combination model shows maximum yearly recharges of more than 3 m/year, more than the maximum annual precipitation. This is an indication that the numerical equations for the root zone model are not solved correctly for all parameter values.

The impulse response functions for the different models show some minor differences. The time to peak (T_{peak}) ranges from 178 days for the combination model to 280 days for the linear model. These times to peak are higher than the calculated maximum time to peak of 70 days for this well (see Section 3.8.4). For this observation well the difference in the time to peak (ΔT_{peak}) is not calculated, as this well is used as the reference. The number of days where 99% of the response has occurred (T_{99}) is around 2500 days for all models.

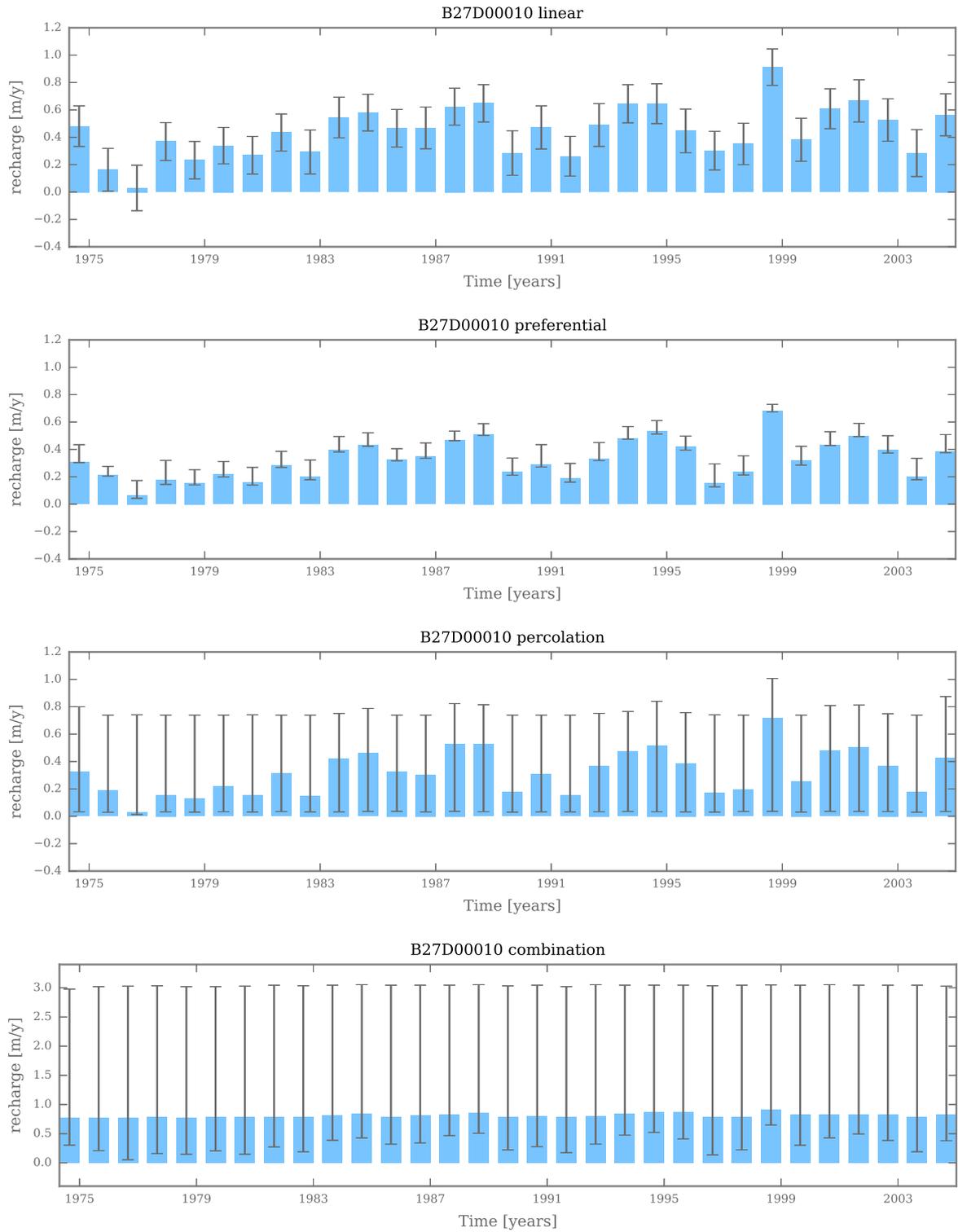


Figure 4.4: Average yearly recharge fluxes and the 2.5% and the 97.5% percentiles (the whiskers) to provide information on the uncertainty in the recharge calculation.

4.1.2. 20 YEAR CALIBRATION PERIOD

The simulated groundwater levels for all four recharge models calibrated on a 20-year period are shown in Figure 4.5. The models show improved performance in the calibration period compared to the results presented in the previous section. The peaks and lows are better captured by all models, most notably around 1987 and 1992. The models performance in the validation period decreases compared to the results shown in the previous section. The groundwater levels are consistently underestimated in the validation period. The preferential and the combination model simulate similar groundwater levels. These are the same because the calibrated percolation flux is negligible in the combination model.

The residual series for all four recharge models, calibrated on a 20 year period, are shown in Figure 4.6. The residuals are somewhat smaller during the calibration period, but larger during the validation period compared to the residuals of the model calibrated on a 30-year period (see Figure 4.2). The linear model performs slightly better than the non-linear models, except for the low in 1977 and the peaks around 1985. The residual seem to have a small upward trend over the entire simulation period.

Compared to the models calibrated on a 30-year period, all models estimate a lower value of α (see Table 4.1). This indicates that a smaller part of the error depends on the error made in a previous time step in these models. The time steps between the groundwater level observations are shown in Figure 4.7. In the validation period some large time steps are present. A large value of α is needed to relate the errors when the time step is large.

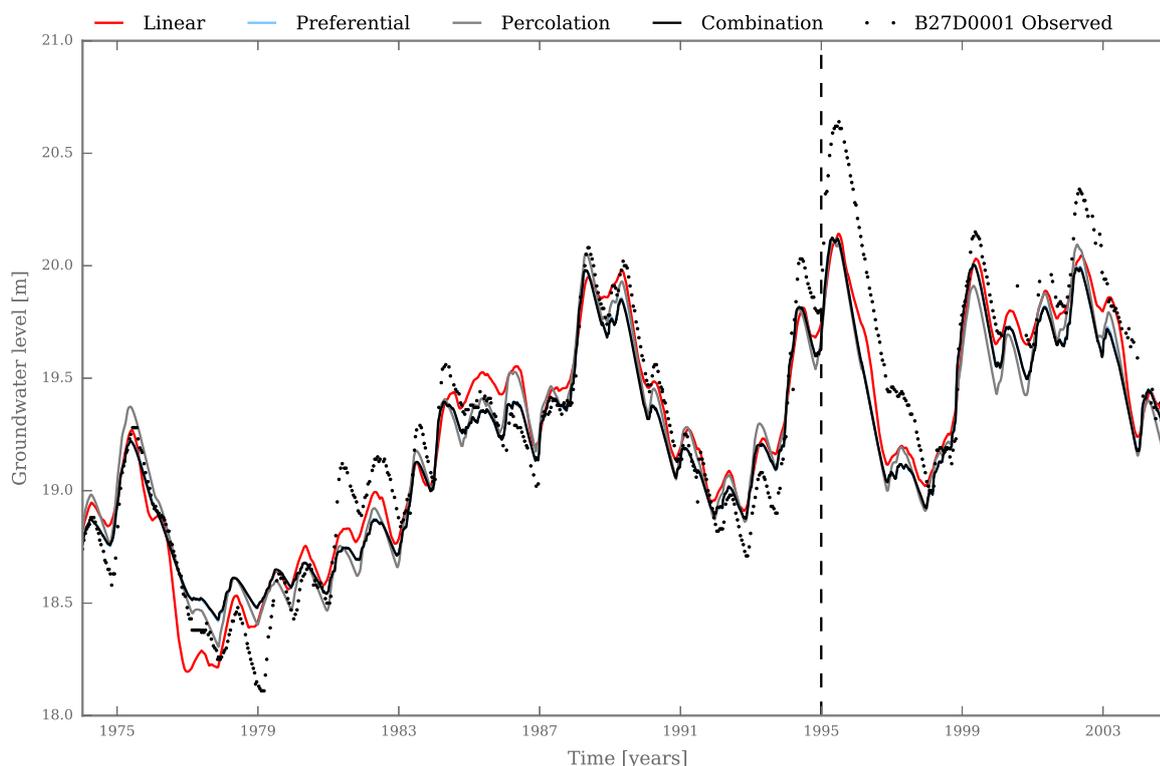


Figure 4.5: Observed and simulated groundwater levels for observation well B27D00010 using four recharge models. The vertical dashed line denotes the end of the calibration period (1974-1995).

The recharge series simulated by the different models again show large differences in terms of

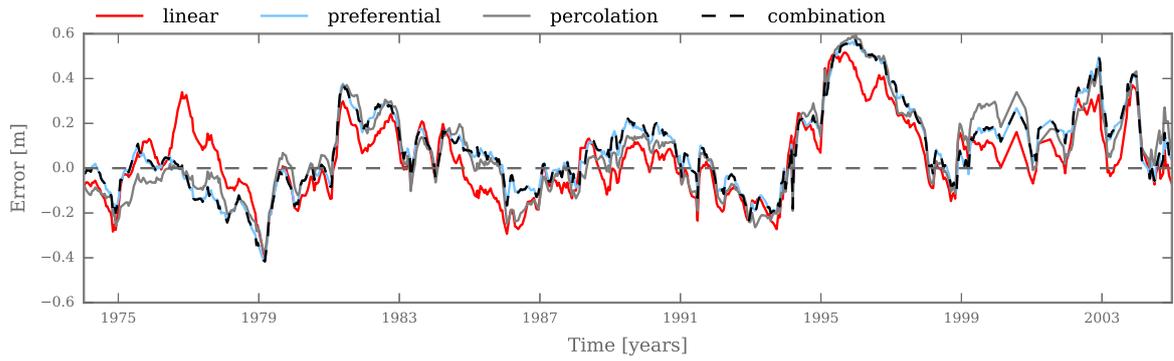


Figure 4.6: Residual series for the four recharge models for observation well B27D00010 calibrated on a 20 year period (1974-1995).

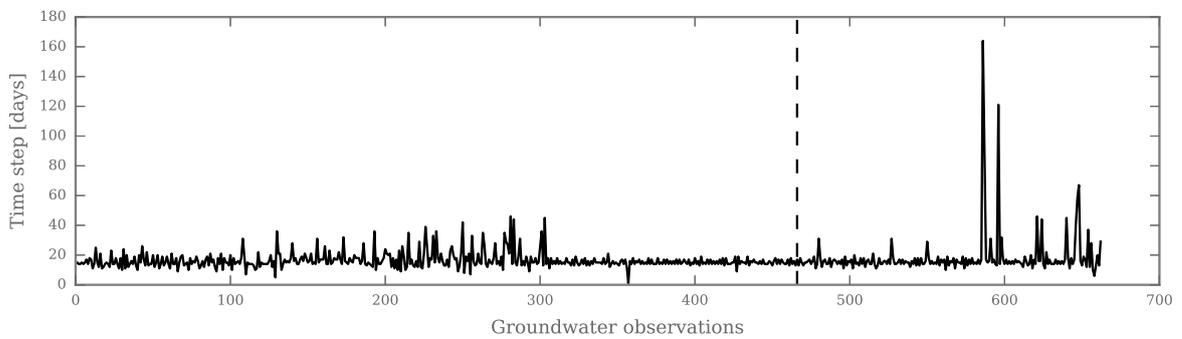


Figure 4.7: Time steps between the observations for observation well B27D00010. The vertical dashed line denotes the end of the calibration period.

the recharge pattern (see Figure C.1 in the Appendix). The uncertainty in the recharge is smallest for the percolation model, followed by the preferential model, the linear model, and finally the combination model. The percolation and combination model perform better when the uncertainty in the recharge for the models calibrated on a 30-year period is compared (Figure 4.4).

The long term average recharges are similar to those of the models calibrated on a 30-year period presented in the previous section, except for the linear model. The non-linear models all calibrated recharge fluxes of around 0.30 m/year for both calibration periods. The recharge of the linear model decreases by over 35% from 0.45 m/year to 0.29 m/year. This is caused by a 37% increase in the factor f (see Table 4.1) and an increase of the base level d of the linear model from 17.31 to 18.11 metres. The parameters f and d show a high correlation ($r > 0.95$) for this model, compromising the estimation of these parameters. High correlations are also found between the parameter d and the parameters of the non-linear recharge models, but they are generally lower ($r \sim 0.85$).

The time to peak of all the non-linear models decreased (121-140 days, see Table 4.8) while the time to peak of the linear model increased. The difference in the time to peak between the linear and the non-linear models is around 150 days. The response times T_{99} are similar to those of the models calibrated on a 30-year period (~ 2500 days), except for the percolation model (~ 3000 days).

4.2. MEDIUM DEEP OBSERVATION WELL B27C00490

The medium deep observation well B27C00490 has an unsaturated zone of approximately 29 m. The groundwater level series shows mostly the low frequency fluctuations from the precipitation and evaporation (see Figure 2.2). The maximum root zone storage S_{rmax} for this location is estimated at 0.23 m using the method presented in Section 3.7.7, the lowest of all locations. The model performance is evaluated for a 30 and a 20-year calibration period.

4.2.1. 30 YEAR CALIBRATION PERIOD

The groundwater levels simulated by all four models calibrated on a 30-year period are shown in Figure 4.8. The simulated groundwater levels are similar to each other, with the largest differences in the peaks and the lows. The first peak in 1975 is underestimated by all models, with the linear model performing slightly better than the non-linear models. The groundwater levels for all models drop too fast after the peak, indicating that the models miss a certain slow component or drain too fast. The periods of decreasing groundwater levels are the periods where the models show the largest differences, as can be observed in the residual series shown in Figure 4.9.

In the period 1980-1990 the groundwater levels are consistently overestimated (see Figure 4.9). The local peak around 2000 is missed by all models and none of them simulates the downward movement of the groundwater level after the peak. This peak is also present in the groundwater series of the shallow well (see Figure 4.1), but is simulated correctly for that well by all the models. This suggests a peak in the precipitation data for the shallow well that is not present in the precipitation data for this well.

The recharge patterns are very similar to those of the shallow observation well (see Figure C.3 in the appendices). The percolation model shows a continuous recharge, where the other models show a rather event-based recharge. The combination model is a mixture of the percolation and preferential model. The larger recharge events of the preferential model can also be observed in the combination recharge. The uncertainty in the recharge is largest for the combination model, followed by the linear, the preferential, and finally the percolation model (see Figure C.4 in the appendices).

The response times T_{99} for the models are between 3000 and 4000 days, meaning the groundwater levels depend on the recharge that occurred in the previous ten years. These response times are larger than the response times found for the shallow well. The times to peak (T_{peak}) are between 551 and 644 days, much higher than the estimated maximum times to peak in Section 3.8.4 based on the thickness of the unsaturated zone. The linear model has the highest value for T_{peak} and the percolation model the lowest. These times to peak are also much higher than the T_{peak} of the shallow well. The difference in the T_{peak} between the shallow and the medium deep well (ΔT_{peak}) is much higher than the ΔT_{peak} estimated in Section 3.8.4.

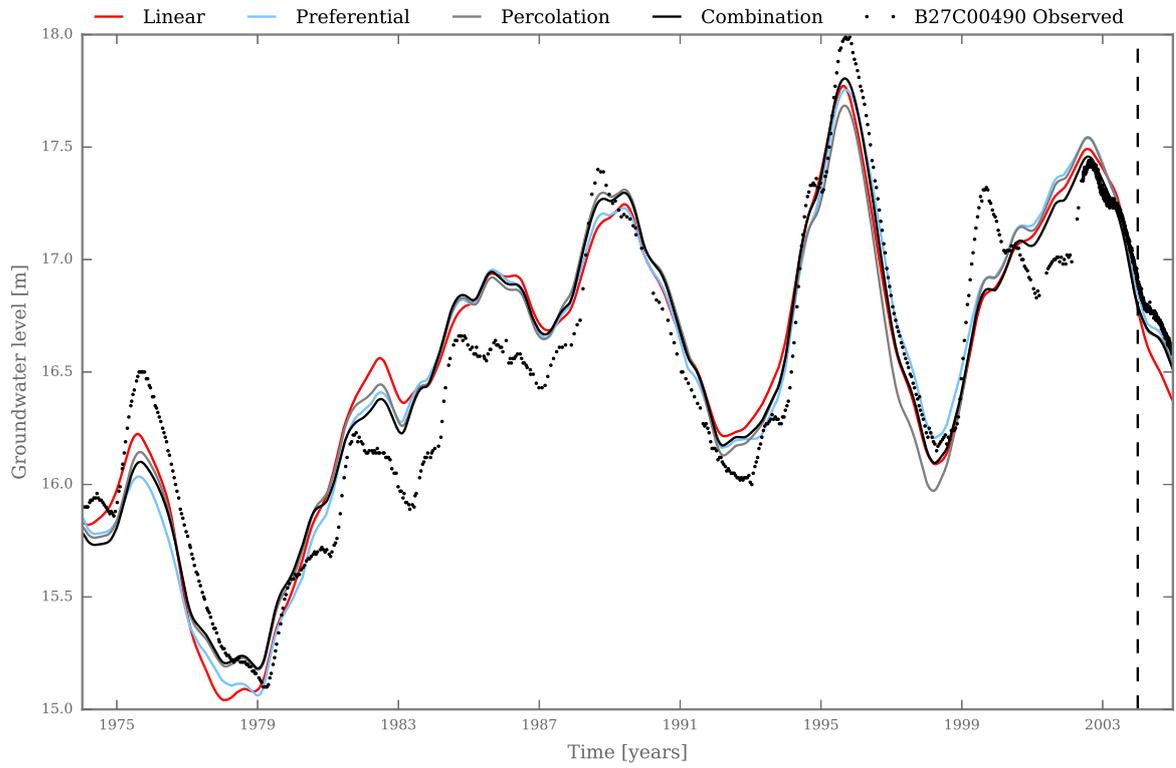


Figure 4.8: Observed and simulated groundwater levels for observation well B27C00490 using four recharge models. The vertical dashed line denotes the end of the calibration period (1974-2004).

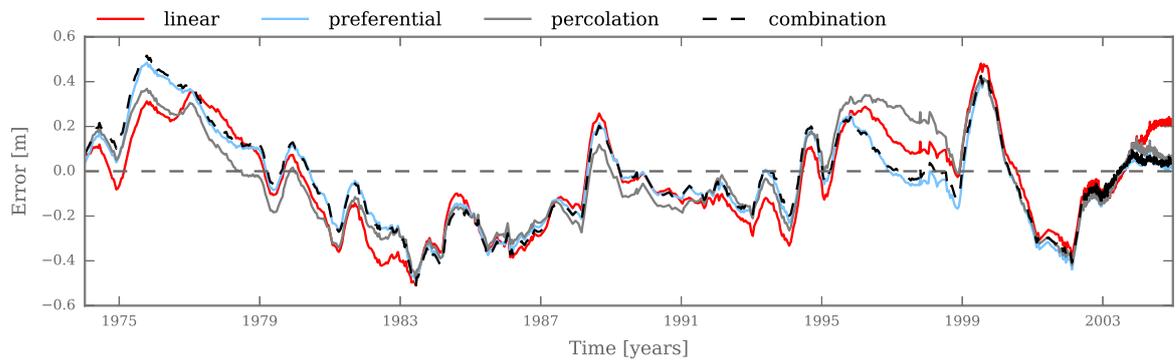


Figure 4.9: Residual series for all four recharge models calibrated on a 30-year period for observation well B49C00490.

Observation Well B27C00490									
1974-2004	A	a	n	f	α	β	γ	K_p	d
Linear	6.05e-4	409.44	2.57	0.95	388.23	-	-	-	13.10
Pref.	7.68e-3	550.78	2.10	-	342.37	3.23	-	-	12.22
Perc.	7.04e-3	469.58	2.17	-	345.94	-	2.64	3.16e-3	11.48
Comb.	5.16e-3	506.28	2.19	-	363.87	3.35	5.89	1.43e-3	12.07
1974-1994									
Linear	2.93e-5	198.55	3.41	1.04	1000*	-	-	-	14.58
Pref.	1.42e-4	201.17	3.15	-	1124.09	9.52	-	-	14.03
Perc.	1.06e-3	315.35	2.65	-	437.5	-	12.59	3.30e-3	12.39
Comb.	1.54e-3	277.21	2.60	-	517.11	26.56	1.99	2.04e-3	12.66

Table 4.2: Overview of the parameters values for the four recharge models calibrated on two different periods for observation well B27C00490. *To obtain reasonable results the parameter α is constrained to 1000 days.

4.2.2. 20 YEAR CALIBRATION PERIOD

The results of the four models calibrated on a 20-year period are shown in Figure 4.10. The results show more differences between the models, compared to the results of the models calibrated on a 30-year period (Figure 4.8). The peak in 1975 is captured better by all models. The decline of the groundwater level after this peak is too fast in all the models and the levels start rising too early after the low. The decline after a peak being too fast is observed throughout the entire period considered. The effect can clearly be observed in the residual series (see Figure 4.11), which exhibit a peak after each peak in the groundwater levels.

The residual series for the models are shown in Figure 4.11. The groundwater levels are over-estimated by all models in the period 1978-1987, most notably by the linear and the preferential models. In the following period the linear and the preferential models outperform the combination and percolation models again. The residual series for all models seems to have a rising trend over the period 1980-2004.

The recharge patterns simulated by the four models show interesting differences (see Figure C.3 in the appendices). The combination model is clearly a mixture of the preferential model and the percolation model. Only a couple of the larger recharge events cause recharge through preferential flow in the combination model. The largest recharge event between 1997 and 2005 in the preferential model occurs in April 2000. The largest recharge event for the combination model in the same period occurs in December 1999. This effectively shows the effect a non-linear root zone model can have on the recharge extremes. The uncertainty in the annual recharges is small for all models. The preferential model has the smallest uncertainty in the recharge simulation.

The response times of the models for this well drastically decrease when a shorter calibration period is used. In the previous section response times between 3000 and 4000 days were found for this observation well. The models with a 20-year calibration period simulate response times between 1500 and 2500 days. The time to peak of the percolation model is similar to the model calibrated on a 30-year period. The time to peak for the other models decreased between 150 and 200

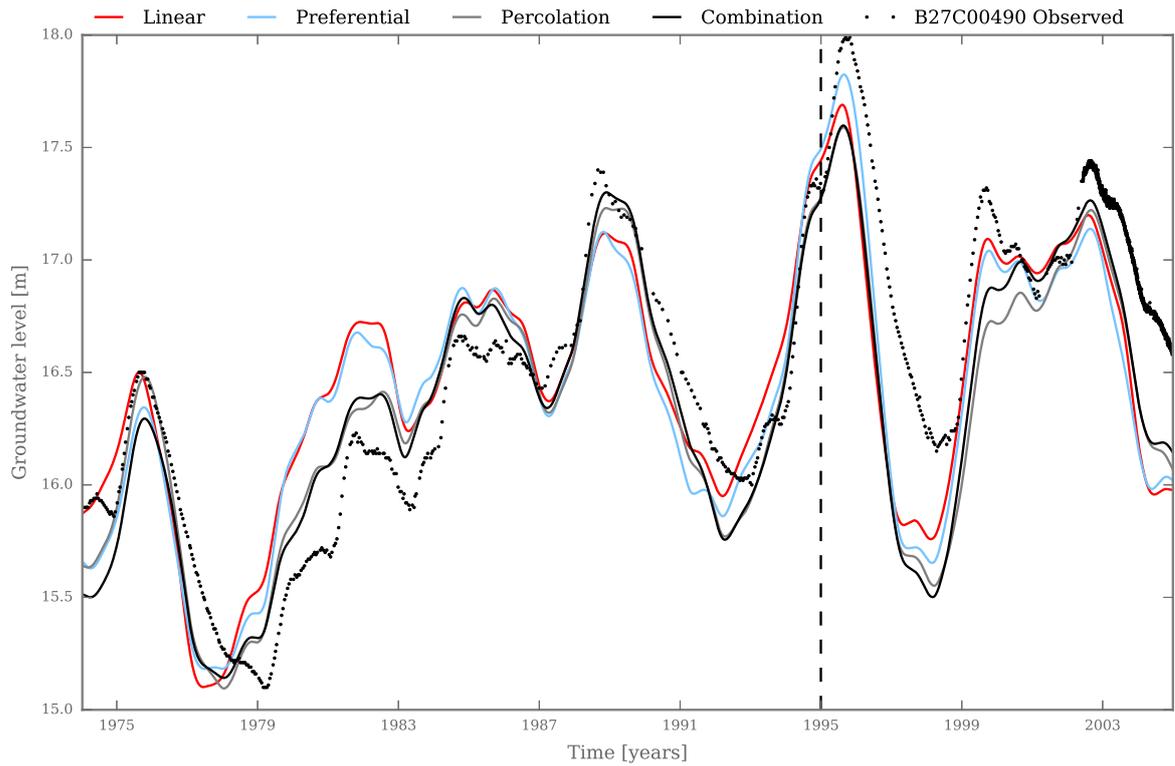


Figure 4.10: Observed and simulated groundwater levels for observation well B27C00490 using four recharge models. The vertical dashed line denotes the end of the calibration period (1974-1995).

days compared to the models calibrated on a 30-year period. Possibly this is the result of different characteristics of the groundwater levels in these periods. It has to be noted that the linear model did not deliver reasonable results without constraining the α parameter, which was constrained to $\alpha = 1000$ days for this model.

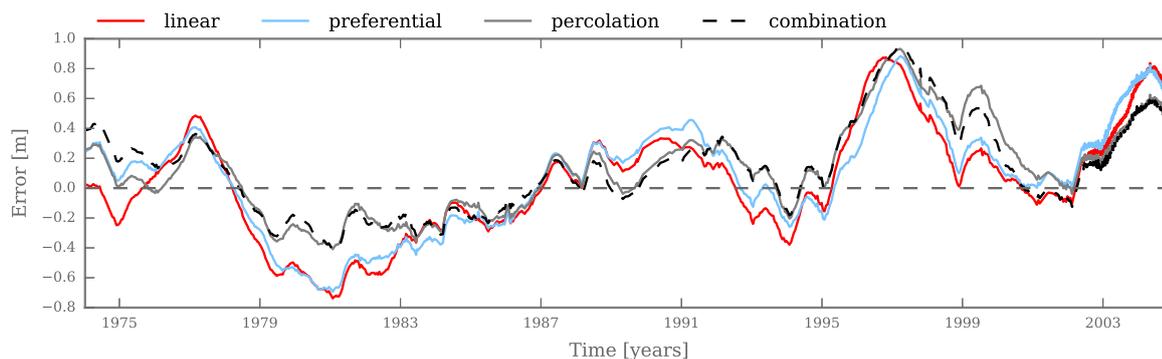


Figure 4.11: Residual series for all four recharge models calibrated on a 20-year period for observation well B27C00490.

4.3. DEEP OBSERVATION WELL B33A01130

The deep observation B33A01130 well is located near Apeldoorn and has an unsaturated zone of 49 m, the thickest considered in this study. The maximum root zone storage S_{rmax} is estimated at 0.26 metres for this area. The yearly fluctuations present in the input time series are almost absent in the groundwater time series. Similar to the analysis of the previous two observations wells, a 30 and 20-year calibration period is used.

4.3.1. 30 YEAR CALIBRATION PERIOD

The groundwater levels simulated by the four models calibrated on a 30-year period for the deep well are shown in Figure 4.12. The simulated groundwater levels follow the observed levels reasonably well, especially in the second half of the period. The first peak is underestimated by all models, but the linear model simulates the peak best. The timing of the following low around 1979 is off for all models as the modelled groundwater levels start rising a year too early. In the period 1980-1985 all models consistently overestimate the groundwater levels. There appears to be a trade-off for all models to which peaks and lows are modelled correctly, as none of the models consistently outperforms all others. For example, the peak in 1995 is modelled correctly by the linear model and is overestimated by the non-linear models, but for the following low the reverse is true.

The residual series for all four models are shown in Figure 4.13. The pattern in the residual series is somewhat similar to the residuals of the models for the medium deep well (see Figure 4.9). The residuals show periods of consistent underestimation and overestimation. From 1980 to 2004 there is a clear upward trend in the residual series.

The recharge pattern simulated by all four models is shown in Figure C.7 in the Appendices. The recharge patterns are similar to those simulated for the medium deep well (see Figure C.3). The combination model shows more recharge events caused through preferential flow mechanism, compared to the combination model for the medium deep well. The uncertainty in the recharge simulation is large for all models except for the combination model (see Figure C.8). The linear model simulated negative annual recharges of up to 0.3 m/year.

The response times for this observation well are smaller than for the shallow observation well. The response times T_{99} are around 2000 days for all models. The time to peak is estimated to be between 421 and 467 days and is larger than the time to peak of the shallow observation well, but

Observation Well B33A01130									
1974-2004	A	a	n	f	α	β	γ	K_p	d
Linear	5.46e-5	206.66	3.26	1.22	117.69	-	-	-	26.27
Pref.	9.19e-4	254.28	2.68	-	385.92	5.60	-	-	25.08
Perc.	3.69e-3	322.80	2.37	-	234.96	-	4.01	5.32e-3	23.93
Comb.	4.69e-3	306.65	2.38	-	300.82	25.65	0.90	1.14e-3	23.66
1974-1994									
Linear	1.53e-5	162.29	3.63	0.93	3284.89	-	-	-	26.05
Pref.	8.15e-5	149.79	3.36	-	3317.55	5.85	-	-	25.90
Perc.	1.21e-3	276.02	2.63	-	413.06	-	11.26	6.26e-3	24.26
Comb.	1.11e-3	221.88	2.75	-	500.81	26.79	1.10	1.21e-3	24.16

Table 4.3: Overview of the parameters values for the four recharge models calibrated on two different periods for observation well B33A01130.

smaller than the times to peak of the medium deep well. This is surprising, as the unsaturated zone is much thicker and a time to peak that is larger than that of well B27C0490 is expected.

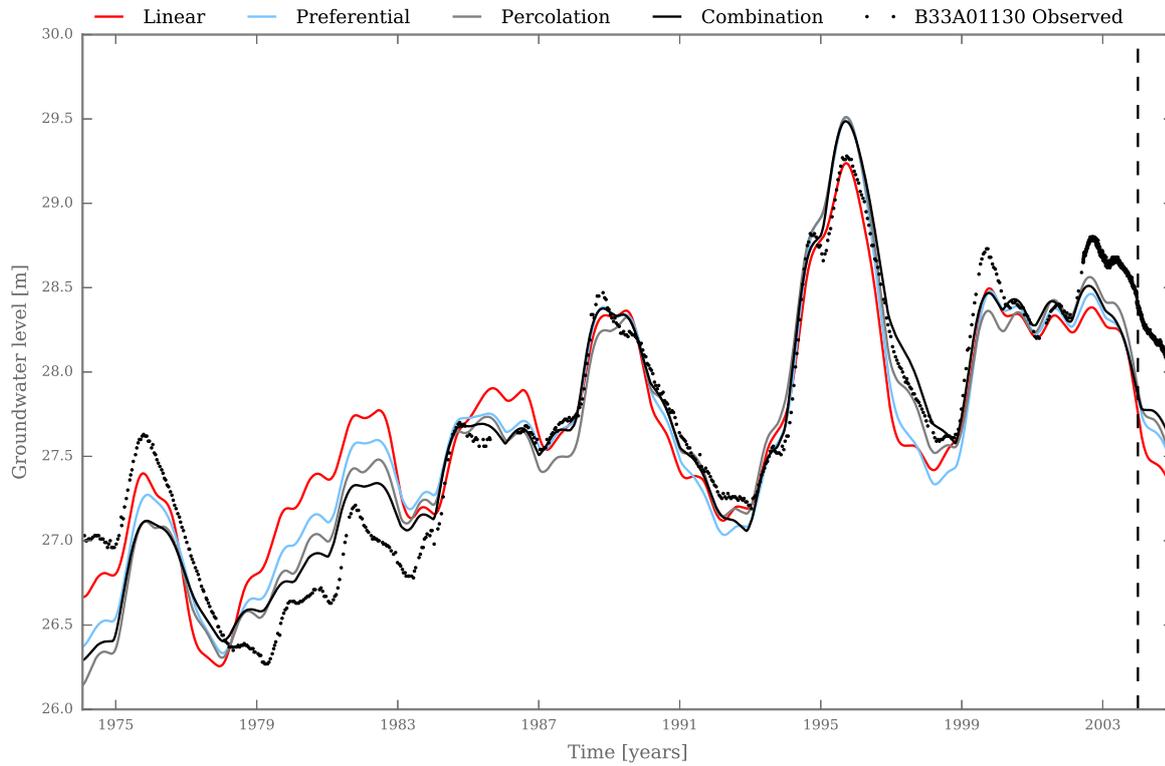


Figure 4.12: Observed and simulated groundwater levels for observation well B33A01130 using four recharge models. The vertical dashed line denotes the end of the calibration period (1974-2004).

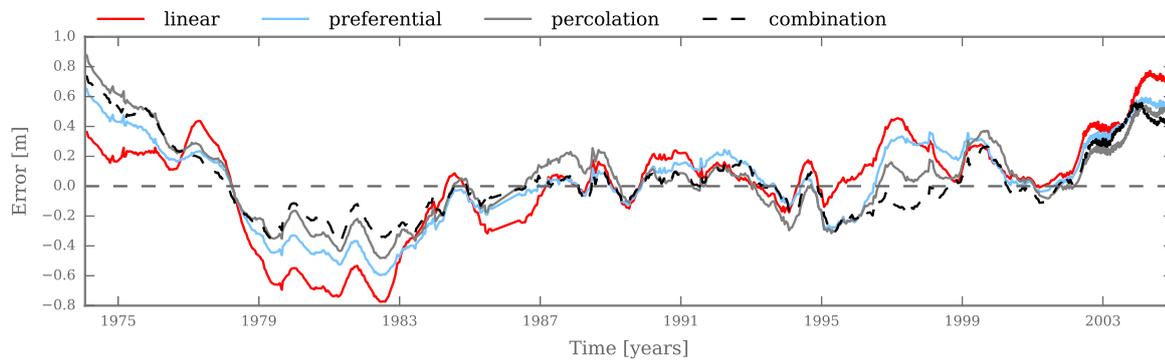


Figure 4.13: Residual series for all four recharge models calibrated on a 30-year period for observation well B33A01130.

4.3.2. 20 YEAR CALIBRATION PERIOD

The groundwater levels simulated by the four models calibrated on a 20-year period are shown in Figure 4.14. The largest differences are observed between the linear and the non-linear models, with the linear model simulating higher groundwater levels for almost all peaks and lows. This can be clearly observed in the residual series, shown in Figure 4.15. The linear and the preferential models capture the first peak better than the models calibrated on a 30-year period. These models perform less well in the period 1980-1985. The non-linear models, in particular the percolation and combination model, perform better in the calibration period. In the validation period, however, the linear model performs better than the non-linear models. The residual series again show a rising trend over the period 1980-2004.

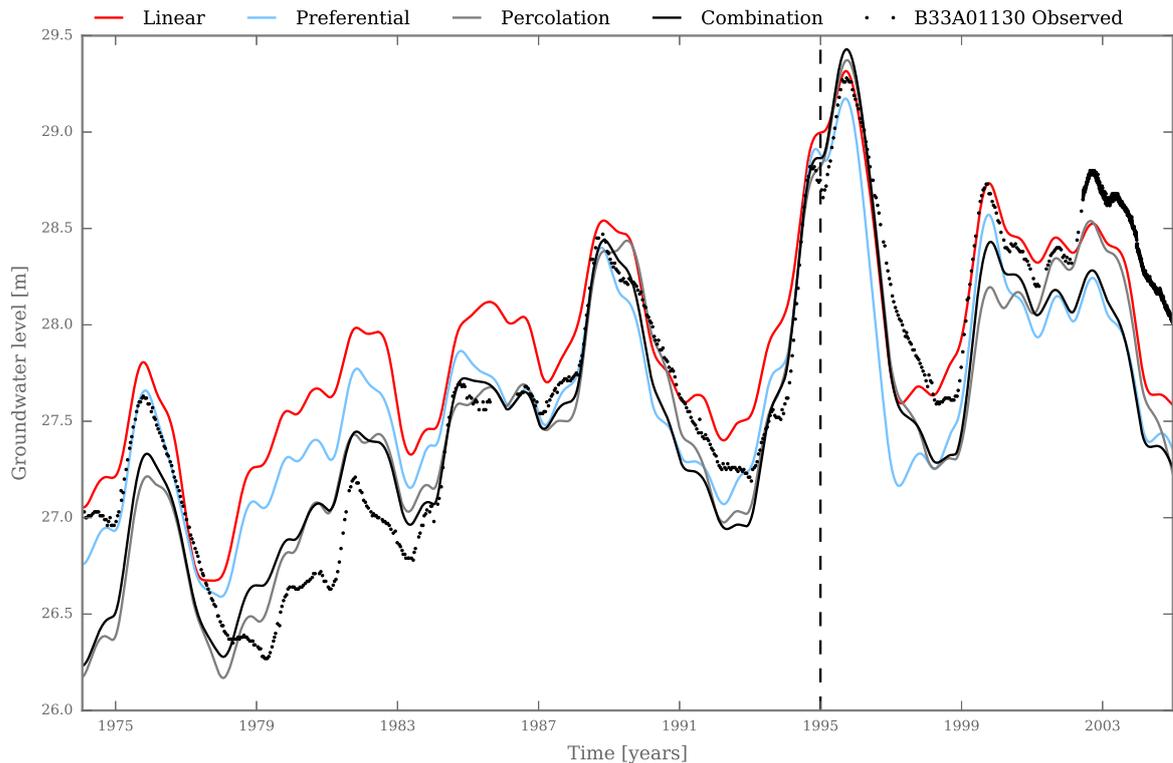


Figure 4.14: Observed and simulated groundwater levels for observation well B33A01130 using four recharge models. The vertical dashed line denotes the end of the calibration period (1974-1995).

The response times T_{99} of the linear and preferential model are around 1500 days while the combination and percolation models have response times of around 2000 days. At the same time, the linear and preferential model have significantly higher values for α of around 3000 days, while the combination and percolation models have values for α between 400 and 500 days. The error in the models with smaller response time thus depends on the error made in time steps that are farther back in time.

The percolation and combination models have a base elevation d around 24 metres, compared to values of d around 26 metres for the other models. The average yearly recharges of all four models are more or less similar, and the difference in the base elevation can only be explained through the shape and size of the impulse response function. This is also confirmed by the correlation between the parameters A , a and n and the base elevation d between 0.4 and 0.6. Longer response times may help in removing the upward trend from the residual series.

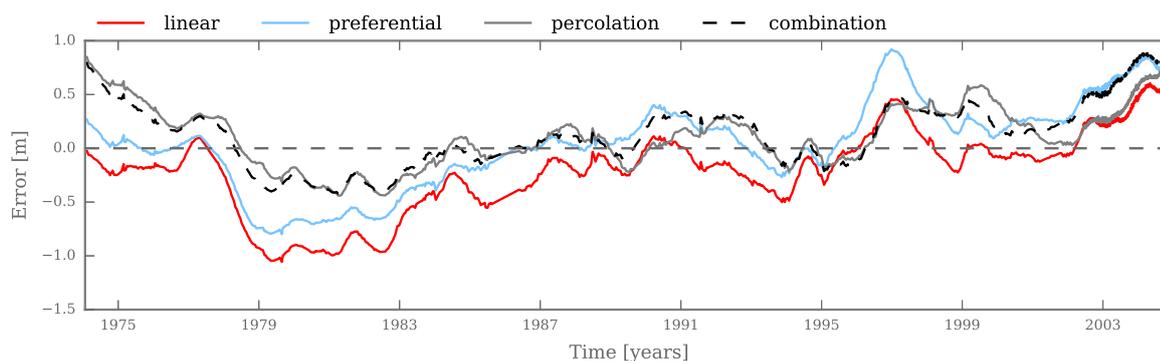


Figure 4.15: Residual series for all four recharge models calibrated on a 20-year period for observation well B33A01130.

4.4. SEPARATE PERCOLATION AND SATURATED ZONE RESPONSES

In this section the results of the models combined with a double-response function (as explained in Section 3.5) are presented. This model structure is based on the hypotheses that the dispersion and retardation of the percolation zone are described by a Gaussian distribution function. The response of the saturated zone is modelled by an exponential decay function. The models are calibrated on a 30 and 20-year period. Parameter values of the impulse response functions for all models can be found in Table 4.4. The results are described for each observation well.

4.4.1. SHALLOW OBSERVATION WELL B27D00010

The results of the four models with a double-response function calibrated on a 30-year period are shown in Figure 4.16. Compared to the models with a single response function (see Figure 4.16), there are larger differences in the simulated groundwater levels between the models. The percolation and the linear model have not improved and show a similar behaviour as the models with a single response function. The preferential and combination model perform better, especially in simulating the peaks and some of the lows. The combination model overestimates the groundwater levels after the first peak, but captures the lows around 1987 and 1999 better. The combination and preferential models both have small values for σ (see Table 4.4), resulting in sharp peaks in the groundwater levels.

The results for the models calibrated on a 20-year period are shown in Figure 4.17. The combination model shows a decreased performance compared to the model calibrated on a 30-year period, and the sharp peaks are replaced by smoother peaks. The parameter values for μ and σ are almost zero (see Table 4.4) and the smoothing is the result of the recharge that is primarily calculated as percolation flow. The linear and percolation model have a similar behaviour to the model with a single response function. The preferential model performs the best, and is better able to capture the peaks and lows.

The time to peak is almost zero for all models and is never higher than 20 days. This is much smaller than the models with a single response function discussed in Section 4.1. The response times T_{99} are between 2000 and 4000 days. Finally, the values for A , determining the peak of the exponential decay response, show a large variation between 2 and 6 (see Table 4.4). This shows that the groundwater level responds very differently to recharge in each model.

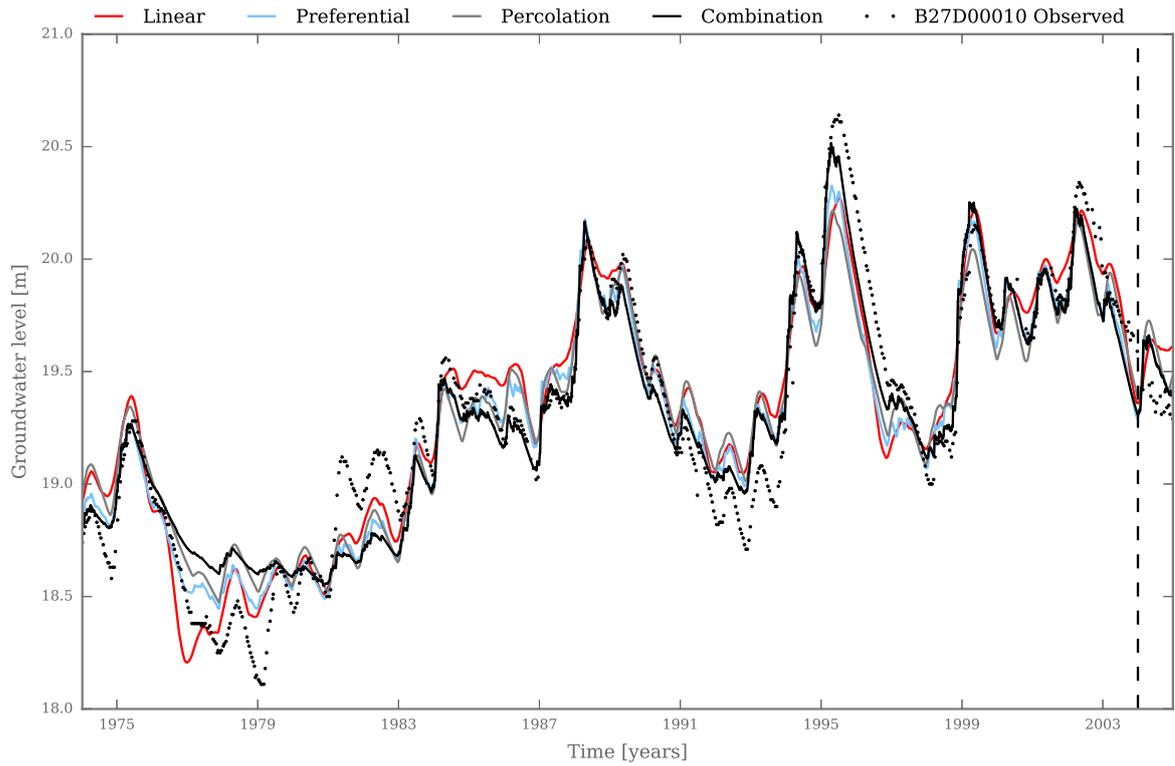


Figure 4.16: Observed and simulated groundwater levels of the four recharge models with a double-response function for observation well B27D00010. The vertical dashed line denotes the end of the calibration period (1974-2004).

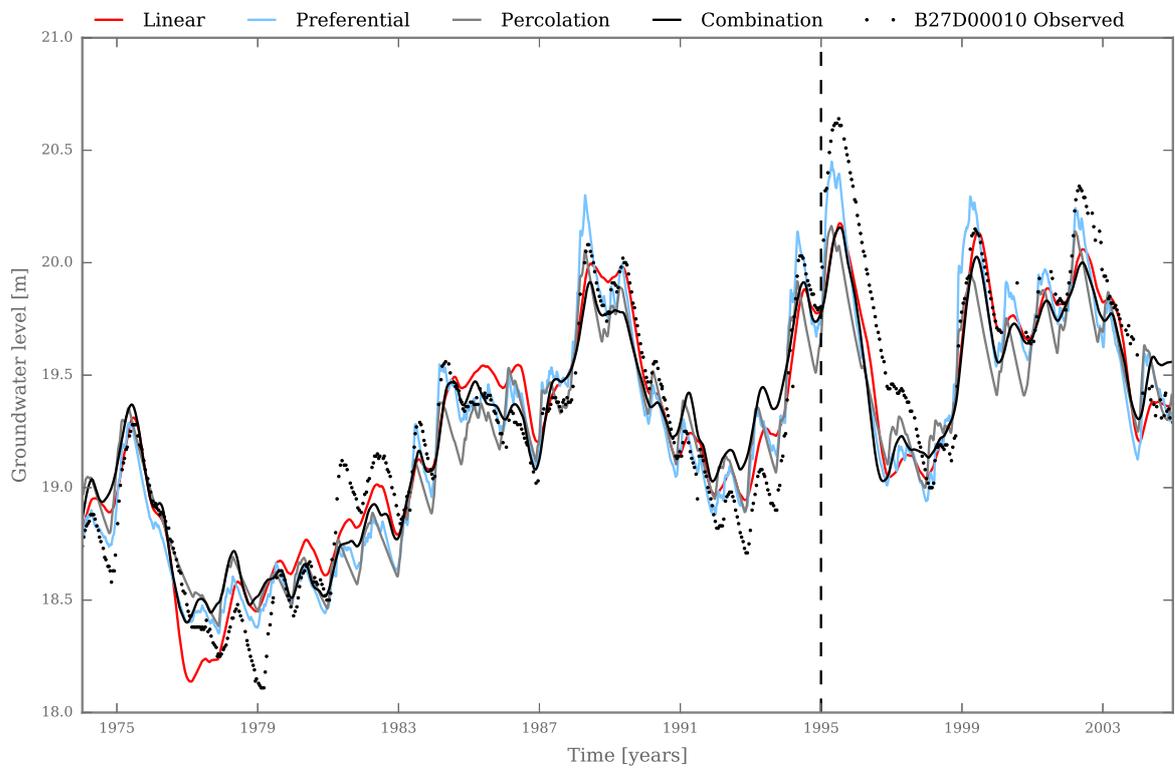


Figure 4.17: Observed and simulated groundwater levels of the four recharge models with a double-response function for observation well B27D00010. The vertical dashed line denotes the end of the calibration period (1974-1995).

4.4.2. MEDIUM DEEP OBSERVATION WELL B27C00490

The groundwater levels simulated with the four different recharge models and a double-response function calibrated on a 30-year period are shown in Figure 4.19. The results show a strong improvement in the model performance for all models, compared to the models with a single response function. The models do no longer overestimate the groundwater levels in the period 1980-1987 and the rise of the groundwater level in 1979 is also timed better by all models. The peak in 2000, which was completely ignored by the models with a single impulse response (see Figure 4.8), is now found back in the modelled groundwater levels albeit still underestimated.

The step and block response for the percolation zone of the different models are shown in Figure 4.18. The left tail of the Gaussian distribution is cut off, causing an instantaneous response of the groundwater level to recharge. This means at least part of the recharge travels quickly through the percolation zone. It could be that preferential flowpaths exist in the percolation zone or that the water added to the top of the percolation zone instantaneously forces water out at the bottom of the percolation zone.

The results of the models calibrated on a 20-year period are shown in Figure 4.20. The models show large improvements compared to the models with a single response function for the same calibration period. The low around 1999 is captured much better by all models. The improvement can also be observed in the various diagnostic statistics for the model (see Table 4.9), which almost all select the models with a double-response function as the best. The preferential model still overestimates the levels in the period 1980-1987. None of the recharge models consistently outperforms the others in capturing all the peaks and lows, but overall the combination model performs best.

The largest difference between these models and the models with a single response function is in the response times T_{99} and the time to peak T_{peak} . The models with a single response function have values for T_{99} between 1500 and 2500 days, where the models with a double-response function have response times of over 5000 days. The times to peak show a strong decrease, from between 551-644 days for models with a single response function to between 134 and 208 days for the models with a double-response function. The difference in the T_{peak} between these models is approximately one year. This possibly explains why some peaks are completely missed by the model with one response function. The time to peak as calculated by the models with two response functions are more in line with the expected values for T_{peak} based on the thickness of the unsaturated zone.

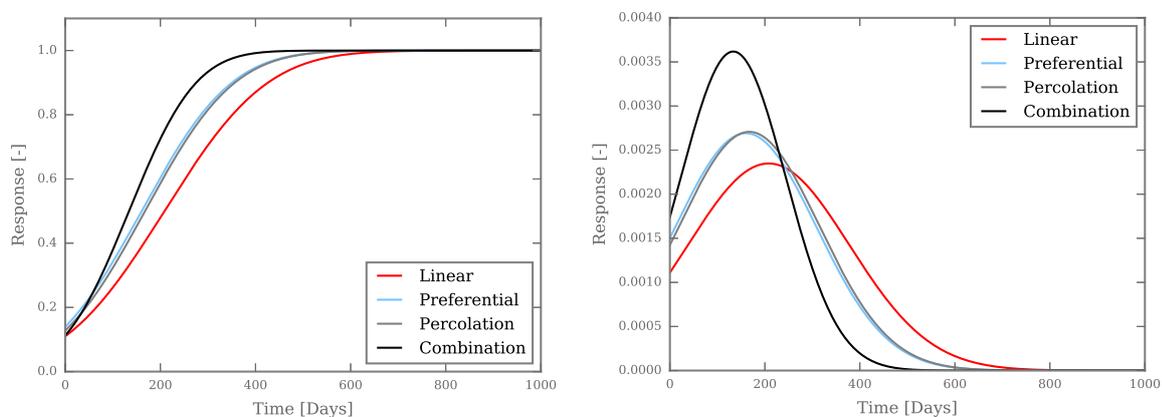


Figure 4.18: Step (left) and block response (right) for the percolation zone for the models calibrated on a 30-year period.

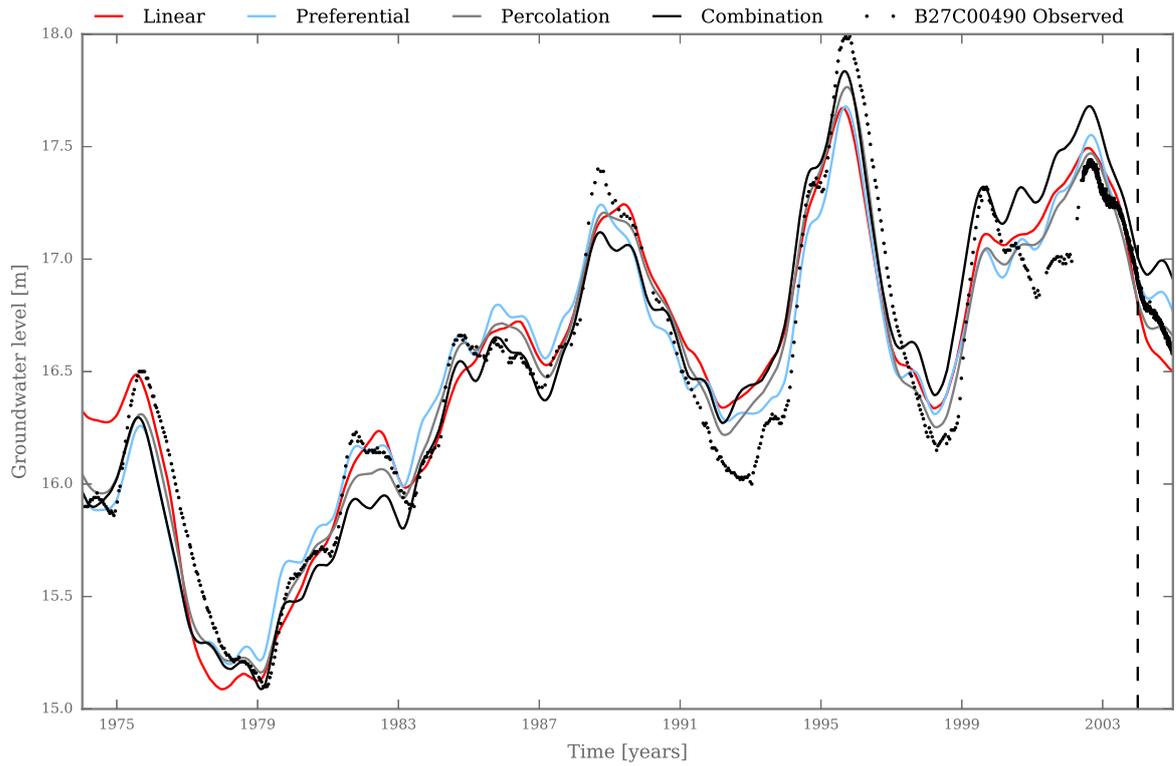


Figure 4.19: Observed and simulated groundwater levels for observation well B27C00490 using four recharge models with a double-response function. The vertical dashed line denotes the end of the calibration period (1974-2004).

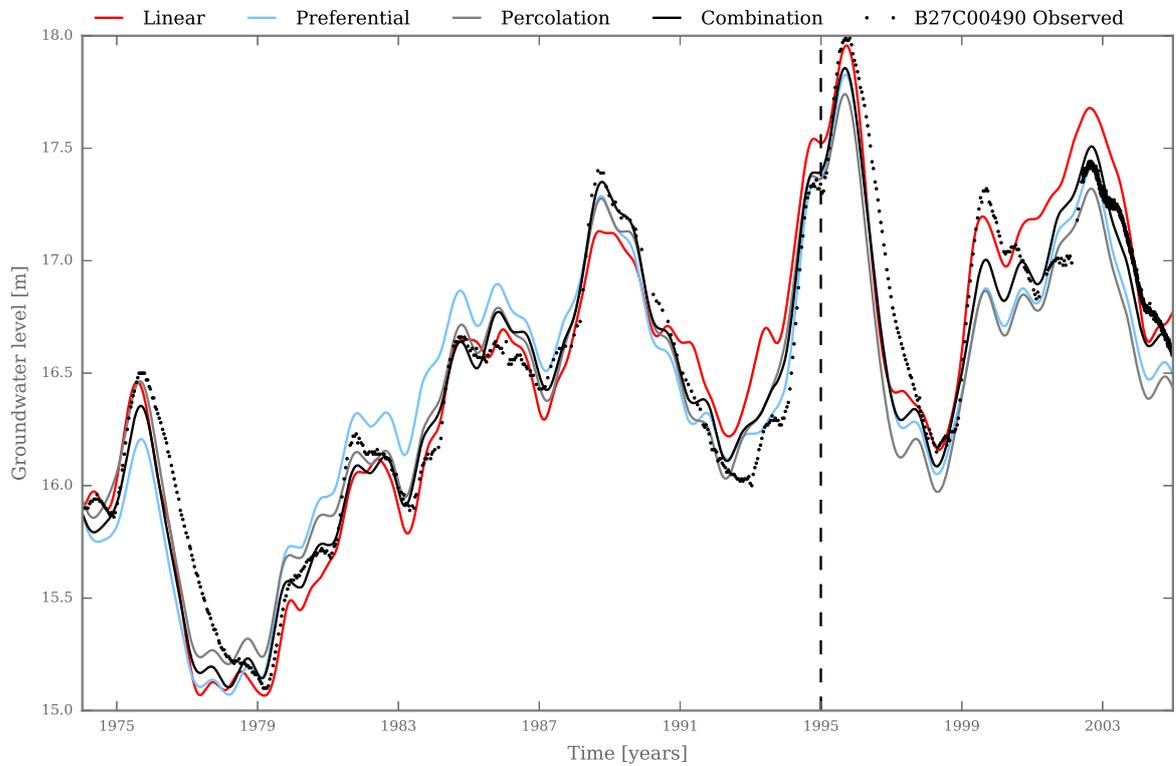


Figure 4.20: Observed and simulated groundwater levels for observation well B49C00490 using four recharge models with a double-response function. The vertical dashed line denotes the end of the calibration period (1974-1995).

4.4.3. DEEP OBSERVATION WELL B33A01130

The results of the four recharge models combined with a double-response function calibrated on a 30-year period are shown in Figure 4.21. All models show a large improvement in their performances, especially in the final 20 years of the simulation period. The peak in 1976 is underestimated by all models and the groundwater levels start rising too quickly around the following low. In the entire period after 1985, all models perform significantly better than the models with a single impulse response function. The peaks and the lows are generally captured well, with the non-linear models outperforming the linear model. The low after the 1995 peak is overestimated by all models but the following peak is captured well, where it was completely ignored in the models with a single response function.

The results of the models calibrated on a 20-year period are shown in Figure 4.22. The models again show a significant improvement to the models with a single response function. Compared to the models with a double-response function using a 30-year calibration period, the models show improved performance in the period 1980-1985. In the validation period the models show a larger deviation than the same models calibrated on a 30-year period. This is especially true for the linear model, where the average deviation increases from 0.09 in the calibration period to 0.41 metres in the validation period. The non-linear models perform better and do not exhibit this overestimation as profoundly.

The response times T_{99} for these models is significantly longer than those of the models with a single response function, with values for T_{99} between 4000 and 5000 days. In the first five years, the models underestimate the groundwater levels, followed by a sudden overestimation. Although it is less pronounced with these models, the residual series still show an upward trend in the period 1980-2004.

B27D00010				
	A	a	μ	σ
Linear	5.27	622.94	0.00	112.51
Preferential	2.60	730.65	17.28	4.46
Percolation	4.02	889.20	0.00	36.73
Combination	3.93	635.41	13.46	0.00
B27C00490				
	A	a	μ	σ
Linear	2.92	1936.04	208.42	169.85
Preferential	3.84	1658.63	160.41	148.20
Percolation	3.89	1532.55	167.63	147.34
Combination	3.37	2255.01	133.99	110.26
B33A01130				
	A	a	μ	σ
Linear	2.94	921.37	176.43	88.04
Preferential	3.31	1055.87	136.22	89.42
Percolation	3.44	1067.29	140.45	99.87
Combination	4.49	975.17	120.12	79.23

Table 4.4: Overview the impulse response parameters values for all recharge models with a double-response function calibrated on a 30-year period for three observation wells.

Observation Well B27D0001

1974-1994	SWSI	EVP (%)	RMSE	Avg Dev	AIC	BIC
Linear	0.94	99.23	0.14	0.015	-4831.26	-4806.38
Preferential	0.90	99.14	0.14	-0.011	-4893.91	-4869.03
Percolation	0.90	98.85	0.16	0.014	-4841.16	-4812.14
Combination	0.90	99.17	0.14	-0.011	-4887.66	-4854.49
1995-2004						
Linear	0.84	96.91	0.24	-0.169	-	-
Preferential	0.92	96.39	0.30	-0.241	-	-
Percolation	0.90	95.35	0.32	-0.25	-	-
Combination	0.93	96.33	0.30	-0.24	-	-

Observation Well B49C00490

1974-1994	SWSI	EVP (%)	RMSE	Avg Dev	AIC	BIC
Linear	0.47	91.14	0.32	0.14	-5848.66	-5824.48
Preferential	0.44	89.36	0.32	0.09	-5941.55	-5917.37
Percolation	0.44	99.00	0.18	-0.04	-5879.38	-5851.16
Combination	0.43	97.17	0.22	-0.01	-5922.47	-5890.23
1995-2004						
Linear	0.43	66.65	0.54	-0.48	-	-
Preferential	0.44	77.34	0.56	-0.51	-	-
Percolation	0.43	87.89	0.36	-0.31	-	-
Combination	0.43	96.02	0.46	-0.44	-	-

Observation Well B33A01130

1974-1994	SWSI	EVP (%)	RMSE	Avg Dev	AIC	BIC
Linear	0.53	90.56	0.48	0.36	-6098.40	-6073.61
Preferential	0.47	91.08	0.35	0.13	-6211.94	-9187.16
Percolation	0.51	95.13	0.28	-0.05	-6156.05	-6127.14
Combination	0.48	95.73	0.27	-0.04	-6211.36	-6178.31
1995-2004						
Linear	0.39	82.75	0.38	-0.324	-	-
Preferential	0.39	84.59	0.65	-0.627	-	-
Percolation	0.40	70.82	0.47	-0.42	-	-
Combination	0.40	63.96	0.66	-0.62	-	-

Table 4.5: Diagnostic statistics for the four recharge models with a single response function, calibrated on the period 1974-1994 and validated from 1995-2004.

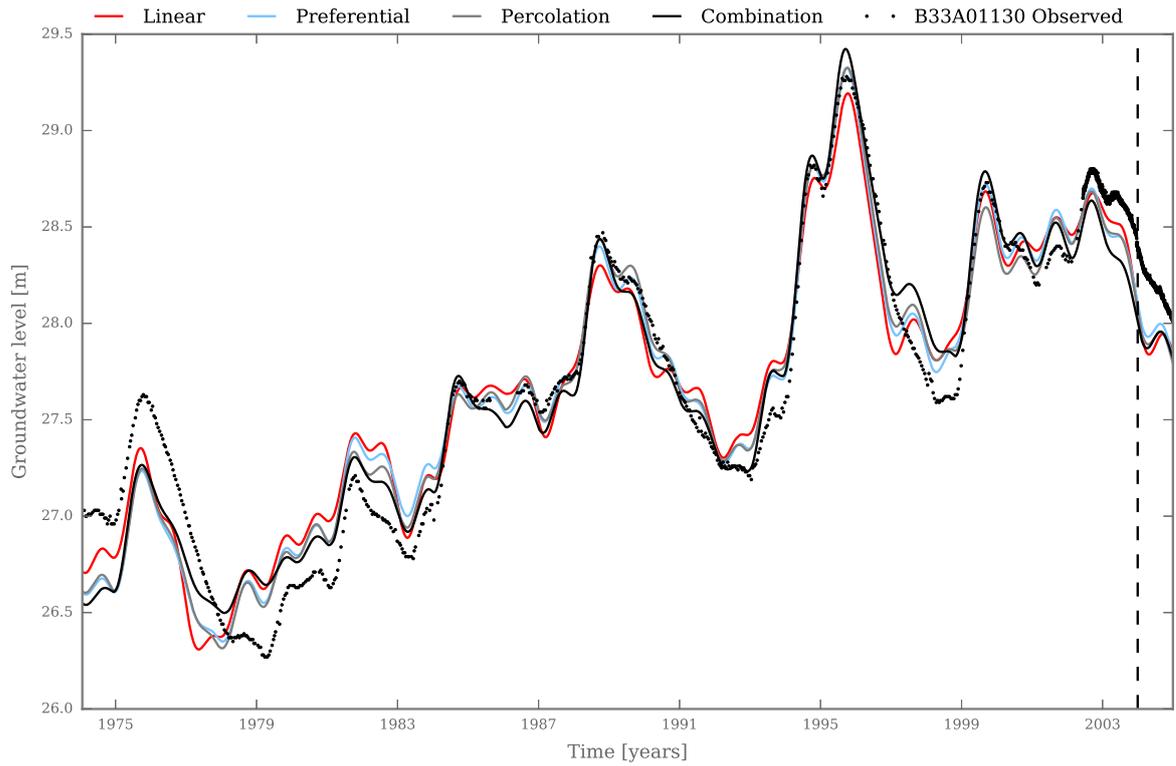


Figure 4.21: Observed and simulated groundwater levels for observation well B33A01130 using four recharge models with a double-response function. The vertical dashed line denotes the end of the calibration period (1974-2004).

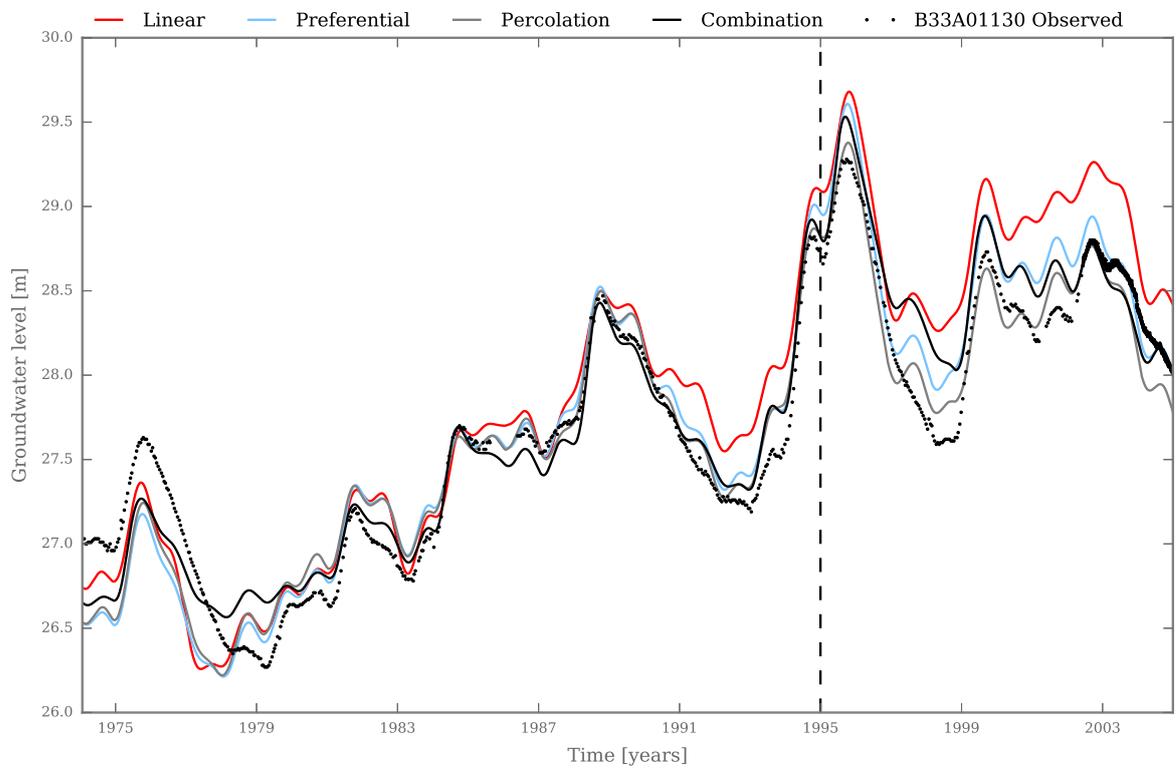


Figure 4.22: Observed and simulated groundwater levels for observation well B33A01130 using four recharge models with a double-response function. The vertical dashed line denotes the end of the calibration period (1974-1995).

4.5. DISCUSSION

In this section the results presented in this chapter are discussed in the following order. First, the recharge models are discussed, comparing the performance of the different models for all observation wells. Second, the results of the models with a single and a double response function are discussed. Each of the previous two subsections consists of a model comparison based on the diagnostic statistics and on a qualitative exploration of the model performance. Third, the methodology followed in this chapter is discussed, focussing on the parameter estimation and correlations.

4.5.1. THE RECHARGE MODELS

DIAGNOSTIC STATISTICS

An overview of the diagnostic statistics is provided in Table 4.5 for all wells of all models with a single response function calibrated on a 20-year period and validated on a 10-year period. Model performance is compared for each observation well. The best performing model for each statistic is highlighted in grey, unless the absolute differences in the values are smaller than two decimals. The need to use multiple statistics when evaluating model performance is clear, as none of the models consistently outperforms all other models in both the validation and the calibration period. However, when model parsimony (number of parameters) is taken into account, as is done by the Akaike and Bayesian Information Criteria, the preferential model is always selected as the best performing model for the 20-year calibration period.

The preferential model is preferred by four out of six statistics in the calibration period for the shallow well B27D00010. The linear model is unanimously selected as the best performing model in the validation period. The differences between the models are small, making it rather difficult to distinguish between models. For the medium deep well B27C00490, the non-linear models all have two statistics selecting them as the 'best' model in the calibration period, while the percolation model is selected by two out of four in the validation period. For the deep well B33A01130, the preferential and the combination model both have three out of six statistics selecting them as the best model in the calibration period. In the validation period, the linear model has the best performance.

Tables B.2 and B.1 in the appendix provide an overview of the diagnostics statistics for all four recharge models combined with a single or a double response function calibrated on a 20 and a 30-year period. Using only the statistics from the calibration period, a total of 24 diagnostic statistics is available for comparison of each recharge model. Table 4.6 provides a summary of the number of times a recharge model is selected as the best for each well. For the shallow observation well B27D00010 the preferential model is selected as the best model by 7 out of 24 statistics, followed by the combination model with 5 out of 24 statistics. The percolation model is selected by 11 out of 24 statistics as the best model to simulate the groundwater levels at the medium deep well B27C00490. The deep groundwater levels at the deep well B33A01130 are best simulated with the combination model, according to 12 out of 24 of the statistics. It is noted that no weighing of the statistics is used in this analysis.

QUALITATIVE EXPLORATION

Apart from evaluating the model performance by examining the diagnostic statistics, the performance is evaluated in a more qualitative way by exploring the simulated recharge series. The

Model	B27D00010	B27C00490	B33A01130
Linear	2	0	0
Preferential	7	3	4
Percolation	1	11	4
Combination	5	5	12

Table 4.6: Summary table of the diagnostic statistics provided in table B.2 and B.1. The values constitute the number of times a model has the highest performance for a statistic.

fundamental idea behind the root zone models presented in this study, is that they are a better representation of the physical processes. As was shown in Figure 4.3, the simulated recharge patterns vary greatly for the shallow observation well depending on the root zone model. This variation in the recharge pattern is also found for the deeper observation wells. The recharge patterns reveal great differences, while relatively small differences in the simulated groundwater levels are observed.

The similarity in the simulated groundwater levels can be explained by the long response times of the groundwater system, where the daily recharge is less important than the long term recharge variation to explain the groundwater level fluctuations. The yearly recharge fluxes are in fact very similar in terms of magnitude and recharge patterns, as can be observed in table 4.7 and by examining the graphs in the appendices. The average yearly recharges calculated with the non-linear model are in the same order of magnitude, while the linear model simulates a much higher or lower average yearly recharge depending on the length of the calibration period.

The results of the recharge simulations showed that the linear model is the sole model that calculates negative recharge fluxes, which are highly unlikely considering the thickness of the unsaturated zone. The root zone depth of the vegetation present at this location (coniferous and deciduous forests), is generally not more than 2 meters, and the capillary rise in sandy soils is not sufficient to cause an upward flux. The estimated recharge fluxes for the linear model are therefore physically implausible as recharge of the saturated groundwater. None of the non-linear models simulates negative groundwater recharge, as imposed by the model structure.

The linear model simulates recharge that occurs throughout the year, while the recharge events of the non-linear models occur primarily in the winter period (the areas shaded in blue in Figure 4.3). A study on the recharge mechanisms on the Veluwe using chloride and $\delta^{18}\text{O}$ isotopic tracers (Gehrels et al., 1998) concluded that most of the recharge happens in the winter months, with incidental recharge through preferential flow in the summer. This recharge distribution is better represented by the non-linear models. The preferential and the percolation flow models differ in the magnitude of the recharge fluxes, and the number of recharge events. The percolation model simulates a more continuous recharge with a smaller amplitude while the preferential model simulates individual, but larger recharge events. However, determining which of these models represents the actual recharge mechanism is difficult and would require more fieldwork.

The uncertainty in the estimate of the recharge fluxes is not consistent for each root zone model, and changes with the observation well that is modelled and the length of the calibration period. The linear model has a relatively stable confidence region of around 0.4 meter per year for all models. The uncertainty of the non-linear models shows a larger variation between the different models. However, the uncertainty of one of the non-linear models is always smaller than the linear model. Generally, the model that has the highest performance also has the lowest uncertainty in

Model	B27D00010	B27C00490	B33A01130
Linear ₃₀	0.45	0.40	0.26
Linear ₂₀	0.29	0.34	0.42
Preferential ₃₀	0.32	0.37	0.37
Preferential ₂₀	0.30	0.35	0.37
Percolation ₃₀	0.32	0.41	0.43
Percolation ₂₀	0.29	0.32	0.38
Combination ₃₀	0.33	0.38	0.39
Combination ₂₀	0.30	0.39	0.39

Table 4.7: Overview of the average annual recharge [m/year] calculated for each model calibrated on a 30 and a 20-year period using a single response function. The length of the calibration period is added in subscript to the model name.

the recharge. This is the preferential model for the shallow well B27D00010, the percolation model for the medium deep well B27C00490 and the combination model for the deep well B33A01130.

4.5.2. COMPARISON BETWEEN A SINGLE AND A DOUBLE RESPONSE FUNCTION

DIAGNOSTIC STATISTICS

An overview of the diagnostic statistics for the models with a single and a double-response function, calibrated on a 20-year period, is provided in Table 4.9. The best performing model per statistic is highlighted in grey for each observation well. The primary interest lies in the difference between the models with a single or a double-response function and not in the applied recharge model. The results suggest that the models with a double-response function significantly improve model performance for the medium deep and deep observation wells. Comparison of the results for the shallow well B27D00010 is less straightforward.

The linear or the percolation model in combination with a single response function is preferred for well B27D00010 by all six statistics. The combination model with a single response function performs best by four out of six statistics, and the preferential model with a double-response function is preferred by three out of six. A large improvement is observed for the models with a double-response function for the two deeper wells B33A01130 and B27C00490. The models with a double-response function perform best judging by virtually all the diagnostics statistics calculated for the medium deep and the deep wells.

QUALITATIVE EXPLORATION

The effect of using a double-response function is also evaluated in a more qualitative way by exploring the time to peak (T_{peak}) of the different models. In general, the shallow observation well is expected to have the shortest time to peak, and the deeper wells are expected to have a time to peak that is around 80 days later than the shallow well ($\Delta T_{peak} \approx 80$ days) based on the estimated made in Section 3.8.4. The times to peak of the models are computed for a single and a double-response function for all three observation wells and different calibration periods (see Table 4.8). There are large differences in the values for T_{peak} between the models, ranging from 0 to 644 days. In general, the lowest values are found for the shallow well B27D00010, followed by the deep well B33A01130 and finally the medium deep well B27C00490. The linear models almost always have the

Single response	Shallow	Medium deep	Deep
linear ₃₀	280	644	467
linear ₂₀	293	478	426
preferential ₃₀	179	603	426
preferential ₂₀	136	431	354
percolation ₃₀	220	551	440
percolation ₂₀	121	521	450
combination ₃₀	178	603	421
combination ₂₀	140	444	389
Double response			
linear ₃₀	0	208	176
linear ₂₀	0	227	180
preferential ₃₀	17	160	136
preferential ₂₀	18	159	140
percolation ₃₀	0	168	140
percolation ₂₀	11	163	148
combination ₃₀	13	134	120
combination ₂₀	0	156	119

Table 4.8: Time to peak (T_{peak} in days) for all models with a single and a double response function calibrated on two different periods. The length of the calibration period is added in subscript to the model name.

largest times to peak, while the differences between the non-linear models are much smaller.

The models with a single response function simulate higher values for T_{peak} than the models with a double-response function. The differences range from 100 to over 400 days between the time to peak calculated by the different models. For the two deeper observation wells, it seems that the time to peak is almost shifted a full year. The difference in the peak between observation wells, ΔT_{peak} , also shows large differences between the single and double response function. The models with a single response function show values for ΔT_{peak} between the shallow and the deeper wells of 250 and 450 days. The models with two response functions show much smaller values for ΔT_{peak} between 100 and 250 days, where larger values are primarily the simulated by the linear model. The values for ΔT_{peak} between the two deeper observation wells are small for both models, which is in line with what is expected.

SELECTION OF THE RESPONSE FUNCTION

Based on the diagnostic statistics, the models with a double response function perform best on the deeper observation wells. The time to peak (T_{peak}) and the differences in the times to peak between deeper wells and the shallow well (ΔT_{peak}) are also better represented by the double response function. For the shallow observation well B27D00010 the single and double response function perform equally well. When model parameter parsimony is taken into account, the models with a single response function are preferred for shallow well.

Observation Well B27D0001

1 IRF	SWSI	EVP (%)	RMSE	Avg Dev	AIC	BIC
Linear	0.94	99.23	0.14	0.02	-4831.26	-4806.38
Preferential	0.90	99.14	0.14	-0.01	-4893.91	-4869.03
Percolation	0.90	98.85	0.16	0.01	-4841.16	-4812.14
Combination	0.90	99.17	0.14	-0.01	-4887.66	-4854.49
2 IRF						
Linear	0.96	99.13	0.15	0.03	-4806.85	-4777.82
Preferential	0.87	99.05	0.15	0.01	-4998.69	-4969.67
Percolation	0.93	97.91	0.18	0.01	-4839.17	-4806.00
Combination	0.89	97.30	0.19	0.00	-4875.17	-4837.85

Observation Well B27C0049

1 IRF	SWSI	EVP (%)	RMSE	Avg Dev	AIC	BIC
Linear	0.47	91.14	0.32	0.14	-5848.66	-5824.48
Preferential	0.44	89.36	0.32	0.09	-5941.55	-5917.37
Percolation	0.44	99.00	0.18	-0.04	-5879.38	-5851.16
Combination	0.43	97.17	0.22	-0.01	-5922.47	-5890.23
2 IRF						
Linear	0.42	97.85	0.21	-0.01	-5953.26	-5925.04
Preferential	0.41	97.04	0.23	-0.04	-5973.60	-5945.38
Percolation	0.43	99.52	0.14	0.01	-5983.23	-5951.00
Combination	0.41	98.14	0.21	-0.06	-5989.17	-5952.89

Observation Well B33A0113

1 IRF	SWSI	EVP (%)	RMSE	Avg Dev	AIC	BIC
Linear	0.53	90.56	0.48	0.36	-6098.40	-6073.61
Preferential	0.47	91.08	0.35	0.13	-6211.94	-6187.16
Percolation	0.51	95.13	0.28	-0.05	-6156.05	-6127.14
Combination	0.48	95.73	0.27	-0.04	-6211.36	-6178.31
2 IRF						
Linear	0.50	96.74	0.26	0.09	-6155.42	-6126.50
Preferential	0.45	97.11	0.24	0.00	-6267.81	-6238.89
Percolation	0.47	97.85	0.22	0.00	-6258.19	-6225.14
Combination	0.44	98.95	0.19	0.01	-6328.92	-6291.74

Table 4.9: Diagnostic statistics for the four recharge models using a single (1 IRF) or a double (2 IRF) response function, calibrated on a 20-year period.

4.5.3. MODEL OPTIMIZATION

The estimation of parameters proved to a difficult and troublesome process in obtaining the results presented in this chapter. Correlation between parameters can be significant. For example, in the linear model the factor f and the base elevation of the model d have correlation coefficients up to 0.99. In the models with a single response function, parameters A , a and n were found to have a high correlation as well. Peterson and Western (2014) devoted some of their work to rewrite the response function to reduce correlation between parameters, but their modifications did not help in this study. The problems of parameter correlation apply in a lesser extent to the parameters of the recharge models and a double-response function. It occurred multiple times that the noise model was used to explain all the variation, leaving the deterministic part of the model as a horizontal line.

This problem was solved by applying a local gradient-based optimization scheme, along with better estimations of the initial parameter values and the implementation of parameter boundaries. Where possible, parameters are fixed (e.g. S_{rmax}) or constrained (e.g. f and K_p) such that only physically realistic values are estimated. Despite these modifications, the models are still sensitive to the initial parameter values, sometimes leading to sub-optimal results. These findings can be related to the issue of equifinality that is also observed for many rainfall-runoff models.

5

ADDITIONAL EXPLAINING STRESSES

In the previous chapter the groundwater levels were simulated using evaporation and transpiration as the explaining stresses. Different recharge models and response functions were tested to improve the fit of the models. The residual series of all models still showed periods with deviations in one direct (too high or too low), despite the improvements that were obtained, . The models for the two deeper wells showed an upward trend after 1980. It is suggested that other processes can explain these deviations, resulting in a better fit.

The effect of other processes that may influence the groundwater levels in this area is explored in this chapter. The chapter is structured as follows. Possible influences on the groundwater levels and how these can be incorporated in the model are presented in the next section. The results of the application of the additional explaining variables are presented in the subsequent three sections, one section for each observation well. The chapter concludes with a discussion of the model results.

5.1. ANTICIPATED INFLUENCES ON THE GROUNDWATER LEVELS

Fluctuations in the groundwater levels can be caused by many variables and relating these to groundwater fluctuations can be difficult. The processes can be correlated, making it difficult to separate the effects using a time series model. An example of the difficulties in separating the influences of several explaining series is found in Shapoori et al. (2015). Apart from precipitation and evaporation, Gehrels (1999) identified three major factors that may influence the groundwater levels in the study area: land reclamation, groundwater extractions, and land use and vegetation cover changes. (Gehrels, 1999) tested the first two factors as separate input series in a time series model while the influence of the latter was investigated through a water balance method. In this study, only the effects of land reclamation and groundwater extractions are simulated to try and improve the performance of the time series models.

5.1.1. RECLAMATION

In 1956 and 1967 two large polders named Oostelijk and Zuidelijk Flevoland were reclaimed from Lake IJssel to the Northwest of the Veluwe. The water levels in the polders were lowered several

metres which also lowered groundwater levels in the surrounding area. Gehrels et al. (1994) studied the impact of the 1967 reclamation on the groundwater levels in the Veluwe and found a significant effect of the reclamation, depending on the distance to the polders. The maximum expected decline is around 0.35 m after 25 years at the deep well B33A01130, based on estimations for the central Veluwe. The final lowering due to the reclamation is around 0.6 m at this location. The shallow well B27D00010, located 12 km from Oostelijk Flevoland and 23 km from Zuidelijk Flevoland, is likely to be primarily impacted by the 1956 reclamation due to its location. Observation well B27C00490 is located at approximately 16 km from both polders, and is impacted by both. The decline at the latter two observation wells is likely of the same order of 0.5-1.0 m.

The decline is modelled with the step function:

$$\theta_r(t) = B \cdot (1 - e^{-\frac{t-t_s}{b}}) \quad (5.1)$$

where B and b are parameters that need to be estimated, and t_s is defined as the moment when the reclamation starts to have effect.

5.1.2. GROUNDWATER EXTRACTIONS

Groundwater extractions are also known to cause a decline of the groundwater levels in this area (Berendrecht et al., 2006). Groundwater is primarily extracted for drinking water supply, but also for industrial use and engineering practices. The use of groundwater in the study area increased consistently in the period 1950-1995 but has stabilized since. Information on the groundwater extractions is not centrally and not publicly available, making it difficult to account for all extraction wells for the entire period. Data is obtained for the major well fields.

The impulse response function that is used to simulate the effect of the groundwater extractions is a modified form of the Theis equation, known as the Ferris and Knowles' well formula (Shapiro et al., 2015):

$$\theta_w(t) = \frac{v}{t} e^{\frac{-r^2 u}{t}} \quad (5.2)$$

where r is the distance between the pumping and the observation well and v and u are parameters with no direct physical meaning. Multiple wells can be simulated by superposition. A single value for u and v can be used when the storativity and the transmissivity are assumed to be uniform. Otherwise, each pumping well has its own values for u and v . The daily extraction rate is convoluted with the block response function of equation 5.2. The derivation of the block response function from the impulse response function is similar to that of the recharge response function.

5.2. SHALLOW OBSERVATION WELL B27D00010

In the previous chapter the preferential recharge model with a single response function was selected as the best model to simulate the groundwater levels for this well. In this section, all four recharge models in combination with a single response function are used for the simulations. Gehrels et al. (1994) simulated the groundwater levels for the period of 1960-1995 and found a clear pattern in the residuals. Data is now available for a longer period and groundwater levels are simulated for the period 1960-2004 with all four recharge models.

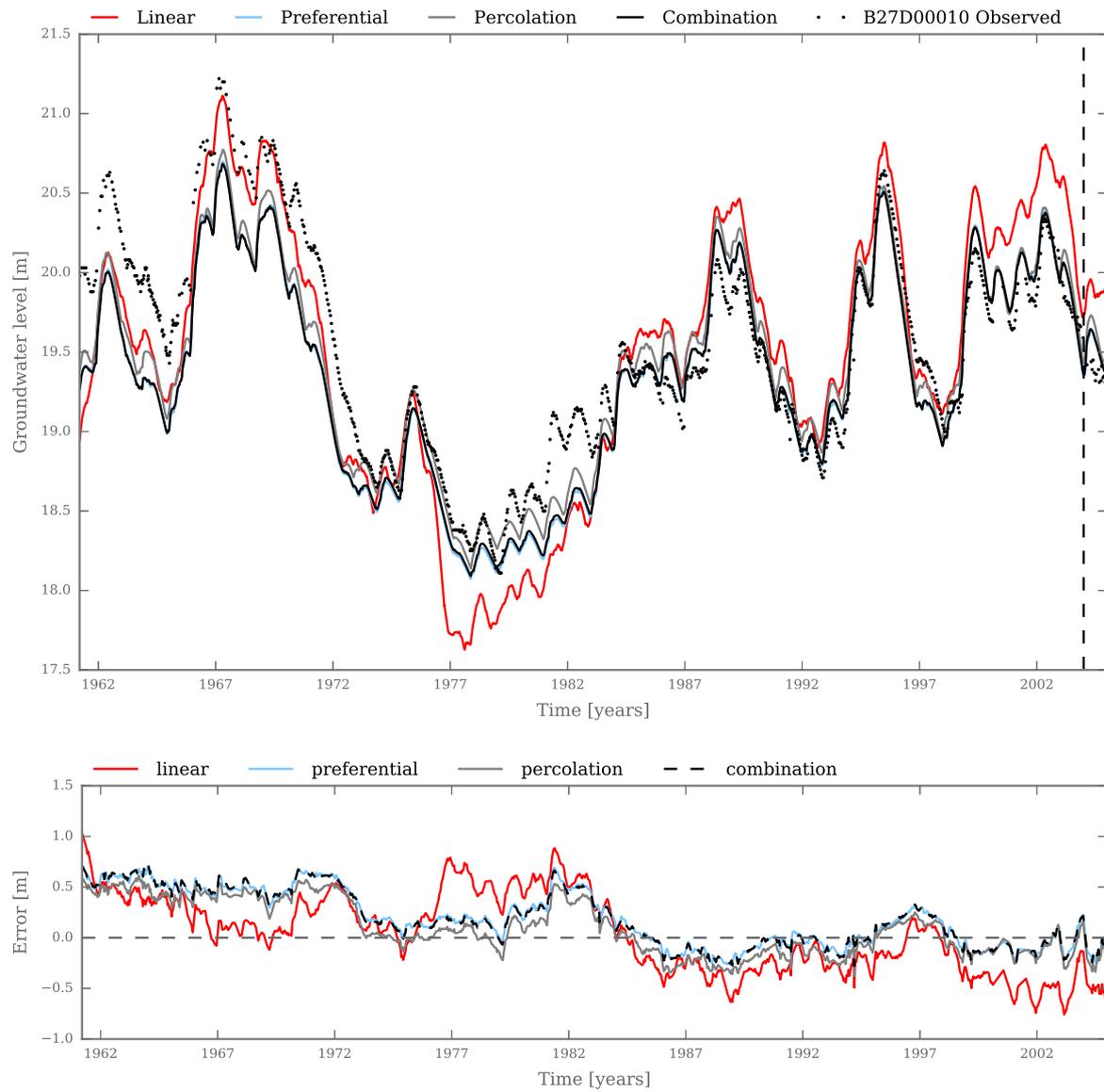


Figure 5.1: Observed and simulated groundwater levels for observation well B27D00010 simulated for the period 1960-2004.

The simulated groundwater levels and the residuals are shown in Figure 5.1. The performance of all models simulated for this longer period decreases compared the models simulated for the period 1974-2004 (see Figure 4.1). The models either accurately simulate the first ten years and overestimate the final period (e.g., the linear model) or underestimate the first period and show a reasonable fit in the final period (e.g., the combination model). The non-linear models are outperforming the linear model in this case, as can also be observed by examining the diagnostic statistics in the Tables B.4 and B.3 in the Appendix. The next subsections show the results of adding the effect of reclamations and groundwater pumping to the model for this longer simulation period.

5.2.1. RECLAMATION

The shallow well B27D00010 is located at approximately 12 km from Oostelijk Flevoland and 23 km from Zuidelijk Flevoland. It is primarily influenced by the reclamation of Oostelijk Flevoland which was completed in 1956. Gehrels et al. (1994) studied the effect of the 1967 reclamation on the groundwater levels in the Veluwe in this area. Based on the analysis of the residual series of a time series model with no additional explanatory stresses, the study concludes that "The influence of Zuidelijk Flevoland does not extend this far", referring to the area where the shallow well is located. Therefore, only the effect of the 1956 reclamation has been modelled in this study. The parameter B , the maximum head increase due to reclamation, has been given an upper limit of $B = 0.0$ m and an lower limit of $B = -2.0$ m for the optimization.

The groundwater levels simulated by the models that include the effect of reclamation, calibrated over the period 1960-2004 are shown in Figure 5.2. The results are similar to those of the models calibrated on a 30-year period 1974-2004 (see Figure 4.1). The groundwater levels simulated by the non-linear models are rather similar, while the levels simulated by the linear model show a slightly different pattern. The low in 1977 is underestimated by the linear model, as has also been observed in Chapter 4. The preferential and the combination model have the best performance independent on the length of the calibration period used (see Table B.3 in the appendices).

The residual series and the simulated effect due to the reclamation are also shown in Figure 5.2. The residual series are very similar to the model calibrated on 30-year period in Chapter 4 (see Figure 4.2). The decline in 2004 due to the reclamation at the location of the shallow well is estimated at around 1.1 m by all models. The ultimate decline due to the reclamation B is estimated at around 1.5 m by the non-linear models and 1.95 m by the linear model.

The groundwater levels simulated by the models that include the effect of the reclamation, calibrated on the period 1960-1995 are shown in Figure C.17 in the Appendices. All four models show a large improvement compared to the results shown in Figure 5.1. The residual series shown in Figure C.17 are similar to those of the models in Chapter 4 calibrated on the period 1974-1995 (see Figure 4.5). In the validation period the groundwater levels are consistently underestimated by all models. The decline due to the reclamation that is simulated is provided in Figure C.17, showing declines of approximately 1.3 m over the simulation period. The fitted value of B is 2.0 m for all models, the boundary that is set for this parameter in the optimization process.

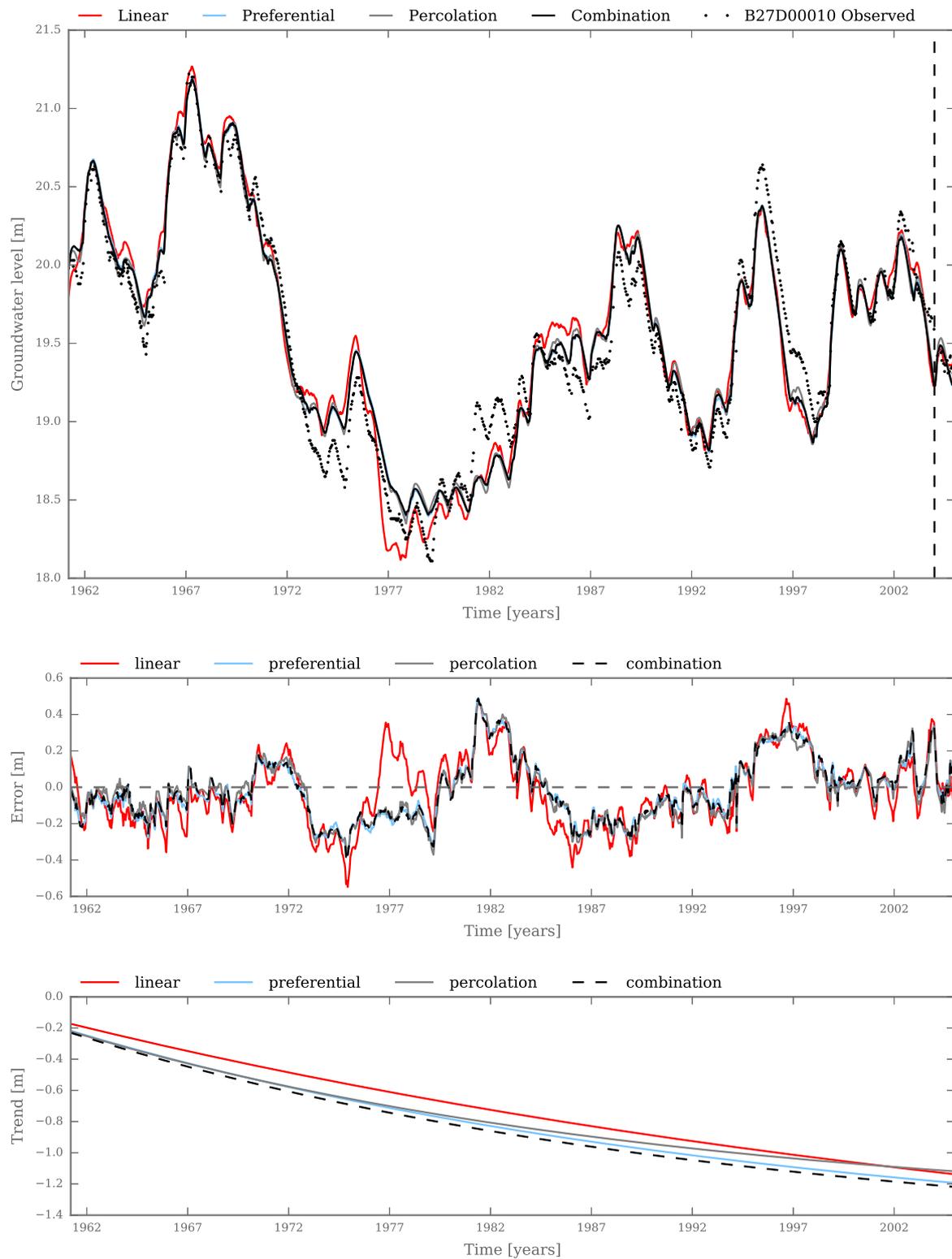


Figure 5.2: Simulated groundwater levels, residuals, and the drawdown for all four recharge models with the effect of the reclamation for well B27D00010. The models are calibrated on the period 1960-2004.

5.2.2. GROUNDWATER PUMPING

Another explanation for the decline in the groundwater levels in this area can be found in the extraction of groundwater. Two major extractions are located in the area, both at approximately 5 km distance. Both wells are owned by the drinking water company Vitens and are called Haere and Epe (see Figure 2.1). It is assumed in this study that these wells are located in a homogeneous aquifer at approximately similar distances. These approximations make it possible to sum the discharges of the two wells and fit a single response function to simulate the effect of the groundwater extractions. The discharges for the individual extractions are shown in Figure 5.3. The reported groundwater extractions in Epe and Haere started in 1958 and 1966 respectively.

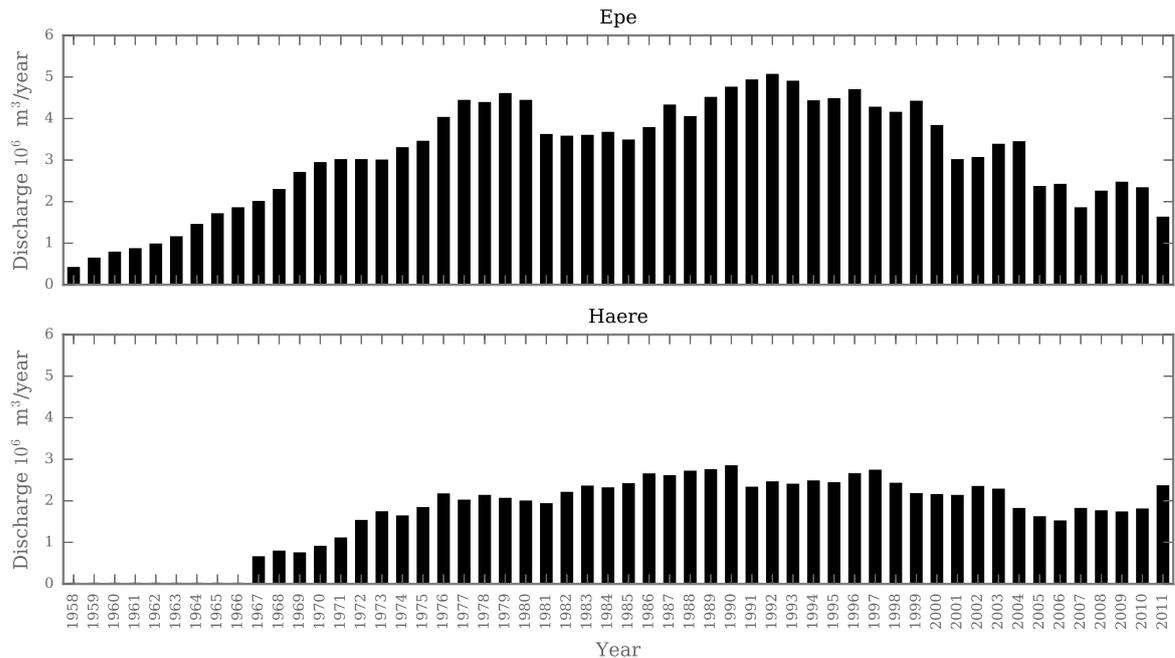


Figure 5.3: Discharges from the Vitens Pumping wells in Haere and Epe in Million m^3 per year. Data is extracted from two reports on these extraction wells Folmer et al. (2012a); Krikken et al. (2014).

The groundwater levels simulated with the models that include the effect of the extractions calibrated on the period 1960-2004 are shown in Figure 5.4. The model performance is significantly improved by adding the effect of the extractions, compared to the simulated groundwater levels shown in Figure 5.1. The simulation of the lows around 1992-1993 is improved in particular. The simulation of the other peaks and lows is not improved by adding the effect of the extractions to the model. The preferential and the combination model again show the best performance when the diagnostic statistics in Table B.3 in the Appendices are examined.

The residual series and the simulated drawdown due to the extractions are also shown in Figure 5.4. The residuals are similar to the model calibrated on 30-year period in Chapter 4 (see Figure 4.2). The trend found in the residuals shown in Figure 5.1 is effectively removed. The effect of the extractions is similar for the non-linear models, but is clearly different for the linear model. The groundwater levels simulated with the non-linear models respond faster to changes in the extraction rate.

The results of the models that include the effect of the extractions, calibrated on the period 1960-1995, are shown in Figure C.18 in the Appendices. The fit of the linear, preferential and percolation

model is comparable to the fit of the models calibrated on a 20-year period in Chapter 4 (see Figure 4.5). Apart from the percolation model, all models underestimate the groundwater levels in the validation period. The percolation and the combination models perform better than the latter two models in the validation period. The percolation and combination models also simulate that the effect of the drawdown levels off, while the others indicate that the drawdown is still increasing. As the percolation model consistently has the lowest residuals, the drawdown simulated by the percolation model is likely to be the best representation of the actual drawdown due to pumping.

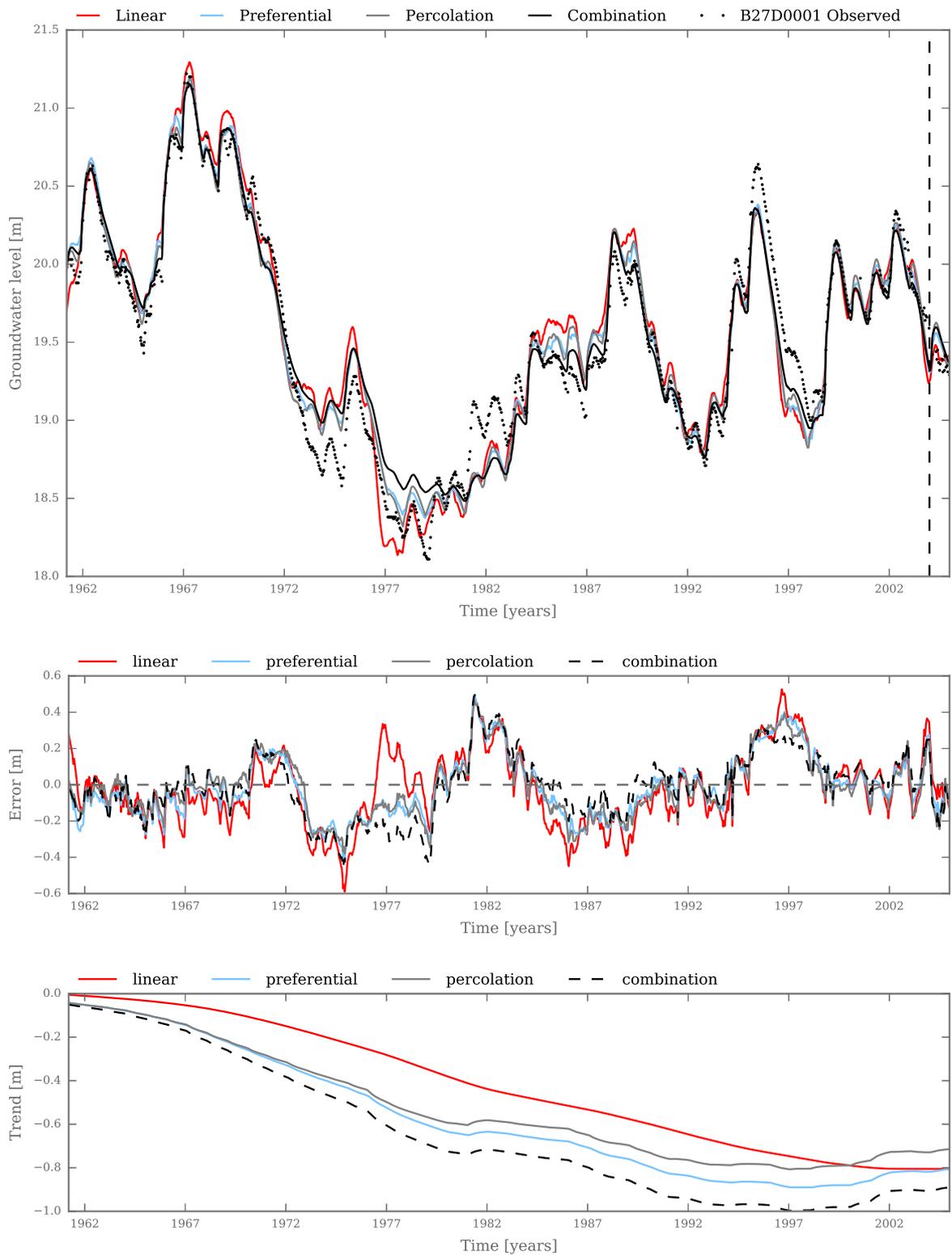


Figure 5.4: Simulated groundwater levels, residuals, and the drawdown for all four recharge models with the effect of extractions for well B27D0001. The models are calibrated on the period 1960-2004.

5.3. MEDIUM DEEP OBSERVATION WELL B27C00490

5.3.1. RECLAMATION

The medium deep well B27C00490 is located at approximately 16 km from both polders, and the groundwater levels are affected by the reclamation of both polders. The groundwater levels at this well are observed since 1974, 20 years after the start of the decline due to the 1956 reclamation. It is likely to be difficult to separate the effect of the the 1956 and 1967 reclamations on the groundwater levels. Therefore, only the 1967 reclamation is taken into account for this observation well. The parameter of the noise model α is constrained to a maximum of $\alpha = 1000$ days to obtain reasonable results. The parameter B , the ultimate rise due to reclamation, has an upper limit of $B = 0.0$ m and a lower limit of $B = -1.0$ m.

The groundwater levels simulated by the models that include the effect of the 1967 reclamation, calibrated on the period 1974-2004 are shown in Figure 5.5. The residual series and the simulated decline is also shown in Figure 5.5. The linear model does not estimate a decline due to the reclamation, while the three non-linear models do. The groundwater levels simulated by the non-linear models are similar to each other, and the largest differences can be found with the linear model. The non-linear models are better at simulating the peaks and the lows in the groundwater levels. The estimated values for B all sit at the lower boundary of $B = -1$ m. Including the effect of the 1967 reclamation in the model does not improve the simulation of the groundwater levels for all models, judging from the diagnostic statistics for the different models (Table B.5 in the Appendix).

The results for the models calibrated on a 20-year period are shown in Figure C.19 in the Appendix. Only the preferential and combination model simulate a decline due to the reclamation for this calibration period. The percolation model, for which no decline is estimated, yields a similar result to the other non-linear models. This is an indication that the addition of the effect of the reclamation does not necessarily improve model performance.

After a peak in the groundwater level, the levels decline too fast. This is more the case for the models with a 20-year calibration period than for the models calibrated on a 30-year period. The response times for the models calibrated on a 30-year period are longer and higher values of σ are estimated. These longer response times in both response function are necessary to simulate declining groundwater levels well.

5.3.2. GROUNDWATER EXTRACTIONS

This location has the least effect of groundwater pumping, as only one nearby extraction well is found at 6 km distance from the observation well (see Figure 2.1). The location is called the Speuld and is owned by drinking water company Vitens. This well field has been closed in 2015 and is replaced by increased extractions in Harderwijk. No discharge rates could be found for this location. The next well with a significant discharge is located at the Amersfoortseweg near Apeldoorn, but the distance to the extraction is assumed to be too large to expect a significant effect on the groundwater levels. A decline in the groundwater level due to local groundwater extractions is not included in the model for this location due to a lack data.

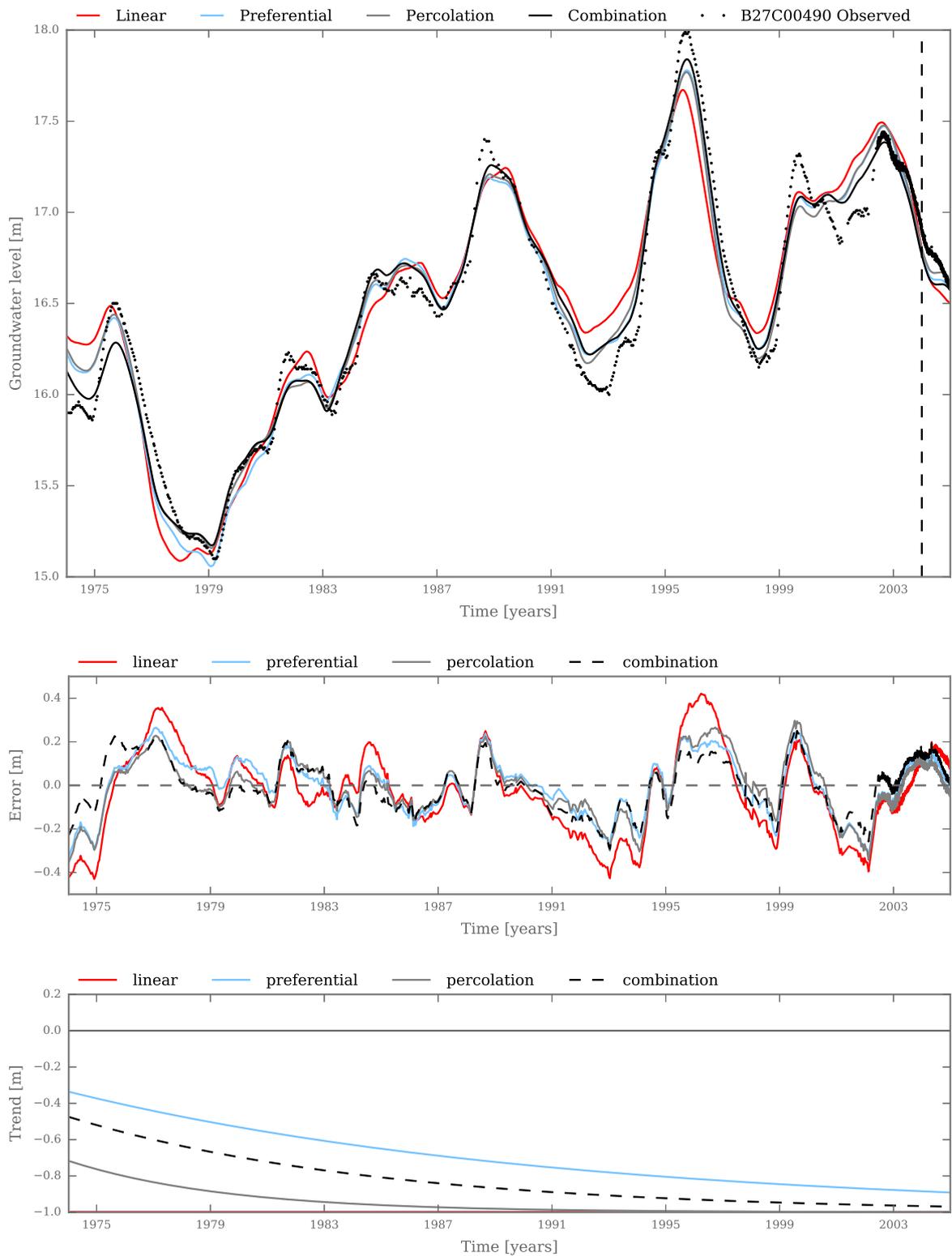


Figure 5.5: Simulated declines due to the reclamation of Zuidelijk Flevoland in 1967 for the models calibrated on a 30-year period.

5.4. DEEP OBSERVATION WELL B33A01130

5.4.1. RECLAMATION

The deep observation well B33A01130 is located at approximately 25 km distance from both Oostelijk Flevoland and Zuidelijk Flevoland. Only the effect of the 1967 reclamation is considered in this model, for the same reasons mentioned for observation well B27C00490. The parameter B has been given a lower limit of $B = -1.0$ m and an upper limit of $B = 0.0$ m for the optimization.

The simulated groundwater levels for the models with the effect of the reclamation calibrated on a 30-year period are shown in Figure 5.6. The four models show similar results, while the percolation model did not simulate a decline due to the reclamation. The effect of adding reclamation to the model is primarily observed in the first few years. The non-linear models still outperform the linear model in the rest of the period. All models calibrated an ultimate rise of $B = -1.0$. This is more than the maximum decline of 0.6 m that was estimated by Gehrels et al. (1994).

For the models calibrated on a 20-year period only the preferential model simulated a decline of the groundwater levels due to the reclamation (see Figure C.20 in the Appendices). the combination model did not simulate a decline due to reclamation, but the simulated groundwater levels did change (see Figure 4.14). The groundwater levels simulated by the other models remain largely similar. Judging from the diagnostic statistics for the models calibrated on the two period, not all models are improved by adding the effect of the reclamation (see Tables B.7 and B.8 in the Appendix).

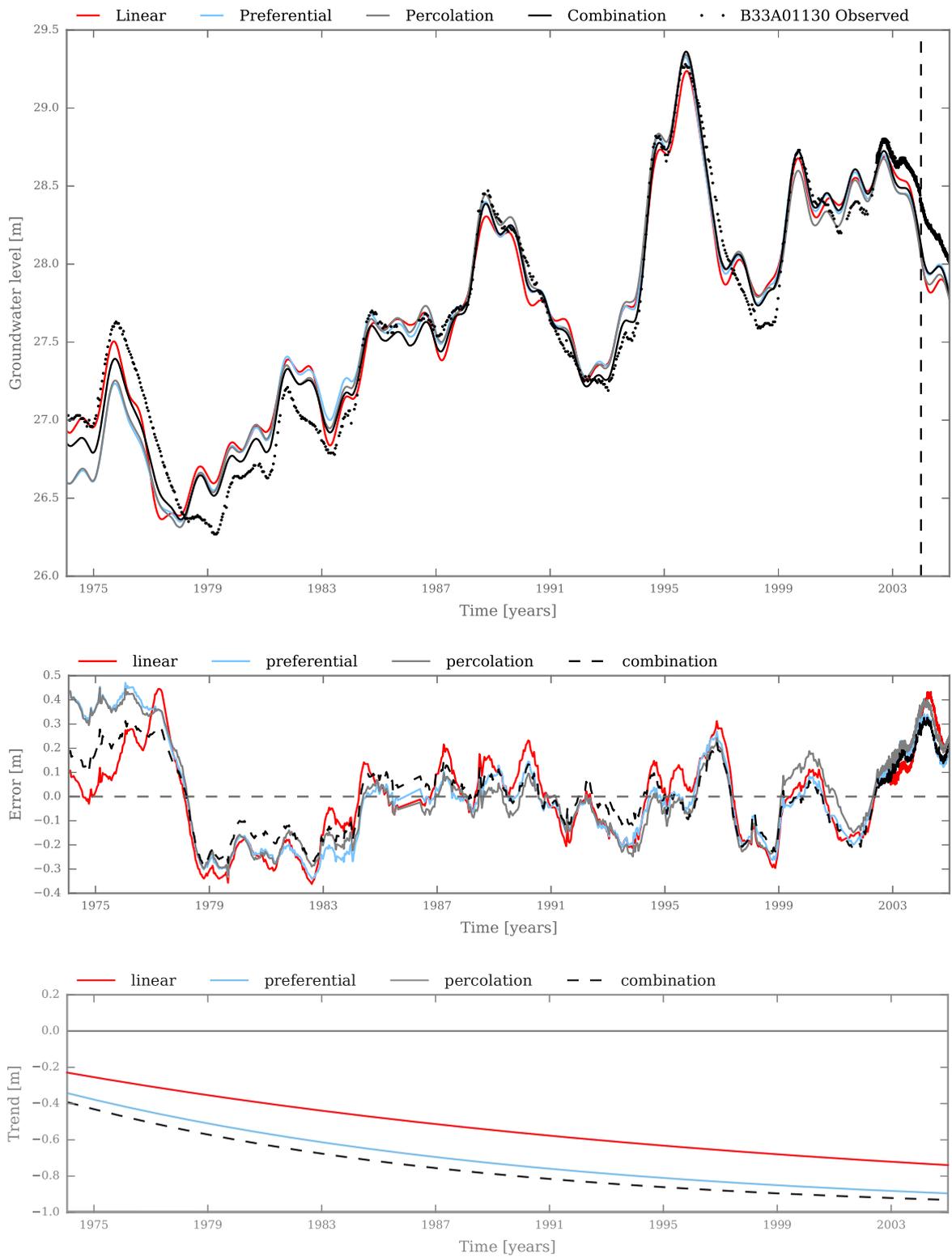


Figure 5.6: Simulated decline of the groundwater levels due to the reclamation of Zuidelijk Flevoland in 1967 for the model calibrated on a 30-year period.

5.4.2. GROUNDWATER PUMPING

For this location three extractions in the vicinity of the observation well are found (see Figure 2.1). Two well fields operated by the drinking water company Vitens and one that is located in the park of the royal palace 't Loo. The extraction well at Hoog Soeren is probably closed as no permits to extract groundwater or detailed discharge series could be found. As far as is known by the author, the discharge at the location Hoog Soeren was relatively small ($400.000 \text{ m}^3/\text{year}$). The pumping well at 't Loo is located around 1 km from the observation well and has a pumping rate of $400.000 \text{ m}^3/\text{year}$. No detailed historic discharges are known for this well either.

The well field at the Amersfoortseweg is located at approximately 2 km distance from the observation well and is important for the drinking water supply in the region of Apeldoorn. It has a varying discharge between 2 and 9 million m^3/year and yearly discharge rates are known from 1954-2010 (see Figure 5.7). Since no detailed discharge data is available for the first two extractions and the discharge at the Amersfoortseweg is relatively large, only the drawdown as a result of the latter is taken into account in the model.

The groundwater levels for the models that include the effect of the groundwater extraction calibrated on a 30-year period are shown in Figure 5.8. The model fit shows considerable improvements, particularly in the first ten years of the period. The low in 1979 is still overestimated by all models, but most of the other peaks and lows are simulated accurately. The improvement of the model when the effect of the groundwater extractions is added can also be observed in the various diagnostic statistics. Judging from Tables B.7 and B.8 in the Appendix most of the statistics are better for the models that include the effect of the groundwater extractions.

The simulated drawdown due to the groundwater extractions shows large differences, particularly between the linear and preferential and the percolation and the combination models (see Figure C.21 in the Appendix). The percolation and the combination models show a drawdown that increases with time. This is compensated by the value for the base level d , which is about a metre higher for the percolation model. The parameters of the well response function show a high correlation ($r > 0.9$) with the base level d , troubling the model optimization. Obtaining results that improve the fit by plausible declines due to the extractions proves to be difficult when the parameters are unconstrained.

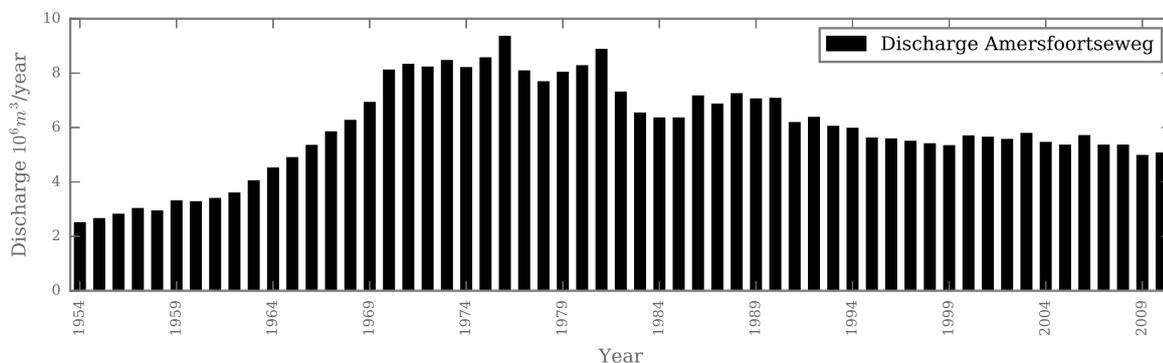


Figure 5.7: Discharge from the Vitens Pumping well at the Amersfoortseweg in Million m^3 per year (Folmer et al., 2012b).

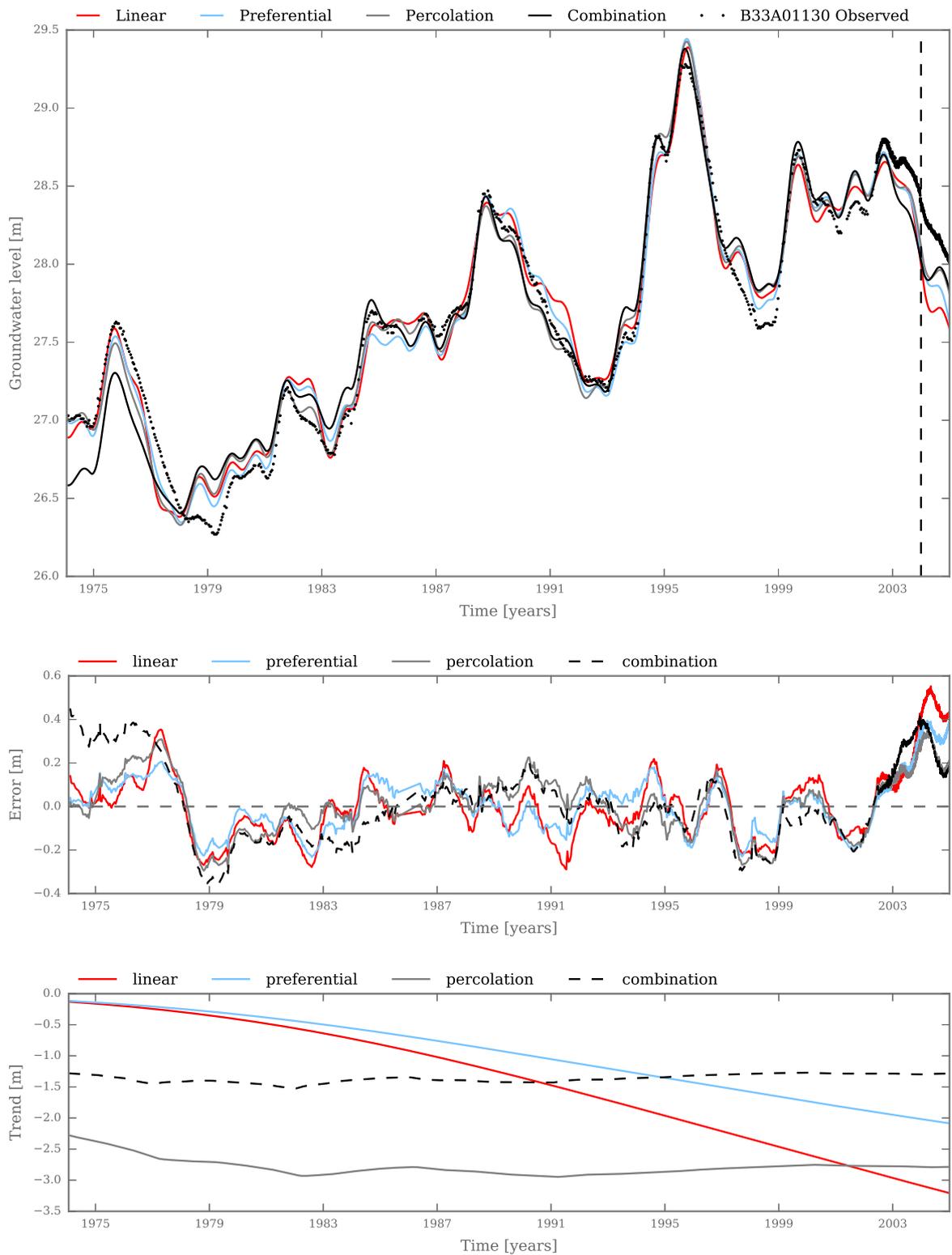


Figure 5.8: Simulated groundwater levels, residuals, and the drawdown for all four recharge models with the effect of extractions for well B33A01130. The models are calibrated on the period 1974-2004.

5.5. DISCUSSION

In this chapter the results of the models with additional explaining stresses were presented. In this section these results are discussed and the stress is selected for each observation well that improved model performance most. The importance of the length of the calibration period is discussed, as well as the difficulties in separating different stresses from each other using time series models.

5.5.1. ADDITIONAL STRESSES FOR THE SHALLOW WELL B27D00010

The groundwater levels for the shallow well B27D00010 are simulated for the period 1960-2004 with a model using a single response function. Compared to the results of the simulation for the period 1974-2004 presented in the previous chapter, model performance decreases significantly and a trend in the residuals is observed. The addition of the effect of the reclamation to the model significantly increases model performance. This is clearly observed in the values of the diagnostic statistics for the models calibrated on the period 1960-2004 (see Table B.3). The root mean squared error (RMSE) decreases from 0.33 m to 0.18 m by adding the effect of the reclamation, averaged over the different recharge models. The average deviation decreases from 0.14 m to 0.02 m by adding the effect of the reclamation. The models calibrated on the period 1960-1995 show similar improvements (see Table B.4)

The addition of the effect of groundwater extractions to the model yields similar improvements as the addition of the effect of land reclamations. The root mean squared error also decreases from 0.33 m to 0.18 m and the average deviation from 0.14 m to 0.02 m. The models calibrated on the period 1960-1995 show similar improvements (see Table B.4). The drawdown simulated by the different models show large differences, while similar performances are observed in the calibration period 1960-1995. In the validation period all models overestimate the groundwater levels, except for the percolation model. It is therefore concluded that the drawdown simulated by the percolation model is the most appropriate.

The difference in the performances of the models with reclamation or extractions as additional stresses are small (see Tables B.3 and B.4). The simulated declines due to land reclamations and due to groundwater extractions show a similar pattern and are correlated. The effects of land reclamations and groundwater extractions are interchangeable, explaining why the performance of the different models is similar. Since the effects are interchangeable and model performance is similar, it is concluded that both additional stresses are suitable. It is recommended however, to add only one of these stresses when no additional information is available to further constrain the models.

5.5.2. ADDITIONAL STRESSES FOR THE MEDIUM DEEP WELL B27C00490

For the medium deep well B27C00490 only the effect of the land reclamation is taken into account, as no data could be found on the groundwater extractions in the vicinity of the well. Model performance does not significantly change when the effect of land reclamation is added to the models (all using a double response function), judging by the diagnostic statistics provided in Tables B.5 and B.6. The performance even decreases for some models, when the effect of land reclamation is added. It is concluded that including the effect of the reclamation does not significantly improve the simulation of the groundwater levels for this well.

5.5.3. ADDITIONAL STRESSES FOR THE DEEP WELL B33A01130

For the deep well B33A01130 both the effect of reclamation and groundwater extractions are added as additional stresses to the models with a double response function. From the models that include the effect of the reclamation, calibrated on the period 1974-2004, only the percolation model does not simulate a decline of the groundwater level due to the reclamation. The root mean squared error averaged for the models that simulated a decline decreases from 0.21 m to 0.15 m. The average deviation decreases from 0.08 m to 0.06 m for these models (see Table B.8). The preferential model is the only model that simulates a decline due to the reclamations, when calibrated on the period 1974-1995. The effect of adding land reclamation is clear, as this model is selected as the best model by four out of six diagnostic statistics (see Table B.7).

The effect of groundwater extractions is considered for this observation well. The addition of the effect of groundwater extractions to the model increases the model performance of all models. This is particularly observed on the first ten years of the simulation period. The increase in model performance is also reflected in the diagnostic statistics (see Table B.8). The explained variance percentage (EVP) is increased for all models, and the root mean squared error (RMSE) is decreased for all models. The simulated drawdown due to extractions is very different for all models, while the simulated groundwater levels are largely similar. The drawdown simulated by the percolation and combination model is likely to be a better representation of the actual drawdown, as the decrease in the extractions can be observed in the declining drawdown.

The models calibrated on the period 1974-1995 show increased performances for all models. The root mean squared error decreases from 0.23 m to 0.13 m, averaged over the different recharge models (see Table B.7). The explained variance percentage is higher for all models compared to the models without the effect of groundwater extractions. The simulated drawdown due to the extractions again show large differences between models.

The models that include the effect of the groundwater extractions are selected as the best models for this well. These models generally have the highest performance judging by the diagnostic statistics provided in Tables B.8 and B.7. Moreover, the first ten years of the simulation period are modelled better by these models compared to the models that include the effect of the reclamations. It has to be noted that model performance could be further increased by improving the simulation of the drawdown when a validation period is used.

5.5.4. LENGTH OF THE CALIBRATION PERIOD

For the calibration of the models presented in this chapter, two calibration periods were used. For the two deeper wells these are 1974-1995 and 1974-2004 and for well B27D00010 these periods are 1960-1995 and 1960-2004. Adding the explaining variables generally increases model performance, but the simulated effect of the reclamation and the groundwater extractions is not always realistic. This is especially the case for the short calibration period, but to a lesser extent also for the long calibration periods. While the model fit is improved in the calibration period, the groundwater levels are consistently underestimated or overestimated in the validation period.

Not all parameters used to model the effect of the reclamation and the groundwater extractions are not constrained in this study, and parameter correlation (particularly with d) troubled the optimization process. For more plausible results and a greater confidence in the model results outside the calibration period, better initial estimations or parameter constraints should be implemented.

Also, the length of the calibration period is shown to be important for the parameter estimation, but no rule on the necessary length of the calibration period could be deduced from these results.

5.5.5. WHAT IS REAL?

In this chapter it was shown how additional influences can be used to improve the fit of the time series model. The fundamental question is which process or combination of processes is causing the groundwater level fluctuations. As also noted by Gehrels et al. (1994) the effect of groundwater extractions in the Veluwe and the effect of the land reclamations are very similar. It is difficult to decompose the effects and simultaneously estimate the parameters for these models when the input series are correlated.

In this study, only one of these variables at a time was added to the model. This means that the effects are probably overestimated, to correct for the absence of one of the processes. The implication is that this model can not be used to evaluate the effect of the groundwater extractions on the groundwater levels. The simulation of the groundwater levels for observation well B27D00010 showed the importance of the period that is used to calibrate the model in identifying the existence of other processes. While precipitation and evaporation data were sufficient to simulate the groundwater levels in the period 1975-2004, additional variables were necessary to explain the groundwater fluctuations over the period 1960-2004.

6

CONCLUSIONS

6.1. INTRODUCTION

The objective of this study was to improve the simulation of deep groundwater levels by time series models with pre-defined impulse response functions. This was attempted by adding a conceptual non-linear root zone model to simulate the recharge series to the model and by testing the use of a separate response function for the percolation zone. The main difference between these models and the classic linear time series models is in the non-linearities of the root zone that are taken into account. The models were tested on three observations wells located in the Veluwe, a largely forested area in the Netherlands characterized by thick unsaturated zones. The effect of groundwater extractions and land reclamations were added to the model to further improve the simulation of the groundwater levels.

6.2. DOES A CONCEPTUAL ROOT ZONE MODEL IMPROVE THE MODEL PERFORMANCE?

Three non-linear root-zone models were developed to simulate the recharge series that are used in the time series model simulating groundwater levels. Two recharge mechanisms were identified and combined in three recharge models: preferential flow, percolation flow and a combination of the two. The performance of these models is compared to a linear model that is widely applied in time series analysis of groundwater levels. Several diagnostic statistics were used as well as a qualitative exploration of the simulated recharge series to compare model performances.

The recharge patterns simulated by the different recharge models are very different, while the simulated groundwater levels are rather similar. Clearly, the time series model is flexible enough to simulate similar groundwater levels with different input series. The linear model simulated negative recharge fluxes, which cannot be representative of the recharge to the saturated zone considering the thickness of the unsaturated zone. The linear model simulated groundwater recharge that occurs throughout the year, while the non-linear models simulate recharge that occurs primarily in the winter period. Based on the results of isotope research in this area (Gehrels et al., 1994) the recharge pattern simulated by the non-linear model is more plausible. It is concluded that the recharge simulated with the linear model is not a correct representation of the actual recharge mechanism in

this area. The non-linear models give more plausible recharge series, but it is not possible to conclude which of the recharge patterns resembles the actual recharge best with the methods and data available in this study.

In general, the non-linear models were better capable of simulating the fluctuations in the groundwater levels. The most prominent example of this is found around the low in 1977 at well B27D00010, where only the linear model largely underestimates the groundwater levels. Judging from the diagnostic statistics the performance of the non-linear models is higher than the linear model. For each observation well, a different non-linear recharge model performs the best. The preferential flow model is selected for the shallow well B27D00010, the percolation model for the medium deep well B27C00490 and the combination model for the deep well B33A01130.

When no information on the exact recharge mechanism is available, it is recommended to use the combination root-zone model in future research. This model is capable of simulating both recharge mechanisms and hence the model will not be constrained to a single process. It will also allow the data to tell which mechanism is dominant, as was for example the case for the shallow well, where the combination and preferential model simulated virtually identical recharges.

6.3. DOES THE USE OF A SEPARATE RESPONSE FUNCTION FOR THE PERCOLATION ZONE IMPROVE MODEL PERFORMANCE?

The use of a separate response function for the percolation and for the saturated groundwater zone in the time series model was tested and compared to the models with a single response function for both zones. This model structure added one extra parameter to be estimated. For the shallow observation well B27D00010 no significant improvement in model performance was observed. The simulation of the deeper groundwater levels at the observation wells B27C00490 and B33A01130 improved significantly as the models were better able to capture the peaks and lows. This is confirmed by the diagnostic statistics for these models, which predominantly selected the models with two response function as the best model. The time to peak (T_{peak}) and difference in the time to peak between wells (ΔT_{peak}) showed a large decrease when two response functions were used. The values obtained with two response functions were closer to the values that were estimated in Section 3.8.4.

6.4. CAN MODEL PERFORMANCE BE IMPROVED BY ADDING STRESSES?

Two additional stresses were used to try to improve the simulation of the groundwater levels: groundwater extractions and land reclamations. The groundwater levels of the shallow well B27D00010 were accurately simulated for the period 1974-2004, but showed a trend in the residual series when simulated for the period 1960-2004. Including the decline due to the reclamation of Oostelijk Flevoland or groundwater extractions in the area removed the trend. Since the two stresses have similar effects on the groundwater levels and the model performance was almost identical, it is concluded that only one of these stresses need to be taken into account.

The model performance for observation well B27C00490 did not improve when the effect of the reclamation of Zuidelijk Flevoland was added to the model. No data on groundwater extractions in the area were found and the land reclamation of Oostelijk Flevoland was not added for this well. For the deep well B33A01130, including the effect of the land reclamation or the groundwater

extractions both improved model performance. The groundwater levels could be simulated more accurately when the effect of groundwater extractions was added to the model, and it is concluded that the effect of the extractions has to be taken into account for this well. For all models and observation wells, the groundwater levels were either overestimated or underestimated when a 10-year validation period was used and reasonable results could only be obtained when the entire period is used for calibration. Without further information to constrain the simulated effects of the two stresses, it is recommended to use the entire simulation period when these stresses are added to the time series model.

7

RECOMMENDATIONS

7.1. THE USE OF LOCAL PRECIPITATION DATA

In the preliminary phase of this study precipitation data from the KNMI weather station in de Bilt was used. This is the same location as is used for the reference evaporation data applied in this study, at approximately 60 km distance from the wells. It was assumed that the groundwater levels are the result of the long term average precipitation and that local precipitation events would not have a significant impact. However, when local precipitation data was used, as is done in this study, large improvements in the model performance for all wells were observed. It is therefore recommended to always apply local precipitation data when available. It has not been tried to apply local evaporation data in this study, but it seems reasonable to expect local deviations in the evaporation data as well. It is recommended to study the impact of using evaporation data from different locations on the simulated groundwater levels in future research.

7.2. VALIDATING THE RECHARGE SIMULATION

It has been shown in this study how different recharge patterns can be used to simulate similar groundwater levels. This result is a clear example of equifinality, where different model structures and parameter sets lead to similar results (Beven, 2006). It is concluded that the linear model simulates recharge patterns that are physically implausible. The non-linear models simulate more plausible recharge fluxes. It is also concluded that it is impossible to select the recharge model that represents the recharge best with the methods used in this study. Validation of the simulated recharge series requires other information. Measuring recharge fluxes with a high temporal resolution is a difficult and often expensive task (Scanlon et al., 2002). Water content measurements in the percolation zone may give insight into the recharge process. However, this method is only feasible if percolation flow is the governing process in this zone. If preferential flow is a dominant process, tracer techniques may be more suitable for the task.

7.3. IS THE PERCOLATION ZONE LINEAR?

Important approximations are made in this study on the linearity and flow mechanisms of the percolation zone. The linearity means that the state (e.g. water content) of the percolation zone has no (significant) effect on the pulse travelling through it. The flow mechanism is assumed to be diffuse, with no preferential flow paths through the percolation zone. There are two findings in this study that indicate violations of these approximations. The models with a double-response function often gave values for σ and μ that cause the groundwater levels to instantly respond to the recharge impulse. This could indicate the existence of preferential flow in the percolation zone, causing (part) of the recharge to reach the groundwater very quickly after a recharge event.

The second reason is the variation in the time to peak ΔT_{peak} for different rainfall events. In this study the peak in 1995 was used to calculate ΔT_{peak} , but in a short analysis of other peaks, different estimates of this value (± 20 days) were found. The time step at which the groundwater levels are observed can possibly explain these differences, but it could also indicate that the velocity with which a recharge impulse travels through the unsaturated zone depends on the state of the percolation zone. The recommendations given in the previous section can help in studying how water travels through the percolation zone.

7.4. IMPROVING PROCESS CONSISTENCY

In conceptual hydrological modelling, recent work has shown the possibility of improving the predictive capacity of models by constraining models based on four criteria: 1) model structure, 2) parameter constraints, 3) model prior constraints and, 4) modelling objective (Hrachowitz et al., 2014). Constraining the model structure can be done by more information on the processes, in this study implemented by identifying the recharge mechanisms that are present the study area. Parameter constraints were implemented by fixing parameters (e.g. S_{rmax}) or by providing plausible parameter ranges. Prior constraints were introduced in this study by estimating the values for T_{peak} and ΔT_{peak} . Model objective constraints were not used in this study, but examples can be found in studies calibrating or validating the recharge models on the measured water content in the unsaturated zone (Gehrels et al., 1998; Berendrecht et al., 2006).

It is in this line of thought that (non-linear) time series models can be developed further to improve their predictive capabilities. The values for T_{peak} and ΔT_{peak} were used as a soft indicator in this study, but were not used as a prior constraint to reject models or constrain the parameters. When a better estimation of the time to peak is available, the impulse response function can be rewritten to include T_{peak} as a parameter that can be constrained (see section A.3 in the Appendix for the proposed impulse response function). Other possible characteristic values for the impulse response function could be proposed for the height of the peak and the response time T_{99} of the system.

The simulation of the recharge series can also be constrained further to improve the process consistency of the recharge calculation. The percolation model in this study has the capability to simulate overland flow, while this process is almost certainly absent in the study area. It is possible to test if overland flow is simulated and to reject models that simulate a significant overland flow flux. Moreover, the actual evaporation is also simulated by the recharge models and these series could be compared to available data (e.g., from remote sensing). Finally, as stated in the second to last paragraph, the water content of the root zone reservoir could be compared to measurements of the water content in the root zone.

7.5. THE NEED FOR OPEN SOURCE FLEXIBLE MODELS

As diverse as the hydrological systems are, as diverse are the conceptualizations of these systems. In other words, many models are possible and the software used to model them should be flexible enough to test different hypotheses. This often means the modeller has to adapt the source code, which can easily be done if the software is written in a programming language with a freely available compiler or interpreter. However, as far as the author is aware, no completely open source time series models are available. Peterson and Western (2014) made their source code available with their publication, but that relies on the commercial software package Matlab (although it might work in the open source package Octave).

The development of such an open-source framework that supports flexible model building and other research needs is recommended and could increase the application of time series models to groundwater levels and the development of methods for it. The software used in this study is written in the open-source programming language Python and is freely available. Future researchers are encouraged to test different conceptualizations of the unsaturated zone for a diverse range of geologies, soil characteristics, vegetation, land cover, depth to water table, and climate using this software, or by publishing their own.

BIBLIOGRAPHY

- Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6):716–723.
- Bakker, M., Maas, K., and Von Asmuth, J. R. (2008). Calibration of transient groundwater models using time series analysis and moment matching. *Water Resources Research*, 44(4):W04420.
- Bakker, M. A. J. and Meer, J. J. M. V. D. (2003). Structure of a Pleistocene push moraine revealed by GPR: the eastern Veluwe Ridge, The Netherlands. *Geological Society, London, Special Publications*, 211(1):143–151.
- Berendrecht, W. L., Heemink, A. W., van Geer, F. C., and Gehrels, J. C. (2006). A non-linear state space approach to model groundwater fluctuations. *Advances in Water Resources*, 29(7):959–973.
- Berendsen, H. (2008). *De vorming van het land*.
- Besbes, M. and De Marsily, G. (1984). From infiltration to recharge: use of a parametric transfer function. *Journal of Hydrology*, 74(3–4):271–293.
- Beven, K. (1993). Prophecy, reality and uncertainty in distributed hydrological modelling. *Advances in Water Resources*, 16(1):41–51.
- Beven, K. (2006). A manifesto for the equifinality thesis. *Journal of Hydrology*, 320(1–2):18–36.
- Bierkens, M. F. P., Knotters, M., and van Geer, F. C. (1999). Calibration of transfer function–noise models to sparsely or irregularly observed time series. *Water Resources Research*, 35(6):1741–1750.
- Box, G. E. and Jenkins, G. M. (1970). Time series analysis, control, and forecasting. *San Francisco, CA: Holden Day*, (3228).
- Breuer, L., Eckhardt, K., and Frede, H.-G. (2003). Plant parameter values for models in temperate climates. *Ecological Modelling*, 169(2–3):237–293.
- De Weerd, M. G. P., Savenije, H. H. G., Euser, T., Hrachowitz, M., Verkade, J. S., Winsemius, H. C., TU Delft: Civil Engineering and Geosciences: Water Management, and TU Delft, Delft University of Technology (2014). Modelling ungauged lowland basins: Does complementary groundwater data add value to topography driven conceptual modelling?
- Droogers, P. (2009). Verbetering bepaling actuele verdamping voor het strategisch waterbeheer. *Definitiestudie. STOWA*.
- Fenicia, E., McDonnell, J. J., and Savenije, H. H. G. (2008). Learning from model improvement: On the contribution of complementary data to process understanding. *Water Resources Research*, 44(6):W06419.
- Fenicia, E., Savenije, H. H. G., Matgen, P., and Pfister, L. (2006). Is the groundwater reservoir linear? Learning from data in hydrological modelling. *Hydrol. Earth Syst. Sci.*, 10(1):139–150.

- Fitts, C. R. (2013). 1 - Groundwater: The Big Picture. In Fitts, C. R., editor, *Groundwater Science (Second Edition)*, pages 1–22. Academic Press, Boston.
- Folmer, I., van Herpen, F., Gommers-Verbeek, L., and Krikken, A. (2012a). Gebiedsdossiers Gelderland Drinkwaterwinning Epe. Technical Report 9X0236, Aequator, Royal HaskoningDHV.
- Folmer, I., van Herpen, F., Gommers-Verbeek, L., and Krikken, A. (2012b). Gebiedsdossiers Gelderland Winning Amersfoortseweg - Apeldoorn. Technical Report 9X0236, Aequator, Royal HaskoningDHV.
- Freeze, R. A. and Harlan, R. L. (1969). Blueprint for a physically-based, digitally-simulated hydrologic response model. *Journal of Hydrology*, 9(3):237–258.
- Gao, H., Hrachowitz, M., Schymanski, S. J., Fenicia, F., Sriwongsitanon, N., and Savenije, H. H. G. (2014). Climate controls how ecosystems size the root zone storage capacity at catchment scale. *Geophysical Research Letters*, 41(22):2014GL061668.
- Gehrels, J. (1999). *Groundwater level fluctuations : separation of natural from anthropogenic influences and determination of groundwater recharge in the Veluwe area, the Netherlands*. PhD thesis, [S.l.]; [S.l.].
- Gehrels, J. C., Peeters, J. E. M., Vries, J. J. D., and Dekkers, M. (1998). The mechanism of soil water movement as inferred from 18o stable isotope studies. *Hydrological Sciences Journal*, 43(4):579–594.
- Gehrels, J. C., van Geer, F. C., and de Vries, J. J. (1994). Decomposition of groundwater level fluctuations using transfer modelling in an area with shallow to deep unsaturated zones. *Journal of Hydrology*, 157(1–4):105–138.
- Goes, B. J. M. (2000). Ground penetrating radar as a tool to improve local groundwater flow models of the ice-pushed ridges in the Netherlands. volume 4084, pages 572–578.
- Hansen, N., Müller, S., and Koumoutsakos, P. (2003). Reducing the Time Complexity of the Derandomized Evolution Strategy with Covariance Matrix Adaptation (CMA-ES). *Evolutionary Computation*, 11(1):1–18.
- Harbaugh, A. W. (2005). *MODFLOW-2005, the US Geological Survey modular ground-water model: the ground-water flow process*. US Department of the Interior, US Geological Survey Reston, VA, USA.
- Hipel, K. W. and McLeod, A. I. (1994). *Time Series Modelling of Water Resources and Environmental Systems*. Elsevier.
- Hrachowitz, M., Fovet, O., Ruiz, L., Euser, T., Gharari, S., Nijzink, R., Freer, J., Savenije, H. H. G., and Gascuel-Oudou, C. (2014). Process consistency in models: The importance of system signatures, expert knowledge, and process complexity. *Water Resources Research*, 50(9):7445–7469.
- Kavetski, D., Kuczera, G., and Franks, S. W. (2006). Calibration of conceptual hydrological models revisited: 2. Improving optimisation and analysis. *Journal of Hydrology*, 320(1–2):187–201.
- Knotters, M. and Bierkens, M. F. P. (2000). Physical basis of time series models for water table depths. *Water Resources Research*, 36(1):181–188.
- Krikken, A., Snijders, J., and de Rooij, G. (2014). Gebiedsdossiers Gelderland Winning De Haere. Technical Report BC7201-101, Aequator, Royal HaskoningDHV.

- Maas, C. (1994). *On convolutional processes and dispersive groundwater flow*. PhD thesis, TU Delft, Delft University of Technology.
- Newville, M., Stensitzki, T., Allen, D. B., and Ingargiola, A. (2014). *LMFIT: Non-Linear Least-Square Minimization and Curve-Fitting for Python*.
- Olsthoorn, T. N. (2008). Do a Bit More with Convolution. *Ground Water*, 46(1):13–22.
- Peterson, T. J. and Western, A. W. (2014). Nonlinear time-series modeling of unconfined groundwater head. *Water Resources Research*, 50(10):8330–8355.
- Savenije, H. H. G. (2004). The importance of interception and why we should delete the term evapotranspiration from our vocabulary. *Hydrological Processes*, 18(8):1507–1511.
- Scanlon, B. R., Healy, R. W., and Cook, P. G. (2002). Choosing appropriate techniques for quantifying groundwater recharge. *Hydrogeology Journal*, 10(1):18–39.
- Schoups, G. and Vrugt, J. A. (2010). A formal likelihood function for parameter and predictive inference of hydrologic models with correlated, heteroscedastic, and non-Gaussian errors. *Water Resources Research*, 46(10):W10531.
- Schwarz, G. (1978). Estimating the Dimension of a Model. *The Annals of Statistics*, 6(2):461–464.
- Seibert, J. (2000). Multi-criteria calibration of a conceptual runoff model using a genetic algorithm. *Hydrology and Earth System Sciences*, 4(2):215–224.
- Shapoori, V., Peterson, T. J., Western, A. W., and Costelloe, J. F. (2015). Top-down groundwater hydrograph time-series modeling for climate-pumping decomposition. *Hydrogeology Journal*, 23(4):819–836.
- Van Geer, F. C., Te Stroet, C. B. M., and Yangxiao, Z. (1991). Using Kalman Filtering to Improve and Quantify the Uncertainty of Numerical Groundwater Simulations: 1. The Role of System Noise and Its Calibration. *Water Resources Research*, 27(8):1987–1994.
- van Geer, F. C. and Zuur, A. F. (1997). An extension of Box-Jenkins transfer/noise models for spatial interpolation of groundwater head series. *Journal of Hydrology*, 192(1–4):65–80.
- von Asmuth, J. R. and Bierkens, M. F. P. (2005). Modeling irregularly spaced residual series as a continuous stochastic process. *Water Resources Research*, 41(12):W12404.
- von Asmuth, J. R., Bierkens, M. F. P., and Maas, K. (2002). Transfer function-noise modeling in continuous time using predefined impulse response functions. *Water Resources Research*, 38(12):1287.
- von Asmuth, J. R., Maas, K., Knotters, M., Bierkens, M. F. P., Bakker, M., Olsthoorn, T. N., Cirkel, D. G., Leunk, I., Schaars, F., and von Asmuth, D. C. (2012). Software for hydrogeologic time series analysis, interfacing data with physical insight. *Environmental Modelling & Software*, 38:178–190.
- Wood, E. F., Lettenmaier, D. P., and Zartarian, V. G. (1992). A land-surface hydrology parameterization with subgrid variability for general circulation models. *Journal of Geophysical Research: Atmospheres*, 97(D3):2717–2728.
- Wösten, J., Veerman, G., Groot, d. W., and Stolte, J. (2001). Waterretentie-en doorlatendheidskarakteristieken van boven-en ondergronden in Nederland: de Staringreeks; vernieuwde uitgave 2001.
- Yi, M.-J. and Lee, K.-K. (2004). Transfer function-noise modelling of irregularly observed groundwater heads using precipitation data. *Journal of Hydrology*, 288(3–4):272–287.

A

DERIVATIONS

A.1. OBJECTIVE FUNCTION REVISITED

In this appendix it is shown that the modification proposed by Peterson and Western (2014) of the sum of weighted squared innovations (SWSI) objective function is mathematically similar to the original SWSI function introduced by von Asmuth and Bierkens (2005). This function is defined as follows:

$$S^2(t, \beta) = \sum_{j=1}^N \left(\frac{\sqrt{\prod_{i=1}^N (1 - e^{-\frac{-2\Delta t_i}{\alpha}})}}{1 - e^{-\frac{-2\Delta t_j}{\alpha}}} v^2(\beta, t_j) \right) \quad (\text{A.1})$$

When calibrating on long time series on with high frequency data, N can become very large while the term $1 - e^{-\frac{-2\Delta t_j}{\alpha}}$ becomes small. This can cause the product operator to near machine precision and return a value of 0.0 for the numerator. To solve this problem, Peterson and Western (2014) proposed to take the natural logarithm of $\sqrt{\prod_{i=1}^N (1 - e^{-\frac{-2\Delta t_i}{\alpha}})}$

$$S^2(t, \beta) = \sum_{j=1}^N \left(\frac{e^{\frac{\ln(\prod_{i=1}^N (1 - e^{-\frac{-2\Delta t_i}{\alpha}}))^{\frac{1}{N}}}{1 - e^{-\frac{-2\Delta t_j}{\alpha}}}} v^2(\beta, t_j) \right) \quad (\text{A.2})$$

which is essentially the same as (A.1) when applying the mathematical rules $\ln(x^y) = y\ln(x)$ and $\ln(ab) = \ln(a) + \ln(b)$. The final equation as proposed by Peterson and Western (2014) is then:

$$S^2(t, \beta) = \sum_{j=1}^N \left(\frac{e^{\frac{\frac{1}{N} \sum_{i=1}^N \ln(1 - \exp^{-\frac{-2\Delta t_i}{\alpha}})}{1 - e^{-\frac{-2\Delta t_j}{\alpha}}}} v^2(\beta, t_j) \right) \quad (\text{A.3})$$

As shown above, the two equations are mathematically similar. Numerically however, the adapted SWSI equation has a lower chance of running into machine precision causing calibration of the time series model to fail. Therefore, this equation is used as the objective function for parameter optimization.

A.2. SOLVING THE SOIL MODEL

In this appendix, the numerical mathematics that have been used to solve the unsaturated zone model are discussed. The unsaturated zone is described by a non-linear differential equation of which no analytical solution exists, and hence has to be solved numerically. Two different unsaturated zone models have been applied in this research. Below we will derive the numerical solution for the percolation model (A.4) in detail, while the solution of the piston flow model (A.5) is given and can be derived in a similar way.

$$\frac{dS}{dt} = (P - I) - K_{sat} \left(\frac{S(t)}{S_{cap}} \right)^\beta - E_p \min\left(1, \frac{S}{0.5S_{cap}}\right) \quad (\text{A.4})$$

$$\frac{dS}{dt} = (P - I) \left(1 - \left(\frac{S(t)}{S_{cap}}\right)^\beta\right) - E_p \min\left(1, \frac{S}{0.5S_{cap}}\right) \quad (\text{A.5})$$

Kavetski et al. (2006) showed that for robust parameter optimization and calibration strategies, smoothness of the objective function is important. Therefore, the above equation is solved using the implicit Euler scheme. For simplicity in the following derivation, we define:

$$f(S, t) = (P - I) - K_{sat} \left(\frac{S(t)}{S_{cap}} \right)^\beta - E_p \min\left(1, \frac{S}{0.5S_{cap}}\right) \quad (\text{A.6})$$

Applying the implicit Euler scheme gives the following:

$$\frac{S^{t+1} - S^t}{\Delta t} = f(S^{t+1}) \quad (\text{A.7})$$

$$S^{t+1} = S^t + \Delta t * f(S^{t+1}) \quad (\text{A.8})$$

This equation has to be solved iteratively, as S^{t+1} is unknown. The Newton-Raphson iteration method is used for this. Since the NR method is a root-finding technique that requires the form $g(S^{t+1}) = 0$, we need to introduce a new equation for $g(S^{t+1})$ using equation (A.8) :

$$g(S^{t+1}) = S^{t+1} - S^t - \Delta t * f(S^{t+1}) = 0 \quad (\text{A.9})$$

The equation for the Newton-Raphson iteration will than be:

$$S_{i+1}^{t+1} = S_i^{t+1} - \frac{g(S_i^{t+1})}{g'(S_i^{t+1})} \quad (\text{A.10})$$

The subscript i is the index for the iteration, hence every iteration the estimate of S^{t+1} is updated. For the first iteration, this requires an initial estimate of S^{t+1} , in this study given applying an explicit euler scheme to solve equation (A.4) (no derivation given here). The derivative of $g(S_i^{t+1})$ depends on the system state, as $f(S, t)$ is not a continuous function:

$$g'(S_i^{t+1}) = \frac{dg_i^{t+1}}{dS^{t+1}} \quad (\text{A.11})$$

$$g'(S_i^{t+1}) = \begin{cases} 1 - \Delta t \left(-K_{sat} \beta \left(\frac{S_i}{S_{cap}} \right)^{\beta-1} \right) & \text{if } S_i^{t+1} \geq 0.5 S_{cap} \\ 1 - \Delta t \left(-K_{sat} \beta \left(\frac{S_i}{S_{cap}} \right)^{\beta-1} - Ep \frac{1}{0.5 S_{cap}} \right) & \text{if } S_i^{t+1} < 0.5 S_{cap} \end{cases} \quad (\text{A.12})$$

The superscript $t + 1$ has been omitted from E and S_i for reasons of readability. Equation (A.10) is generally solved within 3-5 iterations, depending on the error ε that is allowed. However, in some cases the NR method does not find the solution and provides an error in the model. One of these errors is called the zero-division error and results from a value of the derivative that is (very close to) zero. Therefore, if this situation occurs, equation (A.9) is solved using the computationally more expensive bisection method for that specific time step.

The derivative g' for the piston flow model (A.5) is as follows:

$$g'(S_i^{t+1}) = \begin{cases} 1 - \Delta t \left(-(P - I) \beta \left(\frac{S_i}{S_{cap}} \right)^{\beta-1} \right) & \text{if } S_i^{t+1} \geq 0.5 S_{cap} \\ 1 - \Delta t \left(-(P - I) \beta \left(\frac{S_i}{S_{cap}} \right)^{\beta-1} - Ep \frac{1}{0.5 S_{cap}} \right) & \text{if } S_i^{t+1} < 0.5 S_{cap} \end{cases} \quad (\text{A.13})$$

The recharge for percolation and the piston flow model is now calculated by numerical integration as:

$$R^{t+1} = K_{sat} \frac{\Delta t}{2} \left(\frac{S^t + S^{t+1}}{S_{cap}} \right)^\beta \quad (\text{A.14})$$

$$R^{t+1} = (P^{t+1} - I) \frac{\Delta t}{2} \left(1 - \left(\frac{S^t + S^{t+1}}{S_{cap}} \right)^\beta \right) \quad (\text{A.15})$$

A.3. TIME TO PEAK AND THE IMPULSE RESPONSE FUNCTION

When the time to peak is a characteristic value of interest, the impulse response function can be modified to have the time to peak as a parameter. Following the derivation given in (Peterson and Western, 2014), we can derive the time to peak analytically, and use this to rewrite the impulse response function. The peak is where the derivative of equation 3.12 is equal to zero:

$$\frac{d\theta(t)}{dt} = 0 \quad (\text{A.16})$$

$$\frac{d\theta(t)}{dt} = A \left((n-1)t^{(n-2)} e^{-\frac{t}{a}} - \frac{t^{n-1}}{a} e^{-\frac{t}{a}} \right) \quad (\text{A.17})$$

Rearranging equation A.17 and solving for t provides the time to peak:

$$(n-1)t^{(n-2)} e^{-\frac{t}{a}} = \frac{t^{n-1}}{a} e^{-\frac{t}{a}} \quad (\text{A.18})$$

then solving for t provides the time of the peak:

$$t_{peak} = a(n-1) \quad (\text{A.19})$$

This clearly shows how the t_{peak} depends on the parameters a and n . Since these parameters have to be estimated and have no direct meaning, it would be convenient to replace one of these parameters by a more meaningful parameter. Using equation A.19, a in equation 3.12 can be replaced by t_{peak} , a parameter that has a clearer meaning. The impulse and the step response function now become:

$$\theta(t) = At^{n-1} e^{-\frac{t(n-1)}{t_{peak}}} \quad (\text{A.20})$$

$$\theta_s(t) = -At^n \left(\frac{t(n-1)}{t_p} \right)^{-n} \Gamma \left(n, \frac{t(n-1)}{t_p} \right) \quad (\text{A.21})$$

The above equations are applied to study how well the model simulates the expected time to peak and the time shift between boreholes. Possible drawbacks on this method is that the non-linear response is neglected, and the time to peak may not be that constant over time. In determining the expected values, the measurement interval introduces a significant error in the order of weeks. Finally, a faster response in the surrounding area, may result in a groundwater level change that propagates through the saturated groundwater effectively decreasing the time to peak. The method is only applied as a order of magnitude in which the value is expected, and not as a good approximation of the exact value.

B

TABLES

B.1. CHAPTER 4

Observation Well B27D00010: 30-year calibration period

1 IRF	SWSI	EVP (%)	RMSE	Avg Dev	AIC	BIC
Linear	1.06	99.22	0.17	0.01	-6786.13	-6759.28
Preferential	1.01	99.50	0.15	0.00	-6880.86	-6854.02
Percolation	1.00	99.44	0.15	0.01	-6831.55	-6800.23
Combination	1.01	99.51	0.15	0.00	-6873.86	-6838.07
2 IRF						
Linear	1.04	98.84	0.19	0.03	-6741.35	-6710.04
Preferential	1.01	98.99	0.18	0.01	-6944.05	-6912.73
Percolation	1.00	98.27	0.21	0.02	-6736.94	-6701.15
Combination	1.00	99.02	0.18	0.00	-6988.03	-6947.77

Observation Well B27C00490: 30-year calibration period

1 IRF	SWSI	EVP (%)	RMSE	Avg Dev	AIC	BIC
Linear	0.61	99.50	0.17	0.03	-16732.00	-16701.70
Preferential	0.61	99.56	0.17	0.04	-16744.02	-16713.72
Percolation	0.61	99.49	0.17	-0.02	-16757.04	-16721.69
Combination	0.61	99.49	0.17	-0.03	-16740.88	-16700.48
2 IRF						
Linear	0.61	99.68	0.16	0.03	-16759.98	-16724.6
Preferential	0.61	99.83	0.13	-0.01	-16789.21	-16753.86
Percolation	0.61	99.88	0.12	0.00	-16798.02	-16757.61
Combination	0.60	99.64	0.18	0.09	-16766.43	-16720.98

Observation Well B33A01130: 30-year calibration period

1 IRF	SWSI	EVP (%)	RMSE	Avg Dev	AIC	BIC
Linear	0.61	96.78	0.37	-0.19	-17420.48	-17389.96
Preferential	0.60	98.18	0.33	-0.18	-17451.51	-17420.99
Percolation	0.60	98.97	0.28	-0.15	-17438.71	-17403.11
Combination	0.60	98.30	0.32	-0.18	-17473.27	-17432.58
2 IRF						
Linear	0.61	99.66	0.19	-0.06	-17456.39	-17420.79
Preferential	0.60	99.68	0.19	-0.07	-17514.48	-17478.88
Percolation	0.60	99.71	0.19	-0.08	-17513.21	-17472.52
Combination	0.60	99.45	0.23	-0.11	-17537.60	-17491.83

Table B.1: Diagnostic statistics for the four recharge models using a single (1 IRF) or a double (2 IRF) response function, calibrated on a 30-year period. The best diagnostic statistics per model are highlighted in grey.

Observation Well B27D00010: 20-year calibration period

1 IRF	SWSI	EVP (%)	RMSE	Avg Dev	AIC	BIC
Linear	0.94	99.23	0.14	0.02	-4831.26	-4806.38
Preferential	0.90	99.14	0.14	-0.01	-4893.91	-4869.03
Percolation	0.90	98.85	0.16	0.01	-4841.16	-4812.14
Combination	0.90	99.17	0.14	-0.01	-4887.66	-4854.49
2 IRF						
Linear	0.96	99.13	0.15	0.03	-4806.85	-4777.82
Preferential	0.87	99.05	0.15	0.01	-4998.69	-4969.67
Percolation	0.93	97.91	0.18	0.01	-4839.17	-4806.00
Combination	0.89	97.30	0.19	0.00	-4875.17	-4837.85

Observation Well B27C00490: 20-year calibration period

1 IRF	SWSI	EVP (%)	RMSE	Avg Dev	AIC	BIC
Linear	0.47	91.14	0.32	0.14	-5848.66	-5824.48
Preferential	0.44	89.36	0.32	0.09	-5941.55	-5917.37
Percolation	0.44	99.00	0.18	-0.04	-5879.38	-5851.16
Combination	0.43	97.17	0.22	-0.01	-5922.47	-5890.23
2 IRF						
Linear	0.42	97.85	0.21	-0.01	-5953.26	-5925.04
Preferential	0.41	97.04	0.23	-0.04	-5973.60	-5945.38
Percolation	0.43	99.52	0.14	0.01	-5983.23	-5951.00
Combination	0.41	98.14	0.21	-0.06	-5989.17	-5952.89

Observation Well B33A01130: 20-year calibration period

1 IRF	SWSI	EVP (%)	RMSE	Avg Dev	AIC	BIC
Linear	0.53	90.56	0.48	0.36	-6098.40	-6073.61
Preferential	0.47	91.08	0.35	0.13	-6211.94	-6187.16
Percolation	0.51	95.13	0.28	-0.05	-6156.05	-6127.14
Combination	0.48	95.73	0.27	-0.04	-6211.36	-6178.31
2 IRF						
Linear	0.50	96.74	0.26	0.09	-6155.42	-6126.50
Preferential	0.45	97.11	0.24	0.00	-6267.81	-6238.89
Percolation	0.47	97.85	0.22	0.00	-6258.19	-6225.14
Combination	0.44	98.95	0.19	0.01	-6328.92	-6291.74

Table B.2: Diagnostic statistics for the four recharge models using a single (1 IRF) or a double (2 IRF) response function, calibrated on a 20-year period. The best diagnostic statistics per model are highlighted in grey.

B.2. CHAPTER 5**Observation Well B27D00010: 30-year calibration period**

1 IRF	SWSI	EVP (%)	RMSE	Avg Dev	AIC	BIC
Linear	1.23	89.33	0.38	-0.06	-9486.25	-9457.42
Preferential	1.07	97.28	0.33	-0.19	-9587.92	-9559.09
Percolation	1.07	97.50	0.28	-0.10	-9562.99	-9529.35
Combination	1.06	97.13	0.33	-0.19	-9596.15	-9557.71
Reclamation						
Linear	1.27	99.25	0.20	0.03	-9578.83	-9540.38
Preferential	1.17	99.57	0.17	0.02	-9731.36	-9692.92
Percolation	1.17	99.55	0.17	0.02	-9701.54	-9658.29
Combination	1.16	99.57	0.17	0.02	-9736.07	-9688.01
Extractions						
Linear	1.27	99.16	0.20	0.02	-9578.10	-9539.65
Preferential	1.19	99.57	0.17	0.02	-9734.77	-9696.32
Percolation	1.17	99.57	0.17	0.01	-9698.01	-9654.75
Combination	1.16	99.50	0.18	0.01	-9747.42	-9699.36

Table B.3: Diagnostic statistics for the four recharge models using a single response function calibrated on a 30-year period, for well B27D00010.

Observation Well B27D00010: 20-year calibration period

1 IRF	SWSI	EVP (%)	RMSE	Avg Dev	AIC	BIC
Linear	1.14	94.08	0.34	-0.03	-7509.84	-7482.34
Preferential	1.08	63.12	0.54	-0.09	-7614.66	-7587.16
Percolation	0.99	96.59	0.30	-0.00	-7595.45	-7563.38
Combination	0.99	92.58	0.37	-0.07	-7615.76	-7579.10
Reclamation						
Linear	1.16	99.57	0.18	0.03	-7587.04	-7550.39
Preferential	1.08	99.71	0.16	0.02	-7714.02	-7677.36
Percolation	1.06	99.69	0.16	0.02	-7686.46	-7645.22
Combination	1.08	99.71	0.16	0.02	-7707.67	-7661.85
Extractions						
Linear	1.16	99.52	0.18	0.02	-7595.87	-7559.21
Preferential	1.11	99.59	0.18	0.02	-7730.03	-7693.37
Percolation	1.06	99.72	0.16	0.02	-7686.36	-7645.13
Combination	1.05	98.67	0.24	0.05	-7742.11	-7696.29

Table B.4: Diagnostic statistics for the four recharge models using a single response function calibrated on a 20-year period, for well B27D00010.

Observation Well B27C00490: 30-year calibration period

2 IRF	SWSI	EVP (%)	RMSE	Avg Dev	AIC	BIC
Linear	0.61	99.68	0.16	0.03	-16759.98	-16724.6
Preferential	0.61	99.83	0.13	-0.01	-16789.21	-16753.86
Percolation	0.61	99.88	0.12	0.00	-16798.02	-16757.61
Combination	0.60	99.64	0.18	0.09	-16766.43	-16720.98
Reclamation						
Linear	0.61	99.68	0.16	0.03	-16754.38	-16708.93
Preferential	0.61	99.89	0.12	-0.01	-16800.81	-16755.36
Percolation	0.61	99.89	0.12	0.00	-16797.01	-16746.51
Combination	0.61	99.93	0.11	0.02	-16812.62	-16757.07

Table B.5: Diagnostic statistics for the four recharge models using a double response function calibrated on a 30-year period, for well B27C00490.

Observation Well B27C0049: 20-year calibration period

2 IRF	SWSI	EVP (%)	RMSE	Avg Dev	AIC	BIC
Linear	0.42	97.85	0.21	-0.01	-5953.26	-5925.04
Preferential	0.41	97.04	0.23	-0.04	-5973.60	-5945.38
Percolation	0.43	99.52	0.14	0.01	-5983.23	-5951.00
Combination	0.41	98.14	0.21	-0.06	-5989.17	-5952.89
Reclamation						
Linear	0.42	97.85	0.21	-0.01	-5947.22	-5910.94
Preferential	0.41	97.00	0.23	-0.04	-5967.48	-5931.21
Percolation	0.43	98.70	0.19	0.06	-5951.91	-5911.60
Combination	0.41	99.68	0.13	0.04	-6005.96	-5961.62

Table B.6: Diagnostic statistics for the four recharge models using a double response function calibrated on a 20-year period, for well B27C00490.

Observation Well B33A01130: 30-year calibration period

2 IRF	SWSI	EVP (%)	RMSE	Avg Dev	AIC	BIC
Linear	0.61	99.66	0.19	-0.06	-17456.39	-17420.79
Preferential	0.60	99.68	0.19	-0.07	-17514.48	-17478.88
Percolation	0.60	99.71	0.19	-0.08	-17513.21	-17472.52
Combination	0.60	99.45	0.23	-0.11	-17537.60	-17491.83
Reclamation						
Linear	0.61	99.82	0.16	-0.05	-17465.74	-17419.97
Preferential	0.60	99.89	0.15	-0.06	-17544.80	-17499.03
Percolation	0.60	99.72	0.19	-0.08	-17506.75	-17455.89
Combination	0.60	99.90	0.15	-0.07	-17572.09	-17511.06
Extractions						
Linear	0.61	99.85	0.16	-0.07	-17503.43	-17457.65
Preferential	0.60	99.92	0.14	-0.07	-17585.81	-17540.04
Percolation	0.60	99.93	0.14	-0.06	-17540.14	-17489.28
Combination	0.60	99.59	0.21	-0.10	-17527.17	-17471.23

Table B.7: Diagnostic statistics for the four recharge models using a double response function calibrated on a 20-year period, for well B33A01130.

Observation Well B33A01130: 20-year calibration period

2 IRF	SWSI	EVP (%)	RMSE	Avg Dev	AIC	BIC
Linear	0.50	96.74	0.26	0.09	-6155.42	-6126.50
Preferential	0.45	97.11	0.24	0.00	-6267.81	-6238.89
Percolation	0.47	97.85	0.22	0.00	-6258.19	-6225.14
Combination	0.44	98.95	0.19	0.01	-6328.92	-6291.74
Reclamation						
Linear	0.50	96.76	0.26	0.09	-6149.38	-6112.20
Preferential	0.46	99.54	0.15	0.02	-6271.17	-6233.99
Percolation	0.46	97.92	0.22	-0.01	-6256.03	-6214.72
Combination	0.45	96.67	0.25	0.00	-6257.18	-6211.73
Extractions						
Linear	0.50	99.45	0.17	-0.05	-6166.70	-6129.52
Preferential	0.45	99.82	0.12	0.00	-6269.37	-6232.19
Percolation	0.46	99.87	0.11	0.01	-6257.07	-6215.76
Combination	0.46	99.89	0.11	0.02	-6277.43	-6231.99

Table B.8: Diagnostic statistics for the four recharge models using a double response function, calibrated on a 30-year period, for well B33A01130.

C

FIGURES

C.1. CHAPTER 4

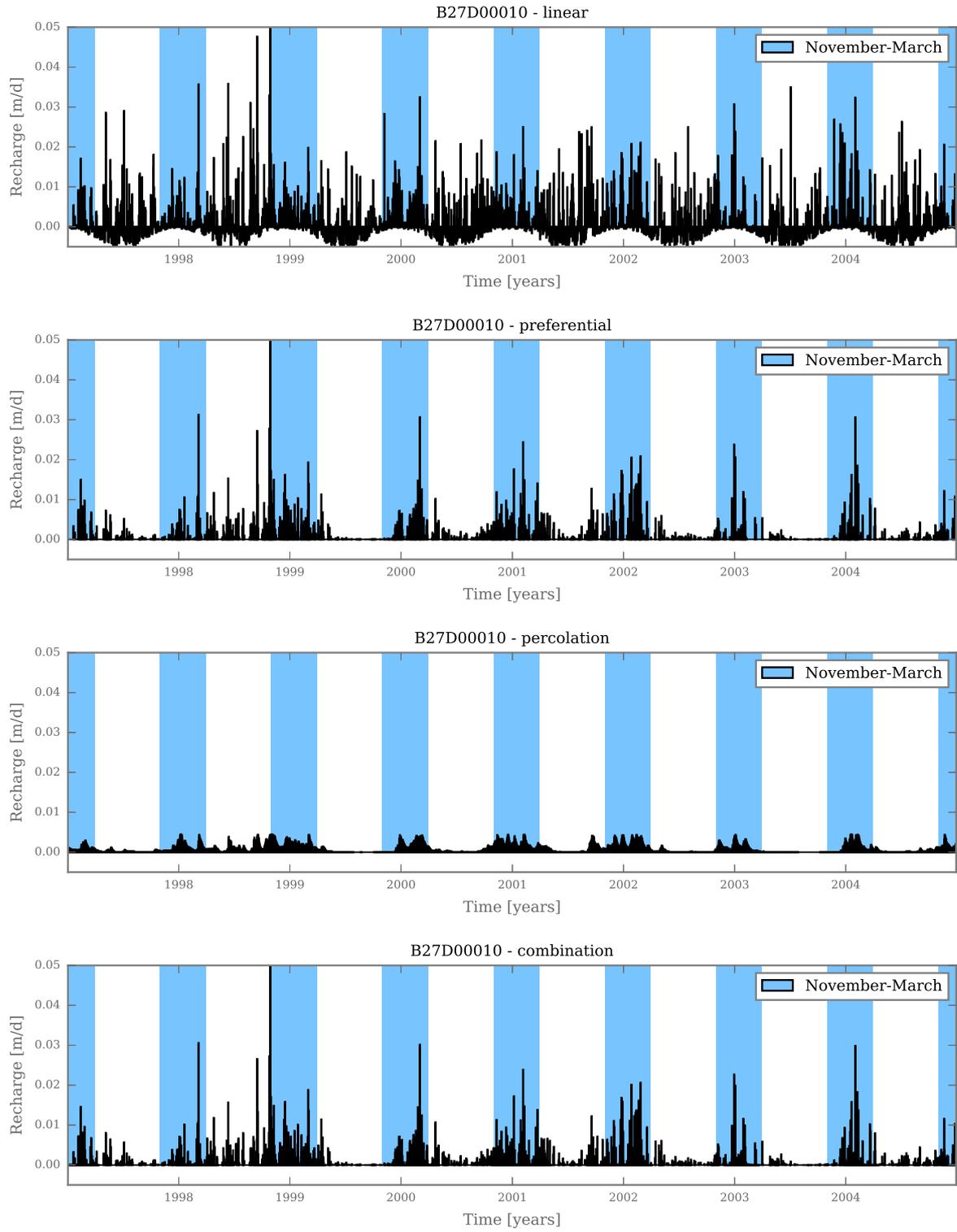


Figure C.1: The recharge time series for the linear, preferential, percolation and combination model for observation well B27D00010 in the period 1997-2004. The calibration period is 1974-1995.

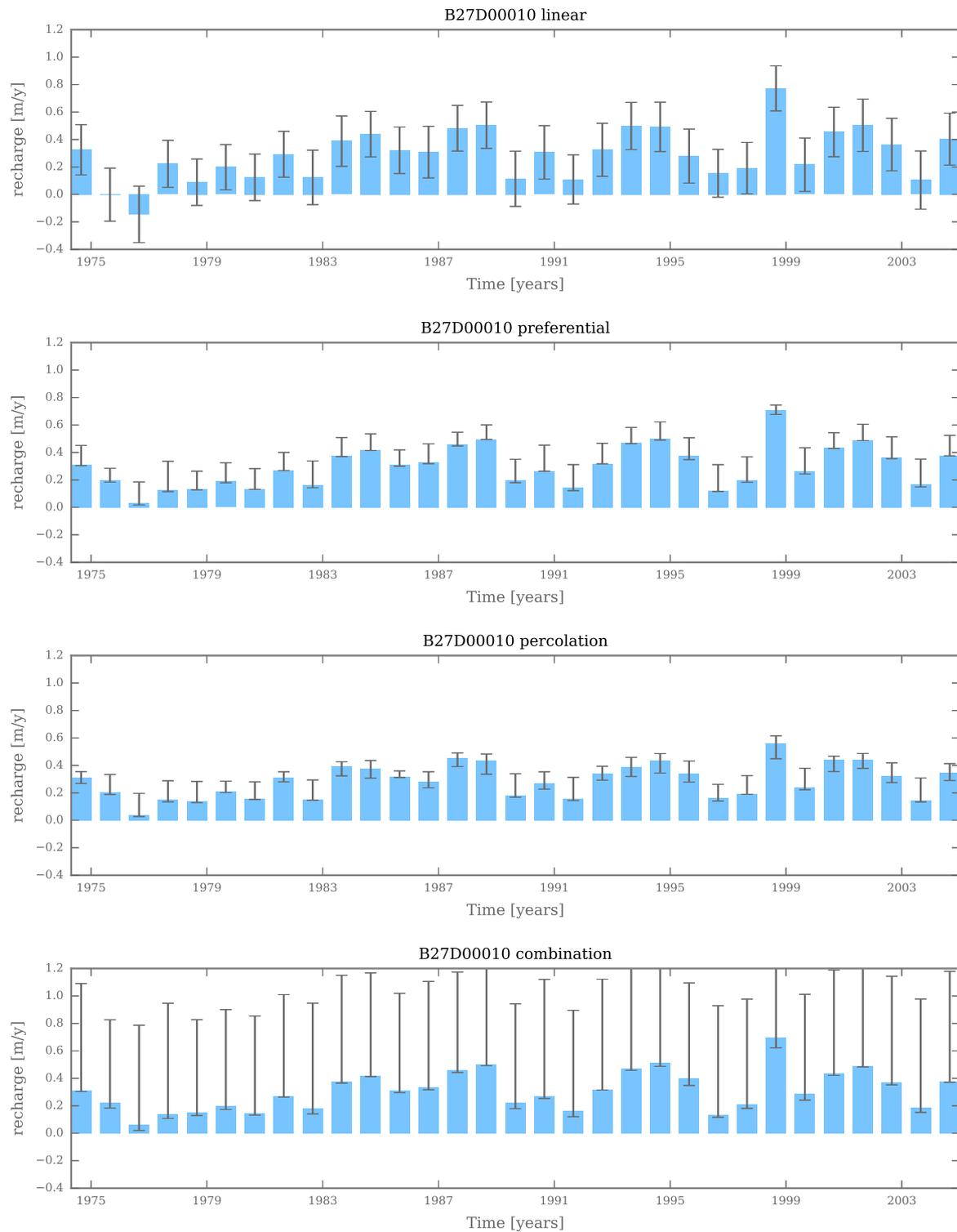


Figure C.2: Average yearly recharge fluxes and the 2.5% and the 97.5% percentiles (the whiskers) to provide information on the uncertainty in the recharge calculation.

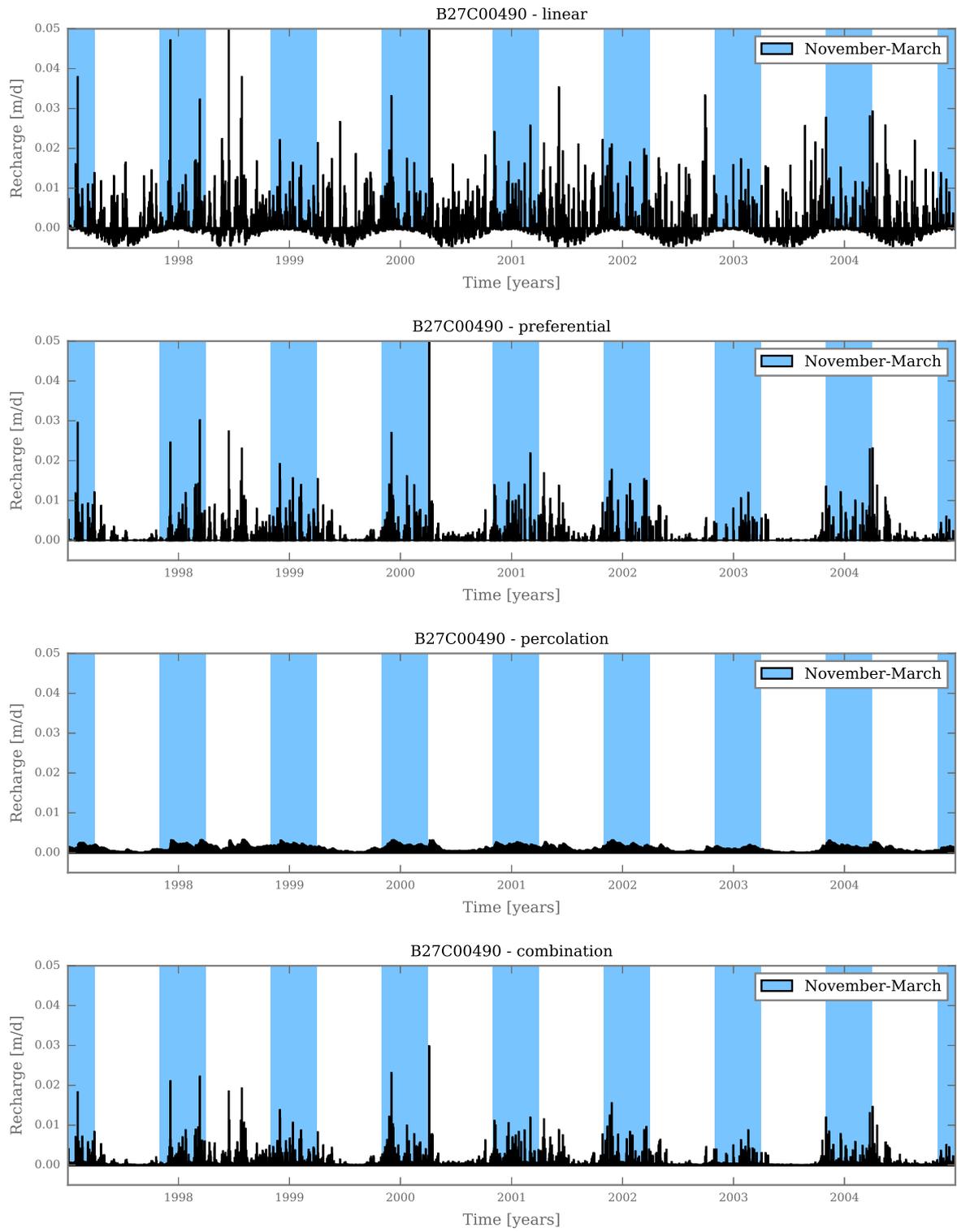


Figure C.3: The recharge time series for the linear, preferential, percolation and combination model for observation well B27C00490 in the period 1997-2004. The calibration period is 1974-2004.

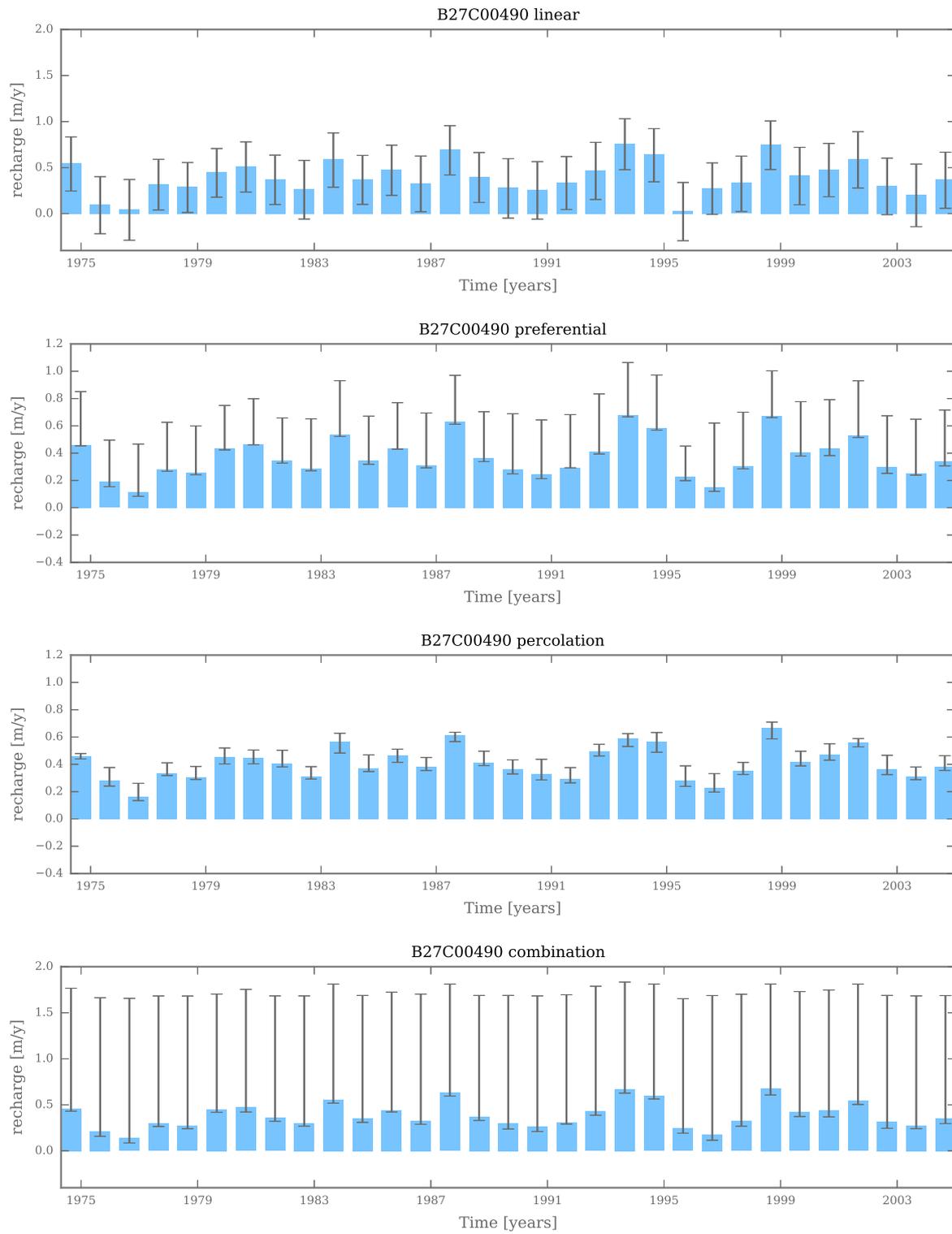


Figure C.4: Average yearly recharge fluxes and the 2.5% and the 97.5% percentiles (the whiskers) to provide information on the uncertainty in the recharge calculation.

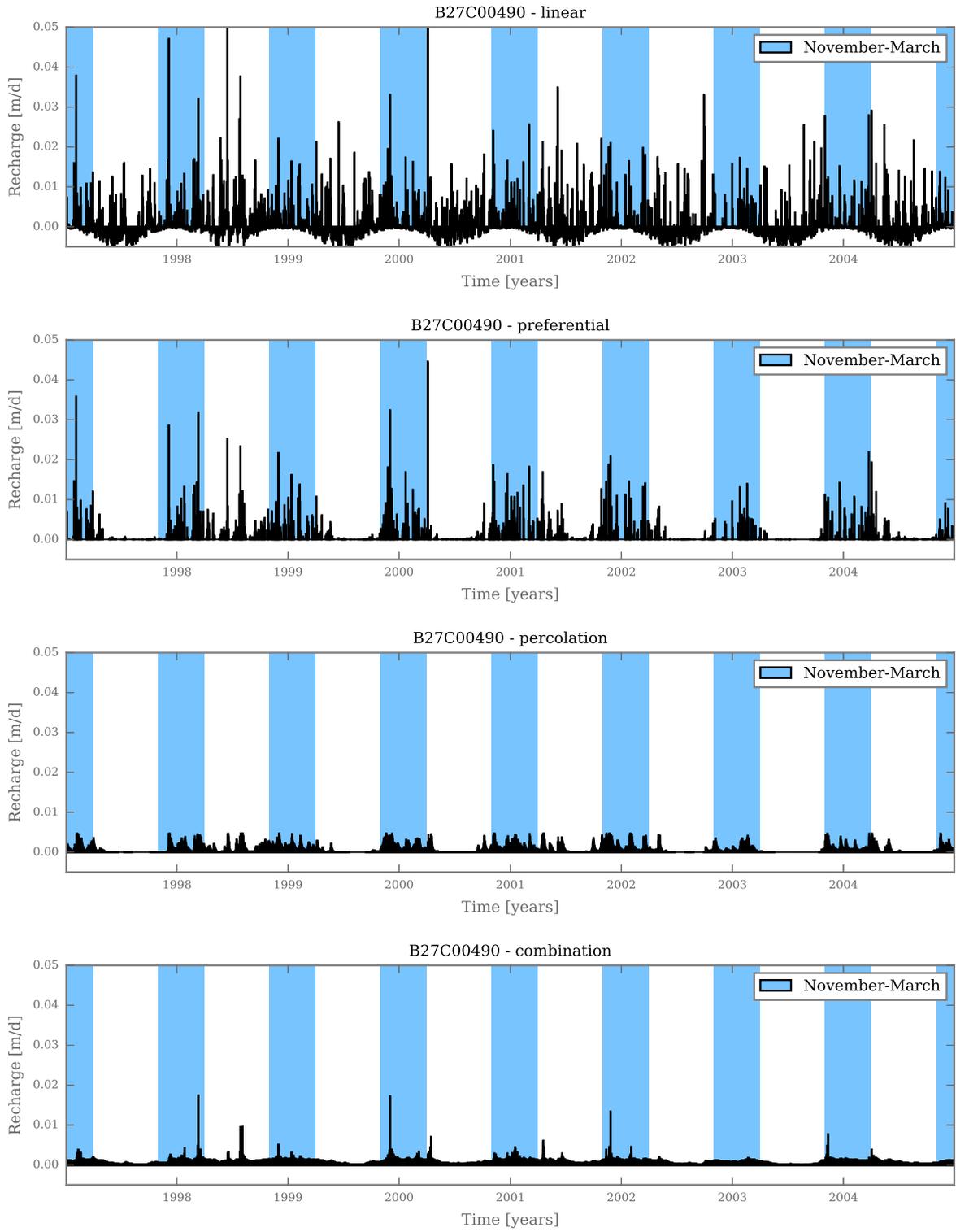


Figure C.5: The recharge time series for the linear, preferential, percolation and combination model for observation well B27C00490 in the period 1997-2004. The calibration period is 1974-1995.

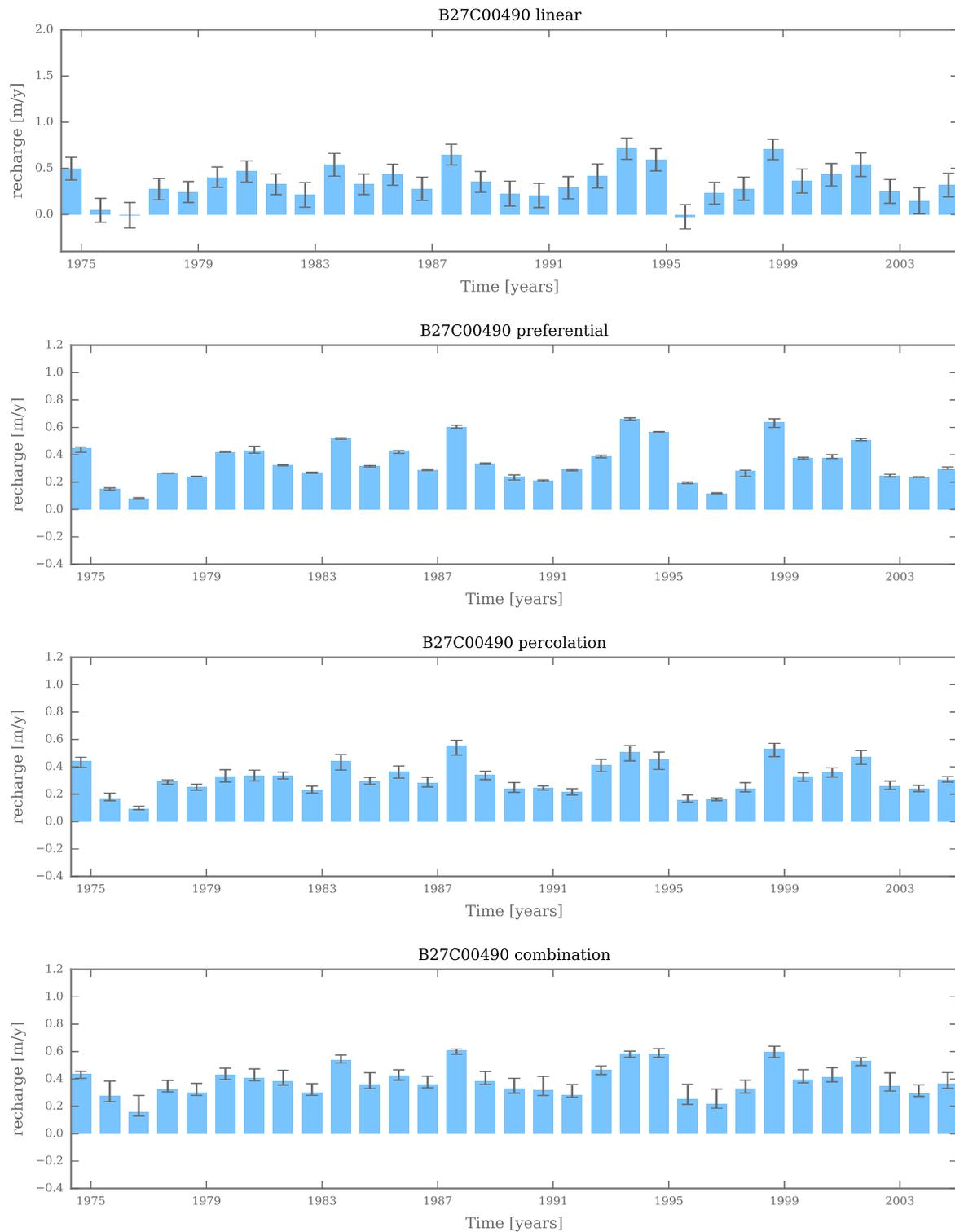


Figure C.6: Average yearly recharge fluxes and the 2.5% and the 97.5% percentiles (the whiskers) to provide information on the uncertainty in the recharge calculation.

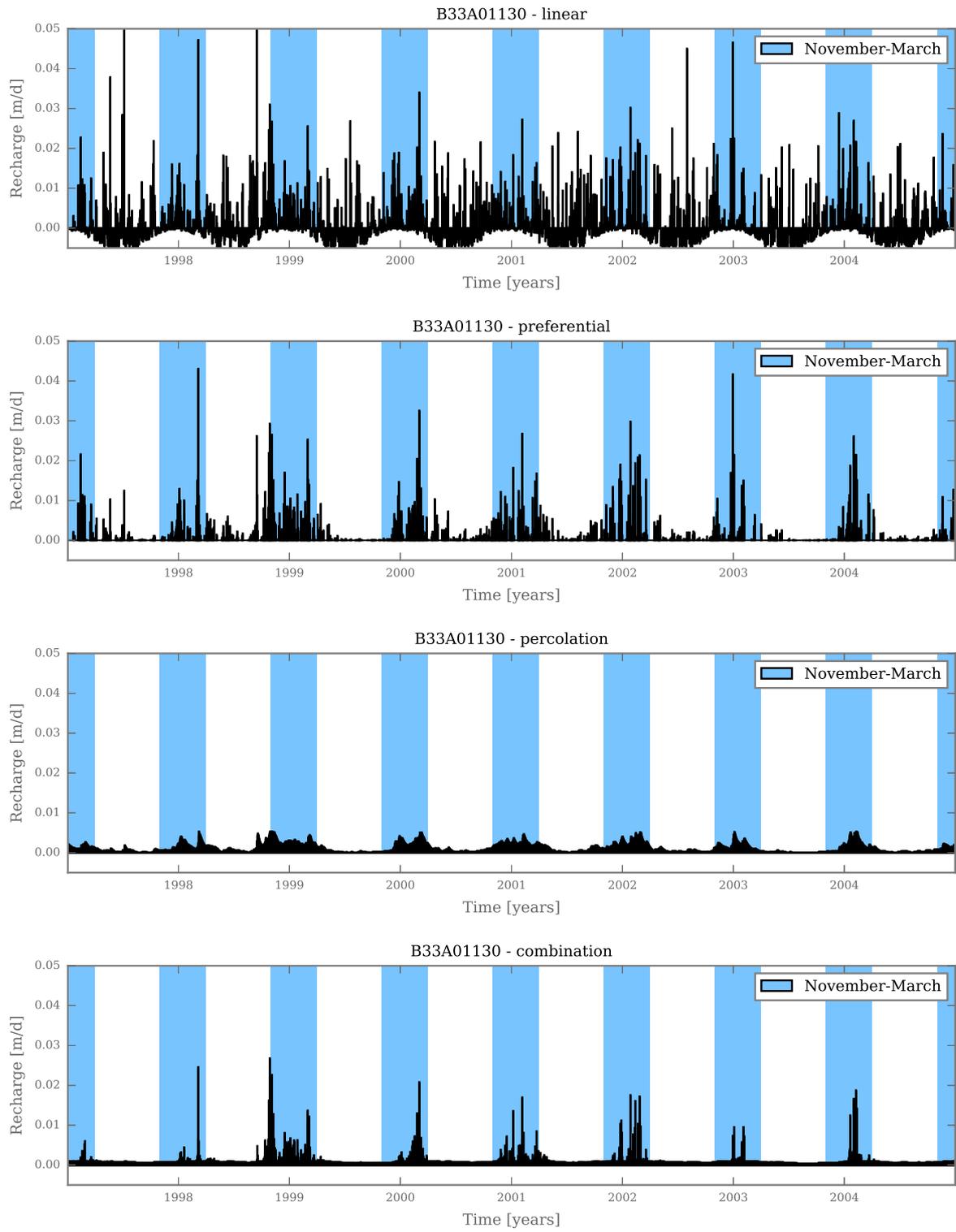


Figure C.7: The recharge time series for the linear, preferential, percolation and combination model for observation well B33A01130 in the period 1997-2004. The calibration period is 1974-1995.

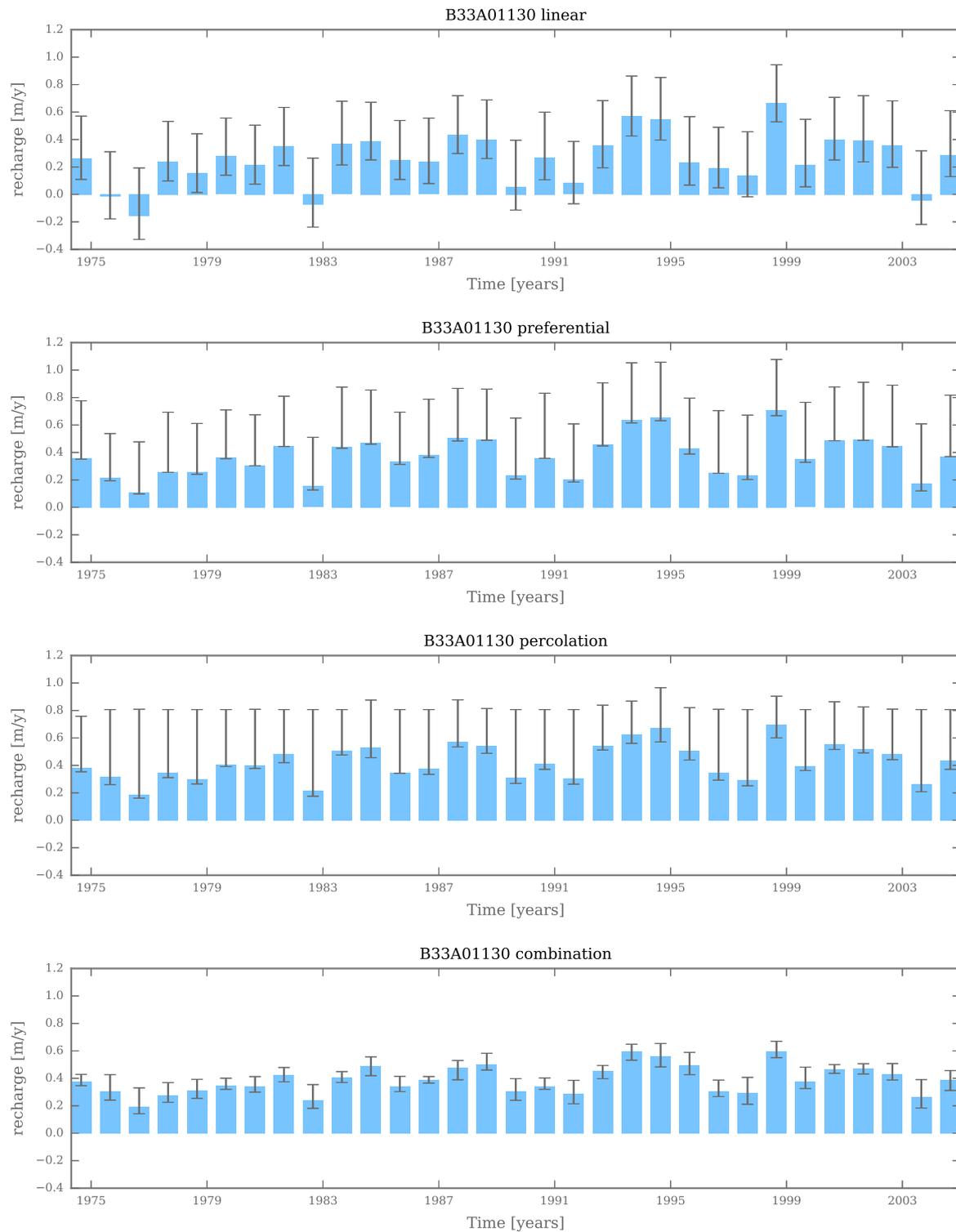


Figure C.8: Average yearly recharge fluxes and the 2.5% and the 97.5% percentiles (the whiskers) to provide information on the uncertainty in the recharge calculation.

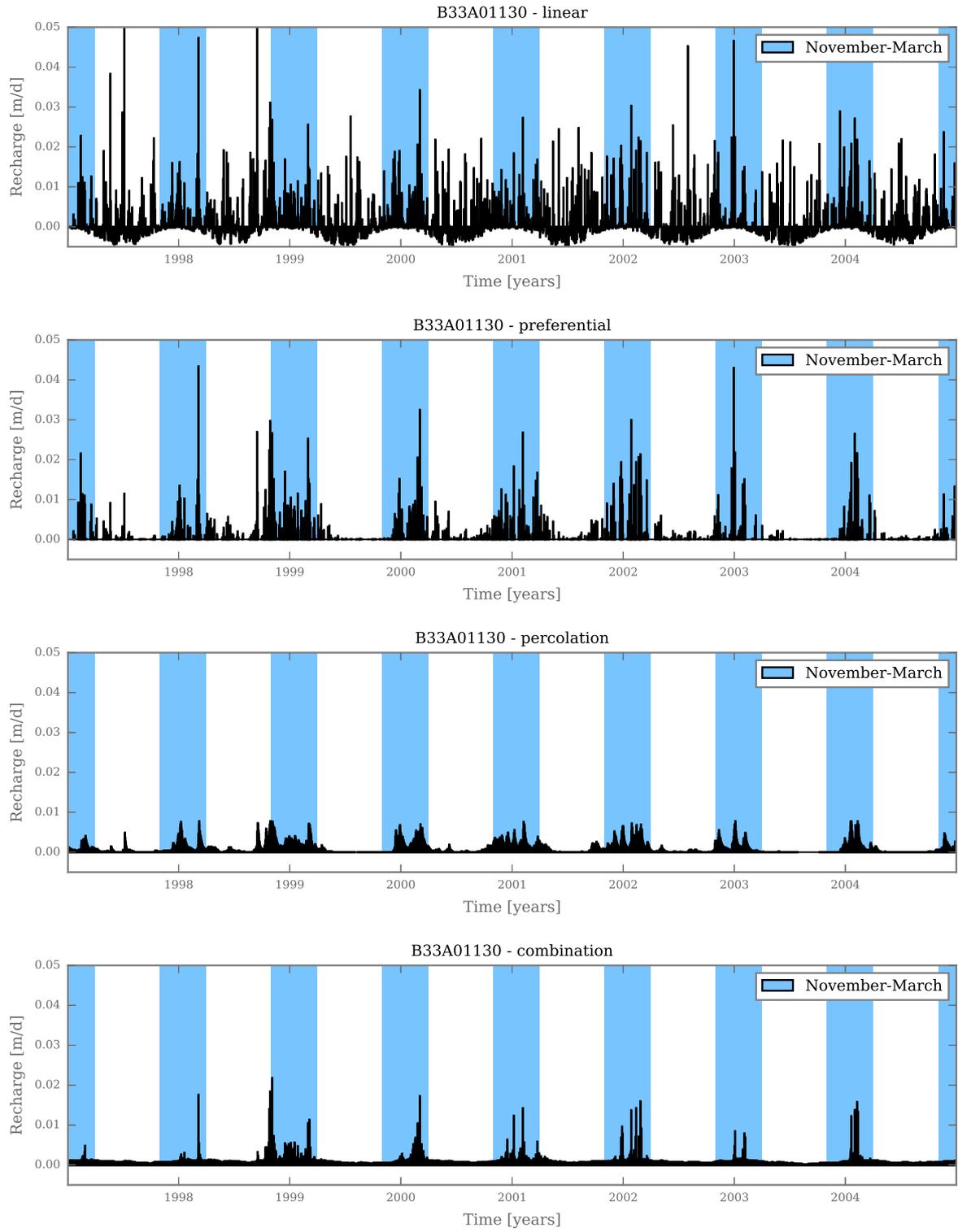


Figure C.9: The recharge time series for the linear, preferential, percolation and combination model for observation well B33A01130 in the period 1997-2004. The calibration period is 1974-1995.

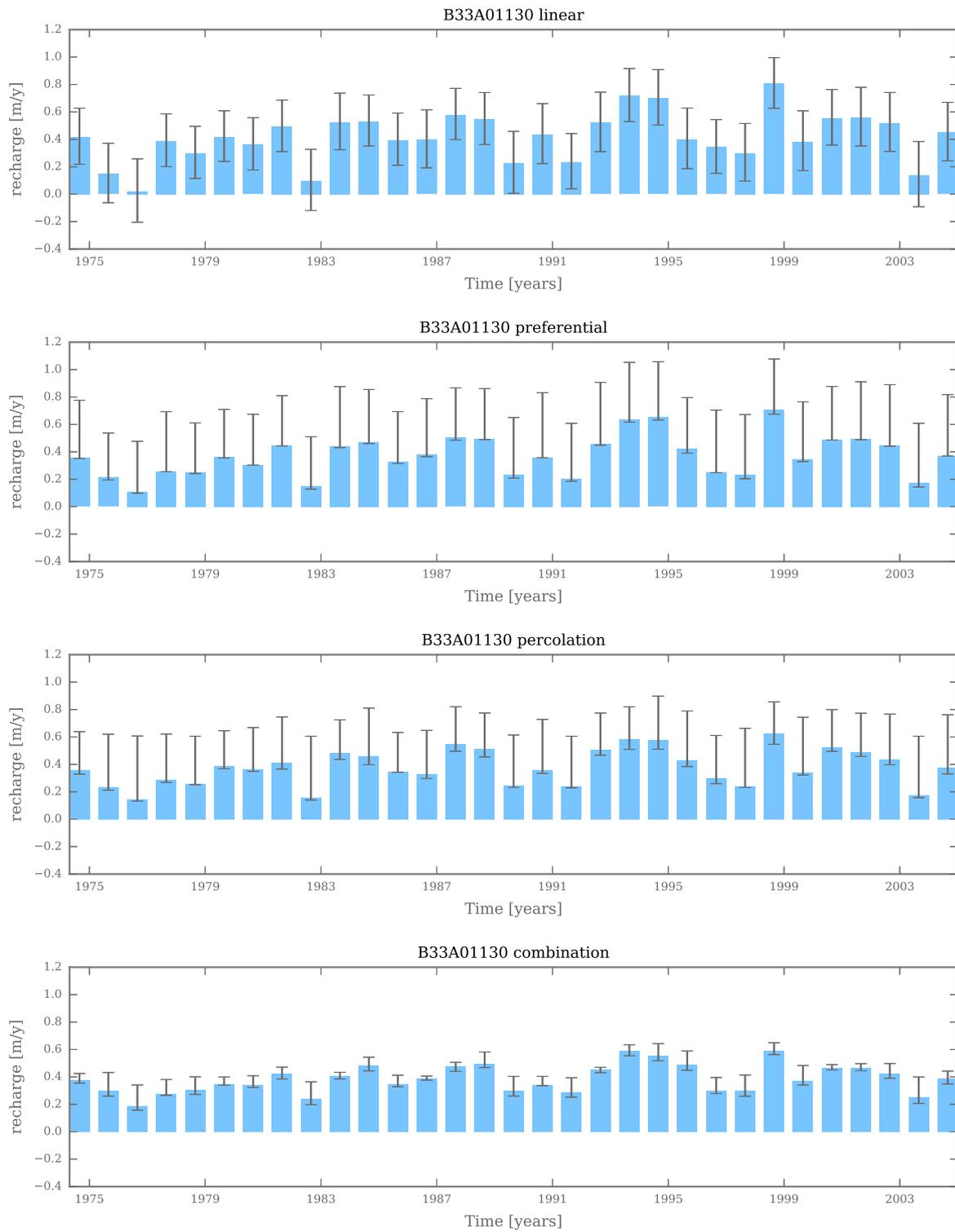


Figure C.10: Average yearly recharge fluxes and the 2.5% and the 97.5% percentiles (the whiskers) to provide information on the uncertainty in the recharge calculation.

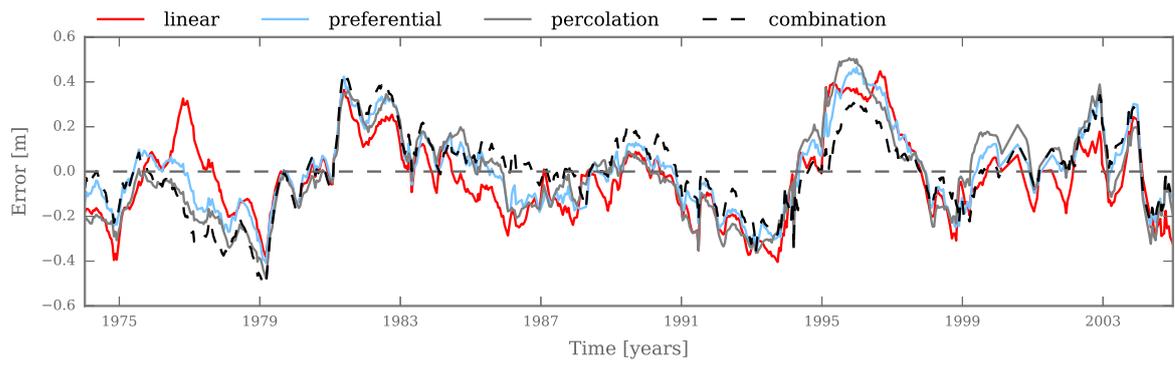


Figure C.11: Residual series for the models with a double response function calibrated on the period 1974-2004 for well B27D00010.

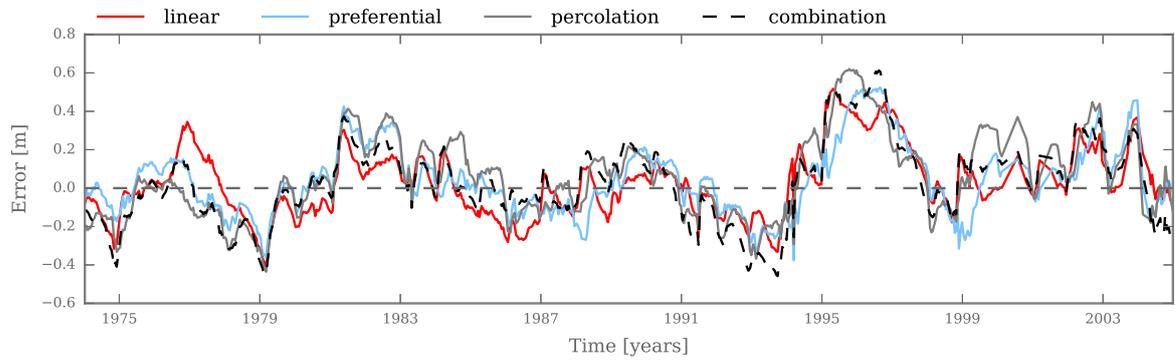


Figure C.12: Residual series for the models with a double response function calibrated on the period 1974-1995 for well B27D00010.

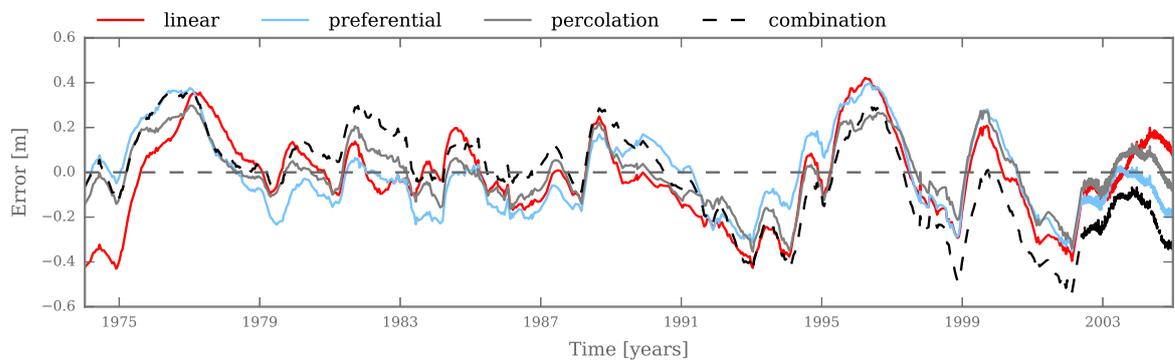


Figure C.13: Residual series for the models with a double response function calibrated on the period 1974-2004 for well B27C00490.

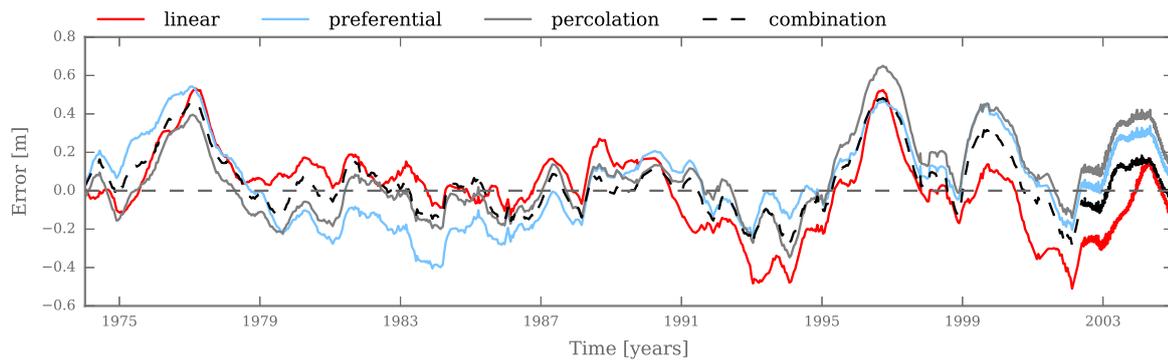


Figure C.14: Residual series for the models with a double response function calibrated on the period 1974-1995 for well B27C00490.

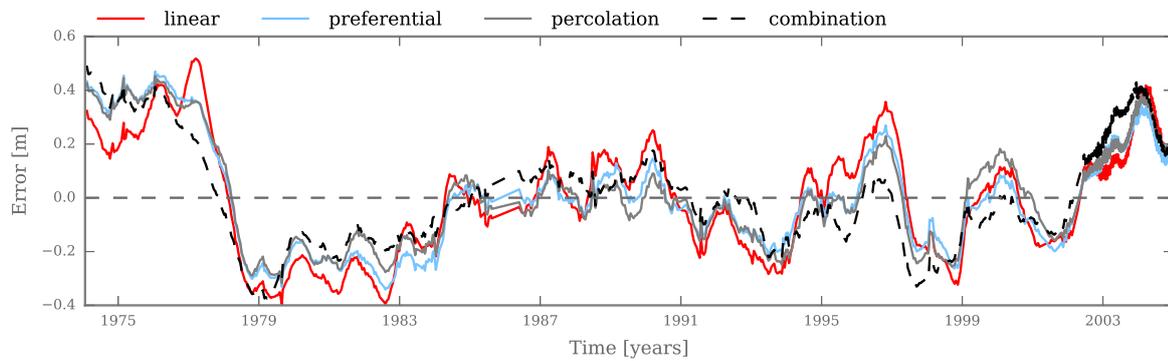


Figure C.15: Residual series for the models with a double response function calibrated on the period 1974-2004 for well B33A01130.

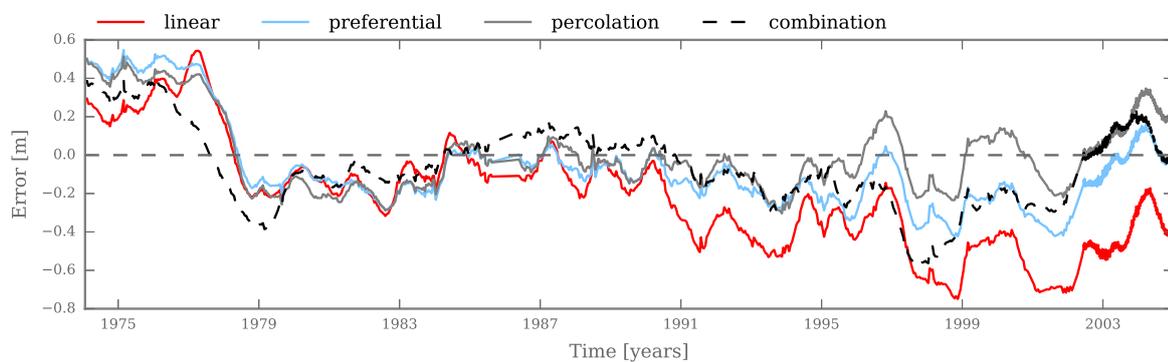


Figure C.16: Residual series for the models with a double response function calibrated on the period 1974-1995 for well B33A01130.

C.2. CHAPTER 5

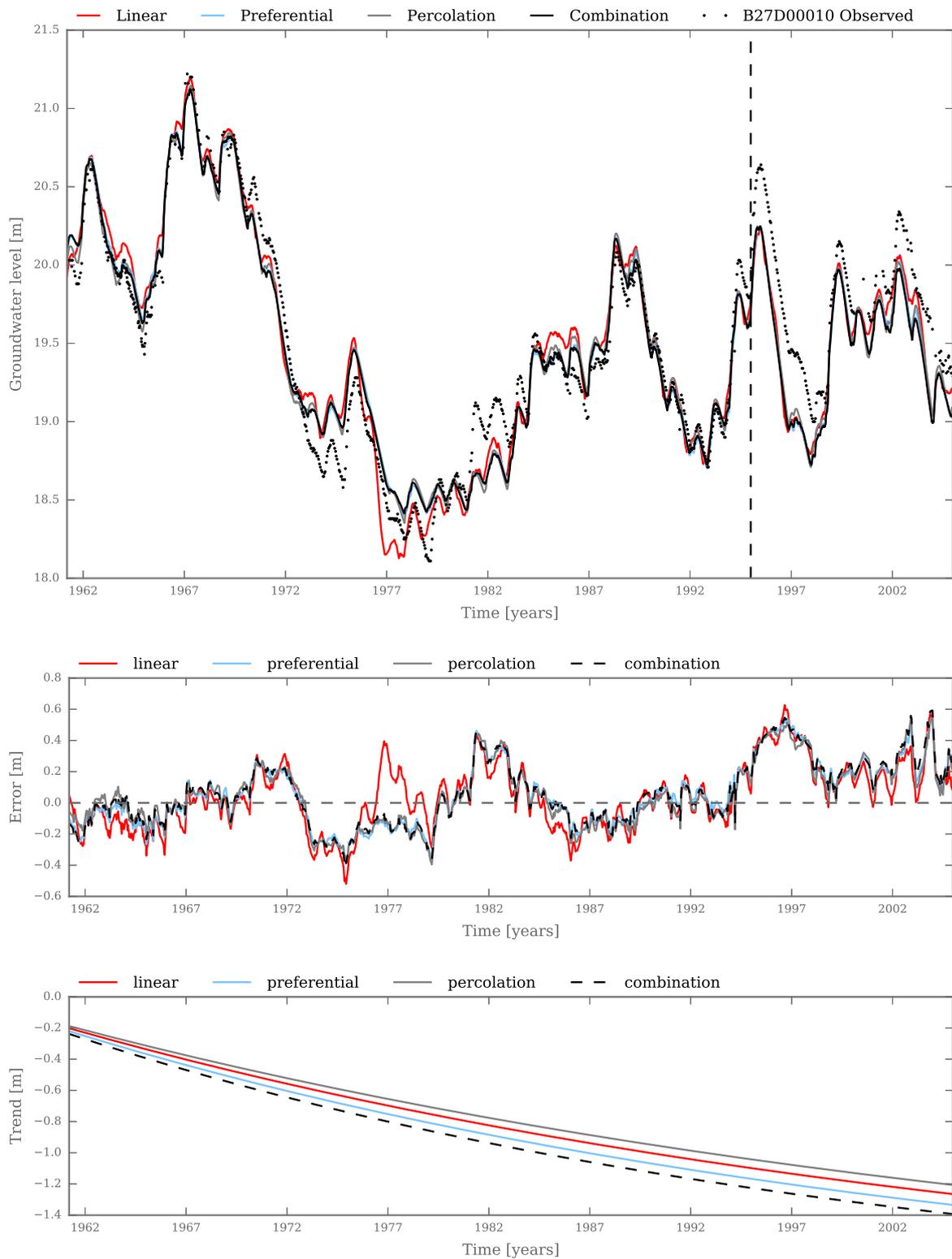


Figure C.17: Simulated groundwater levels, residuals, and the drawdown for all four recharge models with the effect of the reclamation for well B27D00010. The models are calibrated on the period 1960-1995.

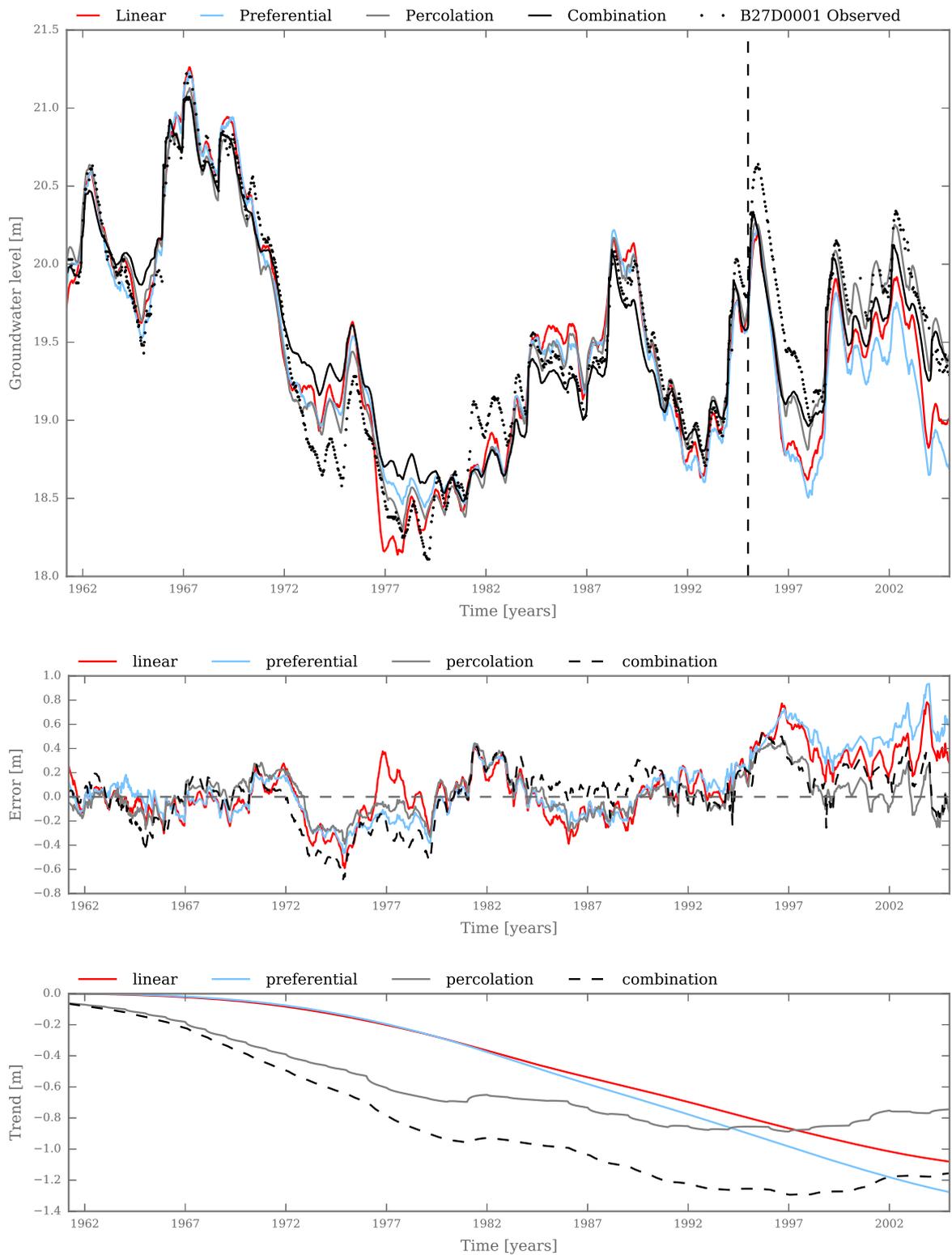


Figure C.18: Simulated groundwater levels, residuals, and the drawdown for all four recharge models with the effect of extractions for well B27D0001. The models are calibrated on the period 1960-1995.

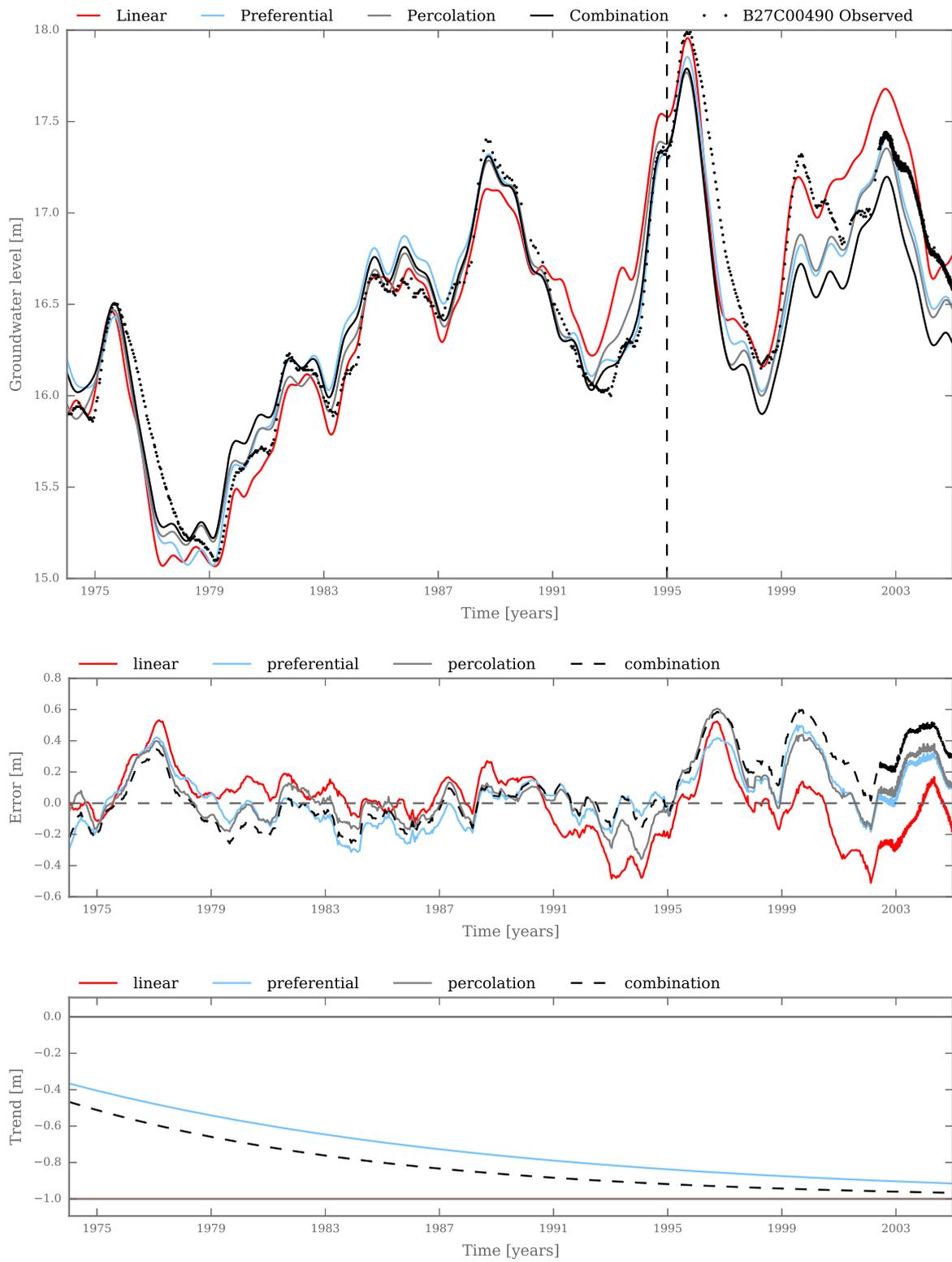


Figure C.19: Simulated groundwater levels, residuals, and the drawdown for all four recharge models with the effect of the reclamation for well B27C00490. The models are calibrated on the period 1974-1995.

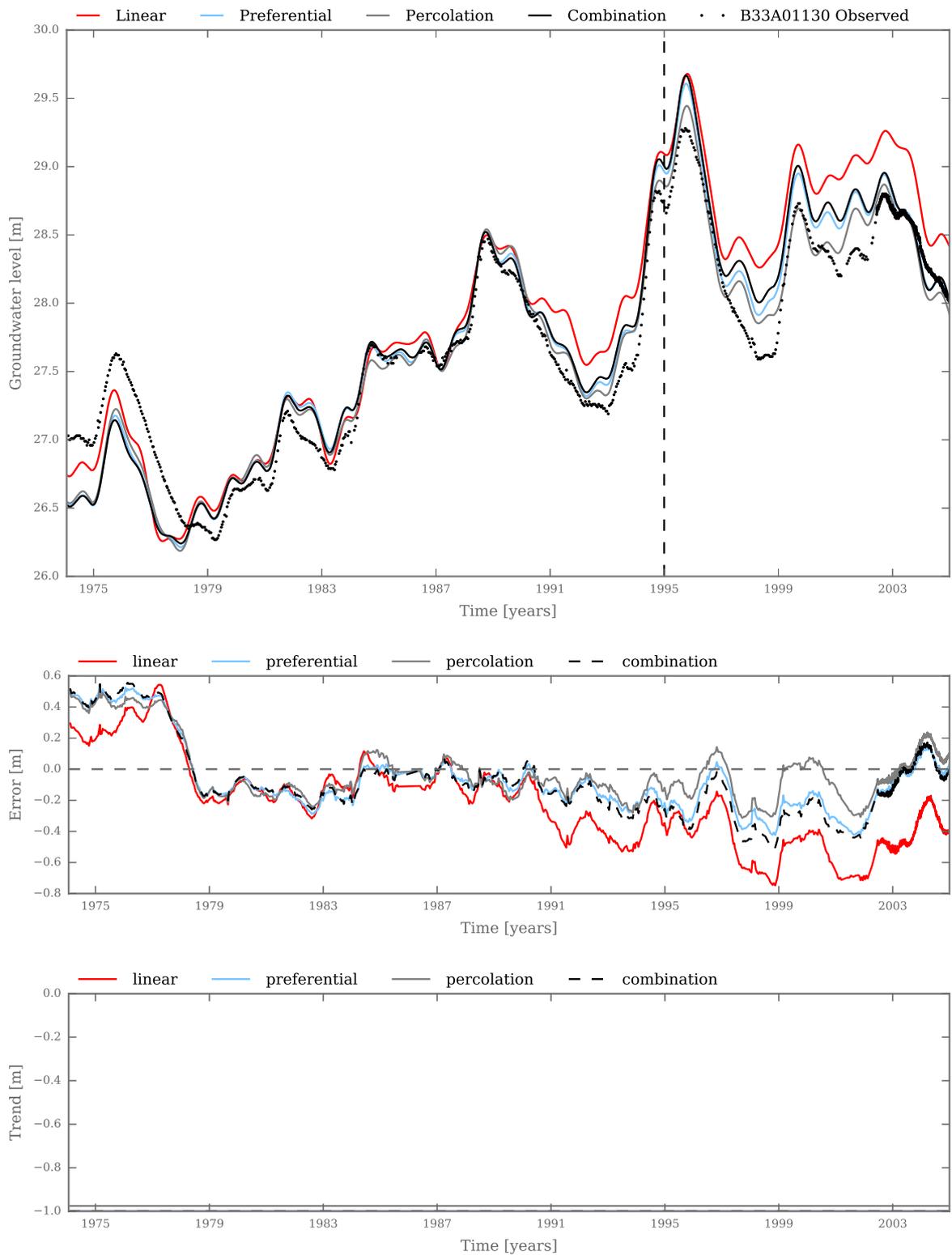


Figure C.20: Simulated groundwater levels, residuals, and the drawdown for all four recharge models with the effect of the reclamation for well B33A01130. The models are calibrated on the period 1974-1995.

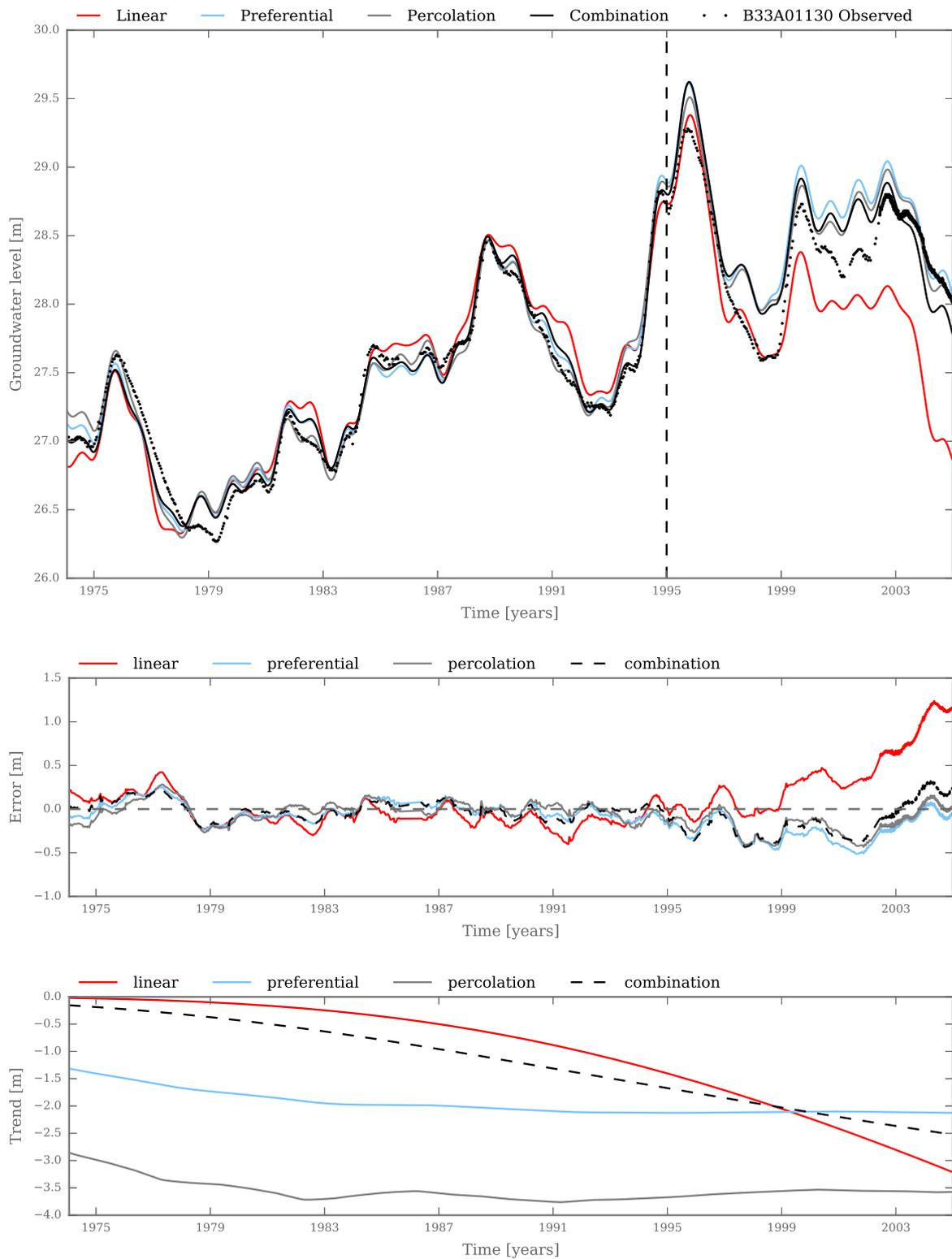


Figure C.21: Simulated groundwater levels, residuals, and the drawdown for all four recharge models with the effect of extractions for well B33A01130. The models are calibrated on the period 1974-1995.