

Time-series Analysis Using a Transfer Function Noise Model

Interpolation of Groundwater Levels from File Hajdar, Gotland

Master's thesis in the Master's programme Infrastructure and Environmental Engineering

SARAH ALI
KARISMA PATEL

DEPARTMENT OF ARCHITECTURE AND CIVIL ENGINEERING
DIVISION OF GEOLOGY AND GEOTECHNICS

CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2021
www.chalmers.se

MASTER'S THESIS ACEX30

Time-series Analysis Using Transfer Function Noise Model

Interpolation of Groundwater Levels from File Hajdar, Gotland

SARAH ALI & KARISMA PATEL



CHALMERS
UNIVERSITY OF TECHNOLOGY

Department of Architecture and Civil Engineering
Division of Geology and Geotechnics
CHALMERS UNIVERSITY OF TECHNOLOGY

Göteborg, Sweden 2021

Time-series Analysis Using Transfer Function Noise Model
Interpolation of Groundwater Levels from File Hajdar, Gotland
SARAH ALI
KARISMA PATEL

© SARAH ALI & KARISMA PATEL, 2021.

Examensarbete ACEX30
Institutionen för arkitektur och samhällsbyggnadsteknik
Chalmers tekniska högskola, 2021

Technical report no xxxx:xx
Department of Architecture and Civil Engineering
Division of Geology and Geotechnics
Chalmers University of Technology
SE-412 96 Göteborg
Sweden
Telephone + 46 (0)31-772 1000

Cover:

The cover shows a plot of the simulations created using a LH model, see figure 4.28.

Department of Architecture and Civil Engineering
Göteborg, Sweden 2021

Time-series Analysis Using a Transfer Function Noise Model
Interpolation of Groundwater Levels from File Hajdar, Gotland

*Master's thesis in the Master's Programme
Infrastructure and Environmental Engineering*

SARAH ALI AND KARISMA PATEL

Department of Architecture and Civil Engineering
Chalmers University of Technology

Summary

In File Hajdar, an area in Northern Gotland, the bedrock consist mainly of limestone and the groundwater levels exhibit large fluctuations between seasons. For over 40 years groundwater levels have been recorded with varying frequency resulting in heterogeneous time-series containing considerable intervals with no recorded data. The purpose of this thesis was to apply an appropriate method that could simulate the time-series to accurately depict the complex fluctuation patterns exhibited by the groundwater. A Transfer Function Noise (TFN) model was selected on the account of it being a prediction method used in the hydrogeological field. The chosen method was evaluated based on its sensitivity to the quality and quantity of data. The model was systematically evaluated in conjunction with site characteristics. This was done for the purpose of establishing an optimal simulation scenario that could potentially be applied to other areas with similar geological composition. The model acquired moderately high goodness of fit values with the majority of the adjusted R^2 -values being over 0.5. Groundwater pumping was also considered and marginal improvements could be observed in the simulations, with more representative pumping data further improvements can be achieved. Overall, the fit of the simulated time-series compared to the recorded time-series varied between bores, the fitted model consistently overestimated and/or underestimated the measured maximum and minimum groundwater levels. This was assumed to be a consequence of the model not accounting for all of the hydrological processes in the area, such as the considerable surface run-off, which provides scope for further research and improvements. The sensitivity analysis concluded that the TFN model had the ability to simulate time-series using shorter time-span while still producing high goodness-of-fit measures, but was determined not function optimally for heterogeneous time-series'. Furthermore, the TFN model was compared with common interpolation methods i.e. Nearest Neighbor and Cubic Spline. These methods were considered to be more beneficial to use for time-series that contain smaller intervals of missing data.

Keywords: interpolation, Pastas, TFN model, Gotland, limestone, Cubic Spline, Nearest Neighbor, Impulse response function

Tidsserieanalys genom tillämpning av en Transfer Function Noise modell
Interpolation av grundvattennivåer i File Hajdar, Gotland

*Masters uppsats inom Masters programmet
Infrastruktur och miljöteknik*

SARAH ALI OCH KARISMA PATEL

Institutionen för arkitektur och samhällsbyggnadsteknik
Chalmers Tekniska Högskola

Sammanfattning

I File Hajdar, ett område i norra Gotland, består berggrunden huvudsakligen av kalksten där grundvattennivåerna uppvisar stora variationer mellan årstiderna. I över 40 år har grundvattennivåer registrerats med varierande intervall vilket har resulterat i heterogena tidsserier som innehåller utdragna perioder utan uppmätt grundvattendata. Syftet med denna uppsats var att tillämpa en lämplig metod som kan simulera tidsserier och på så sätt skildra de komplexa fluktuationsmönstren som uppvisas av grundvattnet. En *Transfer Function Noise (TFN)* modell, valdes på grund av att den är en prognosmetod som tillämpas inom den hydrogeologiska disciplinen. Den valda metoden utvärderades utifrån dess känslighet mot både datakvaliteten och- kvantiteten. Modellen utvärderades systematiskt i kombination med områdets egenskaper. Detta gjordes i syfte för att etablera ett optimalt simuleringsscenario som potentiellt skulle kunna tillämpas på andra områden med liknande geologiska kompositioner. Modellen uppnådde måttligt höga *goodness of fit*-värden där majoriteten av de justerade R^2 -värdena var över 0,5. Grundvattenpumpning inkluderades i modellen och marginella förbättringar kunde observeras i simuleringarna; med mer representativa pumpdata kan ytterligare förbättringar uppnås. Generellt varierade passningen mellan den originella och simulerade tidsserien för diverse borrhål. Modellen överskattade och/eller underskattade de uppmätta maximum och minimum grundvattennivåerna konsekvent. Detta antogs vara en följd av att modellen inte tog hänsyn till samtliga hydrologiska processer i området, till exempel den betydande ytavrinningen, vilket kan ses som ett förbättringsområde för vidare forskning. Känslighetsanalysen som utfördes menade att modellen hade förmågan att simulera tidsserier bestående av kortare tidsintervall samtidigt som den producerade en bra passform, dock fungerade modellen inte optimalt för heterogena tidsserier. Vidare, utvärderades TFN modellen genom jämförelser med tidsserier bestående av vanligare metoder; *Nearest Neighbour* och *Cubic Spline*. Dessa metoder ansågs vara mer fördelaktiga att använda för tidsserier bestående av kortare intervall av avsaknad data.

Nyckelord: interpolering, Pastas, TFN-modell, Gotland, kalksten, Cubic Spline, Nearest Neighbour, impulsrespons funktion

Contents

1	Introduction	1
1.1	Aim and Research Questions	2
2	Background	3
2.1	Groundwater and Evapotranspiration	3
2.2	Particularities of Groundwater Level Time-series	4
2.3	Karst Landscapes	5
2.4	Approaches for Imputing Missing Values in Hydrological Time-series .	6
2.4.1	Linear and Polynomial Interpolations	6
2.4.2	Nearest Neighbor	7
2.4.3	Methods Based on Distance Weighting	7
2.4.4	Methods Based on Kriging	8
2.4.5	Neural Network	9
2.4.6	Regression	10
2.4.7	Autoregression	11
2.4.8	Transfer Function Noise Model	12
3	Method and Data	14
3.1	General Strategy	14
3.2	Site Description	15
3.2.1	Geological and Hydrogeological Conditions	15
3.2.2	Groundwater Data	17
3.3	Data Processing	18
3.3.1	Groundwater Level Data	18
3.3.2	Pumping Data	19
3.3.3	Climate Data	20
3.4	TFN Model	21
3.4.1	Specifics of Chosen TFN Model	22

3.4.2	Strategy to Determine Model Performance	24
3.4.2.1	Coefficient of Determination and Pearson's Correlation	26
3.4.3	Assessment of Data Requirements	26
3.4.4	Alternative Interpolation Methods	29
4	Results	30
4.1	Bores and Climatic Data	30
4.2	Simulated Time-series Using Recharge	35
4.2.1	Time-series Containing Poor Data Quality and/or Disturbances	41
4.2.2	Evaluation Through a Moving Calibration Block	45
4.2.3	Evaluation Through the Removal of Automatic Measurements	47
4.3	Simulated Time-series Using Recharge and Pumping	49
4.3.1	Evaluation of Recharge and Pumping Based Simulations . . .	52
4.4	Comparisons with Cubic Spline and Nearest Neighbor	54
5	Discussion	58
5.1	Choice of Method	58
5.2	Model Performance and Data Requirements	59
5.2.1	AR(1) Noise Model	59
5.2.2	Evaluation of Recharge Based Results	60
5.2.3	Evaluation of Recharge and Pumping Based Results	63
5.2.4	Comparison to Alternative Interpolation Methods	64
5.2.5	Relationships Between StressModels and IRFs	65
5.3	Uncertainties	66
5.3.1	Data and Assumptions	66
5.3.2	Model and Functions	67
5.4	Further Research	68
6	Conclusion	70
7	References	71

8	Appendices	77
A	Appendix	77
B	Appendix	79
C	Appendix	80
D	Appendix	82
E	Appendix	86
F	Appendix	89
G	Appendix	90
H	Appendix	92
I	Appendix	94
J	Appendix	96
K	Appendix	98
L	Appendix	100

Acknowledgements

We dedicate this thesis to ourselves, we are very proud of the achievements and hard work we put into our education as well as in completing this work. In particular, we would like to thank one another for the support, encouragement and patience we expressed to each other throughout the writing of this thesis.

To our supervisors, Jakob Eng at Golder and Ezra Haaf at Chalmers University, whom have provided us with help and guidance throughout this journey, we want to say thank you. We also thank Markus Giese at Gothenburg University for his additional support and Hanna Rydholm, for her help in defeating Python.

We are very grateful for our strong support systems. These have been trying conditions, especially considering the global pandemic, and we could not thank our families enough for their continuous love and support.

A final thank you goes out to Friedlieb Runge for making this thesis possible.

Sarah Ali & Karisma Patel, Gothenburg, 2021.

List of Figures

2.1	The downward movement of water through the soil. M21 Groundwater, Fourth Edition, Copyright © 2014 the American Water Works Association. All rights reserved.	3
2.2	A system of neurons containing the input layer, middle layer and output layer of the network	9
3.3	The adopted general methodological procedure	14
3.4	Map over the area of interest, with File Hajdar quarry highlighted in purple, <i>Västra</i> quarry is not highlighted but can be found to the north of Bogevik, and west of Slite.	15
3.5	An illustration of the flow pattern retrieved from a report by Golder with permission (Cementa AB, 2017A).	16
3.6	An illustration of the 16 simulation scenarios used to produce the simulated time-series.	25
3.7	Flow chart to illustrate the methodological approach applied for the assessment of data requirements.	27
4.8	Three time-series with groundwater observations from Gotland, for a) BH43, b) SGU11001 and c) BH2006. The dark points represent recorded observations.	31
4.9	A map of File Hajdar including approximate location of all commercial and municipal points of groundwater extraction, as well as and the location of the nine bores evaluated in this thesis.	33
4.10	In the order of from top to bottom. Graph a) presents the precipitation and b) the calculated ET_o data.	35
4.11	Simulated results for BH43 with AR(1) noise model in the order of from top to bottom, with following simulation scenarios a) SM1; IRF Gamma, b) SM1; IRF Exponential, c) SM2; IRF Gamma and d) SM2; IRF Exponential.	36
4.12	Simulated results for SGU11001 with AR(1) noise model in the order of from top to bottom, with following simulation scenarios a) SM1; IRF Gamma, b) SM1; IRF Exponential, c) SM2; IRF Gamma and d) SM2; IRF Exponential.	37
4.13	R_{adj}^2 for simulated results using SM1 and SM2, which applied either a Gamma IRF or an Exponential IRF, simulated with the AR(1) noise model.	38

4.14	Simulated results for BH43 without AR(1) noise model in the order of from top to bottom, with following simulation scenarios a) SM1; IRF Gamma, b) SM1; IRF Exponential, c) SM2; IRF Gamma and d) SM2; IRF Exponential. These simulations were done without any noise model.	39
4.15	Simulated results for SGU11001 without AR(1) noise model in the order of from top to bottom, with following simulation scenarios a) SM1; IRF Gamma, b) SM1; IRF Exponential, c) SM2; IRF Gamma and d) SM2; IRF Exponential.	40
4.16	R_{adj}^2 for simulated results using SM1 and SM2, which applied either a Gamma IRF or an Exponential IRF, simulated without the AR(1) noise model	41
4.17	Simulated results in the order of from top to bottom, a) Slite 9 and b) BH96, contains a large gap of years with no measurements.	42
4.18	Simulated results in the order of from top to bottom, a) BH2006 and b) Slite 39, contains a disturbance and has few measurements.	43
4.19	Simulated results for BH98, measurements begin in late 80's.	44
4.20	Assessment of BH43: A moving calibration block of 2 years	45
4.21	Assessment of BH43: A moving calibration block of 7 years	46
4.22	The year-to-year Pearson coefficient of correlation for two and seven year moving block simulations BH43	46
4.23	Comparison of two and seven year moving block simulations, the dependent time-series and a "complete" simulation.	47
4.24	Systematic percentile removal of automatic measurements for BH43. In the order of from top to bottom, a) 100%, b) 75%, c) 50%, d) 25% and e)5%	48
4.25	BH43: Correlation Between Simulated and Recorded Groundwater Levels For Each Individual Year	49
4.26	R_{adj}^2 for simulated results using both pumping StressModels and recharge but without the AR(1) noise model	50
4.27	BH86: In the order of from top to bottom, a) Recharge Based Simulation, b) Recharge and Pump Using WellModel, and c) Recharge and Pump Using LH Model.	51
4.28	BH80: In the order of from top to bottom, a) Recharge Based Simulation, b) Recharge and Pump Using WellModel, and c) Recharge and Pump Using LH Model.	52

4.29 Four different simulations for BH86, out of these, three were simulated with one of the 11 pump manually removed. 53

4.30 A comparison between the original data and NN, CS & Pastas' TFN model simulations for BH86 54

4.31 In the order of from top to bottom, a) Every two weeks of the time-series are manually removed for b) Every six weeks of the time-series are manually removed 55

4.32 *R*-values for Pastas' TFN-model, NN and CS 56

4.33 Groundwater Regime: NN, CS, Pastas' TFN model and original data, for 1980 to 2003 57

8.34 A map containing all bores in the area 77

8.35 a) All available IRFs' plotted against block response b) All available IRFs' plotted against step responses 79

8.36 a) Average extraction per year for all municipal pumps b) Average extraction per year for Cementa AB pumps 82

8.37 Recorded time-series containing a) BH2006 b) BH80 c) Slite 39 83

8.38 Recorded time-series containing a) SGU11001 b) Slite 9 c) BH96 84

8.39 Recorded time-series containing a) BH43 b) BH86 c) BH98 85

8.40 Diagnostic hypothesis tests with AR(1) Noise Model - BH98 92

8.41 Diagnostic hypothesis tests without AR(1) Noise Model - BH98 93

8.42 A moving calibration block of 2 years through the time-series of BH86 94

8.43 A moving calibration block of 7 years through the time-series of BH86 94

8.44 The year-to-year Pearson coefficient of correlation for two and seven year moving block simulations BH86 95

8.45 The coefficient of correlation for all simulated years with a calibration period of 2 and 7 years for BH86 95

8.46 The coefficient of correlation for all simulated years with a calibration period of 2 and 7 years for BH43 95

8.47 Systematic percentile removal of automatic measurements for BH86 of a) 100% b) 75% c) 50% d) 25% and e) 5% 96

8.48 BH86: Correlation between simulated and recorded groundwater levels for each individual year 96

8.49 BH43: Coefficient of determination between simulated and recorded groundwater levels for each individual year 97

8.50	BH86: Coefficient of determination between simulated and recorded groundwater levels for each individual year	97
8.51	Diagnostics of the influences of pumps for BH80 in WellModel	98
8.52	Diagnostics of the influences of pumps for BH86 in WellModel	98
8.53	Diagnostics of the influences of pumps for BH80 in LH Model	99
8.54	Diagnostics of the influences of pumps for BH86 in LH Model	99
8.55	A comparison between the original data, NN, CS & TFN simulation for BH43	100
8.56	a)Every two weeks of the time-series is calibrated or simulated for b)Every six weeks of the time-series is calibrated or simulated for . . .	101

Abbreviations

AET Actual Evapotranspiration

ANN Artificial Neural Network

AR Autoregressive

ARIMA Autoregressive Integrated Moving Average

ARMA Autoregressive Moving Average

CS Cubic Spline

ET Evapotranspiration

ET_o Reference Evapotranspiration

FAO Food and Agriculture Organization

HBV Hydrological Bureau's Water Balance Department

IDW Inverse Distance Weighting

IRF Impulse Response Function

LH Looped Hantusch

LM Linear regression Model

MA Moving Average

NN Nearest Neighbour

RA Regression Analysis

P-M Penman–Monteith

PET Potential Evapotranspiration

SGU Geological Survey of Sweden

SM StressModel

SMHI Swedish Meteorological and Hydrological Institute

TFN Transfer Function Noise

1 Introduction

According to the United Nations groundwater serves as a drinking-water supply to nearly half of the world's population of whom 2,5 billion people are dependent on its availability to satisfy their basic daily water needs (UN, 2015). Groundwater is therefore a critical and precious resource that, if negatively disturbed, may have a significant impact on large populations. The main external influences that affect the groundwater are hydrological processes such as precipitation and evapotranspiration as well as discharge and recharge to and from larger bodies of water (USGS, 2016). Anthropological processes, such as pumping, are also important processes affecting the groundwater. If an aquifer is located in an area made up by highly permeable material the groundwater levels may be particularly sensitive to these external influences resulting in rapid and large fluctuations of water levels, a karst aquifer is one such example.

The groundwater level time-series can be used for detecting changes, patterns and evaluating abnormal changes in the groundwater levels; which may provide insight into the hydrological conditions of an area. The knowledge gained can then be used for predicting groundwater levels (Bakker & Schaars, 2019). A groundwater level time-series, especially one compiled over many years, may contain periods with missing data where measurements have not been logged. In such circumstances, a time-series analysis can be applied for the estimation of missing data. Accurate estimations of time-series data can be challenging as groundwater levels are the product of a complex hydrogeological system (Giese et al., 2020). There are numerous predicting methods available for the infilling of time-series, ranging from simple univariate interpolation methods to more complex algorithmically driven models.

In this thesis time-series data from File Hajdar, an area in Northern Gotland, will be estimated. With a bedrock consisting of mostly limestone, the hydrogeology in this area can be described as complex, with groundwater levels fluctuating up to 20 meters or more within a year. Approximately 60 observation boreholes can be found in File Hajdar of which the majority have been in function since the late 1970s. These bores provide time-series varying in quality and quantity, many of which include large periods of missing data. The groundwater is a critical resource for the inhabitants of Slite, which is a town near File Hajdar. A changing climate, along with the commercial and municipal extractions of groundwater as well as longstanding quarry operations, have had an impact on the groundwater in File Hajdar. Thus a time-series analysis will be conducted using a viable method for the infilling of the heterogeneous groundwater time-series from Gotland. The complex hydrogeology of the area, which is affected by climatic variations and anthropogenic effects, combined with the large variation in quality of the time-series data present challenging conditions for the infilling of time-series containing irregular data.

1.1 Aim and Research Questions

The aim of this study is to determine a suitable method, through a literature study and subsequent evaluation, for the infilling of groundwater levels from File Hajdar. This is a challenge as the time-series from this area are characterized by large seasonal fluctuations and contain heterogeneous data in terms of measurement frequency. The chosen predicting method should therefore be capable of capturing the inherent qualities of the time-series data.

The following research questions will be investigated as a complement to the aim.

- What are the weaknesses and strengths of the chosen models?
- What are the main uncertainties within the chosen method and how accurate can the results of the predicting method be considered?
- To what extent is the chosen model sensitive to the quality and quantity of the groundwater time-series data?
- Are anthropogenic effects detected in the groundwater levels from File Hajdar, if yes, how accurately are they represented in the model?
- As the site has a complex hydrogeology, can knowledge be gained from the predicting method which can be applied to other areas with similar geological composition?
- How does the chosen model compare to other less time-consuming and resource efficient methods?

2 Background

This chapter aims to provide an introductory background for the forthcoming groundwater time-series analysis. The chapter begins with an overview of groundwater and its components and a brief introduction to hydrological time-series analysis. This is followed by a literature overview of a few of the available predicting approaches or methods used to estimate groundwater levels, each method will be explained briefly followed by a literary example.

2.1 Groundwater and Evapotranspiration

Groundwater is an integral part of the hydrological cycle and is formed mainly due to precipitation, which includes snow and hail (AWWA, 2014). As the precipitation reaches the uppermost soil layer, referred to as the soil zone, it starts to infiltrate downwards. The soil zone is a part of the unsaturated zone, in which part of the precipitation will be absorbed by the roots from the vegetation, some of that water will later return to the atmosphere through transpiration. The precipitation then continues downwards through the intermediate zone to the saturated zone where it becomes groundwater (SMHI, 2018). The different soil-zones are presented in Figure 2.1.

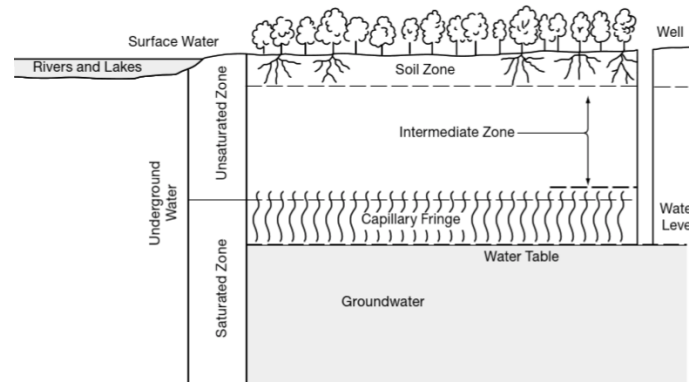


Figure 2.1: The downward movement of water through the soil. M21 Groundwater, Fourth Edition, Copyright © 2014 the American Water Works Association. All rights reserved.

Evapotranspiration (ET) is a vital part of the hydrological cycle and is defined as the sum of the evaporation occurring on all land surfaces and open water bodies along with the transpiration from plants. The following factors affect the rate of the ET; the surface or area of interest, how much water there is available in the area and the climatic variations (Britannica, n.d.). There are different calculation methods available to estimate ET, these are distinguished by different parameters. One such method, developed by Penman and Monteith, combines the energy balance and the mass transfer method in order to obtain the Penman-Monteith (P-M) equation to calculate PET (Allen et al., 1998). The equation was further developed by the Food and Agriculture Organization (FAO) of the United Nations to facilitate its use. The

altered version calculates a type of ET referred to as Reference Evapotranspiration (ET_o), using a fixed hypothetical crop reference for the surface; grass (Allen et al., 1998). According to several studies, the P-M equation yield the most accurate results, particularly in arid and humid climates (Allen et al., 1998). Another variant of ET is the Actual Evapotranspiration (AET), which measures the actual quantity of water removed through evaporation and transpiration when the area is assumed to have a restricted water supply (Park & Allaby, 2017). AET can be derived through the product of ET_o and a crop coefficient; Kc (Thomas, 2021).

2.2 Particularities of Groundwater Level Time-series

A times-series graph presents changes in values of a measured variable over time (SU, 2018). These may consist of measured observations for a short time-period or contain large amount of historical observation data that span decades. The observed data can be measured with either equidistant or irregular time steps (Cryer & Chan, 2008). Mapping data observations over time is very common and is applied in many fields. In hydrogeology, time-series are mainly used to evaluate changes in groundwater levels (Heudorfer et al., 2019). Not only do time-series provide an overview of groundwater level fluctuations over time, they can also be utilized to make predictions regarding future groundwater levels (SGU, 2021). When using time-series containing historic data to make predictions regarding future groundwater scenarios, it may be advantageous to use one that is continuous and without any large gaps (Lepot et al., 2017). A time-series analysis can otherwise be conducted for the purpose of estimating or interpolating periods of missing data.

For estimation purposes, research shows that the quantity of available data, in terms of years as well as measurement frequency, is critical in obtaining accurate estimations (Lepot et al., 2017). There exists no praxis on exactly how many years of data is needed, according to Heudorfer et al. (2019) continuous weekly or daily data for a minimum of 8 years can be suggested as the minimum requirement. For groundwater level estimations, however, the available data and its quality are the driving factor in determining the type of predicting method or model that should be used for the infilling of groundwater levels (Kasiviswanathan et al., 2016).

The characteristics of a time-series are relevant, particularly when conducting a time-series analysis. The quality and quantity of data in the time-series, whether the time-series have equidistant time-steps, what the long term trend is or if seasonal variations can be observed, are all factors that need to be considered (Hyndman & Athanasopoulos, 2018). Trends describe whether the groundwater observation increase or decrease with time, generally the term is used to describe long-term changes (Hyndman & Athanasopoulos, 2018). A time-series may exhibit seasonal and/or cyclic variations, while the former relates to the seasonal changes over a year, the latter describes non-seasonal variations. Groundwater time-series exhibit seasonal variations in data, with groundwater levels decreasing during the dry season and increasing during the wet season of a year.

A time-series analysis will result in residuals and produce, what is known as, white noise, or simply; noise. Residuals can be described as the difference between a interpolated, or predicted value, and the corresponding observed value (Hyndman & Athanasopoulos, 2018). A time-series consisting solely of residuals can be used and evaluated to gain an understanding of the performance of a predicting model. If the time-series show that the residuals are uncorrelated i.e random, have a combined average value of zero or are normally distributed, the model can be considered to produce accurate results (Hyndman & Athanasopoulos, 2018). If an analysis produces a time-series consisting of residuals that comply with the previously mentioned criteria then residual time-series can be considered to consist of noise; random variables that are independent of each other (McLeod, 2019). Though they are random, they occur due to other latent influences.

2.3 Karst Landscapes

Karst landscapes are formed through the erosion of exposed carbonate rocks, such as limestone (Goldscheider et al. 2020). Limestone is a sedimentary carbonate rock consisting of calcium carbonate and can be either biological or chemical (Marshak, 2015). Precipitation containing carbon dioxide infiltrate the limestone through fractures, causing the karstification of the rock. Carbonate rocks dissolve as a consequence of the acidity of the water that percolates through the fissures, both at and under the surface, widening the fractures and increasing the calcium carbonate content in the stone eventually resulting in larger fractures and underground networks of caves (Marshak, 2005). The process of karstification is exceedingly prevalent in carbonate rocks to some degree, particularly in landscapes consisting of limestone (Lolcama, Cohen & Tonkin. 2002).

As a result of the structure of karst landscapes, the aquifers in such environments generally display immediate responses to infiltration of precipitation (Goldscheider et al. 2020). Therefore, groundwater levels and spring discharges in such areas exhibit large and rapid fluctuations in groundwater levels. Furthermore, the unique interconnected structure of karst landscapes make them especially vulnerable to contamination, as contaminants could easily travel through the fractures (Goldscheider et al. 2020). The interconnected quality of karst systems make the landscape sensitive to invasive external disturbances, such as mining operations, which could exacerbate the naturally occurrences of sinkholes (Lolcama et.al. 2002). The heterogeneous character of the landscape often entail a complex transport pathway for the water (Mosthaf et al., 2018). Aquifers in such environments are therefore especially difficult to protect and maintain sustainably (Goldscheider et al. 2020).

The complexity of such a landscape can make it challenging to estimate groundwater levels. Mackay, Jackson and Wang (2014) constructed a lumped numerical model, AQUIMOD, for estimating groundwater levels. The study utilized time-series data from different geological settings consisting of different materials such as limestone, chalk, sandstone and greensand. The model performed well when estimating the fluctuations in groundwater within chalk sandstone and greensand.

However, it was least efficient in capturing the fluctuations of the groundwater levels in limestone. The models inability to capture the large fluctuations could be explained by transportation of water in limestone occurring through its fractures (Mackay et al., 2014).

2.4 Approaches for Imputing Missing Values in Hydrological Time-series

For the infilling of groundwater levels prediction methods or models can be used, which either interpolate or simulate values for gaps in time-series (Lepot, et.al, 2017). Interpolation is a numerical analysis tool that uses existing data to interpolate missing data. One can differentiate between univariate and multivariate interpolations. Univariate interpolation uses a single variable function, while multivariate uses multiple variables (Phillips, 2003). Simulations, on the other hand, do not interpolate between existing measurements, but rather use the data as a basis for its estimations. Prediction methods that simulate time-series are for instance Machine Learning methods. These are based on the principle of training sequences of algorithms to operate in a specific manner when analyzing data (Wang et al., 2017). This chapter will present a compiled lists of different methods and will be used as a foundation for the choice of a predicting method.

2.4.1 Linear and Polynomial Interpolations

Linear interpolation is one of the simplest forms of interpolation and is applied by creating a linear function between two distinct data points (Phillips, 2003). This a type of univariate interpolation, meaning it uses a function consisting of only one variable. Gnauck (2004) compared several univariate interpolation methods for interpolating time-series data. He found that, in comparison, linear interpolation provided better results, however, the time-series he used were stationary and as such, linear interpolation may not be optimal for non-stationary data. The degree of accuracy of this method is highly dependent on the quantity of time-series data available thus it may not be ideal for data sets with large amounts of missing data, especially if the intervals between data points are very large.

Contrarily, polynomial interpolation uses higher degree polynomials to create a curve that passes through existing data points, this type of interpolation can be univariate or multivariate (Salomon 2006). A univariate polynomial interpolation will result in a curve, whilst a multivariate polynomial interpolation will produce a surface. There are many variants of polynomial interpolation, one such is spline interpolation. Though both are based on the use of higher degree polynomials, it is important to differentiate between the two. Salomon (2006) defines a spline as following:

“A spline is a set of polynomials of degree k that are smoothly connected at certain data points. At each data point, two polynomials connect, and their first derivatives

(tangent vectors) have the same values. The definition also requires that all their derivatives up to the $(k - 1)$ st be the same at the point.” (s.140).

Spline interpolation consist of piece-wise polynomial functions that combined create a spline. The higher degree polynomial functions interpolate data between two existing data points or intervals and by increasing the number of piece-wise polynomial functions one increases accuracy. One of the other interpolation methods Gnauck (2004) compared was a from of spline interpolation known as the Cubic Spline interpolation. Though he found linear interpolation to be the superior interpolation method for his data set, he does note that the cubic spline interpolation is better suited for non-linear data.

2.4.2 Nearest Neighbor

According to Campozano et al. (2014) interpolation using Nearest Neighbor (NN) is based on the principle of applying proximity to interpolate new data. The authors state that for a spatial analysis of data, this method can use two different approaches. It can either use the geometrical distance to the nearest control station or it can use the control station with the highest correlation to the station with the missing values. A control station can in the hydrogeological field for instance be a well that records data. The obtained data from the control station is then used to interpolate missing values (Campozano et al., 2014). In time-series data on the other hand, this method can be applied by using the closest neighboring value to the interval or point of interpolation (Lepot, Aubin & Clemens, 2017).

More often NN is used as a comparative tool when evaluating other interpolation methods, this is because the values it produces are often not reliable enough to be used independently (Lepot, Aubin & Clemens, 2017). Bárdossy & Pegram (2014) have evaluated NN and compared it to a new copula-based method in a study that aimed to perform interpolation of daily and monthly rain gauge data in the semi-arid environment of the Southern Cape region in South Africa. The results of the study regarding the daily data show that the copula-based methods performed better in point estimations compared to NN.

2.4.3 Methods Based on Distance Weighting

Methods based on distance weighting are non-statistical and operate on the principle of spatial autocorrelation; the assumption that the data points that are closer together spatially are closer in value than those that are farther apart (Zarco-Perello & Simões, 2017). Thus, when interpolating a data point, the surrounding data points are used, the closer they are to the point of interpolation the more weight they are given relative to their distance, d_i . The weights are thus proportional to the inverse of the space between the existing point and the point being predicted, this is then raised to the power of p (ArcGIS Pro, n.d.). The power value, p , determines the rate at which the weight decreases, if p is too large then only the data points closest to

the interpolated point will carry any weight. If p is 0 then there will be no decrease in weight relative to distance, this will result in the predicted value becoming essentially a moving average (Li & Heap, 2008). For Inverse Distance Weighting (IDW) the power value p is equal to one. When p is equal to two, the method is referred to as Inverse Squared Distance Weighting (Lepot, Aubin & Clemens, 2017).

As this interpolation method is based on proximity it is important to define a searching window; an area within which the relevant data points used for the interpolation are (Zarco-Perello & Simões, 2017). Lam (1983) argues that this is one of the disadvantages of this method, as such an area can be difficult to define. He further states that the choice of weighting to be another disadvantage, as the chosen weighting function may not take into account the unique underlying geological characteristics. Zarco-Perello and Simões (2017) used inverse distance weighting in a comparative study for the interpolation of the sessile community of the Madagascar reef. They found that the distance between data points and data density to be very important to the accuracy of the results, the shorter the distance between samples, the better the predicted results.

Haaf (2020) developed a methodology to estimate groundwater levels at ungauged sites using IDW. Haaf's method is divided into three steps; firstly, duration curves are reconstructed using quantile model estimations. Secondly, the reconstructed duration curves are interpolated and extrapolated for the purpose of making them continuous. Finally, the groundwater time-series for ungauged sites are interpolated using IDW and groundwater time-series from gauged sites. This method has a weighting function based on a dissimilarity measure which expresses the spatial proximity between gauged and ungauged sites.

2.4.4 Methods Based on Kriging

Kriging is a type of geostatistical spatial interpolation tool, and is another method based on the principle of spatial auto-correlation (Li and Heap, 2008). Similar to IDW, kriging also uses weights, but is more complex. While the weights used in IDW are solely depended on the distance, the weights used in kriging take into account the distance, the location of each measured value and the overall arrangement of known measurements (ArcGIS Pro, n.d.). The predicted value is derived by applying a linear weighted average of all measurements neighboring the point of prediction. The weights and range of measured values used in the interpolation of a specific predicted value depend on the degree of correlation between the arrangement of known measurements (Lepot et al., 2017). The term kriging may be described as a form of umbrella term that include many different variants and methods derived from the basic principle of kriging (Li & Heap, 2008).

As it is a geostatistical tool, it has been utilized within the hydrogeological field. In the comparative study conducted by Zarco-Perello and Simões (2017) for the spatial interpolation of a sessile community, IDW was compared to a variant of kriging, known as ordinary kriging. Despite the more complex nature of ordinary kriging, the authors found that overall IDW yielded results that had lower degree of

errors. The authors stress that using spatial interpolation methods provide insight into patterns may not otherwise be understood using non-spatial analysis tools.

2.4.5 Neural Network

A neural network system usually consist of three layers, the first is the input layer which receives the data that is to be analyzed. In the middle, there are different hidden layers of neurons that are connected by links from the input to the output; these are responsible for the computing of the system. The neurons weigh the variables in the input data and decide whether or not to activate the next neuron in line. The computing leads to the output layer which produces the results (Lepot et al., 2017). For a visual representation, see Figure 2.2.

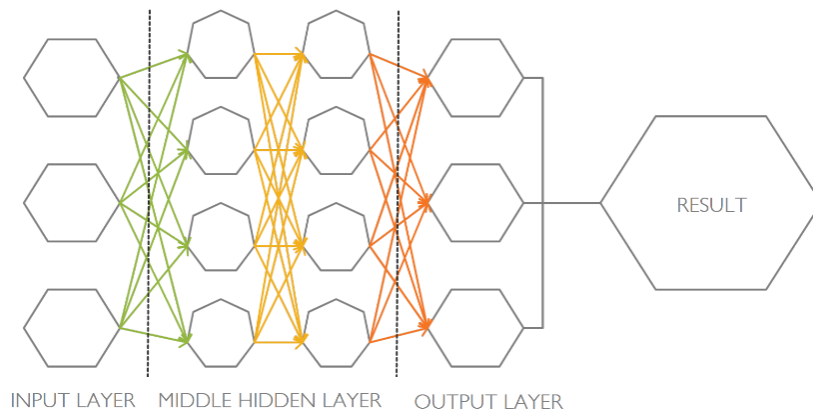


Figure 2.2: A system of neurons containing the input layer, middle layer and output layer of the network

During the training-period algorithms use a continuous iterative process, in order to organize all the information and separate what algorithms understand as the wrong answers from the correct answers. When the input data is processed the algorithms yield an output, that operates as a base for further improvements, they keep improving as a result of the iterative process (Wang et al., 2017).

Neural networks are a machine learning method, which mainly uses supervised and unsupervised learning to train the algorithms (Celebi & Aydin, 2016). Supervised learning is the most common method of the two presented. The algorithms are taught to identify labeled input data samples and predict or classify the outcome (Dey, 2016). Unsupervised learning is when the algorithm is given unorganized input data in order to identify patterns and the properties of said data (Celebi & Aydin, 2016).

Sathya and Abraham (2013) clarify that a few neural networks have the functionality to use either supervised or unsupervised learning in order to teach the algorithm how to operate. The computing process will, however, vary in degree of effectiveness as the algorithm may function differently with one learning method

compared to the other. This will therefore most likely result in different outputs depending on which learning method is used. They further examined a algorithm that can use both supervised and unsupervised learning called Artificial Neural Network (ANN). The algorithm sequence is taught through a supervised approach by weighting the combinations from the input with help of error signals. However, when ANN is taught with an unsupervised approach, a group of neurons are utilized which contain information from earlier sets on how to proceed with the computing (Sathya & Abraham, 2013).

In a study obtained by Jain and Kumar (2006), ANN is implemented for time-series forecasting from the Colorado River, US. The main purpose of the study was to make an overall modelling framework and forecast the monthly streamflow of the river. Jain and Kumar (2006) carried out the study by using both ANN and traditional time-series analysis in order to capture the non-linear trait of complex time-series and therefore produce more accurate results. At first the time-series data used mathematical filters in order to filter out the long term and seasonal variations in the data and modify it before implementing it in ANN. The result of the study concluded that ANN was advantageous to use for complex time-series data as it was able to find the hidden patterns in historical streamflow and get more accurate predictions compared to the other methods presented in the study.

2.4.6 Regression

Regression analysis (RA) is fundamentally used to investigate and model the future relationship between a dependent, or primary, variable and one, or more, independent variables (Chatterjee & Hadi, 2012). When conducting a statistical analysis on a data set using a regression analysis, one tries to fit the regression model to the data set. The resulting curve is an approximate fit rather than an exact one (Lepot et al., 2017).

Similar to many interpolation methods, RA too has several variants to its model. The most common is a linear regression model (LM), which is used to analyze the potentially functional relationship between a dependent variable and a single independent variable, that can represent for instance coordinates or water levels (Li & Heap, 2008). The data set, when using LM, is assumed to be normally distributed and homogeneous. In order to express the relationship, the following linear equation is used, Equation 2.1 (Chatterjee & Hadi, 2012):

$$Y = \beta_0 + \beta_1 X + \epsilon \quad (2.1)$$

Where, Y, is the response variable and X is the predictor variable. β_1 is a coefficient and β_0 a constant, with ϵ being an error factor. Other variants include non-linear regression models such as polynomial regression model, where the predictor variable, X, is modelled as a polynomial. This model can be used for data sets that have a non-linear spread. It has, however, been described as inefficient and non-applicable by researchers (Lepot et al., 2017). There is a risk of over-fitting when using any type of regression model, this is a consequence of the model producing an approxi-

mate fit and therefore generating residuals which in turn may produce uncertainties. Moreover, it is important to consider the amount of data available and any large gaps that the data may have, as this will have a significant impact in achieving optimal results (Lepot et al., 2017).

In Switzerland the authors Figura, Livingstone and Kipfer (2014) made a composite LM in order to predict the groundwater temperature for several aquifers. The inputs for the model were the regional air temperature projections which were calculated for three different greenhouse gas emission scenarios and historical groundwater data. The result of the study showed that the groundwater temperature will increase with different amounts of degrees depending on which scenario that was used. The regression model that was used was most effective when using long periods of time series data which contained at least one noticeable feature of variation, which they did.

2.4.7 Autoregression

The autoregressive model (AR) is a stochastic process; meaning that the model will continuously produce random variables (Lastrapes, n.d.). In order to use the input values and produce an outcome, the method uses past values in a time-series which are called 'lag', to predict the subsequent values (Lastrapes, n.d.). There are different levels of auto-regression; the more lags used, the higher the order of the regression is (Hyndman & Athanasopoulos, 2018). AR can be displayed as the model of order p , AR(p); if the order depends on one lag the AR model is of order one; AR (1).

More often AR can be combined with other several different methods, one such is a moving average (MA), often used in forecasting, similar to AR (Wilks, 2011). The model takes into consideration the past values in order to predict the next value in the series. However, when doing so it adds an error coefficient for the errors that occurs as a consequence of past prediction (Soliman & Al-Kandari, 2010). Levels of degrees are also applicable here, as it takes into consideration the number of lags for p in $M(q)$. Usually this results in values that may not be point-on correct, but circle around an average instead, hence the name moving average (Hyndman & Athanasopoulos, 2018). A combination of AR and MA therefore results in the Auto-regressive Moving Average Model (ARMA) which is used to describe weakly stationary stochastic time series (Lepot et al., 2017). This is done by integrating the two polynomials, AR and MA which also are bound to (p, q) . Where p is the order of the auto-regressive process and q is the order of the moving average process (Eatwell, Milgate & Newman, 1990).

Another frequently used term is I, the integrated model. It is a simple notation which is used in order to get the model to transform non-stationary data into stationary data (Kotu & Deshpande, 2019). The integrated model in combination with AR and MA creates the Autoregressive Integrated Moving Average Model (ARIMA) which used for analyzing and predicting time-series data. The difference between ARIMA and ARMA is that ARIMA uses non-stationary data which has a trend

component (Salcedo, Hernandez & Puente, 2009). ARIMA is bound to the components (p, d, q), where d is the number of time that the raw observations are integrated (Pandey & Bajpai, 2019).

Durdu (2016) used both ARIMA and ANN to model and predict the time-series for water quality in terms of temperature, dissolved oxygen and boron in the Menderes river in Turkey. The author used both models separately and proceeded to merge them into a hybrid to gain their advantages. This is due to ARIMA not performing well on non-linear relationships and ANN not being capable of handling both linear and non-linear patterns independently. The hybrid model was structured with ARIMA which received the data, filtered it and solved the linear part of the problem. The data was then inserted into ANN to be analyzed with an optimized training algorithm that capture the non-linear behavior of the data. The results of each method individually were then compared to the hybrid model. The hybrid model was able to capture the non-linear complexity more representatively and thus provided a more accurate prediction of the water quality.

2.4.8 Transfer Function Noise Model

Transfer function noise models (TFN) are multivariate mathematical functions often used in time-series modelling for irregular groundwater head observations. First introduced by Box and Jenkins (Von Asmuth, Bierkens & Maas, 2002), the model describes the dynamic interconnection between one or more input series and a single output or response series, and can be likened to a black box model (Hipel & McLeod, 1994A; Yihdego & Webb, 2010). The relationship between an input and output series is described as dynamic when it is characterized by constant change (Merriam Webster, n.d.). The resulting output series created by a TFN model therefore consist of a dynamic component and a noise component. Noise is produced in the model and can be described as influences, other than the input series, that also affects the resulting output series. Whilst the dynamic component, which is represented through a transfer function, describes the dynamic influence of an input series on an output series (Hipel & McLeod, 1994A). Different methods may be used to model the noise component within both variants of the TFN model. If the TFN model is stationary, most commonly ARMA is used for modelling the noise-component, if not; it can instead be modelled with the ARIMA method (Hipel & McLeod, 1994B).

An alternative and more recent variant of the TFN model applies impulse response functions (IRF) to model the dynamic component, this version of the TFN model has been widely used in the hydrogeological field (Von Asmuth et.al, 2002). Peterson and Fulton (2019) simulate gross recharge, transmissivity, storativity, and daily usage at multiple production bores with help of a self-developed IRF based TFN model called HydroSight, which applies a Kriging noise function rather than an ARMA noise function. The model was tested on numerous observation boreholes in Australia and achieved results with high accuracy. Another study conducted by Yihdego and Webb (2010) used both a TFN model and an auto-regressive model to model groundwater levels in regions consisting of inter alia limestone. The study

was conducted for the purpose of determining how much the groundwater levels fluctuate as a result of land use and climatic impact. They concluded that a substantial amount of the fluctuations are the result of climatic variations, for instance around 90% of the fluctuations in groundwater levels from areas containing limestone could be explained by climatic variations.

3 Method and Data

This chapter covers the methodology for the predicting of the groundwater level time-series from File Hajdar as well as the site description. The chosen method is the open source package; Pastas, which is a TFN model. The infilling of the time-series and subsequent evaluations procedure will be presented in this chapter, in addition to the motivations and assumptions made during the process.

3.1 General Strategy

The methodological procedure adopted in this thesis is presented in Figure 3.3.

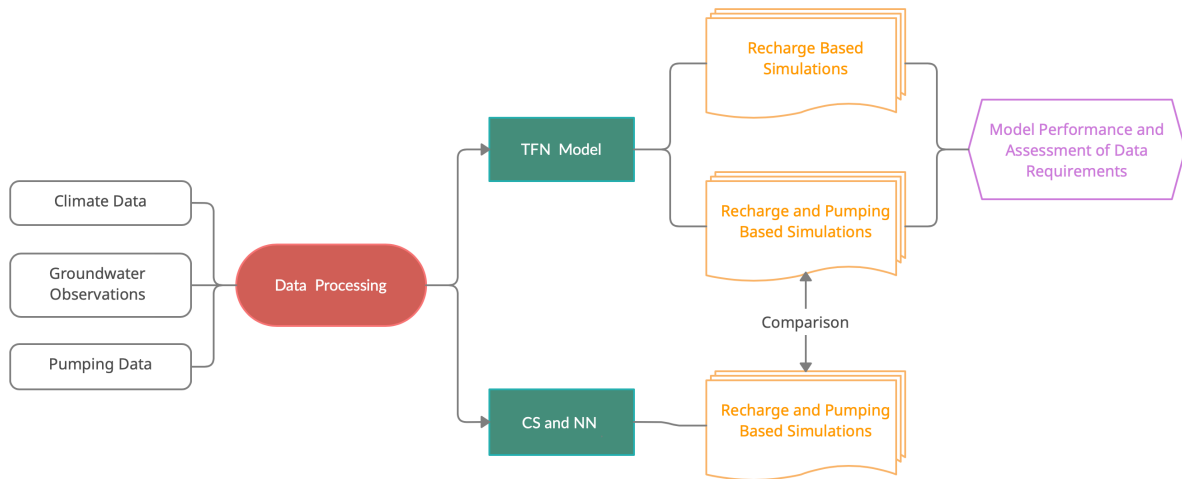


Figure 3.3: The adopted general methodological procedure

Firstly, the data processing component entailed the calculation and/or compilation of all data used within the thesis, such as the Et_o , precipitation, groundwater level time-series' and and pumping data. The processed data was thereafter used in the Pastas' TFN model to simulate groundwater time-series. The simulation created with TFN model were either recharge based or recharge and pumping based. Recharge in this case is referring to primarily rainfall recharge, as the Pastas' TFN model only accounts for precipitation and ET. Thereafter, the model performance and data requirements for the model were assessed. Finally, the simulated time-series derived through the TFN model were compared to interpolated time-series using Nearest Neighbor (NN) and Cubic Spline (CS), these interpolations did not require any climatic data and were thus only based on recorded groundwater time-series.

3.2 Site Description

The groundwater data originates from File Hajdar, a territory that spans from the town of Slite to Lake Tingstäde, see Figure 3.4.

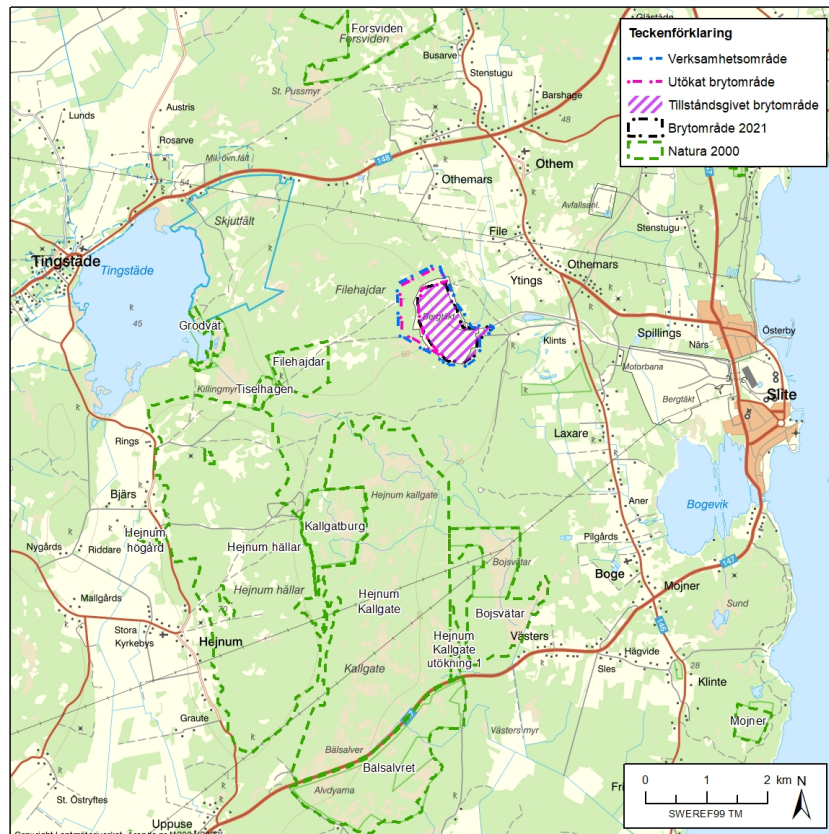


Figure 3.4: Map over the area of interest, with File Hajdar quarry highlighted in purple, *Västra* quarry is not highlighted but can be found to the north of Bogevik, and west of Slite.

There are two major quarries operated by Cementa AB in this area; *Västra* quarry and the File Hadjar quarry, the former is located close to Slite and the latter is situated between the lake and Slite. There exists a third quarry close to *Västra* quarry, but is not assumed to have an influence on the groundwater levels. *Västra* quarry has been operational since 1960s whilst the File Hajdar quarry commenced operation in 1981. This study examines boreholes that can be found between the lake and File Hajdar quarry, see Figure 3.4. The territory is host to nature reserves and water conservation areas, which can be found southeast of the lake and south of the File Hajdar quarry, large parts of the territory is covered with vegetation.

3.2.1 Geological and Hydrogeological Conditions

Gotland has a sedimentary bedrock consisting largely of limestone, other sedimentary deposits found in smaller quantities include marl and sandstone. A very thin

layer of soil and vegetation cover most of the limestone, hence the conventional "soil zone", demonstrated in figure 2-1, is not entirely applicable to the geology of File Hajdar (SGU, 2013). The sedimentary rocks on Gotland are horizontally layered and have a slight inclination of between 0.15 to 0.3 degrees in the south-east direction (Cementa AB, 2017A). Groundwater is found between the horizontal layers and vertical fractures of the limestone as well as in the the underlying marl. The groundwater is separated by denser layers with low effective porosity. Precipitation permeates the denser layers through a limited number of vertical fractures, which vary in size and degree of permeability, see Figure 3.5 for a schematic (Cementa AB, 2017A). In total, six groundwater bearing layers have been identified in the territory that spans between *Västra* quarry and File Hajdar quarry (Holmén, 2017). According to J. Eng (verbal communication, 12-03-2021), a hydrogeologist that has long worked with the Cementa AB, the File Hajdar area can in broad strokes be considered to consist of a single confined aquifer. The flow between the layer can be described to have a leaky character, he does stress that this is naturally a simplification. File Hajdar belongs mainly to the Aneråns catchment area which is 21.5 km^2 , the surface run-off empties out into the Baltic Sea (Cementa AB, 2017B).

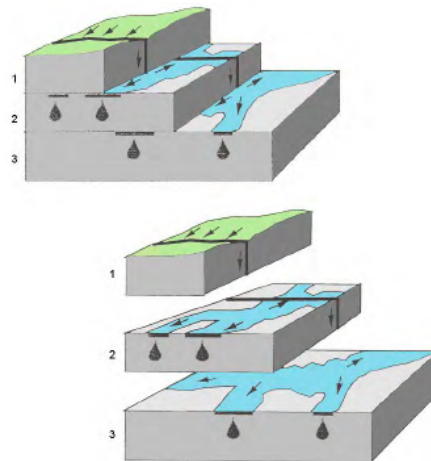


Figure 3.5: An illustration of the flow pattern retrieved from a report by Golder with permission (Cementa AB, 2017A).

The vertical fractures at surface level are mostly smaller in size but they allow for water to be quickly transported through the limestone (Djurberg, 2016). October to April can be described as the wet season, during which precipitation falls more frequently. During this season, the limited vertical fractures and their size tend to cause the ground to become fully saturated, due to the heavier and more frequent precipitation (Cementa AB, 2017A). This limits the fractures ability to allow for water to infiltrate which in turn creates larger amounts of surface run-off. These conditions are prevalent during the months between October to April, where much of the precipitation that isn't absorbed by vegetation or evaporated ends up in the Baltic Sea through surface run-off (Djurberg, 2016). The fully saturated condition of the surface during the wet season induce higher surface level pressures due to the limited pore capacity of the ground. During warmer months, or the dry season,

the surface level pressures abate due to groundwater recharge being low as a result of the high evapotranspiration and lower amounts of precipitation (Cementa AB, 2017A). In addition to the surface run-off, groundwater discharge and recharge to and from Tingstäde lake, as well as the Baltic Sea, also occurs albeit in a slower and smaller capacity, with the net quantity varying over the year and being dependent on numerous factors (J. Eng, verbal communication 25-03-2021).

Groundwater recharge in Gotland is highest during the wet season, when evapotranspiration is low and precipitation - in the form of rain and melted snow - is high (Djurberg, 2016). The recharge rate in the territory is believed to vary greatly due to the complex hydrogeological conditions in the area. The rate of recharge is dependent on a variety of factors; such as whether or not the limestone is covered in a soil layer, the quantity and size of the vertical fractures as well as the capacity and fullness of the groundwater reservoir (Djurberg, 2016). The estimated recharge rate in File Hadjar may on average be low as around 10 mm/year due to the limited amount of vertical fractures (Cementa AB, 2017A). Whilst in an area south-east of File Hajdar, known as the cross-zone, the recharge rate during a typically wet year is estimated to be around 250 mm/year. It is believed that the cross-zone has an elevated capacity for permeation, either due to the presence of an uncharacteristically larger number of vertical fractures or through several prominent vertical fracture zones. The cross-zone is believed to be 550 meters in depth and at least 2 kilometers in width. The county extracts its water supply from this area through seven production wells drilled in a linear formation (Cementa AB, 2017A).

3.2.2 Groundwater Data

The area has since the late 1970s been subject to logging of groundwater levels. From the 1960s to early 2000s the groundwater data was collected by Geological Survey of Sweden (SGU), thereafter Cementa AB handled the data collection. The approximate location and number of observation bores used for groundwater data collection over the years can be found in Appendix A. According to a geologist at Cementa AB, A. Birgesson (verbal communication, 26-02-2021), many of the older boreholes used for groundwater data collection were originally created for prospecting and later re-purposed by SGU. This is important to consider in terms of the impact this has on the the groundwater time-series. As the depth of many borehole was originally dimensioned for prospecting, this may therefore result in time-series' that are, for instance, upper and/or lower bounded. Majority of the older boreholes that were used for prospecting still exist in the area but are no longer in use.

During the period in which SGU collected groundwater data, measurements were taken manually and with irregular frequency, for the majority of 1970s and 1980s data was collected roughly twice a month. Later, the frequency progressively decreased and became more sporadic, varying from once a month to every few months. SGU stopped collecting groundwater data altogether during the early 2000s, Cementa AB resumed the data collection for a number of the boreholes during the

2010s. In many of these they installed automatic measuring stations, which provide hourly data collection. Cementa AB also drilled new boreholes in conjunction with the resumed data collection. Information regarding hydrogeological conditions, such as transitivity and hydraulic conductivity is available, however, it is difficult to compile in such a manner that accurately represents the complex structure and character of the limestone. These parameters vary greatly over the territory, as well as between the groundwater water bearing layers.

The territory from which the groundwater time-series originate is close to the town of Slite as well as its water treatment plant, it is therefore important to consider groundwater extraction from the area. As mentioned, the municipality has seven production wells in the cross-zone, from which drinking water is extracted. There are also a considerable amount of private wells in Slite and the surrounding territory. Moreover, the two quarries in the area also pump groundwater and the precipitation that is collected in the quarries in order to proceed with their operations. Cementa claims to have made efforts to reduce their pumping of groundwater after 2011 (Cementa AB, n.d.). This means that groundwater was extracted at higher rates in the years prior to 2011, however, pumping data for these years was not accessible. These sources of extraction combined have had a considerable effect on the groundwater levels over the years.

3.3 Data Processing

Pastas uses dependent and independent time-series to simulate groundwater levels. The dependent time-series contain recorded groundwater observations and can therefore have missing data or be irregular in measurement frequency (Collenteur et al., 2020). The independent time-series should not contain any missing data and must have regular daily time steps and begin preferably 10 years before the first borehole observation data.

The following independent time-series have been compiled:

- Daily precipitation in mm.
- Daily ET_o data in mm.
- Daily pumping rates in cubic meters per day.

3.3.1 Groundwater Level Data

The site description and the map identifying the bores in File Hajdar, found in Appendix A, convey the extensive spread and amount of bores found in the area. As the groundwater data has been collected by different parties over the years, the very first step in the simulation process was to systematically locate, retrieve and organize groundwater data for each and every borehole. By charting the data an overview of all pertinent boreholes in the area was obtained, likewise the boreholes

geographical positions and their relative distances to each other and other areas of importance was also acquired.

Following the data collection, time-series for each bore was created and assessed. This aided in identifying boreholes that exhibited unusual changes in groundwater levels, such bores were further researched in an attempt to determine the cause of the irregularity. The bores determined to have significant disturbances from external factors were excluded from the infilling process, however, if the disturbed parts of the data could be identified, it was manually removed from the time-series.

In the site description it was mentioned that the logging of groundwater commenced during late 1970s, many time-series therefore contain observation data from the 1970s. However, the ET_o used in this study could only be calculated for 1980 to 2020, which restricts the use of the observation data from 1970s. The dependent time-series' have to be compiled in unit meters above sea level. However, some bores from File Hajdar have recorded observation levels in unit meters below the sea level, these bores located at a lower elevation have therefore recorded groundwater levels in negative units. As the Pastas' TFN model cannot calibrate time-series' containing negative values, these were raised by a few meters in order to circumvent this issue. Furthermore, it should be noted that the automatic device used for recording groundwater levels since the 2000s provide hourly data. The TFN model requires a daily time-step, therefore time-series' containing hourly recorded values were averaged to represent daily groundwater levels.

3.3.2 Pumping Data

There are eleven major points of groundwater extraction in the territory, seven of these are the drinking water production wells operated by the municipality through a water treatment plant; P1, P2, P3, P3A, P4, P5 and P6. Though the plant has been operational since 1974, only pumping data from 1998 to 2019 was accessible. It has been verbally communicated by J. Eng (04-06-2021), that groundwater was believed to be extracted at higher rates prior to 1998, but pumping data was not accessible to confirm this. As the pumping data was only available for a limited number of years, for the years prior to 1998 an average daily value based on all the available data was used instead. The pumping data was supplied in cubic meters of extracted groundwater per month. As the TFN model requires daily input data for all of the independent time-series, the extraction rate is required to be in unit cubic meters per day. Therefore, an average daily pumping rate based on the monthly volume of extracted groundwater was calculated, it should be noted that this implies a constant daily rate of extraction during all days of the year.

The File Hajdar quarry and the *Västra* quarry have two operational pumps each, which are utilized to remove the precipitation and groundwater that collects in the quarry. The pumping rates are partially reflective of the precipitation pattern, as water is pumped more frequently during the wet season. The water is removed at higher rates during most of the year with the exception of the dry season. Monthly pumping rates in cubic meters for all four pumps are available for the years 2011

between 2015, despite groundwater having been extracted by the quarries for much longer. As this data was also provided in monthly volume of extracted groundwater, it was revised to an average daily rate, based on the monthly data. It should be noted that this implies a constant daily extraction rate, which is not representative as the water is pumped when the need arises, such as during rainfall events, as to minimise collection of water in the quarry. Historically, the pumping rate can be said to have increased as the quarries got deeper and larger. The acquired pumping data does not reflect this, as it only represents a period of four years. *Västra* quarry has been in operation since 1960s and therefore an assumption was made that pumping rates have more or less plateaued from 1980 and onwards, thus the derived average daily values were used between 1980-2020 to represent the daily pumping rates at *Västra* quarry. The same cannot be assumed for File Hajdar quarry as it commenced its operations in 1982 and would therefore have not pumped much in the beginning. To reflect the increase in pumping over the years for File Hajdar quarry, the average pumping rates were linearly increased from zero, in 1982, to the derived average values in 2005. An assumption is made that the pumping rates are plateaued even for File Hajdar after 2005. It is important to note that the independent time-series that contain the pumping data from both quarries therefore consists of average daily data, for a 40 year period, derived using only four years of pumping data in total. Yearly average variations in groundwater extraction from municipal and commercial pumps can be found in Appendix D.

3.3.3 Climate Data

The daily precipitation data was downloaded from the Swedish Meteorological and Hydrological Institute (SMHI) for the years 1971 to 2020, the precipitation includes both rain and snow and it was recorded by a station in Hejnum, located south of Lake Tingstäde. Hejnum station started recording precipitation during 1971; for the year 1970, for which there is no data, the TFN model extended the time-series with an average of the precipitation values. The data needed for the calculation of ET_o was acquired from SMHI and the Copernicus database (Hersbach, et.al, 2018). With precipitation being an exception, there was an extensive lack of meteorological data for the area of File Hajdar therefore data from a station in Visby was used instead. Visby is a town based on the western coast of Northern Gotland, as opposed to File Hajdar which is on the eastern coast, Visby is located roughly 37 kilometers from Slite.

As P-M is estimated to provide the most accurate ET_o (Allen et al., 1998), this method was chosen to calculate the ET_o for northern Gotland. Rather than using the original version of the P-M method, the FAO-56 version was used instead. This version doesn't require the heat flux factor, G , which further reinforced the choice of calculation method as heat flux data was unavailable for northern Gotland. ET_o was calculated for the years 1980 to 2020, ideally, ET_o data should start 10 years before the first borehole observation point, i.e 1970. This, however, was not possible as reliable solar radiation data for the years prior to 1980 could not be acquired. The TFN model corrects for the shorter span of the ET_o data by extending it 10

years. The extension contains average values of the calculated ET_o based on the years 1980 through 2020.

The FAO-56 version requires climatic data in the form of daily maximum and minimum temperatures, average daily wind speeds [m/s], relative humidity and incoming solar radiation. SMHI measured the wind-speed at 10 meters from the ground, whereas FAO-56 requires the wind-speed to be measured at an elevation of two meters from the ground, this was corrected for using an equation provided by the guide. It is important to note that the wind speed data downloaded from SMHI had days with missing measurements over the 40 year period for which ET_o is being calculated. For those days, the measurements from the previous day and the day after are averaged and used instead. The humidity data from SMHI from Visby had an irregular measurement frequency and contained large intervals with no recorded data, hence could not be considered adequate enough to be used. Substitutions and corrections were made according to the guide whenever humidity data was needed for the ET_o calculations. Finally, incoming solar downward radiation data was acquired from Copernicus' ERA5 database (Hersbach, et.al, 2018). The ERA5 database contains estimated historic weather and climate data which is based on past observations and numerical analysis. As the solar radiation data is estimated this contributes with an inherent uncertainty.

ET_o , seen in Eq.(8.21), can be calculated by the addition of ET_{rad} [$\frac{mm}{day}$] and ET_{wind} [$\frac{mm}{day}$], which are derived using Eq. (8.19) and (8.20).

$$ET_{rad} = DT * R_{ng} \quad (3.1)$$

$$ET_{wind} = PT * TT(e_s - e_a) \quad (3.2)$$

$$ET_o = ET_{wind} + ET_{rad} \quad (3.3)$$

where; DT, PT and TT are auxiliary terms used to facilitate the calculation procedure. R_{ng} is the revised net radiation, e_s [kPA] is the mean saturation vapor pressures and finally, e_a [kPA] represents the actual vapor pressures.

The step-by-step calculation procedure of daily ET_o [$\frac{mm}{day}$] can be found in Appendix E. All of the equations used for the calculation of ET_o , including Eq.(8.21), (8.19) and (8.20), are supplied by a University of Florida guide for the calculation of FAO-56 ET_o (Zotarelli et al., n.d.).

3.4 TFN Model

A TFN model is a multivariate mathematical function, used in modelling irregular groundwater time-series. The model is appropriate to use within the hydrogeological field, and more specifically to model groundwater levels, due to its ability to describe and capture the dynamic nature of groundwater recharge. The Box and Jenkins'

TFN model used in the context of hydrogeology, is comprised of Equation 3.4 and (3.5) (Von Asmuth et.al, 2002):

$$h_t = h_t^* + n_t + d \quad (3.4)$$

$$h_t^* = \Theta(B)p_t \quad (3.5)$$

Where, h_t are the observed groundwater head at time, t and h_t^* the predicted groundwater head at time, t . The n_t are the produced residuals and d represents the predicted groundwater level with only local drainage. The equation for h_t^* contains the actual transfer function, $\Theta(B)$, with p_t representing the precipitation surplus at time, t . In the transfer function Θ represent the weights and n_t provides the noise. The time, t , is discrete in this equation, and as such does not capture the continuity of the fluctuations that groundwater levels may exhibit (Von Asmuth et.al, 2002).

An alternative TFN model created by Von Asmuth et.al (2002) uses an impulse response function (IRF) to model the groundwater levels. These response functions are predefined and continuous over time, which is what differentiates IRF based TFN models from the Box and Jenkins model. IRF is an analytical expression that describes a dynamic system's reaction to a sudden impulse or change. This version of the TFN model fundamentally has the same equation as the one presented above, but components of the equation are calculated differently. In the case of h_t^* , the equation contains a convolution integral and is as follows Equation 3.6(Von Asmuth et.al, 2002):

$$h_t^* = \int_{-\infty}^t \theta(t - \tau)p(\tau)d\tau \quad (3.6)$$

Where, the IRF is represented by $\theta(t)$ in the convolution integral and $p(\tau)$ being the precipitation surplus.

3.4.1 Specifics of Chosen TFN Model

The Pastas package was used in this study to produce the simulations, the package uses Python and is created for inter alia analysing and estimating groundwater time-series, and is built by Collenteur et al., (2020). The model is not limited to estimating groundwater levels but can also be used for other time-series based processes, for instance it is used as a data preparation method for draught analysis (Brakkee, van Huijgevoort & Bartholomeus, 2021). This sub-chapter will cover the specifics of this model and how it functions.

Pastas uses an IRF based TFN model that has the option to apply either an AR(1) or an experimental ARMA noise model, the latter can at the time of writing only be applied on time-series that have equidistant time-steps and is therefore not used in the simulations for File Hajdar. The purpose of the AR(1) model is to model and reduce noise. However, if the model determines certain data to be noise, when it is not, then the modelling may not be accurate. Collenteur et al. (2020)

explicitly state, that the AR(1) noise model may not provide accurate results for high frequency data, meaning hourly or daily measured data, which is what many of the dependent time-series from File Hajdar contain.

To create the simulations Pastas' TFN model uses one or more independent time-series containing hydrological data that affect the groundwater levels, such as precipitation or evapotranspiration, and a dependent time-series containing the observed groundwater levels. It then proceeds to calibrate the whole data series, within a specified simulation period, to thereafter simulate a complete and continuous groundwater levels time-series. The TFN model offers the option to specify the time period of the simulation, therefore shorter periods of groundwater time-series can be produced.

The TFN model functions by applying the independent time-series' as stresses to the model, each separate stress contributes in simulating groundwater levels. These are applied through a StressModel, which in turn uses an IRF to model the groundwater level's response to the different stresses. The Pastas package includes a few predefined IRFs, these can be found in Appendix B . The choice of IRF used in the TFN model depends on the type of stress being applied. Collenteur et al. (2020) recommends using the Gamma IRF when simulating recharge as it is a general function, not limited to solely recharge, and can be used to model different types of stresses. There are additional IRF options available such as the Exponential IRF, which can be used to simulate stresses that appear to have a nearly immediate effect on the groundwater level. Additional simulation alternatives include adding linear trends in the simulation.

There are several available options for StressModels in the TFN model, for recharge models a standard StressModel (SM1) or alternatively StressModel2 (SM2) can be used. The difference between the two is based on SM1 simulating recharge by subtracting ET from precipitation, whilst SM2 estimates the AET by introducing an additional factor, f , and multiplying it with the ET_o . What exactly the factor f is or how its estimated, is not specified, but its purpose is to convert ET_o to AET (Collenteur et al. 2020). According to Collenteur et al. (2020), SM2 may provide a better fit but does not necessarily translate into a better model. In addition to recharge, the model also offers the option to simulate the effects of pumping, by applying a Hantush IRF in separate StressModels. The Hantush IRF utilized to model the influence of pumping is based on the Hantush well function for leaky aquifers which is dictated by ten assumptions, these are listed in Appendix B. If a borehole has multiple pumping wells in its vicinity these can be modelled using two additional StressModels, either a WellModel or by looping multiple wells using a "for-loop" (Collenteur et al., 2019). The difference between the two options is that the WellModel considers the distances between the observation well and the pumps whereas the looped option, henceforth referred to as the LH model, does not.

The Pastas package includes diagnostics tests and hypothesis tests as well as various goodness-of-fit measures that aid in evaluating the residuals and the fit of the simulated groundwater levels. The diagnostic tests provide information on auto-correlation, the distribution of the residuals and noise. To improve the knowledge

gained from the diagnostic tests, hypothesis tests can be utilized to test a simulation for the assumptions of noise and residuals. These are based on rejecting or accepting a null-hypothesis (H_0). The H_0 hypothesis may state that the residuals are uncorrelated, have homogeneity of variance or are normally distributed (Collenteur et al., 2019). If the hypothesis is rejected, it can be determined that data does not comply with one of the criteria for noise, see sub-chapter 2.2. There are several different hypothesis tests included in the Pastas package which test for normality or autocorrelation. It's important to note that the reliability of these tests are dependent on several restrictions, such as the sample size of the data or if the dependent time-series has equidistant time-steps. After each simulation Pastas presents statistical figures and goodness-of-fit estimations used to evaluate the degree of accuracy of the simulated time-series compared to the observed groundwater levels. A goodness-of-fit tool supplied by the TFN model is the adjusted coefficient of determination (R_{adj}^2), which will be discussed more in depth later in the chapter.

The package was operated in Jupyter Notebook, which supports Python. The scripts used in this thesis can be found in Appendix C.

3.4.2 Strategy to Determine Model Performance

The assessment of the TFN model was mainly centered around the noise component, the IRFs and the StressModels used. This was done for the purpose of evaluating model performance as well as acquiring knowledge that may be applied to areas with similar geology. The results of all simulations were assessed in consideration with the information acquired of the bores' location, history and distances relative to each other, for the purpose of determining whether a single response-function or StressModel presented superior results in terms of R_{adj}^2 . The site-specific factors taken into account for each bore were: distance to body of water, distance to quarry and distance to municipal pumps as well as considering the vegetation and elevation surrounding each borehole. Information was acquired from SGU regarding the soil stratigraphy in File Hajdar, mainly through soil profiles from bores drilled in the area, although there are only a few and many weren't located near any observation bores. Nevertheless, when possible, these soil profiles were used as a complement to the general information acquired regarding the stratigraphy of the area.

A systematic strategy was adopted for simulating the groundwater level time-series. Firstly, the dependent time-series' of each bore included in the study was assessed to determine if any linear trends could be observed throughout the time-series'; those that exhibited an obvious linearity during a specific time period were applied a linear trend. When simulating time-series using only precipitation and ET_o , i.e. recharge, both StressModels, SM1 and SM2, were evaluated and used in conjunction with either Gamma and Exponential IRF. The first was included as an IRF to be tested due to it being a standard function to model any stress, especially recharge. The Exponential IRF was included as it models an immediate reaction to a given change, which is an appropriate IRF to consider for the area due to the quick transportation of water in the limestone. To further improve the

simulations an additional parameter was added to the recharge based simulations; pumping. It was only applied as a stress for time-series belonging to bores that have a municipal or commercial pump in its vicinity. To model the influence of pumping, the Hantush IRF was applied through both WellModel and LH model. The systematic strategy consisted of combining different StressModels with different IRFs, resulting in numerous simulation scenarios. A flow chart of the various simulation scenarios can be found in Figure 3.6. The simulations that only use ET_o and precipitation as independent variables are referred to as recharge based scenarios, whereas the simulations scenarios that include an additional variable of pumping are referred to as recharge and pumping based scenarios.

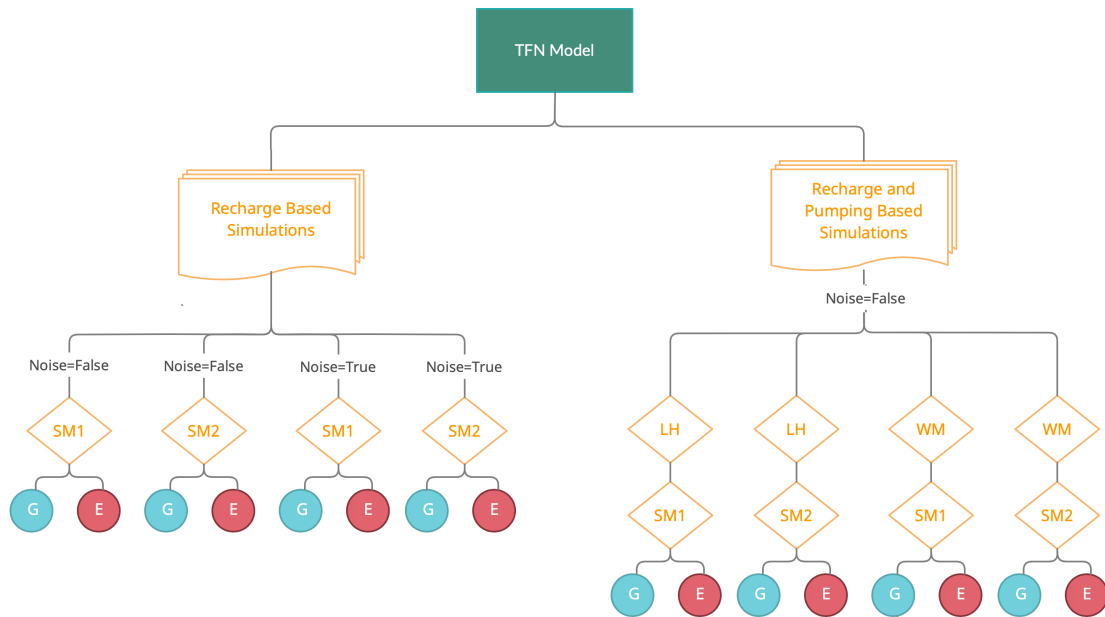


Figure 3.6: An illustration of the 16 simulation scenarios used to produce the simulated time-series.

As can be seen from Figure 3.6, eight simulation scenarios were created for recharged based simulations, of which four were calibrated with the AR(1) noise model and the rest without. In Figure 3.6, G and E denote Gamma IRF and Exponential IRF, respectively. For modelling pumping as a stress, the recharge scenarios were used in conjunction with the StressModels for pumping; WellModel and the LH model, denoted as WM and LH, respectively, in Figure 3.6. Through not explicitly stated in Figure 3.6, the both StressModels for pumping utilize the Hantush IRF. Modelling pumping resulted in an additional eight scenarios, these were simulated without the AR(1) noise model.

The concept of noise was briefly introduced in the literature study, noise can be described as influences, other than the input, that interfere and impact the output of a TFN model (Hipel & McLeod, 1994B). Noise are random variables present in the time-series. Diagnostics and hypothesis tests were utilised to assess the noise

and residuals of the simulations, although the results of these tests were analysed critically as they presuppose certain conditions for the data sets. The purpose of the tests were to check for auto-correlation in the residuals, if significant auto-correlation was found then it meant that the model could be improved. As previously mentioned, the AR(1) noise model may perform poorly on high frequency data, it was therefore expected that it may not provide accurate modelling of the noise for the File Hajdar data set, as many dependent time-series from the area contain periods of daily recorded data. Therefore, the time-series' were simulated with and without the noise model, for all recharge based simulation scenarios.

3.4.2.1 Coefficient of Determination and Pearson's Correlation

Mainly three goodness-of-fit measures were used to assess the fit of the simulated time-series; the coefficient of determination (R^2), the Pearson coefficient (R) and an adjusted coefficient of determination (R_{adj}^2) produced by the TFN model. The R^2 -value can be defined as the degree to which the variability in the dependent time-series can be explained or predicted by the simulation, the value ranges from 0 to 1 (Moles & Terry, 2005). If the R^2 value is 1, then all of the variability is explained by the independent time-series, the opposite is true if the R^2 value is 0. Whereas the R -value, referred to as Pearson's coefficient of correlation, estimates the strength of the linear relationship between the simulated or predicted value and the measured groundwater levels (Hyndman & Athanasopoulos, 2018).

An R^2 -value measures how well a simulated time-series performs, in terms of fit, against the historical time-series. This tool is appropriate to use for estimating the success of infilling (Hyndman & Athanasopoulos, 2018). The addition of independent time-series' will generally increase the R^2 -value automatically, irrespective of whether the independent time-series provides cause improvements. To combat this phenomenon, R_{adj}^2 can be utilized instead, which will not increase solely based on the inclusion of additional independent time-series. The TFN model yields an R_{adj}^2 , for each simulation it produces. When evaluating the simulations R_{adj}^2 -values were used as well as R -values that were manually calculated using Microsoft Excel. As R_{adj}^2 provides information of how well the simulations fits historical data, with consideration to the amount of independent time-series, this was an appropriate goodness-of-fit measure to use for evaluating the simulated time-series. It should be noted that squaring the manually calculated R -values will provide R^2 , which was used in some capacity as a complement to the R -values. The Pastas' TFN model offers additional predictability or goodness-of-fit measures, such as Akaike information criterion, Bayesian information criterion and Root mean square error, however, these were not used to evaluate the simulations.

3.4.3 Assessment of Data Requirements

Both the simulated time-series' and the TFN model were evaluated to assess their accuracy and to determine the requirements of the model in terms of quality and

quantity of data. A sensitivity analysis was utilized to estimate how sensitive the model was to changes in quantity and length of data. The recharged based simulations were evaluated differently and more extensively than the recharge and pumping based simulations. This was done as the former, which includes only the essential independent time-series, is the foundational basis of the model whereas pumping is included as an additional component that may further improve the simulations. See Figure 3.7 for a visual representation of the adopted methodology.

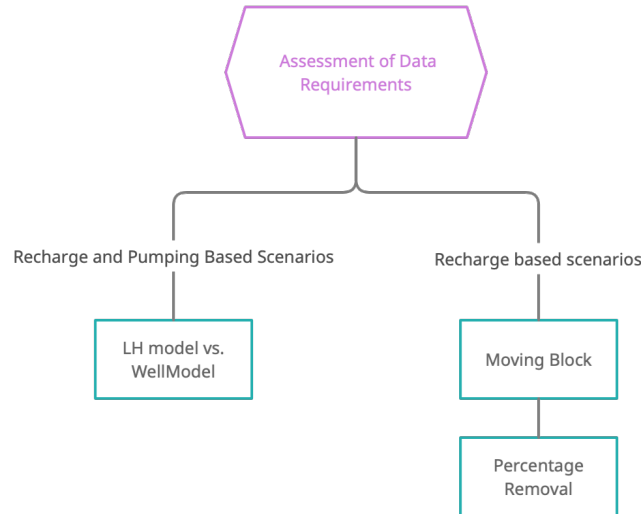


Figure 3.7: Flow chart to illustrate the methodological approach applied for the assessment of data requirements.

For recharge based simulations, a two-pronged approach was adopted to assess the sensitivity of the model in terms of the used quality and quantity of recorded groundwater data. The first approach was based on using a shorter calibration period, or block, to simulate a consecutive year and progressively moving the block through the dependent time-series'. Two blocks, consisting of two and seven years respectively, were used to create two separate simulated time-series. The principle consisted of using two years of recorded data to simulate the next consecutive year, i.e. data for 1980 and 1981 would be calibrated by the TFN model and used to simulate, or extrapolate, the time-series for the entire year of 1982. Next, the recorded data for 1981 and 1982 was used, to simulate a time-series for 1983. This process was repeated for each consecutive year culminating in a year-to-year simulated time-series created by a moving block, meaning the time-series consisted of individually simulated years. The same process was repeated for a seven year moving block. The simulated time-series for this block was shorter, starting from 1987, as the years 1980 through 1986 were used for the calibration. This approach was applied to two bores that contained large amount of semi-consistent measurements as well as daily recorded data. The resulting time-series were compared to each other as well as the dependent time-series, R^2 -values and R -values were manually derived for each year to determine how the simulation varies in accuracy when a shorter calibration period is used for a simulation.

The second approach consisted of systematically removing a certain percentage of the automatically recorded groundwater data in order to determine to what degree the reduction of data quantity affects the accuracy of the overall simulated time-series. Only the automatically recorded daily data was removed, between 2000 to 2020, the manually recorded data was left untouched. Removing the already limited amount of manually recorded data would result in a time-series with almost no recorded groundwater levels for the years 1980 through 2000, which was not desirable. This approach is meant to examine the effects a few years of additional daily data can have on the overall accuracy of a simulation. In Table 1, the percentages and the resulting approximate measurements per month is presented.

Table 1: Percentage of Original Daily Data Remaining with a Corresponding Value for the Resulting Measurements per Month.

Percentage Remaining	Approximate Measurements per Month
100%	30-31
75%	22-23
50%	15-16
25%	7-8
5%	1-2

As can be seen from Table 1, this approach resulted in five simulations, out of which four contained a reduction in data quantity. For context, the approximate number of remaining measurements per month is also provided in the table. When 95% of the data is removed, the measurement frequency resembles that of the manually recorded data during the 1980s and 1990s, meaning approximately 1-2 measurements per month.

Simulated time-series belonging to bores disturbed by an external influence and/or containing large intervals with no measurements were analysed based on visual fit and goodness-of-fit measure. The simulations for these bores were produced to evaluate how well the model performs when presented with time-series' of such conditions. Only the simulation scenario for each bore that achieved the highest R_{adj}^2 -value was visually presented in a graph.

When evaluating the recharge and pumping based simulations, the StressModels, i.e. LH model and WellModel, for pumping were assessed, these simulate the influence of groundwater extraction on the groundwater levels through the use of a Hantush IRF. As there were 11 pumps that could potentially influence the groundwater levels, diagnostic tests were used to determine which pumps each StressModels used as influences. This was done for the purpose of determining which pumps had an impact on the time-series, to what degree and if the influence of the included pumps can be considered reasonable.

As the LH model does not account for distances between the pumps and the bores, pumps determined by the diagnostic test to be of influence were manually removed in succession before simulating the time-series'. This was done to assess if the influences of the pumps were weighed reasonably by the LH Model and what

impact the removal of pumps had on the simulation. Only bores in proximity to the pumps were included in this analysis.

3.4.4 Alternative Interpolation Methods

To further assess the performance of the TFN model, the simulated time-series were compared to interpolated time-series created by two additional methods; Nearest Neighbour (NN) and Cubic Spline (CS). Neither methods are algorithmically based and are therefore less complex than machine learning methods and/or models. Both methods only use groundwater time-series to produce the interpolations without incorporating any additional data. NN was chosen as a method for comparison due to it being an exceedingly simple and straightforward interpolation method. CS was chosen as it is a piece-wise polynomial; appropriate for interpolating time-series of oscillating nature and suitable for non-linear data.

Interpolation with NN consisted of using the initial value in the dependent time-series continuously until the next consecutive recorded value in the time-series appeared, after which that value was used instead. This process was repeated for every gap, in order to achieve an interpolated time-series with a daily time-step. Whereas, CS interpolates between the recorded values using a piece-wise cubic spline function. The curve used a first and second derivative to remain smooth and continuous across its curvature. This process was also repeated for every new sets of values introduced in the dependent time-series. Time-series interpolated using NN were created with Jupyter Notebook whilst the CS interpolations were created using a function in Microsoft Excel. These interpolation methods were used on time-series' that contained a large amount of daily recorded data. In order to analyse the performance of all three methods relative to each other, only the period during which there is daily recorded data was used. During this period, 95% of the recorded daily data was removed, for the purpose of creating artificial gaps in the time-series that could be interpolated. Thereafter, these were evaluated against the complete dependent time-series. The estimated time-series by the TFN model, NN and CS were compared using the goodness-of-fit measurements; R and R^2 .

To estimate how well CS and NN interpolated for larger intervals of missing data, artificial gaps were manually created in a time-series containing three years of daily recorded data. The intervals consisted of numerous variations in gap length, resulting in every 2, 4, 6 or 8 weeks, respectively, being manually removed. The same process was repeated for a TFN simulation. The altered time-series were interpolated and simulated to thereafter be compared with the dependent time-series to study how the increasingly larger gaps change the goodness-of-fit, both visually and by the R^2 -value. As a complement to this analysis, groundwater regimes were produced through seasonal averaging. The interpolated time-series were averaged every day for every year, yielding a time-series that contained averaged daily values for one year. This was followed by creating weekly time-steps, by averaging the daily values per week, resulting in a time-series with weekly indices. The same process was repeated for a TFN simulation and the dependent time-series.

4 Results

This chapter covers the results of the data processing and in detail introducing the bores chosen to be simulated. Thereafter, the TFN simulations and corresponding evaluations are presented. The chapter concludes with a comparison between a TFN simulation and time-series interpolated using alternative methods.

4.1 Bores and Climatic Data

Prior to producing the simulated time-series using the TFN model, a total of 57 boreholes were evaluated. During the evaluation it was determined that the manual measurement frequency for most of the observation data during the 1980-2000 varied between 1-3 times a month. As previously mentioned, the measurement frequency for the majority of boreholes decreased during the 1990s and by mid-2000 all measuring had stopped completely. It was resumed for a handful of bores during the mid-2010 using an automatic measuring instrument which logged the water levels hourly. Thus, all time-series - with the exception of those belonging to new bores drilled during the 2010s have, at the very least, a period 10 years with no measurements. For the majority of these boreholes this period can vary between 10-30 years.

In Figure 4.8 the time-series' containing observation data from three boreholes are presented for the purpose of highlighting the large variation in data quality.

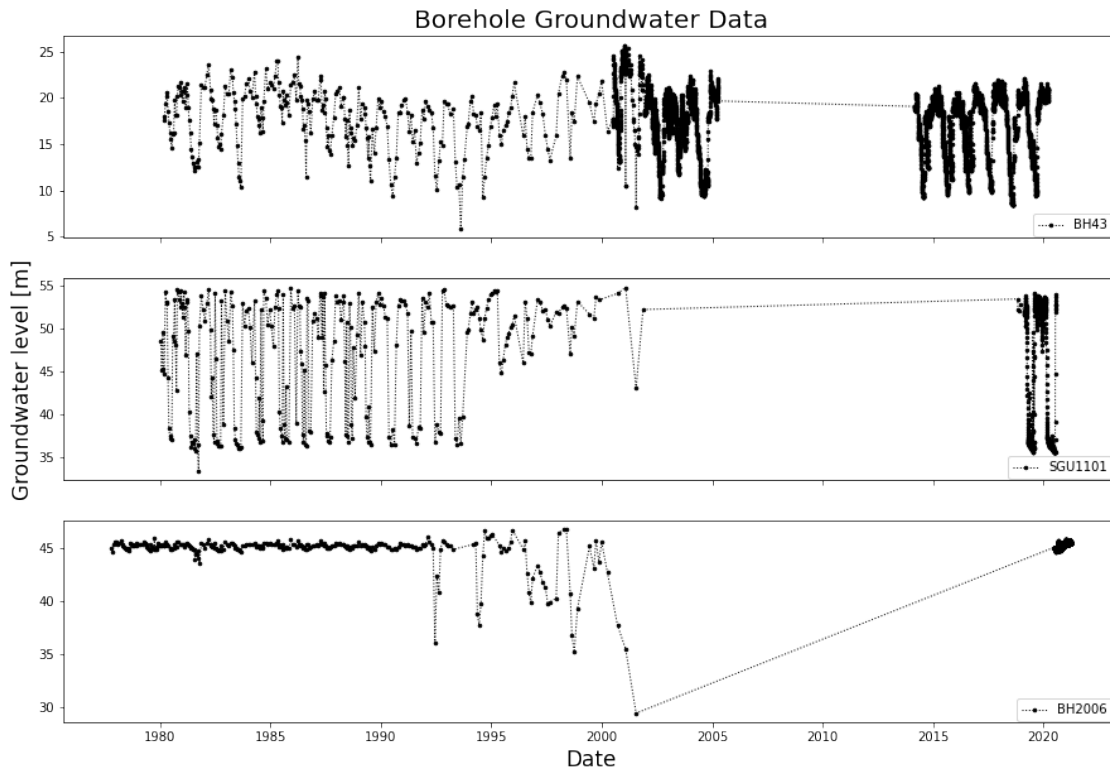


Figure 4.8: Three time-series with groundwater observations from Gotland, for a) BH43, b) SGU11001 and c) BH2006. The dark points represent recorded observations.

BH43, Figure 4.8 a), is one of the few bores that have been manually measured consistently from the late 1970s until 2000 and automatically measured from 2000 to mid-2005. After a period of inactivity, the measurements were resumed again using an automatic measuring device from 2015 and on-wards. Bore SGU11001 in contrast, presented in Figure 4.8 b), contains a long period of inactivity and presents a steadily decreasing measurement frequency from 1995 and onward. This bore also happens to be one of few boreholes that is both upwards and downwards bounded.

A number of time-series from the area of File Hajdar were either upper and/or lower bounded. Whether a time-series is bounded depends on numerous factors, for instance if a time-series is bounded upwards it can be the result of pressures build-up in the borehole during rainier periods which causes the groundwater to overflow and spill onto the surface. If a bore is lower bounded, it could be due to the short length of the borehole which doesn't penetrate the aquifer deep enough. This can be problematic as the time-series will not accurately capture the changes in groundwater levels.

Both SGU11001 and BH43 present the large variations in groundwater levels that are characteristic of File Hajdar. The groundwater levels are higher during the precipitation, or recharge, heavy months of the year and lower during the warmer

months when ET_o is high and precipitation is low.

Finally, the last bore presented is BH2006, see Figure 4.8 c), which contains both a long period of missing data and abnormal changes in groundwater levels caused by an external disturbance - a tin can. Not only do the time-series vary in measurement frequency and quality, the time-series also exhibit irregular variations, unrelated to seasonality or cyclic variations, in the overall trend, see Figure 4.8 a) for instance.

Out of the 57 boreholes evaluated nine were chosen to be interpolated including the three bores presented in the figure above. The choice of bores were based primarily on the length, completeness and quality of data. For evaluation purposes time-series containing shorter length, disruptions and poorer quality were included. It is important to note that for relevant bores, automatically measured hourly data during the 2000s was averaged to acquire daily recorded data, a table with information on which time-series contain such data can be found in Appendix A.

Each bore's defining characteristics are presented in Table 2 and the dependent time-series plots for all nine bores can be found in Appendix D along with information specific to each time-series. As an official site visit was not possible, aerial images and information gained from official reports were used to gain insight into the area surrounding the bores in terms of vegetation, elevation, distances between bores and bodies of water, these can be found in Appendix A.

In Figure 4.9 a map with each bore's relative location is presented, bores with a lower elevation in comparison to the surrounding area are distinguished by a purple color, whereas bores located on approximately the same elevation as the surrounding area are designated a blue color. Pumps are denoted by a green color and as the location of the seven municipal pumps are classified the large green square represents their approximate location. Each bores' radial distances to the pumps can be found in Appendix A, it should be noted that the location of the commercial pumps' are approximate.

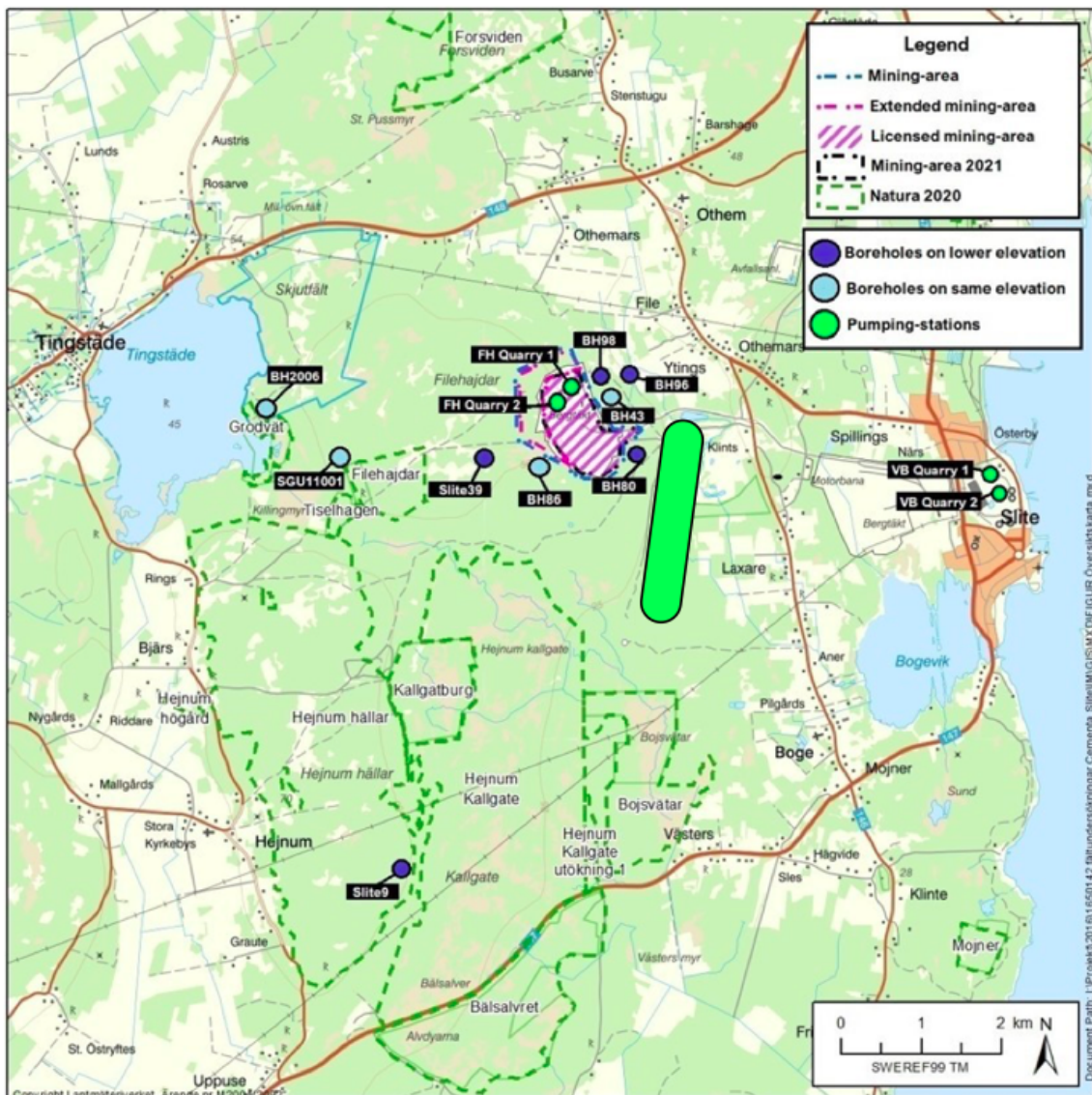


Figure 4.9: A map of File Hajdar including approximate location of all commercial and municipal points of groundwater extraction, as well as and the location of the nine bores evaluated in this thesis.

Table 2: A table presenting the individual characteristics of each borehole.

Borehole	Characteristics
BH43	Found to the northeast of the File Hajdar quarry on a relatively flat surface. Aerial images show sparse vegetation surrounding the bore, with arable land in close proximity. The borehole does not neighbor any large bodies of water.
BH86	Found just southwest of the File Hajdar quarry, this bore appears to be surrounded by a small amount of vegetation. It has no large bodies of water in close proximity and it is situated on a relatively flat surface.
BH98	Vegetation is determined to be sparse in the area where BH98 is located and the limestone surrounding the bore is exposed. It located in the middle of a slight decline and is close to the File Hajdar quarry and arable land. The bore does not neighbor any large body of water.
BH96	Located in close proximity to both the File Hajdar quarry and the municipal pumps, this bore is found to be on land with sparser vegetation. It is situated near the same arable land as BH43 and appears to be located at the end of a small hill.
BH80	This bore neighbors the File Hajdar quarry and appears to be located in the middle of a steep decline, to the west of the municipal pumps. No large bodies of water can be found near the bore and it is surrounded by sparse vegetation.
Slite 39	Located west of the File Hajdar quarry, the bore appears to be situated in a small pit and is surrounded by sparse vegetation, where the limestone is partly exposed. The bore does not neighbor any large body of water.
Slite 9	This bore is isolated from the rest as it is found at the southern most part of the examined territory. Information about the soil layers show that the layers consist only of limestone SGU (n.d.). It is situated at the end of a steep hill with a significant declination.
SGU11001	The land surrounding the bore is determined to have some vegetation and is located on semi-flat area. It is found close to a large body of water; Lake Tingstäde.
BH2006	This borehole is located in close proximity to a large body of water - Lake Tingsstäde. The soil stratigraphy for the area shows that the first soil layers consist of sand sediment in varying thickness, followed by limestone (SGU, n.d.). This bore is not in close proximity to the quarry or any municipal pumps and is located on a flat surface.

Precipitation and calculated ET_o are presented in Figure 4.10 a) and b), respectively. As mentioned, ET_o is only calculated for 1980 and onwards, although as the TFN model requires ET_o for at least 10 years prior to the first recorded groundwater level, averaged values are calculated and by the model during the calibrations, these

cannot be observed in Figure 4.10 b).

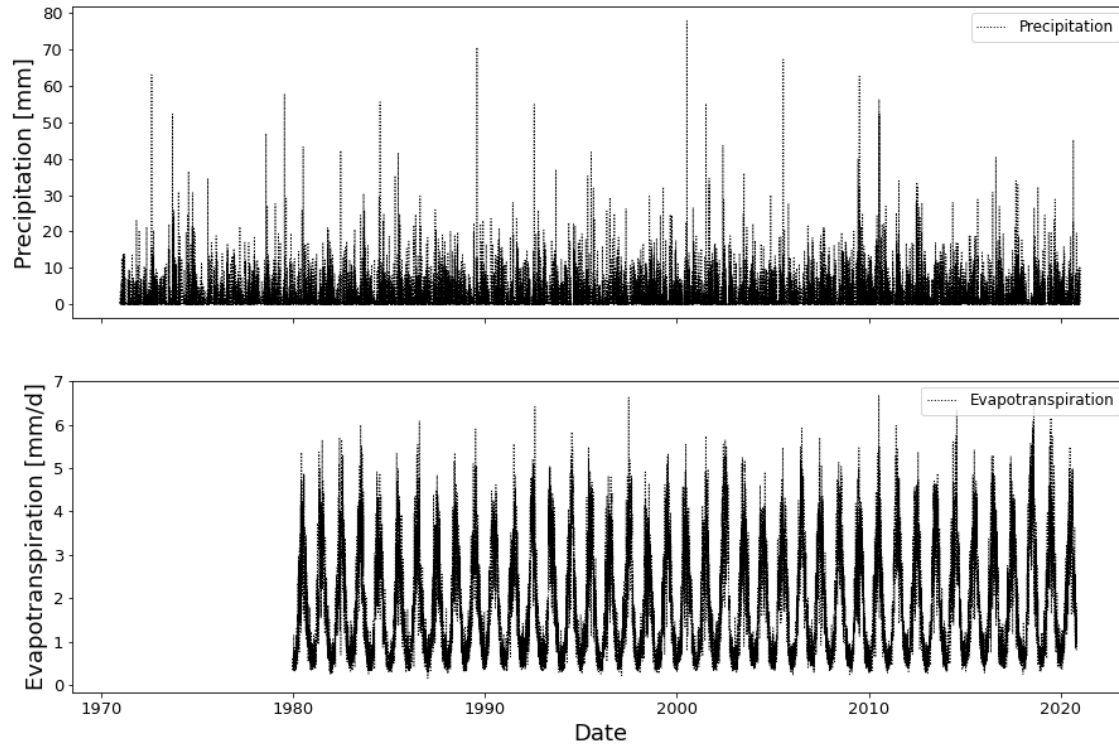


Figure 4.10: In the order of from top to bottom. Graph a) presents the precipitation and b) the calculated ET_o data.

4.2 Simulated Time-series Using Recharge

This sub-chapter presents the results of simulations using only recharge as a stress. Changes in linear trend for shorter periods can be observed in the dependent time-series for a few of the bores, for these a supplementary trendline has been added in the TFN model. The trendlines have been observed to marginally improve the fit, a list of the time-series with an added trendline can be found in Appendix F, which also present the increase in R_{adj}^2 before and after the addition of the trendline for each bore.

The results for BH43 are presented in Figure 4.11, the simulated time-series are coloured red and the original groundwater time-series is black. Four scenarios are simulated, each with a different combination of recharge IRFs and StressModel. The standard AR(1) noise model was used for the simulations.

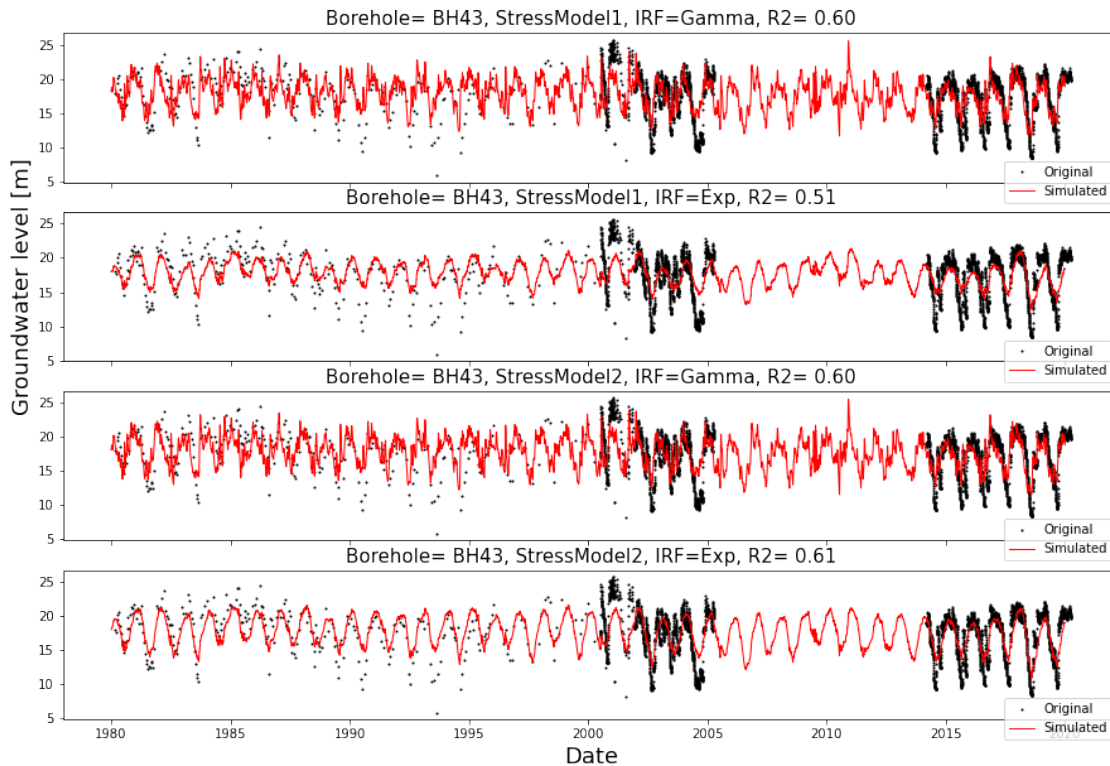


Figure 4.11: Simulated results for BH43 with AR(1) noise model in the order of from top to bottom, with following simulation scenarios a) SM1; IRF Gamma, b) SM1; IRF Exponential, c) SM2; IRF Gamma and d) SM2; IRF Exponential.

As can be observed from Figure 4.11 b), the lowest R_{adj}^2 -value was acquired for the scenario that applies SM1 using the Exponential IRF. In addition to being smoother, the simulation underestimated the groundwater levels and did not manage to reach the large variations in fluctuations that is characteristic of all observation data from File Hajdar. The Exponential IRF provided marginally improved results when applied using SM2, see Figure 4.11 d), but was still smooth in curvature and under-performed. The simulations that used Gamma IRF applied through SM1 and SM2 provided identical results, see Figure 4.11 a) and c), both in terms of visual fit and R_{adj}^2 . Gamma IRF, unlike the Exponential IRF, is not smoothed.

The recharge based simulation scenarios for SGU11001 are presented in Figure 4.12, the dependent time-series for this bore is upper and lower bounded.

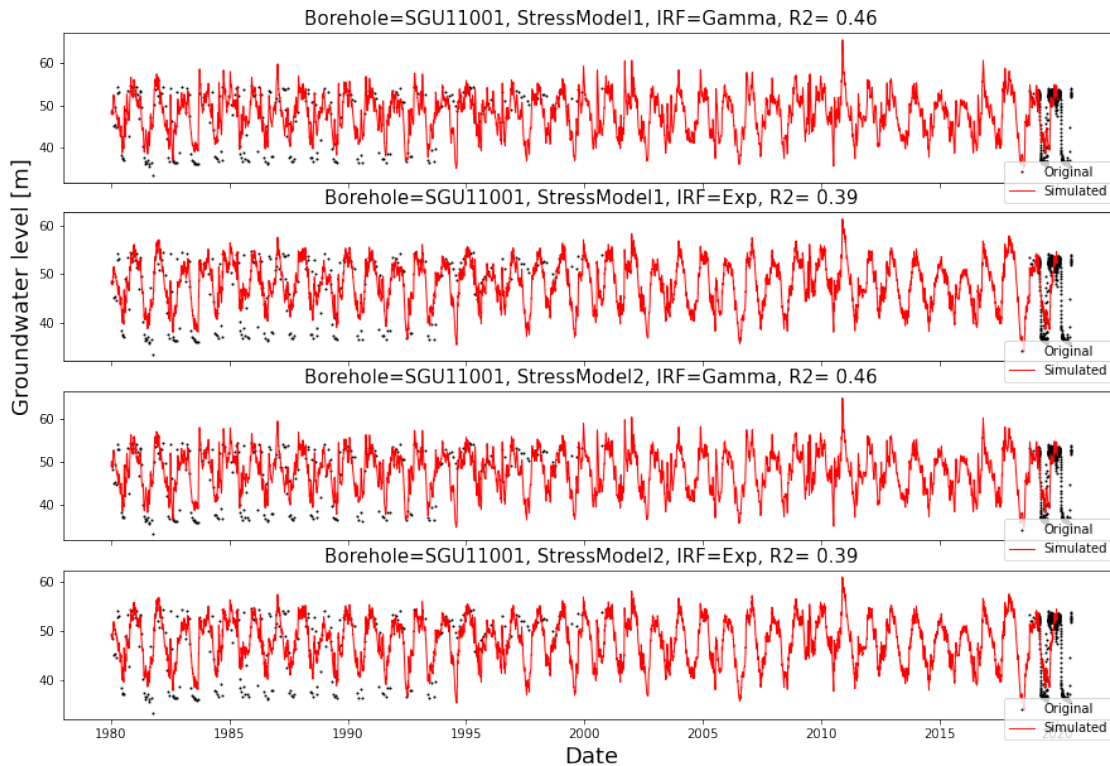


Figure 4.12: Simulated results for SGU11001 with AR(1) noise model in the order of from top to bottom, with following simulation scenarios a) SM1; IRF Gamma, b) SM1; IRF Exponential, c) SM2; IRF Gamma and d) SM2; IRF Exponential.

The simulation for SGU11001 are differ considerably compared to BH43, as the dependent time-series for this bore has both less data and is bounded upwards and downwards. The simulations produced a marginally higher R_{adj}^2 -value using the Gamma IRF, see Figure 4.12 a) and c). However, no difference between the StressModels can be observed as identical R_{adj}^2 -values were acquired for both SM1 and SM2. There was little difference in R_{adj}^2 between all four scenarios with little variation in the visual fit. The TFN model does not consider any boundness that may be present in a time-series, therefore the simulations for this borehole seemingly overestimates the groundwater levels during the precipitation heavy months.

In Figure 4.13 a compiled grouped bar chart of all R_{adj}^2 -values for all four scenarios is presented for each respective borehole, the values can also be found in a table format in Appendix G. Each time-series was studied and evaluated with respect to StressModels, IRF and R_{adj}^2 , similar to those done for BH43 and SGU11001.

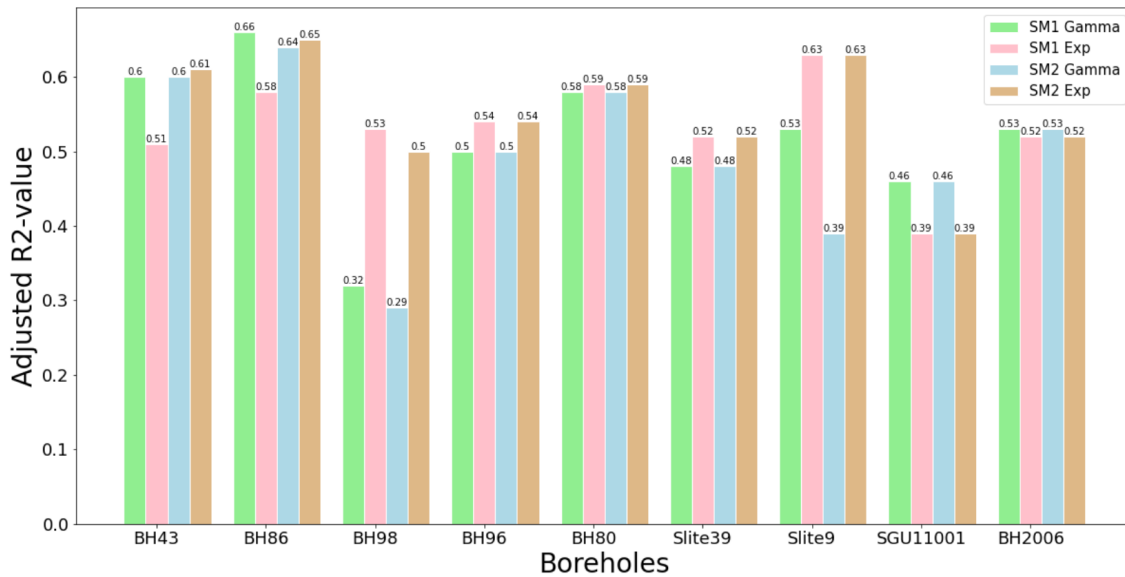


Figure 4.13: R_{adj}^2 for simulated results using SM1 and SM2, which applied either a Gamma IRF or an Exponential IRF, simulated with the AR(1) noise model.

To evaluate the residuals and noise both diagnostic and hypothesis tests were used. The diagnostic tests indicated the presence of auto-correlation in the data. The auto-correlation plot presented spikes indicating auto-correlation. Moreover, a grouped bar chart was utilized to check the distribution of the data residuals, it showed that the residuals were not normally distributed, implying that one of the assumptions of noise is not followed. Hypothesis tests were used as a complement to the diagnostic tests, these showed that the null hypothesis was rejected for all of the time-series with the exception Slite 39. Meaning the residuals are independent, have homogeneity of variance, or normally distributed (Collenteur et al., 2019). The rejection of H_0 , which checks for normal distribution and auto-correlation, corroborates the information acquired from the diagnostic test; the data does not comply with one of the assumption of noise. Therefore, auto-correlation can be assumed to be present in the residuals indicating that the model can be improved. For an example of the diagnostic plots and hypothesis tests, see Appendix H.

As the AR (1) noise model may not function accurately for high frequency data, the recharge based scenarios were simulated without the noise model, resulting in a considerable increase in both visual fit and R_{adj}^2 -values. In Figure 4.14 and Figure 4.15 the simulated time-series for BH43 and SGU11001, respectively, are visualized without the noise model.

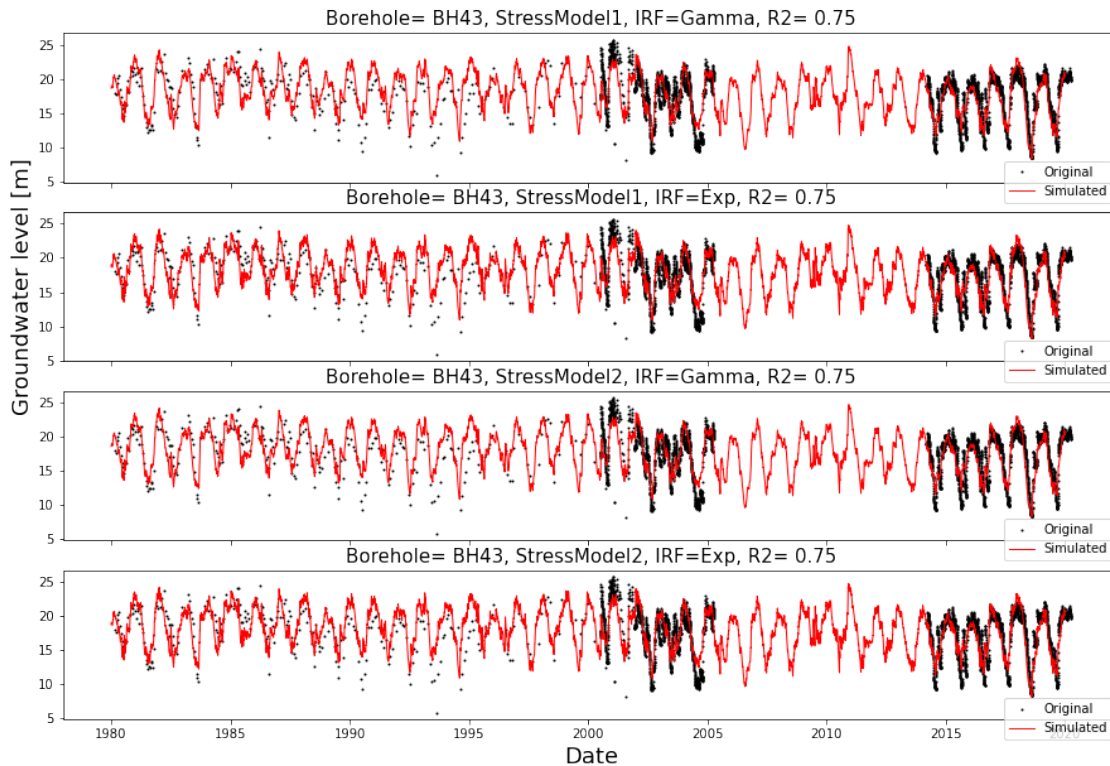


Figure 4.14: Simulated results for BH43 without AR(1) noise model in the order of from top to bottom, with following simulation scenarios a) SM1; IRF Gamma, b) SM1; IRF Exponential, c) SM2; IRF Gamma and d) SM2; IRF Exponential. These simulations were done without any noise model.

As can be seen from graphs a) to d) in Figure 4.14 for BH43, smoothing no longer occurs for any the scenarios. Without the noise model, the simulations manage to get closer to reaching the extremes of the fluctuations.

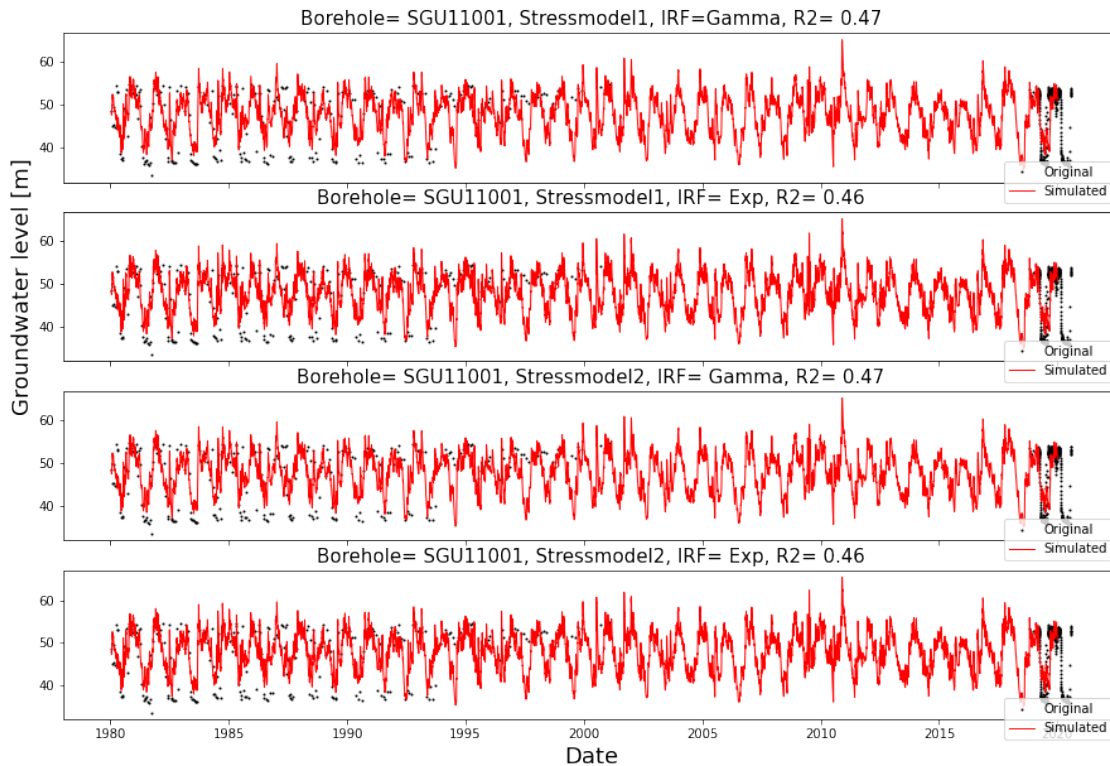


Figure 4.15: Simulated results for SGU11001 without AR(1) noise model in the order of from top to bottom, with following simulation scenarios a) SM1; IRF Gamma, b) SM1; IRF Exponential, c) SM2; IRF Gamma and d) SM2; IRF Exponential.

The simulations for SGU11001 without the noise model did not present an increase in fit equal to that of BH43. The scenarios using Gamma IRF did not show any considerable improvement with the noise model deactivated, see Figure 4.15 a) and c), the simulations calibrated using the Exponential IRF showed a marginal improvement. The visual differences between each scenarios is almost indistinguishable, see Figure 4.15 b) and d). This bore contains a short period of recorded daily data during the late 2010s.

In Figure 4.16 the R_{adj}^2 -values for each simulation scenario for all bores, without the AR(1) noise model, is presented, the values can also be found in table-format, see Appendix G.

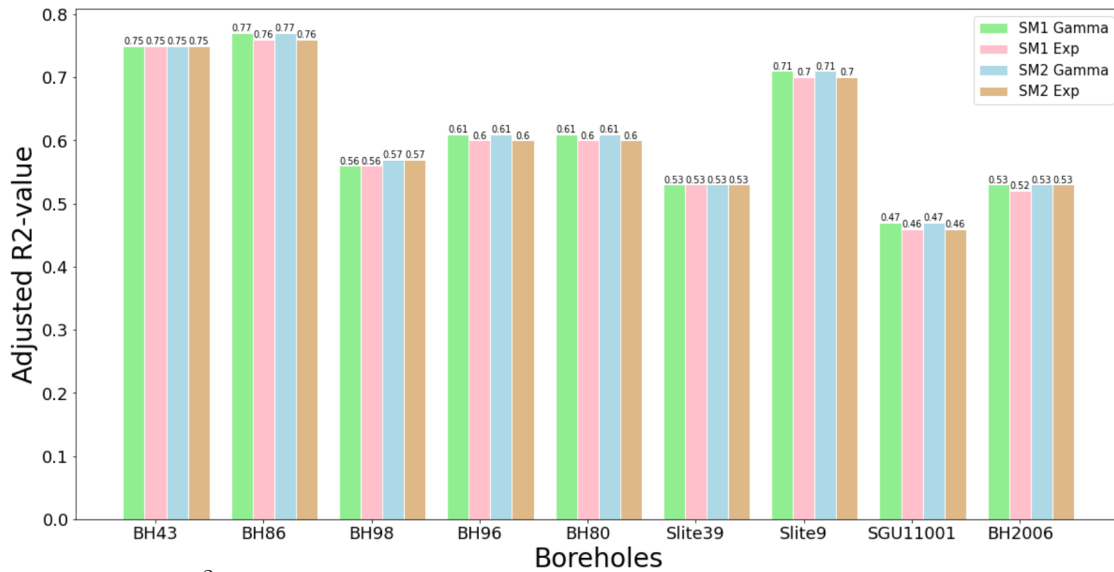


Figure 4.16: R_{adj}^2 for simulated results using SM1 and SM2, which applied either a Gamma IRF or an Exponential IRF, simulated without the AR(1) noise model

Overall, the R_{adj}^2 -values for all bores that contained daily recorded data increased considerably, the difference in R_{adj}^2 for bores without daily recorded data was smaller. Moreover, the difference in R_{adj}^2 -values between both StressModels and both IRFs decreased for each individual bore, with the value being nearly identical for all four scenarios, for most bores.

Residuals and noise were evaluated for the simulations without the AR(1) noise model. The results of the diagnostic tests indicated significant auto-correlation, more so than when simulated with the noise model, the residuals were not normally distributed. The H0 was rejected for all bores. It can therefore be assumed that the data does not comply with one of the assumptions of noise; leading to the conclusion that the simulation model can be improved. For an example of the diagnostic plots and hypothesis tests, see Appendix H.

4.2.1 Time-series Containing Poor Data Quality and/or Disturbances

The simulations presented in this sub-chapter belong to bores that contain external disturbances or large gaps which may affect the interpolation. The time-series graphs presented in this sub-chapter are simulated using the StressModel and IRF scenario that achieved the highest R_{adj}^2 -value, see Figure 4.16 in previous sub-chapter. The simulations were calibrated without the noise model for all time-series in this study, unless stated otherwise.

In Figure 4.17 the simulations for Slite 9 and BH96 are presented.

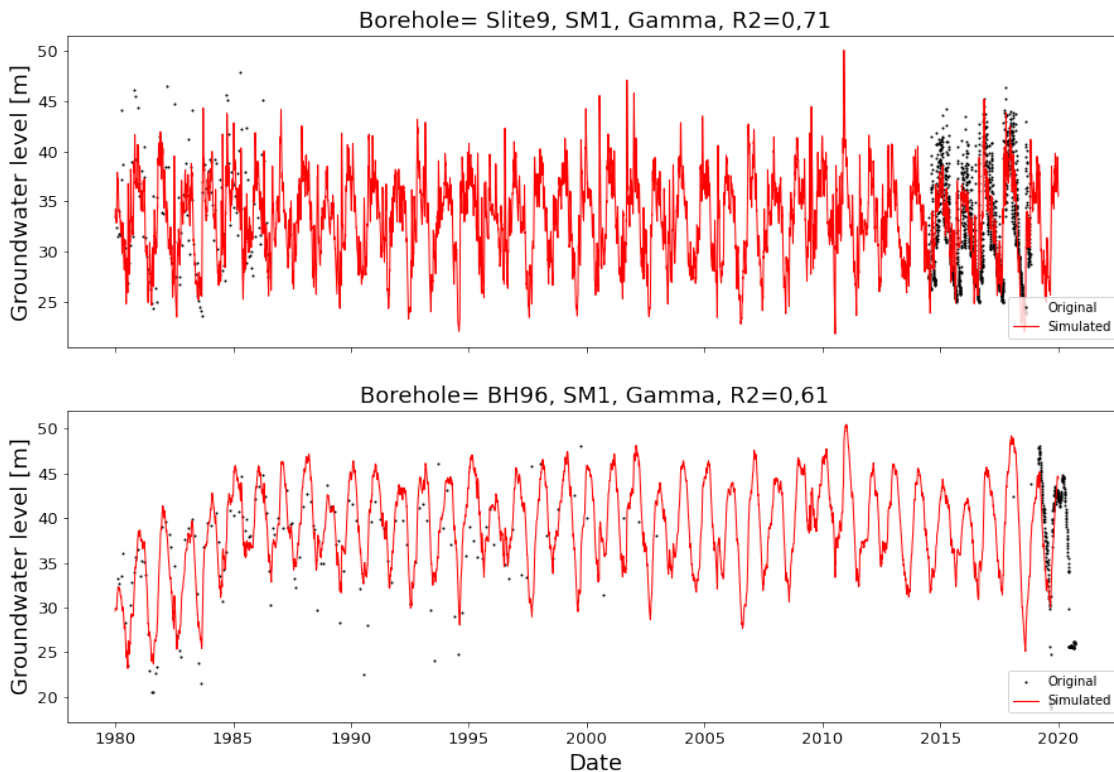


Figure 4.17: Simulated results in the order of from top to bottom, a) Slite 9 and b) BH96, contains a large gap of years with no measurements.

The original time-series for the bores contain large periods with no recorded measurements. For Slite 9, see Figure 4.17 a), groundwater levels ceased being recorded in 1986 and were resumed in 2014, totaling in a gap of 28 years with no recorded groundwater levels. It is apparent that the model underestimates the groundwater levels during months where recharge occurs at a higher rate, this is true for the years between 1980 to 1986. The same can be concluded for the simulated values produced after 2014, despite the dependent time-series consisting entirely of daily recorded data for this time period. The model simulates the groundwater levels during the dry season with a higher accuracy than compared to during wet season. This bore is quite isolated in its location and isn't in close proximity to any points of groundwater extraction. The simulated values for the gap are entirely produced by the independent time-series and general fluctuation patterns of the years before and after the gap.

A different pattern can be detected for BH96, Figure 4.17 b), for which the dependent time-series contains a period with no measurements from 2004 until 2018. During 2000 through 2004, there are only four instances of recorded groundwater levels, with the measurement frequency being once a year for this period and for the year 2018, there are only two instances of recorded measurements. The resulting gap is therefore approximately 14 years, with the measurement frequency at the

beginning and end of the gap being lower than what is usual for this bore. From 2019 and onwards the dependent time-series contains daily recorded data. A linear trend is manually added between 1980 to 1986, as an obvious upward linear trend could be observed during these years. In general, the simulations for this time-series consistently overestimate the groundwater levels during the wet and dry seasons. Additionally, a vague U-shaped can be detected in the dependent time-series between 1987 and 2000, which the model does not capture and cannot be manually corrected for. Though the simulations for both bores use the same modeling scenario, SM1 with Gamma IRF, the simulated time-series for bore BH96 is smoothed out in comparison to the time-series for Slite 9.

In Figure 4.18 the simulated time-series for BH2006 and Slite 39 are presented.

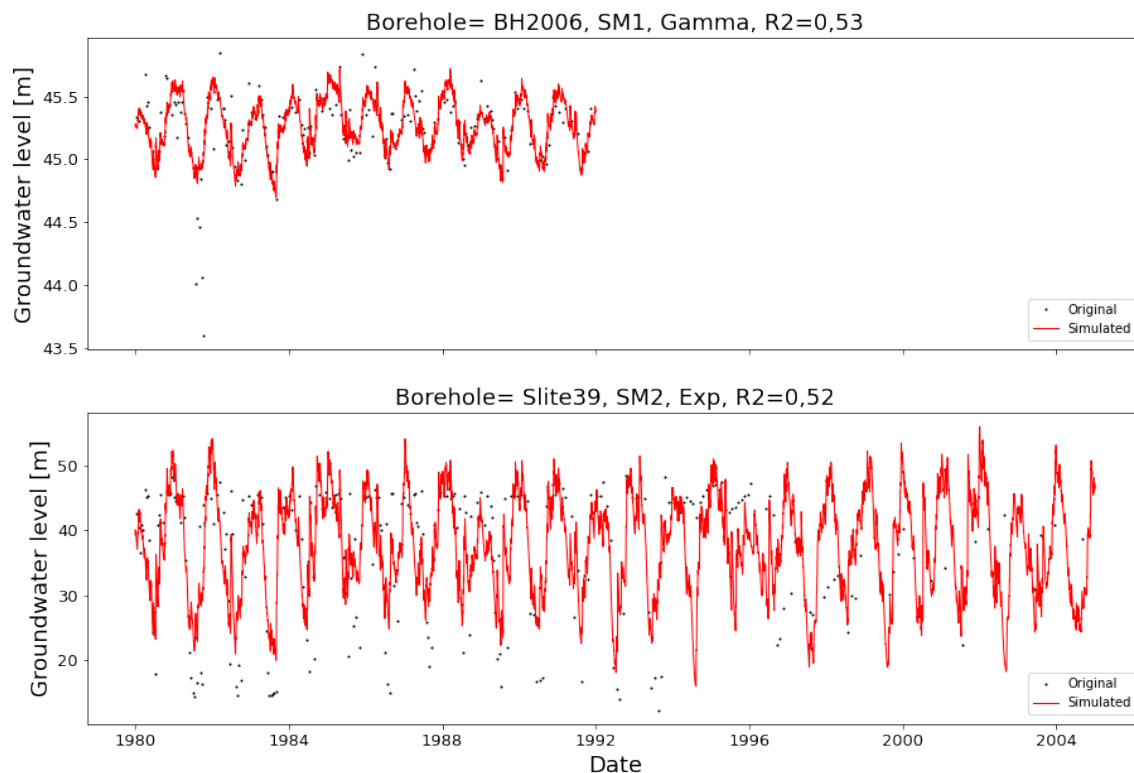


Figure 4.18: Simulated results in the order of from top to bottom, a) BH2006 and b) Slite 39, contains a disturbance and has few measurements.

The simulated time-series for BH2006, Figure 4.18 a), is distinguished as being affected by an external disturbance, whereas Slite 39, see Figure 4.18 b), has an upper boundary. Moreover, the dependent time-series for both bores contain fewer years of recorded data, compared to other bores presented in this thesis. The simulated time-series for both bores are created with the noise model activated as these bores do not contain daily recorded data. The dependent time-series for bore BH2006 was disturbed due to a blockage, the exact year in which the blockage occurred cannot be

determined. Therefore, recorded values between 1992 and early 2000 were removed prior to the simulation, observe the differences between Figure 4.8 c) and Figure 4.18 a), for changes made in the dependent time-series. The logging of groundwater levels was resumed for this bore in late 2020, however, these groundwater levels were not included in the simulations as not enough data had been gathered at the time of calibration. This bore is located close to the lake, the discharge to and from it could explain the shorter yearly fluctuations in groundwater levels. The groundwater levels for BH2006 fluctuate less than a meter during each year, which is unusual for the area of File Hajdar. The dependent time-series for it contains only 12 years of recorded data, but despite the short time span, the simulation manages to capture the fluctuations well.

Groundwater levels for Slite 39 were only recorded from 1980 until 2004, see Figure 4.18 b), with the measurements becoming increasingly sporadic after 2000. The logging of groundwater levels for this bore has not been resumed, thus the dependent time-series contain only manually collected measurements over a 24 year period. The simulation for this bore exhibit similar tendencies as the one for SGU11001, another bounded bore. The model overestimates the groundwater levels during the wet season, due to the boundary, and underestimates the groundwater levels during the dry season.

Lastly, Figure 4.19 presents the time-series for borehole BH98, where the groundwater levels were recorded starting 1989.

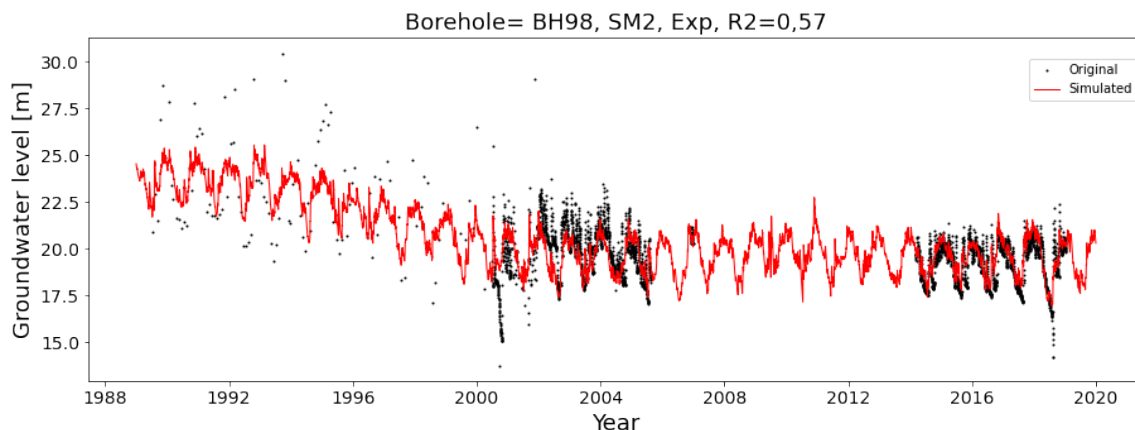


Figure 4.19: Simulated results for BH98, measurements begin in late 80's.

BH98 exhibits a clear downward trend from 1989 to 2000, a manual linear trend-line was therefore added to the model between these years. After the year 2000, no discernible change in the general trend can be identified. There is a distinct change in the range, or scope, of the fluctuations in the dependent time-series after the year 2000. The manually recorded groundwater levels prior to 2000 demonstrate larger fluctuations of more than 10 meters, while the automatically recorded daily groundwater levels after 2000 fluctuate approximately five meters between seasons. The shift in both trend and fluctuation range could be attributed to the bores close proximity to the File Hajdar quarry, their operations could have impacted the net-

work of fractures and layers lowering the groundwater levels. The model does not capture the large variations in fluctuations between 1986-2000, hence the less accurate visual fit and lower R_{adj}^2 -value. The model consistently underestimates and/or overestimates the groundwater levels between seasons.

4.2.2 Evaluation Through a Moving Calibration Block

The time-series from File Hajdar vary drastically in data quality and quantity. In general, the data is heterogeneous in quantity, meaning it varies in frequency over the years and doesn't have regular, or equidistant, time-steps. The "moving block" approach was applied to evaluate the data requirements of the TFN model. As the dependent time-series for both BH43 and BH86 contain large amounts of data consisting of manually and automatically measured groundwater levels, these were chosen to be evaluated. The moving block evaluations for BH86 can be found in Appendix I. The simulation uses the scenario that achieved the highest R_{adj}^2 without the noise model; Gamma IRF with SM1 for both bores, see Figure 4.13. The graphs presented for BH43, in Figure 4.20, contains a simulation created using a moving block of two years.

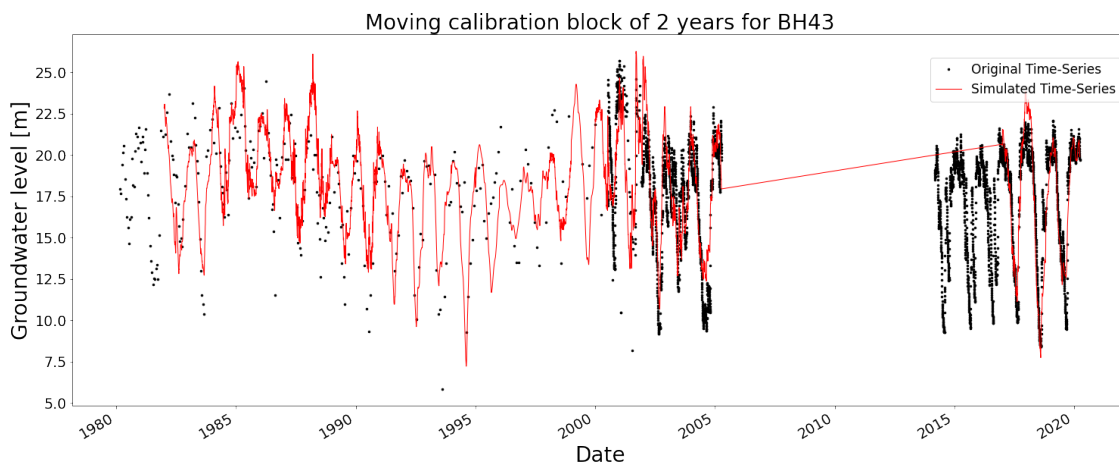


Figure 4.20: Assessment of BH43: A moving calibration block of 2 years

Using only two years for calibration, the model managed to produce simulated time-series that were visually semi-accurate. This is particularly true for the years prior to 2000, which only have manually recorded data, with a measurement frequency of less than three times a month. The simulation also manages to capture the irregular progression of the time-series well, see the almost U-shaped irregularity between 1986 to 2000. It should be noted that as the moving block approach essentially extrapolates the groundwater levels for a forthcoming year based on two years of sporadic measurements, extrapolating for a longer period, such as four or five years, will likely result in a poorer simulations. Moreover, the recorded groundwater

data for the year being modelled is not considered in the calibration. The year-to-year simulations are thus wholly based on the meteorological data, the observed trend and two years prior to the simulation year.

The same evaluation has been done using a 7 year calibration period, the results can be observed in Figure 4.21.

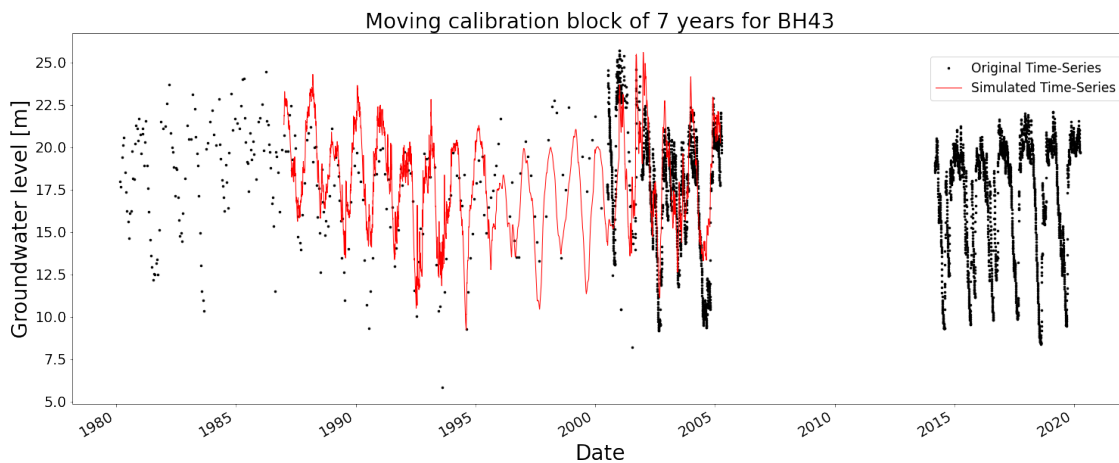


Figure 4.21: Assessment of BH43: A moving calibration block of 7 years

The seven year moving block simulation in Figure 4.21, has a shorter simulation span, from 1987 to 2005. This is due to calibration block being too long to be applied to the initial seven years of the dependent time-series as well as the final years, namely 2015-2020. Contrary to the two year moving block simulation, this simulated time-series does not capture the irregular variations as well, observe the differences between Figure 4.20 and Figure 4.21 during the years 1986-2000.

The year-to-year R -values for the two and seven year moving block simulations can be found in Figure 4.22 in a comparative plot.

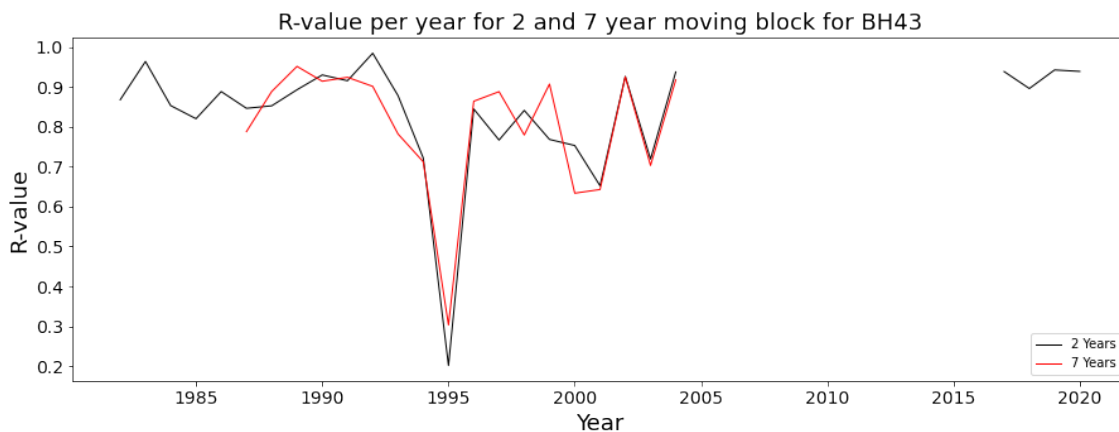


Figure 4.22: The year-to-year Pearson coefficient of correlation for two and seven year moving block simulations BH43

Large differences in R -values cannot be observed between the two blocks containing two and seven year calibration periods, with the latter proving to be almost as effective as the former. When using two-year calibration for the simulations, the performance of the models decreased considerably in between 1995 and 1999, with the lowest goodness-of-fit values acquired for the year 1995, as can be evidenced from the R -values in Figure 4.22. Despite the shorter simulation span the two year moving block show a marked improvement in terms of R -values in Figure 4.22, especially between 1995 to 2000, during which the measurement frequency for the original data decreased considerably. The measurement frequency for BH43 during 1994 and 1995 was 12 times, respectively, for 1996 and 1997 groundwater levels were measured seven times, respectively, and for 1998 eight measurements were taken. This explains the considerable decrease in fit for the two year calibration block during 1995 through 1999. For majority of the years, the simulation overestimates the groundwater levels. The year-to-year simulations during which the moving block contained only daily data showed a marked improvement for both bores. The R^2 -values are presented in Appendix I.

Figure 4.23 presents a comparative plot containing the two and seven year moving block simulations, the dependant time-series as well as the complete TFN simulation for BH43 presented in Figure 4.14a).

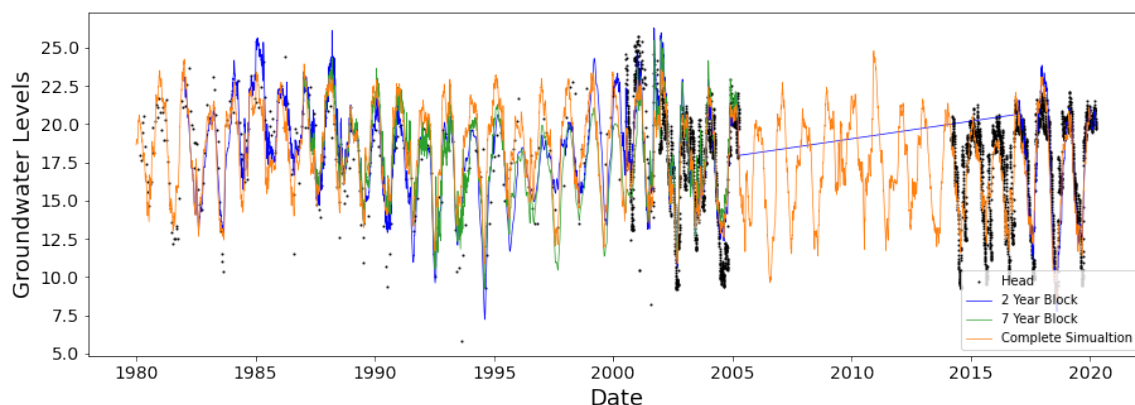


Figure 4.23: Comparison of two and seven year moving block simulations, the dependant time-series and a "complete" simulation.

It can be observed that the two and seven year moving block simulation capture the irregular variations in the dependant time-series with a higher accuracy than the complete simulation, with the two year simulation being superior.

4.2.3 Evaluation Through the Removal of Automatic Measurements

The two bores used for this sensitivity analysis were BH43 and BH86, the evaluation for BH86 can be found in Appendix I. The Exponential IRF in combination with SM2 was used for the BH43 simulations as this simulation scenario achieved the highest

R_{adj}^2 -value without noise model (see Figure 4.16). In Figure 4.24 five simulated time-series are presented for BH43.

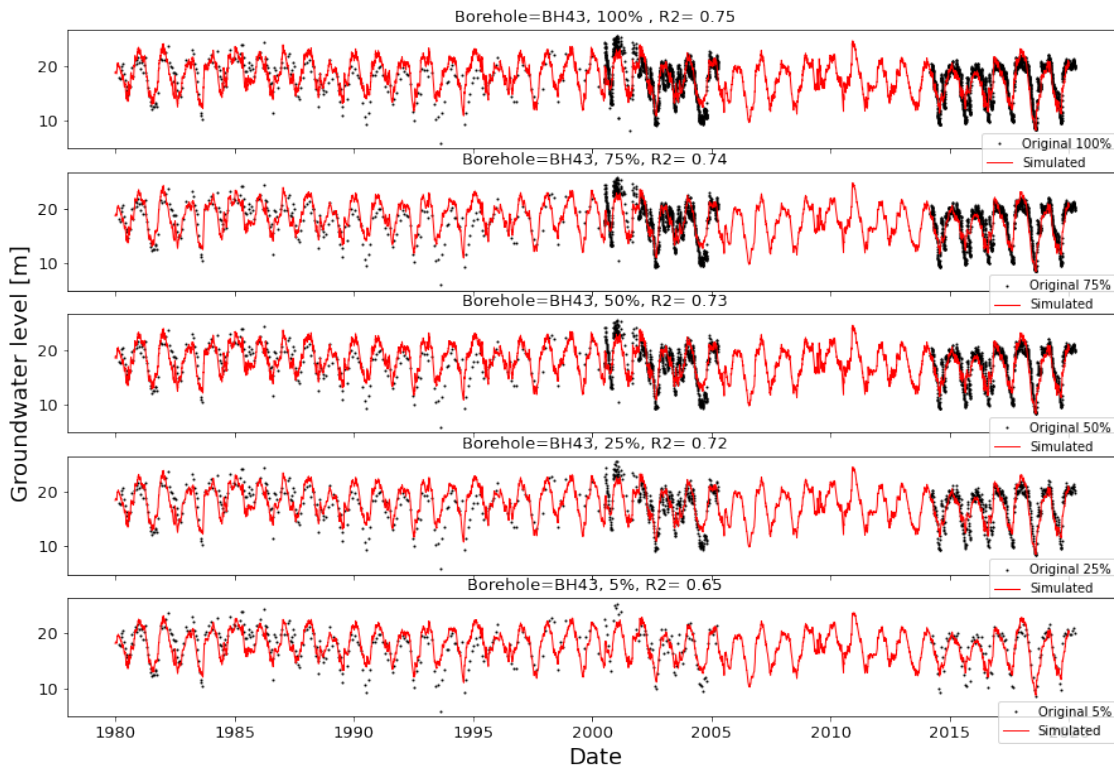


Figure 4.24: Systematic percentile removal of automatic measurements for BH43. In the order of from top to bottom, a) 100%, b) 75%, c) 50%, d) 25% and e) 5%

As can be seen from the plot, Figure 4.24 d), c) and b), the data removal of 25% 50% and 75% respectively, resulted in no considerable decrease in R_{adj}^2 -values. When only 25% of the automatic measurements are used for the simulation, which constitutes around 7-8 recorded measurements per month, the R_{adj}^2 -value decreased with only 0.03 units. In comparison, the drop in R_{adj}^2 -value was larger when 95% of the automatically recorded groundwater levels were removed, Figure 4.24 e), the R_{adj}^2 -value decreased considerably from 0.75 to 0.65. The remaining 5% of the automatic measurements represent around 1-2 measurements per month, which resembles the measurement frequency of the manual measurements between 1980 to 1995. By scrutinizing all five graphs, it is apparent that the model's ability to successfully capture the extreme data points decreases as the measurements are removed, furthermore, the decrease in daily data has a detrimental effect on the simulation as a whole, even between 2005 and 2015, during which there is no recorded data.

To further assess the difference in fit between the five simulated time-series, the correlation, R , between each of the simulated time-series and the original dependent time-series was derived. An R -value was derived for each individual year, for every

simulated time-series and can be seen in Figure 4.25.

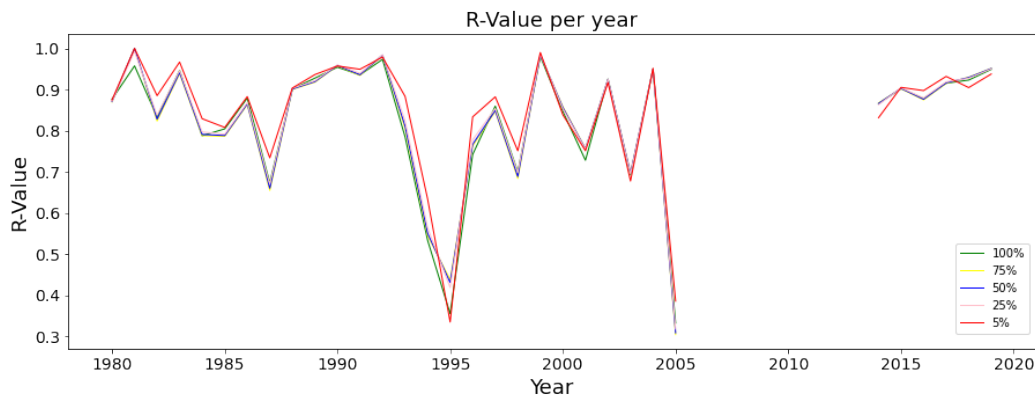


Figure 4.25: BH43: Correlation Between Simulated and Recorded Groundwater Levels For Each Individual Year

For the years 1980-2000, the correlation between the manually recorded groundwater levels and the simulated values seem to increase as the percentage of automatically recorded daily groundwater levels decrease. From the year 2000 and onwards the correlation between the recorded groundwater levels and each individual simulations show little difference. Noticeably, a decrease in automatic daily measurements present a larger difference for the years with manual measurements than the years with automatic measurements. The graph for BH86 can be seen in Appendix J along with the calculated R -value. The R^2 -value for both bores are also presented in Appendix J.

4.3 Simulated Time-series Using Recharge and Pumping

To further improve the simulations, commercial and municipal pumping were added as stresses, these were calibrated using both LH model and WellModel. To reiterate, the LH model loops the independent time-series' containing pumping data to fit the dependent time-series while the WellModel instead considers the effects of each pump based on its relative distance to the borehole being examined. Only time-series belonging to bores in close proximity to the municipal and/or commercial pumps are presented in this chapter, thus bores such as BH2006 and SGU11001 are not included. A comprehensive list of the distances between each borehole and pump can be found in Appendix A. Furthermore, as the bores included in this sub-chapter all contain high frequency data in some capacity, only time-series calibrated without the noise model are presented.

Both recharge and pumping are applied as stresses in the TFN model, resulting in eight simulation scenarios. StressModels and IRFs for recharge are calibrated in conjunction with StressModels and IRF for pumping, for each borehole, however, only the scenario with the highest R_{adj}^2 -value are presented Figure 4.26.

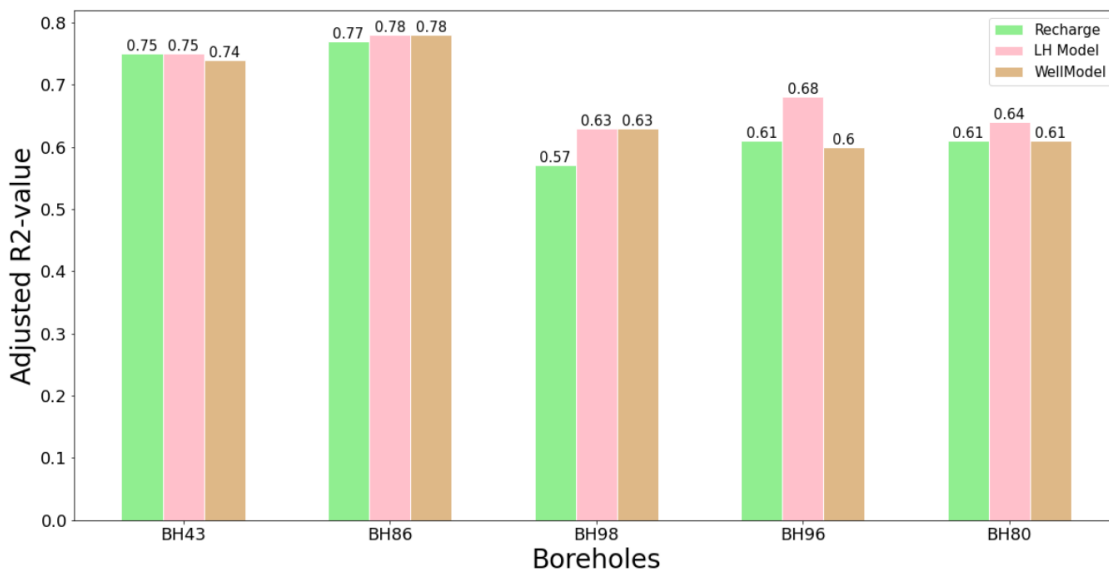


Figure 4.26: R_{adj}^2 for simulated results using both pumping StressModels and recharge but without the AR(1) noise model

Comparing the R_{adj}^2 -values from Figure 4.16 to those presented in Figure 4.26, it is apparent that there are only a marginal increases in R_{adj}^2 -values with the inclusion of pumping as a stress.

Simulations for two bores, both located in close proximity to the municipal and commercial pumps, are presented in Figure 4.27 for BH86 followed by Figure 4.28 for BH80. The first graph in each figure contain a recharge based simulation as seen in the previous sub-chapter for comparative purposes, the following two graphs in the figures present recharge and pumping based simulations created with either the WellModel or the LH model. A complete list of resulting R_{adj}^2 -values for all eight scenarios can be found in Appendix G.

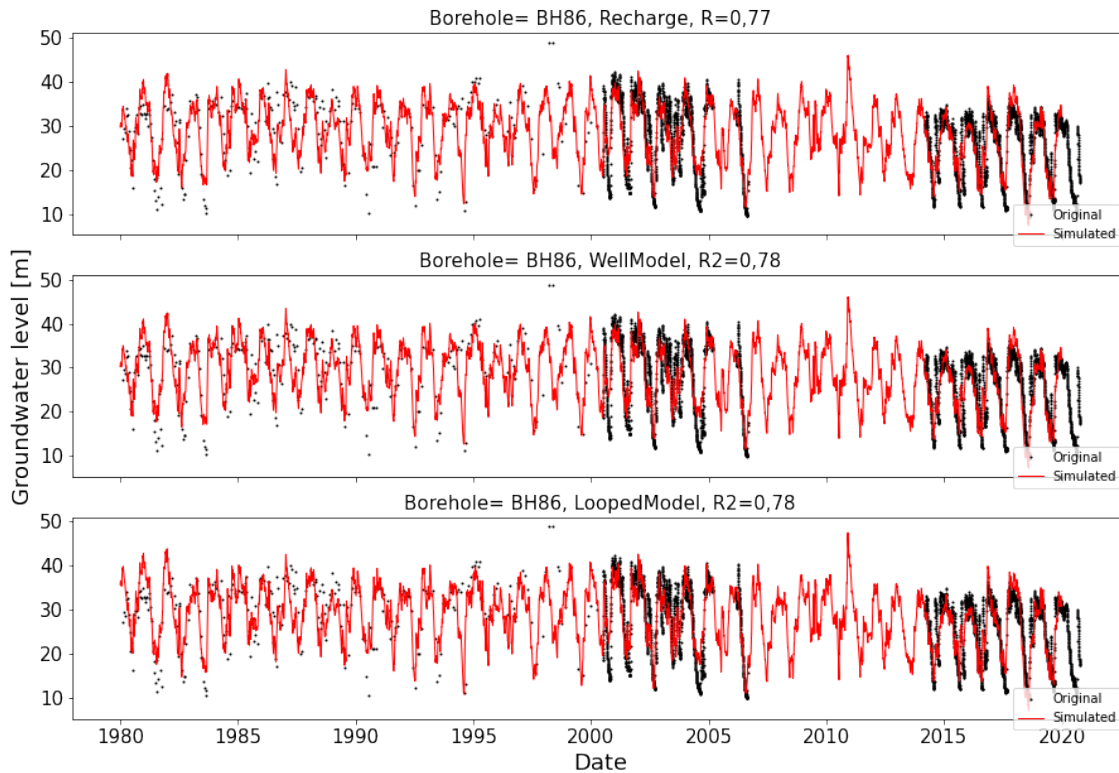


Figure 4.27: BH86: In the order of from top to bottom, a) Recharge Based Simulation, b) Recharge and Pump Using WellModel, and c) Recharge and Pump Using LH Model.

SM1 in conjunction with a Gamma IRF was used for all three simulations. There is no difference in R_{adj}^2 -values between the WellModel and LH model for bore BH86, see Figure 4.27 b) and c). Studying both StressModels used for pumping closely, a slight seasonal difference in fit can be observed between the dry season and wet season. The LH model performs marginally better during the dry season in terms of getting closer to capturing the lower most groundwater levels. WellModel performs marginally better during the wet season in capturing the highest groundwater levels. It overestimates the groundwater levels at times although this is due to the underlying StressModel and IRF used for simulating recharge. Overall, when comparing all three graphs, a slight improvement in visual fit can be observed when pumping is added as a stress, although the inherent discrepancies observed with recharge based simulation remain.

For bore BH80, see Figure 4.28, SM2 and Gamma IRF was used for simulating recharge, this bore has very little automatically recorded daily data therefore optimal to use for comparison with BH86.

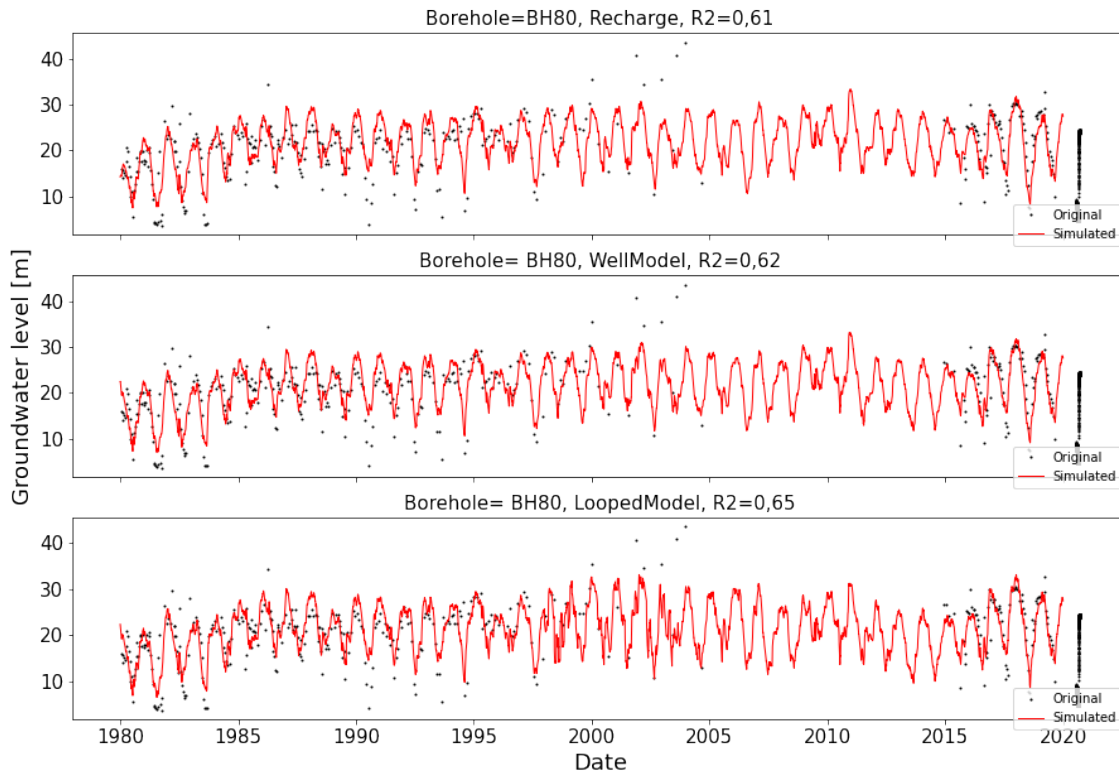


Figure 4.28: BH80: In the order of from top to bottom, a) Recharge Based Simulation, b) Recharge and Pump Using WellModel, and c) Recharge and Pump Using LH Model.

There is no difference R_{adj}^2 -value between the recharged based simulation and the simulation that includes WellModel, see Figure 4.28 a) and b), a slight increase in R^2 for the simulation that uses LH model can be observed, see Figure 4.28 c). All three simulations consistently overestimate the groundwater levels during wet and dry seasons. The simulation produced by the LH model, performs marginally better at reaching lowest groundwater levels, but overestimates them nevertheless.

4.3.1 Evaluation of Recharge and Pumping Based Simulations

Both StressModels for pumping use the the Hantush IRF for modelling of the draw-down, they diverge, however, in their application. The WellModel accounts for distances whereas the LH model does not. The simulated time-series for BH86 and BH80 are evaluated closely, both bores are located close to the File Hajdar quarry.

Diagnostic tests were applied to WellModel and LH model for the purpose of determining which pumps had the largest influence for each simulation, these are presented in Appendix K. The diagnostic tests for BH80 revealed that when applying the WellModel, the model considered the pumps from File Hajdar quarry to have a large influence on the groundwater levels, whereas the municipal pumps were

considered to have a smaller influence and the pumps from *Västra* Quarry extracting no influence at all. For BH86, the WellModel accounted for no influences from the municipal pumps or the *Västra* quarry pumps, with only one of the pumps from File Hajdar quarry extracted the largest influence.

The pumps considered by the WellModel differ from those accounted for by the LH model. The pumps from File Hajdar quarry were determined to have a smaller influence by the LH model, despite being radially closer than the municipal pumps, while the municipal pumps extracted a larger influence on the groundwater levels for both BH86 and BH80. For BH86, one of the pumps from *Västra* quarry was included as an influence, which is noteworthy as the pump is located over five kilometers away. It's important to note that the extracted amount of groundwater from the File Hajdar quarry exceeds that of a municipal pump when compared individually.

The simulated time-series produced by the LH model were scrutinized closer by examining the changes in the simulated time-series when the independent time-series containing pumping data are manually removed. In Figure 4.29 four different simulations can be observed for BH86, the simulations are calibrated from 1980 to 2020 as per usual, but for visual purposes only the simulated groundwater levels for 1997 to 1999 are presented. The simulation calibrated with all 11 of the pumps is included, the remaining have either P6, P5 or the pumps from *Västra* quarry manually removed before calibration.

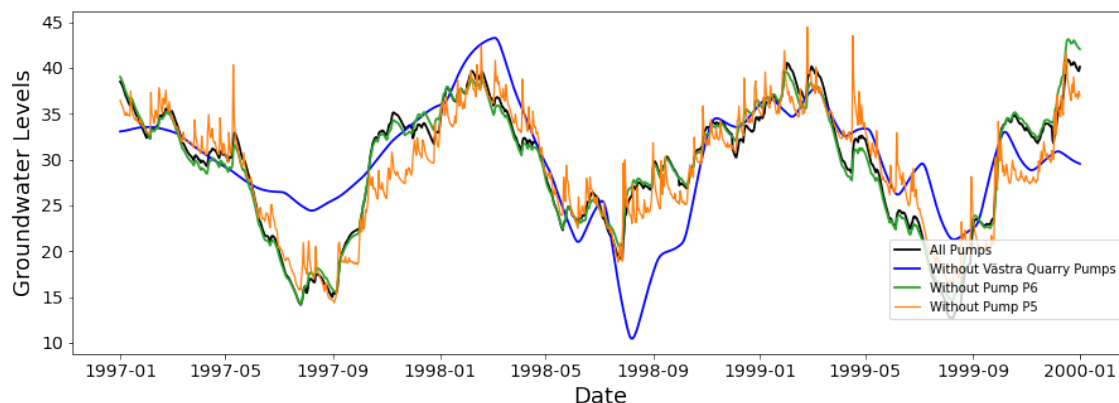


Figure 4.29: Four different simulations for BH86, out of these, three were simulated with one of the 11 pump manually removed.

As evident from Figure 4.27, the exclusion of the *Västra* quarry pumps from the simulation has a considerable impact on the simulated time-series whereas only marginal changes to the simulated time-series can be observed through the removal of P5 and P6. It should also be noted that the exclusion of the *Västra* quarry pumps lowers the R_{adj}^2 -value to 0.53 from 0.78, which is even lower than the R_{adj}^2 -value for the recharged based simulation; 0.77.

Lastly, diagnostic and hypothesis tests were performed for the recharge and

pumping based simulations to evaluate residuals and noise. The tests determined the residuals did not comply with one of the assumptions of noise; entailing that the model can be further improved.

4.4 Comparisons with Cubic Spline and Nearest Neighbor

Time-series were interpolated for bore BH86 and BH43, between 2000 and 2020, as these years contained the largest amount of automatically measured daily data, with the characteristic gap from 2006 to 2014, prevalent in both time-series, being included. The TFN model simulations as well as NN and CS interpolations were compared to the original time-series and to each other; the resulting comparisons for BH86 can be observed in Figure 4.30. It should be noted that the simulation scenario used for both bores in the TFN model was; SM1, Gamma and the WellModel, based on Figure 4.26.

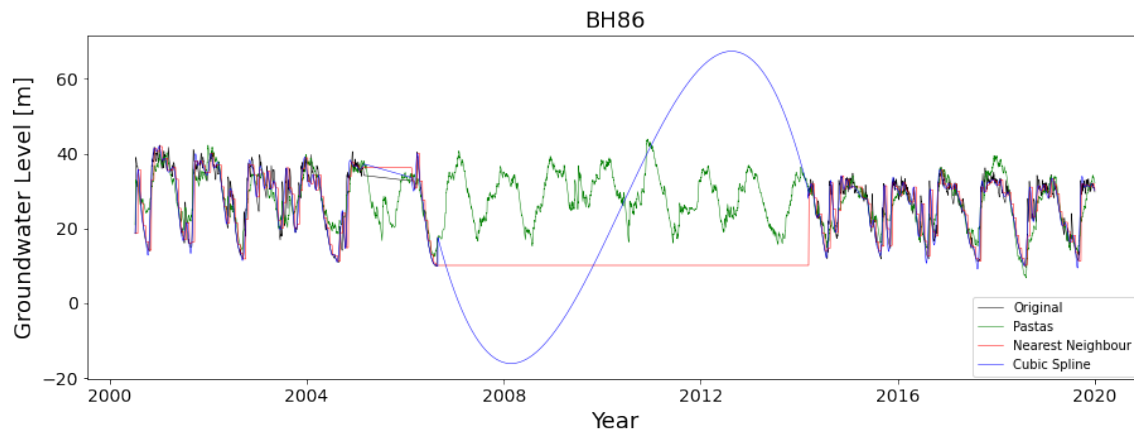


Figure 4.30: A comparison between the original data and NN, CS & Pastas' TFN model simulations for BH86

Only 5% of the daily recorded data was used for the TFN simulations along with the NN and CS interpolations, however, as can be observed in Figure 4.30, 100% of the recorded data is presented in the graph as to more easily provide a visual understanding of the fit. The simulated time-series using the TFN model achieved an estimated R -value of 0.82. CS obtained the closest fit to the original time-series, with the interpolation achieving an estimated R -value of 0.96. The interpolation reached the extreme seasonal fluctuations of the original time-series, but did not fluctuate to the same degree between measurements, acquiring a curve-like shape. As CS produces a piece-wise curve, it may be a reason as to why the interpolation managed to match the curvature of the fluctuations properly. Whereas NN exhibited no fluctuations between the recorded measurements as it repeated the first recorded value for every single day until the next value in the time-series appeared, thus obtaining a step-like appearance. Although, NN managed to capture the extreme fluctuations, it did not achieve the same degree of correlation to the original time-series as the CS interpolation, resulting in a calculated R -value of 0.85. It should

be noted that both the interpolation methods did not perform adequately for the gap in between 2005-2006 and 2006-2014. Whereas the TFN model simulated the values based on both meteorological data and established fluctuation patterns. The same analysis was done for BH43, where CS and NN performed similarly. Manually calculated R^2 -values were derived for both BH86 and BH43, these, along with the graph of BH43, can be found in Appendix L.

Another approach adopted for evaluating the interpolation methods was to remove increasingly larger weekly intervals of daily recorded data between the years 2015-2018. This approach was only applied to the time-series for BH86, with the results being presented in Figure 4.31

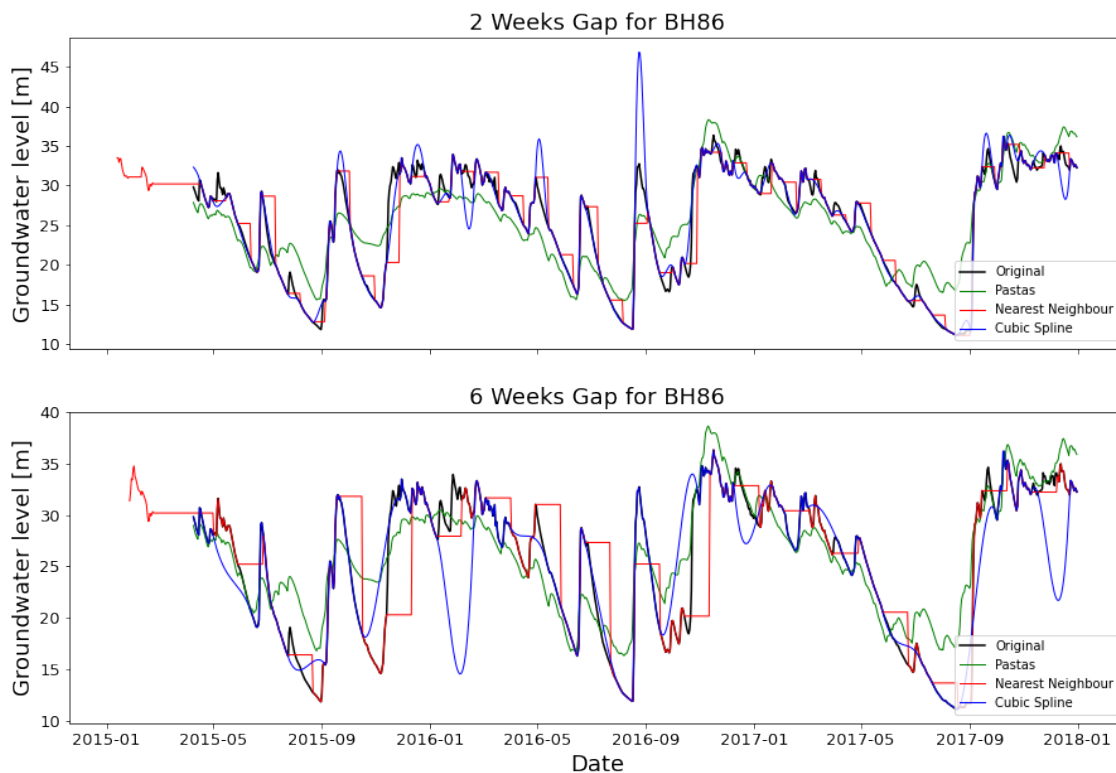


Figure 4.31: In the order of from top to bottom, a) Every two weeks of the time-series are manually removed for b) Every six weeks of the time-series are manually removed

As can be observed in Figure 4.31 a), where every two weeks were removed, CS manages to closely fit the original data but tends to overestimate the groundwater levels between the gaps. Increasing the gaps with an additional four weeks, seen in Figure 4.31 b), only decreased the fit, as the interpolation acquired an increasingly step-like appearance. The periods in the plot where the model achieved a perfect fit was due to the longer stretches of included daily data; where no interpolation occurs. Interpolations using NN achieved increasingly lower fit as the artificial intervals got progressively larger, with the interpolations obtaining a step-like shape, as can be

observed in Figure 4.31 a) and b). The stimulated time-series using the TFN model, in both Figure 4.31 a) and b), is contained towards the middle of the graph and does not change drastically between the different simulation. R -values for the estimated time-series using different gaps can be seen in Figure 4.32, these can also be found as R^2 -values Appendix L.

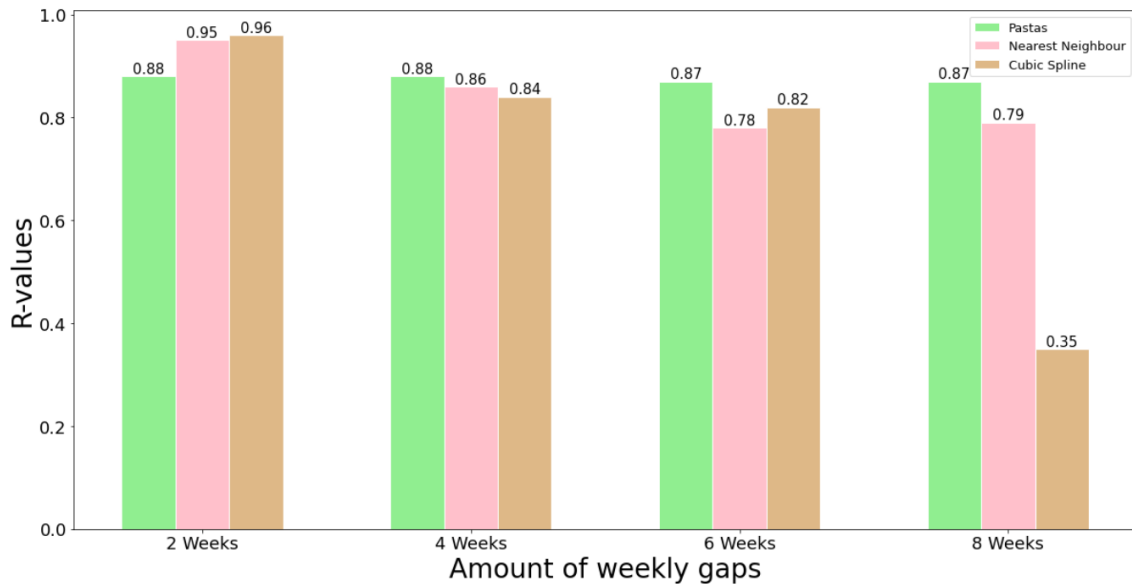


Figure 4.32: R -values for Pastas' TFN-model, NN and CS

When studying the produced result for this comparison it is important to not put too much weight on the calculated R^2 -values, as the methods do not interpolate for the daily recorded data, but rather copies it instead as it is already continuous. It can be seen in Figure 4.32 that when removing two weeks of daily data both CS and NN acquire a higher R -value than the TFN model. However, with increasingly larger gaps the R -value decreases for the interpolation methods but remains approximately the same for the TFN-model. The interpolated time-series are affected rather drastically by the increase of the gap to six weeks, which is also reflected by the changes in R -values. Contrarily, the TFN model simulates continuously with little variations in R -value between the different time-series, irrespective of the amount of data removed. The goodness-of-fit measure decreased considerably, for both CS and NN interpolations, with increasingly larger gaps. The CS exhibits a curved shape between interpolations which can be likened to local variations, NN produces no local variations.

In the groundwater regimes, recorded data between 1980 to 2003 is used. The derived time-series for NN, CS and TFN as well as the seasonally averaged dependent data are compared to the original daily recorded data for the year 2004, see Figure 4.33.

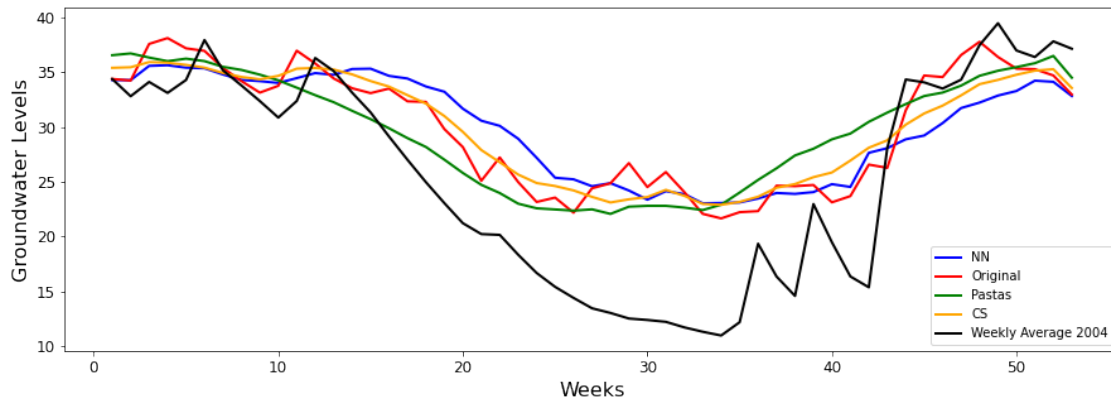


Figure 4.33: Groundwater Regime: NN, CS, Pastas' TFN model and original data, for 1980 to 2003

This analysis shows that none of the seasonally averaged time-series manage capture the steeper decline and values in groundwater levels during the dry period. As can be seen in Figure 4.33 all of three of the predicting methods fluctuate between a 25 meter range and exhibit a similar smooth curvature. The seasonally averaged dependent time-series is the exception, as it is less smooth and exhibits local variations similar to the weekly groundwater levels for 2004.

5 Discussion

This chapter contains a detailed discussion regarding the results from the previous chapter. The chapter begins with providing a motivation for why the TFN model was chosen for estimating groundwater levels from File Hajdar. This is followed by a thorough discussion regarding the performance of the model as well as its sensitivity to the quality and quantity of data. The chapter is concluded with a discussion regarding the inherent uncertainties within the data used in the model as well as the functions of the TFN model itself.

5.1 Choice of Method

The chosen model used for the infilling of the time-series was a TFN model which applies an auto-regressive noise component. The transfer function describes the relationship between the input variables, here, the dependent and independent time series' and output variable, the simulated time series', through the use of a stochastic model (Manzione et al., 2010). The relationship between the independent and dependent time-series, i.e the relationship between precipitation and recharge, is not linear as the process of recharge in a catchment is dynamic in nature and needs to be modeled accordingly. The TFN model transforms the non-linear relationship between the independent and dependent time-series into linearly related equations that attempt to capture the dynamic nature of the hydrogeological processes in order to simulate continuous groundwater time-series. This distinctive ability of the TFN model is the reason for its application in the hydrogeological field (Mohanasundaram, Narasimhan & Kumar, 2016), and one of the main reasons it was chosen for the infilling of groundwater levels from File Hajdar.

The TFN model does not require information on the geology and/or physical processes in the area, which could be considered both an advantage and a disadvantage. A spatial model that attempts to accurately depict and describe the hydrogeology of an area for interpolation purposes would be difficult to compile, it would require a lot of resources and include many simplifications, and may not even produce accurate results (Peterson & Western, 2014). As information required to accurately depict an area of interest is not always readily available, simplifications are therefore unavoidable. Despite the fact, it can be argued that completely excluding the hydrogeological processes can be problematic and limit the success of a model, particularly when an area has a complex geology such as the one studied in this thesis - not considering the surrounding geology can be seen as a simplification.

For the area of File Hajdar, there is information available in regard to the actual hydrogeological parameters and soil layers, but the complex and extremely heterogeneous character of the landscape limits the use of a spatial or 3D models, as the hydrogeological parameters vary vastly over the area (Cementa AB, 2017A). An accurate depiction of a karstic landscape in a prediction model would be difficult to achieve. In a study conducted by Yihdego and Webb (2010), groundwater level fluc-

tuations were estimated using, among other, a TFN model. The groundwater data was retrieved from mainly two sites, with one of the sites containing a limestone aquifer overlain by a volcanic stone aquifer. The study determined that climatic variables alone could explain up to 90% of the fluctuations, this further reinforces the choice of the TFN model. Critical processes, such as precipitation and ET_o , are included in the TFN model and as Yi and Lee (2004) argue, the correlation between precipitation and groundwater levels weighs heavily enough due to the strength of the existing dynamic relationship between the two processes. There are numerous software programs and models that apply the TFN model; the open-source Pastas package was chosen due to its user-friendly interface and because it allows room for improvement.

5.2 Model Performance and Data Requirements

As evidenced by the simulations presented in the previous chapter, the acquired simulations achieved varying results in terms of R_{adj}^2 and visual fit. The dependent time-series from File Hajdar express large variations in groundwater levels between seasons. The model consistently underestimated and/or overestimated the extreme variations, more so when the noise model was activated. This chapter presents a comprehensive analysis of the acquired results, the performance of Pastas' TFN model and an evaluation of the data requirements, this is followed by a discussion regarding inherent uncertainties within the model and data, finally, the possibilities for further research are presented.

5.2.1 AR(1) Noise Model

A considerable improvement in fit is observed for most of the simulations when the noise model is deactivated. It has been previously noted that the Pastas noise model may not function efficiently with high frequency data, i.e. hourly or daily. Therefore the need to simulate the time-series both with and without the noise model was a necessity, for the purpose of determining the difference in resulting simulations. When the noise model was deactivated the largest improvements visually and in R_{adj}^2 -values were observed for time-series that contained daily data, which could be attributed to the noise model's inability to perform efficiently on high frequency data. For time-series without daily data, such as those belonging to BH2006 and Slite 39, the observed changes without the noise model were almost non-existent. A marginal improvement can be seen for Slite 39, this could however be attributed to the bounded nature of the time-series. Although, a higher R_{adj}^2 -value was acquired through the deactivation of the noise model, the influence of the IRFs and the StressModels became almost indistinguishable in terms of R_{adj}^2 -value, resulting in only marginal differences between the different simulation scenarios. Thus, the improvements observed in the simulations through the exclusion of the noise model should be considered a weakness in the TFN model as noise is not accounted for at all. Although, it can be argued that the exclusion of an inaccurate noise model is

preferable to including it.

5.2.2 Evaluation of Recharge Based Results

The main strength of the TFN model is that it creates continuous time-series based on variables related to the climate and land use. Though this model isn't specifically created for karstic landscapes, it has nevertheless proven to produce results with relatively high goodness-of-fit values. For larger intervals of no recorded data, the model provides estimates based on concrete input data, which is in itself quite a remarkable function of the model. Though the model did capture the large fluctuation characteristic of the karstic landscape of File Hajdar, it did not, however, manage to capture the outer most extreme values as it consistently overestimated, and at times underestimated, the groundwater levels. This is particularly true for bounded time-series, the low R_{adj}^2 -values for SGU11001 and Slite 39 were not surprising as the TFN-model does not consider that a time-series may have an upper and/or lower boundary. It can be argued, however, that the simulated time-series' may be more representative when considered from a perspective wherein the dependent time-series has no boundary. Had the bores been dimensioned differently the groundwater levels would have followed a similar pattern of fluctuations, albeit in an overestimated capacity. The model provides a strong starting point for further evaluations and improvements for analysing time-series belonging to bounded bores.

For time-series that contained exceptionally long intervals wherein no groundwater levels were recorded, such as those for Slite 9 and BH96, the model simulated groundwater levels for the intervals based on precipitation and ET_o as well as the fluctuation patterns exhibited by the dependent time-series before and after the gaps. Bore Slite 9 and BH96 contained a gap of 28 and 14 years, respectively. Despite the long gaps the model produced relatively high R_{adj}^2 -value without the application of the noise model. However, these R_{adj}^2 -values are only representative for the simulated years that have corresponding recorded groundwater data. The model has no recorded data to evaluate the simulation against during the 28 and 14 year gaps, so naturally the R_{adj}^2 -values does not provide insight into the accuracy of the simulated values for these years.

As for the borehole with an external disturbance, BH2006, the model yielded a R_{adj}^2 -value of 0.53 for a dependent time-series consisting of sporadically measured values over a 12 year period. It is important to note that as the only bore located in close proximity to a large body of water, the groundwater levels are highly influenced by the lake. One of the central limitation of this study was that discharge and recharge to and from Tingstäde lake was not considered. By taking the lake and its changing levels into account, a more accurate simulation could be acquired for BH2006 and other bores in close proximity to the lake. Though there was no scope to explore it in this thesis, there is potential for further improvements by considering the interchange between the lake and the aquifer.

While studying the simulations an obvious weakness in the model is detected. The TFN model appears to model the fluctuation patterns based on the periods

in the dependent time-series with the largest quantity of data. Observe the simulations for BH98 presented in Figure 4.19, despite the recorded groundwater levels exhibiting larger fluctuations between 1987 and 2000 than from 2001 and onwards, it can be seen in the simulation that there is little variation in the range of the fluctuations, they are consistently the same for the entirety of the time-series. This is a consequence of the large amount of daily recorded data that is included in the time-series for this bore, which could weigh more heavily in the simulation than the manually recorded data. As the shift in trend occurs right before the measurement frequency increased, this could be a possible explanation.

Furthermore, the TFN model appears to have difficulties in modelling short term irregular variations. BH96, for instance, displays an almost M-shaped irregularity, unrelated to any cyclic variation or seasonality, between 1980 and 1998, see Figure 4.17. The model does not recognise this irregular variation and therefore does not correct for it. A manually added linear trend was included to improve the fit of the simulated time-series which increased the overall R_{adj}^2 -value of the simulation. However, the TFN model does not allow for the addition of more than one linear trendline, nor does it have the option to include non-linear trends, thus the M-shape in the time-series could not be replicated. This can be considered both a positive and negative aspect of the model. Positive, because this means that the model creates the simulations based only on the relationship between the independent and dependent time-series. Adding multiple trendlines, based solely on visual variations, manipulates the model and undermines its methodological process and could result in an over-fitted model.

These irregularities could be a consequence of a variety of reasons, those observed in the time-series for both BH96 and BH43 - which exhibits an U-shaped trend between 1985 and 1995 - could be attributed to their close proximity to the File Hajdar quarry. The operations at the quarry could have disrupted the network of fractures, inadvertently affecting the groundwater levels, which cannot be accounted for in the model through the inclusion of an independent time-series. It could also be the result of municipal pumping, which has been indicated to have occurred at higher and more varied rates prior to 1998, but the pumping data was not accessible and therefore this cannot be reflected in the independent time-series for pumping. Most likely, the irregular variations in the time-series can be attributed to being an effect of both municipal pumping and the impacts of the quarry.

The models tendency to consistently overestimate the groundwater levels is prevalent throughout the entirety of the time-series', even during years with daily recorded data. This could be attributed to the StressModels currently included in the Pastas package not being able to model the large oscillations in groundwater levels characteristic of File Hajdar. The StressModels convulses the stresses applied to the model, if the StressModels aren't inherently capable of accurately modelling the stresses to capture the large oscillations then improvements cannot be observed no matter how many stresses may be applied to the model. Recharge occurs at a higher rate between October and April, which causes the ground to eventually become fully saturated, resulting in an inability to allow for any more water to

peculate through the fractures. A considerable amount of excess precipitation ends up in the Baltic Sea as surface run-off. During dry seasons intense precipitation events are an occurrence, although recharge occurs at a much smaller rate, extreme rain events could also result in excess run-off. As the model does not account for run-off or the saturation levels of the ground, this component is not considered in the simulations and could also partially explain why the model overestimates the groundwater levels. This ties back to the central dilemma regarding choice of the prediction method, touched upon in sub-chapter 5.1, of whether to chose a model that attempts to recreate the hydrogeological conditions of an area or choosing a data driven model such as the TFN model.

The data requirements for the TFN model were evaluated through a sensitivity analysis, applied to BH43 and BH86, using two approaches; a "moving block" approach and a systematic percentile removal of daily data. The sensitivity analysis using both approaches showed that the TFN model managed to produce simulations with relatively high R_{adj}^2 and R -values, even with shorter quantity of data. The moving block assessment demonstrated that the two year calibration period achieved similar R -values as the seven year calibration period. This is interesting as the two year calibration period inherently contains less data points, which is especially true during the manually recorded years where the measurement frequency is one to two times a month. This demonstrates that the model is capable of simulating shorter periods of a time-series with smaller quantities of data while achieving relatively high goodness-of-fit measures. Meaning continuous weekly or daily data of a minimum of eight years, as suggested by Heudorfer et.al (2019) for instance, is not a necessity for the TFN model. The model did produce a more accurate fit visually and marginally higher R -values using the seven year moving block. However, the seven year moving block simulations did not visually capture the large range in the fluctuations to the same extent as the two year simulation block. The model has difficulties accurately modelling irregular variations in the general trend, using shorter calibration periods manages to circumvent this issue to an extent. The majority of the two and seven year moving block simulations did not contain daily recorded data until the block reached the year 2000, meaning the considerable weight daily data has on the simulations in general could not be exerted in the shorter two year moving block.

The second approach was implemented for the purpose of evaluating the effects that additional years of daily data can have on the entirety of the simulation. The systematic 100%, 75%, 50% and 25% reduction in daily data displayed little difference in R_{adj}^2 -value for each simulation, further proving that an extensive amount of daily and weekly data is not necessary for this type of prediction method. When 95% of the daily data was removed, which translates to a measurement frequency of approximately twice a month, a considerable visual change and a decrease in R_{adj}^2 could be viewed. This could be explained by long periods of daily data carrying a heavier weight in the simulation. If that is the case, the continuous addition of daily recorded data could have a positive effect on the overall simulation. For many of the time-series from File Hajdar, that consist of fewer years of recorded groundwater levels or have larger periods of missing data, this could prove to be beneficial. Al-

though, as daily recorded data in time-series has previously been discussed to carry considerable weight in the simulations, the continuous addition of daily data may further exasperate the issues associated with heterogeneous data.

5.2.3 Evaluation of Recharge and Pumping Based Results

One of the main limitations this study which affects the pumping based simulations can be attributed to the use of the Hantush function for leaky aquifers used for modelling draw-down. This function is the basis of the IRF used in both LH model and WellModel and presupposes 10 assumptions related to the hydrogeology of the area, which can be found in Appendix B, one of which is that the area assumed to be homogeneous and isotropic. The aquifer in File Hajdar is described as confined, with the flow being of leaky character between the layers. As File Hajdar does not fully conform to the assumptions of a leaky aquifer, the Hantush response function may therefore represent the influences of pumping inaccurately leading to an uncertainty in the results. The addition of an alternative IRF, better suited to the condition of the area, may potentially provide improved results.

The marginal improvements observed in R_{adj}^2 -values through the inclusion of pumping data is a clear indication of anthropogenic effects being relevant to model. As previously mentioned, the R_{adj}^2 -value counteracts spontaneous and unexplained increases in fit when additional independent time-series' are added to the simulation, therefore the observed increase in R_{adj}^2 can be considered justified and a viable effect of the improvement of the model.

The resulting R_{adj}^2 -values varied depending on which StressModel was used for pumping; LH model or WellModel. The LH model provided marginally improved results in terms of R_{adj}^2 -values than the WellModel. Although, as evident from the diagnostic tests, found in Appendix K, the LH model may not be fully representative with regard to the pumps the model considers an influence. When studying the diagnostic tests for BH86, it is evident that the LH model considers the groundwater extraction from *Västra* quarry to be an influence, these are located over five kilometers from the bore, while neglecting to consider more local influences such as the commercial pumps in File Hajdar quarry or the municipal pumps.

Manually removing the independent time-series that represents the pumps from *Västra* quarry, as seen in Figure 4.29, resulted in a considerable decrease in R_{adj}^2 of the simulation. The observed changes in R_{adj}^2 through the inclusion or exclusion of *Västra* quarry pumps, reaffirms that the LH StressModel bases its inclusion of pumping influences on the fluctuation patterns of each specific pump, as to match it optimally to the dependent time-series. Although, this may not be entirely representative, it can be argued that as limestone does not have a radial flow pattern, due to the groundwater being contained in the fractures and between layers of denser material, the influences of pumping may not exclusively depend on distances. This characteristic of the LH model can potentially be considered advantageous for areas with non-radial flow and heterogeneous soil composition. As the LH model does not consider distances, the use of it would stipulate that each simulations be closely

evaluated and that diagnostic tests be utilized to evaluate whether the inclusion of the considered pumps may be reasonable. In the case of BH86, it can be stated that the inclusion of *Västra* quarry pumps aren't reasonable. The simulations produced by WellModel achieved marginally lower R_{adj}^2 -values, the diagnostic tests revealed that pumps from File Hajdar quarry had the largest influence on the simulations.

It is important to consider that without representative pumping data the model will not be able to produce optimal simulations, which may explain why the pumps from *Västra* quarry are considered to influence the groundwater levels when using the LH model. Prior to 2011, the quarry pumped groundwater at a higher rate than they do currently, a disconnect in pumping patterns may therefore be a potential explanation for why the LH StressModel considers *Västra* quarry to be an influence instead of the pumps at File Hajdar quarry or the municipal pumps which are located much closer. Alternatively, the municipality allegedly pumped at a higher rate prior to 1998, the LH model may be considering the *Västra* quarry as an influence to compensate for the lack of representative municipal data. If the pumping data is not representative then the resulting simulations will naturally be affected, with the fit of the produced simulations not being as optimal as they could be.

The residuals were evaluated for the recharge and pumping based simulations, with evidence suggesting that residuals do not comply with criteria for noise. Meaning, there is room for further improvements in the model, as the noise is not completely random. It can be argued that as the data used for the simulation of draw-down was not representative, the addition of better pumping data could further improve the results. Moreover, irrigation and groundwater extracted from private wells were not included in the model due to a lack of data, their inclusion may also contribute in achieving an improved fit.

5.2.4 Comparison to Alternative Interpolation Methods

Between the estimated time-series derived using the TFN model, CS and NN, the CS obtained the most accurate fit in terms of R and R^2 . This is perhaps expected, as where the TFN model used meteorological data and the dependent time-series to simulate groundwater levels, both CS and NN only interpolated between recorded groundwater levels. This is evident when observing the simulated values for both the gaps between 2005-2006 and 2006-2014 in Figure 4.27, where CS produces a piecewise cubic curve, NN a rectangular line and the TFN model estimates groundwater levels.

Even though the interpolations were not able to estimate groundwater levels for the gaps, they nevertheless managed to fit the dependent time-series. However, if the original time-series had fewer measured values or not enough measurements during the seasons the interpolations would not have been able to accurately represent the range in fluctuations. The interpolations are only as good as the underlying recorded data, if the recorded groundwater time-series contain short intervals of missing data then, in particular CS, could be considered a useful tool for simple interpolations. With larger intervals of missing data both interpolation methods would prove to

perform poorly or not at all, as evidenced by the 2006 to 2014 gap. The accuracy of the NN and CS interpolations are entirely the result of the quality and quantity of the dependent time-series.

To further analyze how the interpolation methods responded to changes in data quantity, additional evaluations were done by increasing the gaps between the data points. CS and NN may be more beneficial to use for time-series that contain smaller gaps between measurements; of two weeks or less. Whereas a prediction method that simulates continuously would be more appropriate for analyzing time-series containing larger and more varying gaps, as the model has shown not be sensitive to the removal of larger periods of data.

Studying the groundwater regimes, presented in Figure 4.33, it can be concluded that the seasonal averaging produces smoother curves for all of the estimated time-series. With the seasonally averaged dependent time-series being the only exception, as it is not smooth and does exhibit local variations. None of the estimated time-series show the steep decline in groundwater levels exhibited by the averaged time-series for year 2004. It can be argued that extending the time-series from 2003 to 2019, to include more daily data, would have presented more varied results, with the seasonally averaged dependent time-series likely exhibiting a steeper decline reflective of the one exhibited by the time-series for 2004. It can also be assumed that the rest of the time-series, with the exception of the one produced by the TFN model, would similarly be more reflective of the decline while still retaining their smooth curvature. As the TFN model has shown to continuously overestimate groundwater levels and not been able to capture extreme fluctuations, including more daily data would likely not a large difference in the groundwater regime.

5.2.5 Relationships Between StressModels and IRFs

Solely based on the R_{adj}^2 -results for recharge, there is no extreme superiority in terms of the IRFs or StressModels when analysing the results, although the Exponential IRF achieves marginally higher R_{adj}^2 -values for a majority of the bores but only for simulations that include the noise model. Without the noise model, the different simulation scenarios essentially become obsolete, as only minimal differences can be observed between them.

The main difference between SM1 and SM2 is based on the convulsion of the stresses with regard to the IRF used to simulate the responses. As previously mentioned, SM2 requires the user to define precipitation and ET as separate stresses, which it then convulses using a chosen IRF and an evaporation factor, f . SM2 attempts to estimate the AET through the use of the ET_o and factor f . There were expectations that SM2 would produce higher R_{adj}^2 -values compared to SM1 for the File Hajdar area, as SM2 was created by Collenteur et al., (2020) with the intention of it being an improvement to SM1, however, these expectations were not realized. AET can be estimated by multiplying the reference evapotranspiration by a crop coefficient, K_c , which is crop specific, and therefore, site specific. It is unsure how the factor f in Pastas is estimated and if it is somehow related to the K_c coefficient,

all that is known is that it is not a fixed parameter. Perhaps if the model allowed for the manual input of an actual estimated Kc coefficient rather than the f factor, the difference between SM1 and SM2 for the File Hajdar area would be more prominent, as it would more accurately reflect the site.

Two IRFs used for simulating recharge are evaluated in this thesis; Gamma IRF and Exponential IRF. A site and borehole based assessment of the IRFs was made for the purpose of determining whether specific tendencies could be observed that could relate an IRF to a site or bore specific characteristic. The assessment resulted in a discernible pattern being identified that connected IRFs to elevation. The Exponential IRF produced higher R_{adj}^2 -values for bores that have a lower elevation relative to the surrounding area, the bore could be located in a small curvature or at the end of a decline. This relationship is particularly evident with the noise model activated. Slite 39, BH96, Slite 9, BH80 and BH98 all have a lower elevation in comparison to their surrounding area and time simulations for all five bores achieve higher R_{adj}^2 -values when being modelled with Exponential IRF rather than the Gamma IRF. This could partially be explained by precipitation travelling from higher elevations to lower elevations. The Exponential IRF is used to model an immediate response to a stress, meaning that model assumes the precipitation to have an immediate effect on the groundwater levels. It could be argued then that the Exponential IRF should be optimal for File Hajdar as the limestone allows for the quick transportation of water. However, as the simulations for the remaining four bores produce higher R_{adj}^2 -values using Gamma IRF, this argument is flawed. Therefore, the combination of the geology and differences in elevation is considered to be the reason for the above mentioned bores achieving higher R_{adj}^2 -values whilst the remaining bores did not. Although, as only nine bores in total were evaluated in this thesis, further research and evaluation is required to explicitly determine whether the Exponential IRF is preferable for bores situated in lower elevation.

5.3 Uncertainties

Uncertainties need to be considered when analysing the results of the simulations. These lay in part with the assumptions made before and during the compilation of the results as well within the data used for this study. The functions in the Pastas' TFN model are also a source of uncertainties.

5.3.1 Data and Assumptions

For the compilation of the independent time-series, data was acquired from SMHI and Copernicus. As ET_o needed to be calculated for over a 40 year period, the input data needed for the ET_o calculations could vary in quality. As mentioned, relative humidity data was poor in quality and had to be excluded from the calculations. Moreover, due to a lack of accessible meteorological data for the File Hajdar area, data from a station in Visby was used instead. This introduces a certain uncertainty in terms of the reliability of the ET_o . Wind and humidity are factors that

affect the ET, it can be assumed these may differ between the areas. Although, it should be noted that Slite is also a coastal town, similar to Visby which is located approximately 37 kilometers away. The differences in ET_o between Visby and File Hajdar may therefore not vary that much. Since the precipitation data, a central component of recharge, was acquired from a station close to File Hajdar this can be considered to reduce the uncertainty.

All of the data used for the simulations was examined thoroughly and processed with as minimal assumptions as possible. The effects of pumping was included in an attempt to improve the simulations; it is safe to assume that pumping does have an impact on the groundwater levels but to what degree and how wide spread the impact is can be argued. Private pumping was excluded from the simulations completely due to a lack of accessible data, but as there are a considerable amount of private wells in the area, their exclusion could be considered an uncertainty. Groundwater used for irrigation purposes of any kind were also excluded for the same reason. Only commercial and municipal pumping was included, but its inclusion was restricted by the limited amount of groundwater extraction data available. The municipality provided data for the years 1998-2019, but only pumping data for 2011-2015 could be accessed for the commercial groundwater usage. Therefore, average monthly values were used for the years where pumping data was missing. The averaged values entail a constant daily extraction of water based on the accumulated monthly extraction, thus producing inaccurate independent time-series' that are not representative. Moreover, the pumps at File Hajdar and Västra quarry are used for the removal of both groundwater and precipitation that collects in the quarries. It was not possible to discern how much of the water removed through pumping was groundwater and how much was precipitation.

It is also important to consider the factor of error. As the foundation of this thesis is made up of the data used in the calculations and simulation, it is incredibly important to consider errors in data that may arise from inaccurate measurement or faulty equipment. Errors can be difficult to discern but where obvious errors were apparent corrections were manually made to a time-series, as was done for BH2006, which was blocked by a tin can causing abnormal changes in fluctuations. For the majority of bores and data used in the calculations of ET_o , errors in measurements or errors due to faulty equipment may not be apparent at all. It should be noted that attempts were made to critically examine and question the time-series for each borehole, pumping data and underlying data for ET_o . Finally, despite the work procedure being a reiterative process imbued with constant verification, errors could have unknowingly been made by the authors of this thesis.

5.3.2 Model and Functions

It was earlier established that an explanation as to why the TFN model overestimated groundwater levels could be due to the bores being located in limestone; the ground in File Hajdar tend to get fully saturated rather quickly during wet seasons. Therefore, an overall uncertainty may be attributed to the models inability to ac-

count for run-off produced in the area, in any capacity. As the geology, and therefore indirectly run-off, is not accounted for, it can not be ruled out that this process is not one of the causes for the overestimation of groundwater levels. Furthermore, relating to the geology, an additional uncertainty arises when applying Hantush IRF to the pumping scenarios. The IRF presupposes certain assumptions in regard to the calculation of draw-down for a leaky aquifer, which are not representative of the studied area.

An additional uncertainty with the model concerns the yielded R_{adj}^2 which provides information on how well the simulations fits the dependent time-series. It does so with consideration to the number of independent time-series applied in the model. The value increases if the added variables contributes to the results more than expected and decreases if it does not (Hyndman & Athanasopoulos, 2018). However, if too many factors are added these may manipulate the model and undermine its methodological process, leading to an over-fitted simulation.

Lastly, although deactivating the AR(1) noise model may contribute to improving the fit of the simulated time-series, it can be considered an uncertainty in itself. The noise model is an important part of the Transfer Function Noise model, excluding it completely results in the Pastas' TFN model not modelling or reducing the noise. However, as the AR(1) noise model does not function optimally for high frequency data, an active choice was made to simulate time-series without it. As a consequence of deactivating the noise model, differences between the simulation scenarios were essentially eliminated. The difference in the fit between the simulated time-series with and without the noise model speaks for the impact of a noise model that does not function accurately.

5.4 Further Research

There is scope for further research based on the work done within this thesis. As established, the TFN model does not account for the geological environment when simulating the time-series, which has proven to be especially decisive for the area of File Hajdar due to its complex geology. By incorporating components that reflect the geology either through the creation of a StressModel or an IRF, the model would be able to improve the produced results. This may allow the simulations to more accurately capture the large range in fluctuations that reflect the complex geology of limestone.

Alternatively, the dependent and independent time-series can be processed before being applied in the TFN model, using for instance the Hydrological Bureau's Water Balance Department Light (HBV-Light) model. The main purpose of the HBV-Light model is to depict the area of interest and calculate the runoff during wet seasons (UZH, 2020). Exploring this option could aid in more accurately estimating the actual amount of precipitation that contributes to recharge and the amount that instead becomes run-off, therefore producing more representative simulations. Moreover, a potential correlation was found between the elevation of the bores and

the Exponential IRF. This relationship can be further evaluated and applied to other areas with similar conditions to determine whether this relationship truly is the cause of the correlation. Lastly, there are additional functions or simulation alternatives that can be added to further improve the fit of the TFN simulations. One suggestion is to apply weights to the dependent and independent time-series. This can be done to improve the resulting simulation by either considering the weight of meteorological data more than the recorded groundwater data or vice versa. Another alternative is to include the possibility to apply additional linear and non-linear trends in the model. Such an alternative will make it possible to obtain an improved fit for the simulation, especially when the recorded time-series has a trend that is known to emerge from an external source which has not been accounted for in the model.

6 Conclusion

In conclusion, the Pastas' TFN model was able to capture the dynamic nature of the groundwater levels at File Hajdar with a relatively high goodness-of-fit values, which is note worthy due to the complex nature of karstic landscapes. However, the resulting simulations overestimated the groundwater levels and did not reach the extreme range in fluctuations common for limestone. Regarding the models response to the data quantity when calibrating for a short period of time, the simulations adequately fit the recorded time-series. It was further concluded that the addition of daily recorded data increased the models capacity to improve the simulations in terms of goodness-of-fit. Although, as the model does not adequately take into account the heterogeneity of the dependent time-series, it will weigh the fluctuations of the daily recorded groundwater levels more heavily than the sporadically measured data. When deactivating the noise model the simulations obtained a better fit to the recorded data, but as a consequence, the component that models and reduces the noise is lost.

The inclusion of pumping in the model presented marginal improvements in fit. As the commercial pumping data used in this study consisted of averaged values for a short measurement period, it can be assumed that with more representative data the fit could be further improved to yield more accurate simulations. Moreover, between the evaluated Gamma and Exponential IRF, the latter produced higher goodness-of-fit values for bores situated on a lower elevation relative to the surrounding area. The diagnostic tests showed that the residuals from the pumping based simulations did not comply with at least one of the assumptions of noise, indicating that the model can be further improved.

Finally, the TFN simulations were compared to interpolations produced by Cubic Spline and Nearest Neighbour. It can be concluded that for time-series containing shorter gaps between data, Cubic Spline and Nearest Neighbour would be more beneficial to use. Although, it should be noted that the interpolation methods are highly sensitive and only as good as the underlying recorded data. For larger gaps the TFN model should be considered more representative.

7 References

Allen, R. G., Pereira, S. L., Raes, D. & Smith, M. (1998). *Crop evapotranspiration* (56). Food and Agriculture Organization of the United Nations Rome. Retrieved <http://www.fao.org/3/x0490e/x0490e00.htmContents>

American Water Works Association (AWWA). (2014). M21 Groundwater, Fourth Edition, Copyright © 2014 *the American Water Works Association*. ISBN:9781583219645

ArcGIS Pro. (n.d.). *How inverse distance weighted interpolation works*. Retrieved on 2021-04-02 from: <https://pro.arcgis.com/en/pro-app/latest/help/analysis/geostatistical-analyst/how-inverse-distance-weighted-interpolation-works.htm>

ArcGIS Pro. (n.d.). *How Kriging works*. Retrieved on 2021-01-06 from: <https://desktop.arcgis.com/en/arcmap/10.3/tools/3d-analyst-toolbox/how-kriging-works.htm>

Aqtesolve. (n.d.). *Solution for Leaky Confined Aquifers*. Retrieved on 2021-05-30 from: <http://www.aqtesolv.com/hantush2.htm>

Bakker, M., & Schaars, F. (2019). Solving Groundwater Flow Problems with Time Series Analysis: You May Not Even Need Another Model. *Groundwater*, 57(6), 826-833. <https://doi.org/10.1111/gwat.12927>

Bárdossy, A., & Pegram, G. (2014) Infilling missing precipitation records – A comparison of a new copula-based method with other techniques. *Journal of Hydrology*, 519(A), 1162-1170. <https://doi.org/10.1016/j.jhydrol.2014.08.025>

Brakkee, E., van Huijgevoort, M., & Bartholomeus, R. P.: Spatiotemporal development of the 2018–2019 groundwater drought in the Netherlands: a data-based approach, *Hydrology and Earth System Sciences*. Sci. Discuss. [preprint], <https://doi.org/10.5194/hess-2021-64>, in review, 2021.

Britannica. (n.d.) *Evapotranspiration*. Britannica. Retrieved on 2021-03-05 from: <https://www.britannica.com/science/hydrologic-sciences/Interception>

Campozano, L., Sanchez, E., Aviles, A., & Samaniego, E. (2014). Evaluation of infilling methods for time series of daily precipitation and temperature: The case of the Ecuadorian Andes. *MASKANA*, 5(1), 99–115. DOI: 10.18537/mskn.05.01.07

Cementa AB. (2017A). *Hydrogeologiska fältundersökningar på File hajdar*. (1650142). Golder Associates. www.cementa.se/sites/default/files/assets/document/6d/fb/4._enmodell_-_bilaga_2_hydrogeologiska_faltundersokningar_pa_file_hajdar_med_underbilagor-1.pdf

Cementa AB. (2017B). *PM Ytvatten för ansökan om tillstånd för fortsatt täktverksamhet och vattenbortledning m.m. i Slite, Region Gotland*. (US16022). BergAB. https://www.cementa.se/sites/default/files/assets/document/32/e4/05._pm_ytvatten_med_underbilagor.pdf

- Cementa AB. (n.d.). *Fakta kring Cementas vattenpåverkan på norra Gotland*. Retrieved on 2021-03-21 from: <https://www.cementa.se/sv/fakta-kring-Cementas-vattenpaverkan-pa-norra-Gotland>
- Celebi, E. M. & Aydin, M. (2016). *Unsupervised Learning Algorithms*. Springer. ISBN 978-3-319-24211-8
- Chatterjee, S., & Hadi, S. A. (2012). *Regression Analysis by Example*. (5th edition). Wiley. ISBN: 978-1-118-45624-8
- Collenteur, R. A., Bakker, M., Caljé, R., Klop, S.A., & Schaars, F. (2019) Pastas: open source software for the analysis of groundwater time series. *Groundwater*, 57, 877-885 Doi: 10.1111/gwat.12925.
- Collenteur, R.A., Bakker, M., Caljé, R., Klop, S.A. & Schaars, F. (2020). *Pastas Documentation; Release 0,16.0*. Retrieved on 2021-03-02 from: https://pastas.readthedocs.io/_/downloads/en/latest/pdf/
- Cryer, D. J., & Chan, K-S. (2008). *Time Series Analysis*. (2nd) Springer. e-ISBN: 978-0-387-75959-3
- Dey, A. (2016). Machine Learning Algorithms: A Review. *International Journal of Computer Science and Information Technologies*, 7(3), 1174-1179. ISSN: 0975-9646
- Djurberg, H. (2016). *Gotlands grundvatten och dricksvatten. Förutsättningar och utmaningar inför framtiden*. Region Gotland. Retrieved on 2021-03-15 from: <https://www.gotland.se/94272>
- Durdu, Ö. F. (2009). A hybrid neural network and ARIMA model for water quality time series prediction. *Engineering Applications of Artificial Intelligence*, 23(2010), 586-594. Doi:10.1016/j.engappai.2009.09.015
- Eatwell, J., Milgate. M., & Newman, P. (1990). *The New Palgrave Time Series and Statistics*. (4). Macmillan Publishers ISBN 978-1-349-20865-4
- Figura, S., Livingstone, M, D., & Kipfer, R. (2014) Forecasting groundwater temperature with linear regression models using historical data. *Groundwater*, 53(6), 943-954. <https://doi.org/10.1111/gwat.12289>
- Giese, M., Haaf, E., Heudorfer, B. & Barthel, R. (2020). Comparative hydrogeology – reference analysis of groundwater dynamics from neighbouring observation wells. *Hydrological Sciences Journal* 65, (10), 1685-1706. <https://doi.org/10.1080/02626667.2020.1762888>
- Goldscheider, N., Chen, Z., Auler, A.S., Bakalowicz, M., Broda, S., Drew, D., Hartmann, J., Jiang, G., Moosdorf, N., Stevanovic, Z. & Veni, G. (2020). Global distribution of carbonate rocks and karst water resources. *Hydrogeology Journal* 28, 1661–1677. <https://doi.org/10.1007/s10040-020-02139-5>
- Gnauck, A. (2004). Interpolation and approximation of water quality time series and process identification. *Analytical and Bioanalytical Chemistry*, 380, (3), 484–492.

Doi:10.1007/s00216-004-2799-3.

Haaf, E. M. (2020). Towards Prediction in Ungauged Aquifers – Methods for Comparative Regional Analysis. [Thesis for the Degree of Doctor of Philosophy, University of Gothenburg]. GUPEA. <https://gupea.ub.gu.se/handle/2077/63694>

Heudorfer, B., Haaf, E., Stahl, K., & Barthel, R. (2019). Index-Based Characterization and Quantification of Groundwater Dynamics. *Water Resources Research*, 55(7), 5575-5592. <https://doi.org/10.1029/2018WR024418>

Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J., Nicolas, J., Peubey, C., Radu, R., Rozum, I., Schepers, D., Simmons, A., Soci, C., Dee, D., & Thépaut, J.-N. (2018). *ERA5 hourly data on single levels from 1979 to present*. Copernicus Climate Change Service (C3S) Climate Data Store (CDS). Retrieved on 2021-02-15 from: Doi: 10.24381/cds.adbb2d47

Holmén, J. (2017). *CEMENTA, SLITE, GRUNDTVATTENMODELL Fall 2041D*. (CemSliteGrvmod Ver 8). Retrieved on 2021-03-20 from: www.cementa.se/sites/default/files/assets/document/bb/ed/4._attenmodell.pdf

Hyndman, J. R., Athanasopoulos, G. (2018). *Forecasting: Principles and Practice* (2). <https://otexts.com/fpp2/AR.html>

Hipel, W. K., McLeod, I. (1994A). Chapter 18 Forecasting with Transfer Function-Noise Models. *Developments in Water Science* 45, 617-652. [https://doi.org/10.1016/S0167-5648\(08\)70680-4](https://doi.org/10.1016/S0167-5648(08)70680-4)

Hipel, W. K., McLeod, I. (1994B). Chapter 17 Constructing Transfer Function-Noise Models. *Developments in Water Science* 45, 573-616. [https://doi.org/10.1016/S0167-5648\(08\)70679-8](https://doi.org/10.1016/S0167-5648(08)70679-8)

Jain, A., & Kumar, M. A. (2007) Hybrid neural network models for hydrologic time series forecasting. *Applied Soft Computing*, 7, 585-592 Doi:10.1016/j.asoc.2006.03.02

Kasiviswanathan, K.S., Saravanan, S., Balamurugan, M., & Saravanan, K. (2016). Genetic programming based monthly groundwater level forecast models with uncertainty quantification. *Modeling Earth Systems and Environment* 2, (27). <https://doi.org/10.1007/s40808-016-0083-0>

Kotu, V. & Deshpande, B. (2019). Chapter 12 - Time Series Forecasting. *Data Science*.(2), 395-445. <https://doi.org/10.1016/B978-0-12-814761-0.00012-5>

Lam, N. S. N. (1983). Spatial interpolation methods: a review. *American Cartographer* 10 (2), 129-149. <https://doi.org/10.1559/152304083783914958>

Lastrapes, D. W. (n.d.). *Autoregressive Models*. Encyclopedia. Retrieved on 2021-02-03 from: <https://www.encyclopedia.com/social-sciences/applied-and-social-sciences-magazines/autoregressive-models>

Li, J., & Heap, A.D., (2008). *A Review of Spatial Interpolation Methods for Envi-*

- ronmental Scientists*. Geoscience Australia, Record 2008/23, 137 pp. ISBN 978 1 921498 28 2
- Lepot, M., Aubin, J-B., & Clemens, H. L. R. F. (2017) Interpolation in Time Series: An Introductory Overview of Existing Methods, Their Performance Criteria and Uncertainty Assessment. *Water*, 9(10), 1-20. Doi:10.3390/w9100796
- Lolcama, J.L., Cohen, H.A. & Tonkin, M.J. (2002). Deep Karst Conduits, Flooding, and Sinkholes: Lessons for the Aggregates Industry. *Engineering Geology*, 65 (2-3), 151-157. [https://doi.org/10.1016/S0013-7952\(01\)00122-3](https://doi.org/10.1016/S0013-7952(01)00122-3)
- Marshak, S. (2015). Earth: portrait of a planet (Fifth edition, International student edition). W.W. Norton & Company.
- Merriam-Webster. (n.d.). *Dynamic*. In Merriam-Webster.com dictionary. Retrieved on 2021-05-05 from: <https://www.merriam-webster.com/dictionary/dynamic>
- Mackay, J, D., Jackson, C, R., & Wang, L. (2014). A lumped conceptual model to simulate groundwater level time-series. *Environmental Modelling Software*, 61, 229-245. <https://doi.org/10.1016/j.envsoft.2014.06.003>
- Manziona, R, L., Knotters, M., Heuvelink G, B., Asmuth, J, R., & Camara, G. (2010). Transfer function-noise modeling and spatial interpolation to evaluate the risk of extreme (shallow) water-table levels in the Brazilian Cerrados. *Hydrogeology journal*, 18, 1927-1937. Doi: 10.1007/s10040-010-0654-5
- Mohanasundaram, S., Narasimhan, B. & Kumar, S. G. (2016). Transfer function noise modelling of groundwater level fluctuation using threshold rainfall-based binary-weighted parameter estimation approach. *Hydrological Sciences Journal*, 62 (1), 36-49. <https://doi.org/10.1080/02626667.2016.1171325>
- Moles, P. & Terry, N. (2005). The Handbook of International Financial Terms. *Coefficient of determination*. Oxford University Press. Published online: 2005. <https://doi.org/10.1093/acref/9780198294818.001.0001>
- Mosthaf, K., Brauns, B., Fjordbøge, A. S., Rohde, M. M., Jespersen, H., Bjerg, L. P., Binning, J. P., & Broholm, M. M. (2018). Conceptualization of flow and transport in a limestone aquifer by multiple dedicated hydraulic and tracer tests. *Journal of Hydrology*, 561, 532-546. <https://doi.org/10.1016/j.jhydrol.2018.04.011>
- Park, C. & Allaby, M. (2017). A Dictionary of Environment and Conservation. *Actual Evapotranspiration*. 3rd Edition. Oxford University Press. eISBN:9780191826320
- Pandey, S. V. & Bajpai, A. (2019) Predictive Efficiency of ARIMA and ANN Models: A Case Analysis of Nifty Fifty in Indian Stock Market. *International Journal of Applied Engineering Research*, 14 (2) ISSN: 0973-4562
- Peterson, T. & Fulton, S. (2019). Joint Estimation of Gross Recharge, Groundwater Usage, and Hydraulic Properties within HydroSight. *Groundwater*, 57 (6), 860-876. <https://doi.org/10.1111/gwat.12946>

- Peterson, T. J., & Western, A. W. (2014). Nonlinear time-series modeling of unconfined groundwater head. *Water Resources Research*, 50 (10), 8330- 8355. Doi:10.1002/2013WR014800.
- Phillips, G.M. (2003). *Interpolation and Approximation by Polynomials*. Springer Science+Business Media, New York. Doi:10.1007/978-0-387-21682-9.
- Salcedo. J. O., Hernandez. A. C., & Puente. J. F., (2009). Comparative analysis of time series techniques ARIMA and ANFIS to forecast Wimax traffic. *The Online Journal on Electronics and Electrical Engineering (OJEEE)*, 2(2)., 223-228. <https://doi.org/10.1186/1687-1499-2014-15>
- Salomon, D. (2006). *Spline Interpolation*. In: *Curves and Surfaces for Computer Graphics*. Springer. Springer, New York, NY. https://doi.org/10.1007/0-387-28452-4_5
- Sathya, R. & Abraham, A. (2013). Comparison of Supervised and Unsupervised Learning Algorithms for Pattern Classification. *International Journal of Advanced Research in Artificial Intelligence*, 2(2), 34-38. Doi:10.14569/IJARAI.2013.020206
- Stockholms Universitet (SU). (4 October 2018). *Tidsserier*. Statistiska institutonen. Retrieved on 2021-02-16 from: <https://www.statistics.su.se/forskning/tidsserier>
- Sveriges Geologiska Undersökning (SGU). (2013). *Bedömningsgrunder för grundvatten* (SGU-rapport 2013:01). Retrieved on 2021-03-06 from: <http://resource.sgu.se/produkter/sgurapp/s1301-rapport.pdf>
- SGU. (10 December 2020). *Så beräknar SGU aktuella grundvattennivåer*. Retrieved on 2021-03-01 from: <https://www.sgu.se/grundvatten/grundvattennivaer/om-grundvattennivaer/sa-beraknar-sgu-aktuella-grundvattennivaer/>
- SGU. (3 February 2021). *Grundvattennivåer, tidsserier*. Retrieved on 2021-02-21 from: <https://www.sgu.se/produkter/geologiska-data/oppna-data/grundvatten-oppna-data/grundvattennivaer-tidsserier/> © Sveriges Geologiska Undersökning
- SGU. (n.d.). *Jordlagerföljd, mätstation*. Retrieved on 2021-04-21 from: <https://www.sgu.se/grundvatten/grundvattennivaer/matstationer/>
- SMHI (26 Feb 2018). *Vattnets kretslopp-föreningar hydrologi, meteorologi och oceanografi*. Retrieved on 2021-02-20 from: <https://www.smhi.se/kunskapsbanken/hydrologi/vattnets-kretslopp-forenar-hydrologi-meteorologi-och-oceanografi-1.20615>
- Soliman. A-H. S., & Al-Kandari. M. A. (2010). 9 - Electric Load Modeling for Long-Term Forecasting *Electrical Load Forecasting*, 353-406 <https://doi.org/10.1016/B978-0-12-381543-9.00009-9>
- Sutton, S. R. (1992). Reinforcement learning. *Machine learning*, 8(2), 225-227. ISBN 978-1-4615-3618-5
- Thomas, C. G. (2021). *Re: Difference between the Actual Evapotranspiration (AET)*

- and Potential Evapotranspiration (PET)?. Retrieved on 2021-05-13 from: https://www.researchgate.net/post/Difference_between_the_Actual_Evapotranspiration_AET_and_Potential_Evapotranspiration_PET/60abd70ee092b230a025d004/citation/download
- United Nations (UN). (2015). *Water for a sustainable world*. (WWDR). Retrieved on 2021-04-26 from: www.unesco.org/new/fileadmin/MULTIMEDIA/HQ/SC/images/WWDR2015Facts_Figures_ENG_web.pdf
- United States Geological Survey (USGS). (23 November 2016). *Natural Processes of Ground-water and Surface-water Interaction*. Retrieved on 2021-04-20 from: https://pubs.usgs.gov/circ/circ1139/htdocs/natural_processes_of_ground.htm
- USGS. (n.d.A). *Aquifers and Groundwater*. Retrieved on 2021-04-05 from: https://www.usgs.gov/special-topic/water-science-school/science/aquifers-and-groundwater?qt-science_center_objects=0qt-science_center_objects
- USGS. (n.d.B). *Groundwater: What is Groundwater?* Retrieved on 2021-03-21 from: https://www.usgs.gov/special-topic/water-science-school/science/groundwater-what-groundwater?qt-science_center_objects=0qt-science_center_objects
- University of Zurich (UZH). (25 November 2020). *HBV-light Model*. Retrieved on 2021-05-30 from: <https://www.geo.uzh.ch/en/units/h2k/Services/HBV-Model.html>
- Von Asmuth, J. R, Bierkens M. F. P, & Maas, K. (2002). Transfer function-noise modeling in continuous time using predefined impulse response functions. *Water Resources Research*, 38(12), 23(1)-23(12). Doi:10.1029/2001WR001136, 2002.
- Wang, Z., Zhang, M., Wang, D., Song, C., Liu, M., Li, J., Lou, L., & Liu, Z. (2017). Failure prediction using machine learning and time series in optical network. *Optics Express*, 25 (16), 18553-18656. <https://doi.org/10.1364/OE.25.018553>
- Wilks, D. S. (2011). Chapter 9- Time Series. *International Geophysics*, 100, 395-456. <https://doi.org/10.1016/B978-0-12-385022-5.00009-9>
- Yi, M,-J., & Lee, -K,-K. (2004). Transfer function-noise modelling of irregularly observed groundwater heads using precipitation data. *Journal of Hydrology*, 288 (3-4), 272-287. Doi:10.1016/j.jhydrol.2003.10.020
- Yihdego, D.Y., Webb, J. (2010). Characterizing groundwater dynamics using Transfer Function-Noise and autoregressive modeling in Western Victoria, Australia. *WATER AND GEOSCIENCE*. La Trobe University. ISBN: 978-960-474-160-1
- Zarco-Perello, S., & Simões, N. (2017). Ordinary kriging vs inverse distance weighting: spatial interpolation of the sessile community of Madagascar reef, Gulf of Mexico. PeerJ 5:e4078; Doi:10.7717/peerj.4078.
- Zotarelli,L., Dukes. M. D., Romero, C. C., Migliaccio, K. W., & Morgan, K. T. (n.d.). *Step by Step Calculation of the Penman-Monteith Evapotranspiration (FAO-56 Method)*. AE459. IFAS Extension. University of Florida. Retrieved on 2021-04-06 from: <https://edis.ifas.ufl.edu/pdf>

8 Appendices

A Appendix

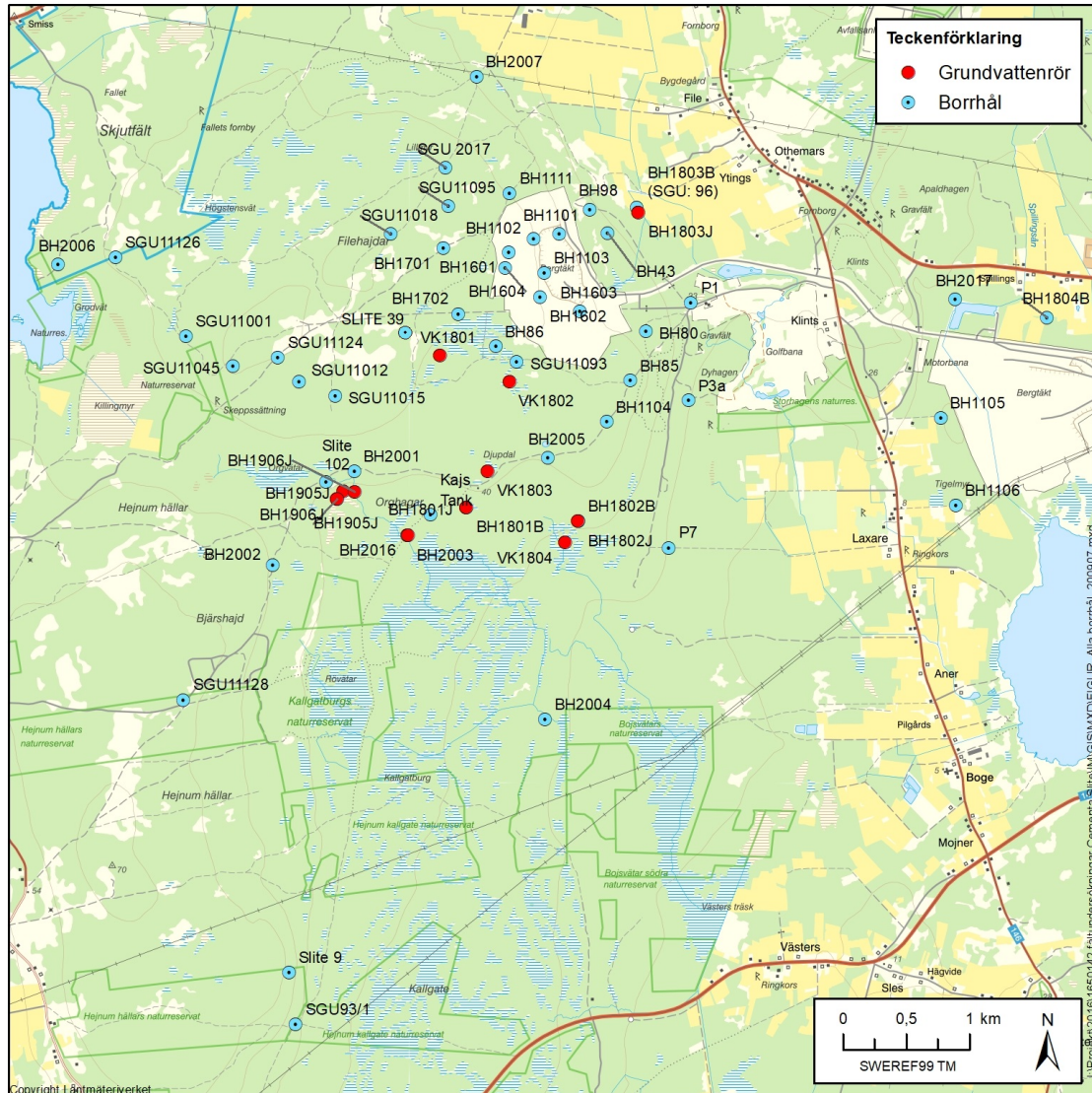


Figure 8.34: A map containing all bores in the area

Table 3: Boreholes containing automatically recorded hourly data

Boreholes	Contains Hourly Data
BH43	Yes
BH86	Yes
BH98	Yes
BH96	Yes
BH80	Yes
Slite39	No
Slite9	Yes
SGU11001	Yes
BH2006	No

Table 4: Distance from all boreholes to different locations

Boreholes	Distance to body of water	Distance to quarry	Distance to municipal pumps
BH43	3400m	300m	1400m
BH86	4900m	250m	350m
BH98	4170m	240m	1140m
BH96	4500m	570m	970m
BH80	4600m	230m	370m
Slite39	2930m	800m	1970m
Slite9	5100m	5260m	4760m
SGU11001	930m	3460m	4900m
BH2006	230m	3500m	5000m

Table 5: Distance from all boreholes to all municipal and commercial pumps

Borehole	P1	P2	P3	P3a	P4	P5	P6	FH1	FH2	VB1	VB2
BH43	950m	1236m	1339m	1453m	1725m	2035m	2150m	513m	580m	4921m	5078m
BH86	1555m	1689m	1617m	1578m	1647m	1762m	1830m	743m	808m	5630	5776m
BH98	1570m	1390m	1623m	1682m	1949m	2252m	2365m	580m	538m	5109m	5267m
BH96	953	1273m	1416m	1558m	1854m	2187m	2309m	820m	866m	4760m	4919m
BH80	371m	527m	541m	671m	890m	1210m	1333m	818m	1066m	4464m	4482m
Slite39	2023m	2126m	2028m	1960m	1968m	2007m	2050m	1219m	1236m	6044m	6186m
Slite9	6708m	5922m	5711m	5526m	5244m	4914m	4811m	5902m	6019m	8509m	8585m
SGU11001	3979m	4130m	4052m	3994m	4001m	4007m	4032m	2973m	2840m	8069m	8214m
BH2006	5002m	5183m	5125m	5084m	5116m	5142m	5173m	3944m	3772m	9128m	9276m

B Appendix

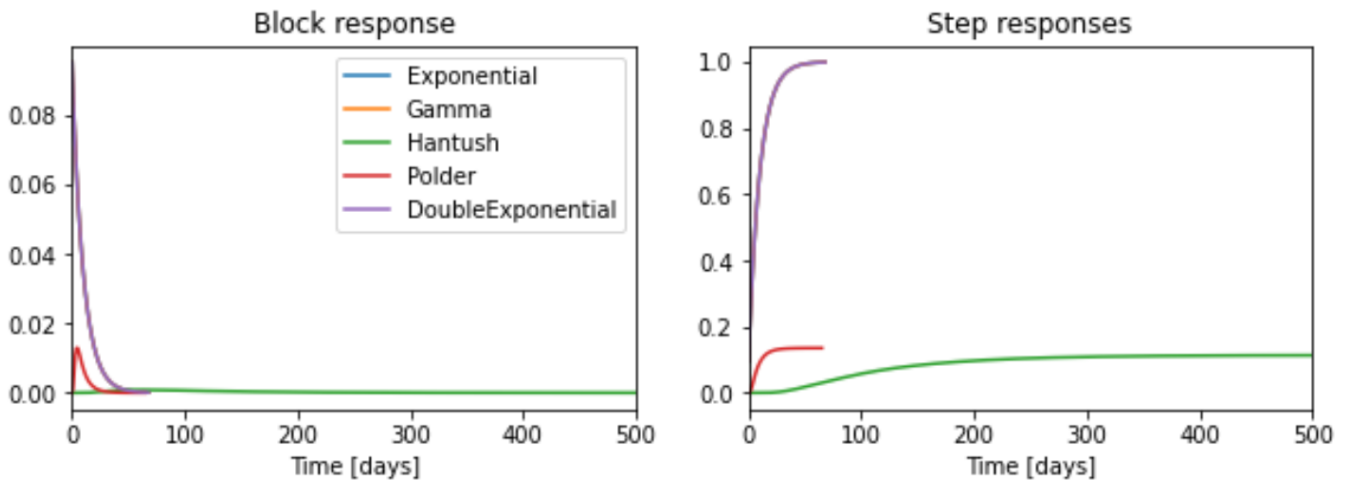


Figure 8.35: a) All available IRFs' plotted against block reponse b) All available IRFs' plotted against step responses

Assumptions for leaky aquifer (Aqtesolve,n.d.)

1. The aquifer is assumed to be leaky and confined.
2. The aquifer is assumed to be of homogenous, isotropic material with a uniform thickness.
3. The area of the aquifer or aquitard is assumed to be infinite.
4. The water flow is assumed to be unsteady.
5. The aquitard consist of water with a vertical flow.
6. The pumping wells over the aquifer-area are either fully or partially penetrating.
7. The water table and piezometric surface is horizontal before pumping has taken place.
8. When pumping has taken place it is assumed to has been pumped at a constant discharge rate over the influenced area where the test has taken place.
9. The well reaches through the aquifer and obtain water from the horizontal flow.
10. The drawdown in the aquifer or aquitard is negligible.

C Appendix

These scripts are created by (Collenteur, 2020), but have been edited or changed for the use in this thesis.

```

import numpy as np
import pandas as pd
import pastas as ps
import matplotlib.pyplot as plt

#Import groundwater, precipitation and evaporation time-series.
head = pd.read_csv('PATH.csv', parse_dates=['date'], index_col='date', squeeze=True)
rain = pd.read_csv('PATH.csv', parse_dates=['date'], index_col='date', squeeze=True)
evap = pd.read_csv('PATH.csv', parse_dates=['date'], index_col='date', squeeze=True)

#Calculate and plot the recharge
recharge = rain - evap
plt.figure()
recharge.plot(label='Recharge', figsize=(10, 4))
plt.xlabel('Time [years]')
plt.ylabel('Recharge (m/year)');

#Create a model object for groundwater data and add recharge using either SM1 or SM2.
m1 = ps.Model(head)

#Either SM1,
sm1 = ps.StressModel(recharge, ps.SPECIFY IRF, name='recharge', settings="evap")
m1.add_stressmodel(sm1)

#or SM2
sm2 = ps.StressModel2([rain, evap], ps.SPECIFY IRF, name='rainevap', settings=("prec","evap"))
m1.add_stressmodel(sm2)

#Linear trend
tm = ps.LinearTrend(start="DATE", end="DATE", name="trend")
m1.add_stressmodel(tm)

# Solve the model
m1.solve(noise=False, tmin='DATE', tmax='DATE')
m1.plot()

#####

Pumping: WellModel

#Create model
m1 = ps.Model(head)

#Import X number of time-series contain pumping data.
PumpX = pd.read_csv('PATH', parse_dates=['date'], index_col='date', squeeze=True)

#Specify pump name and corresponding distances in meters.
extractions = [PumpX,.....]
distances = [DISTANCES,.....]

```

```

# Add extractions in WellModel and specify StressModel for recharge and any linear trends.
wm = ps.WellModel(extractions, rfunc=ps.HantushWellModel, name="extractions",
                  distances=distances, settings="well")
ml.add_stressmodel(wm)

sm1 = ps.StressModel(recharge, ps.SPECIFY IRF, name='recharge', settings="evap")
ml.add_stressmodel(sm1)

#Linear trend
tm = ps.LinearTrend(start="DATE", end="DATE", name="trend")
ml.add_stressmodel(tm)

ml.solve(noise=False, tmin='DATE', tmax='DATE', solver=ps.LmfitSolve)
ml.plot(figsize=(12, 4));

#####

Pumping: LH model.

#Create model and import X number of time-series contain pumping data, specify pump name.
ml = ps.Model(head)
PumpX = pd.read_csv('PATH', parse_dates=['date'], index_col='date', squeeze=True)
extractions = [PumpX,.....]

# Loop and add each extraction separately
for i, ex in enumerate(extractions):
    sm = ps.StressModel(ex, rfunc=ps.Hantush, name=f"ex-{i}", settings="well", up=False)
    ml.add_stressmodel(sm)

# Add the StressModel for the net recharge and Linear trends if needed.
sm1 = ps.StressModel(recharge, ps.SPECIFY IRF, name='recharge', settings="evap")
ml.add_stressmodel(sm1)

#Linear trend
tm = ps.LinearTrend(start="DATE", end="DATE", name="trend")
ml.add_stressmodel(tm)

ml.solve(noise=False, tmin='DATE', tmax='DATE')
ml.plot(figsize=(12, 4));

```

D Appendix

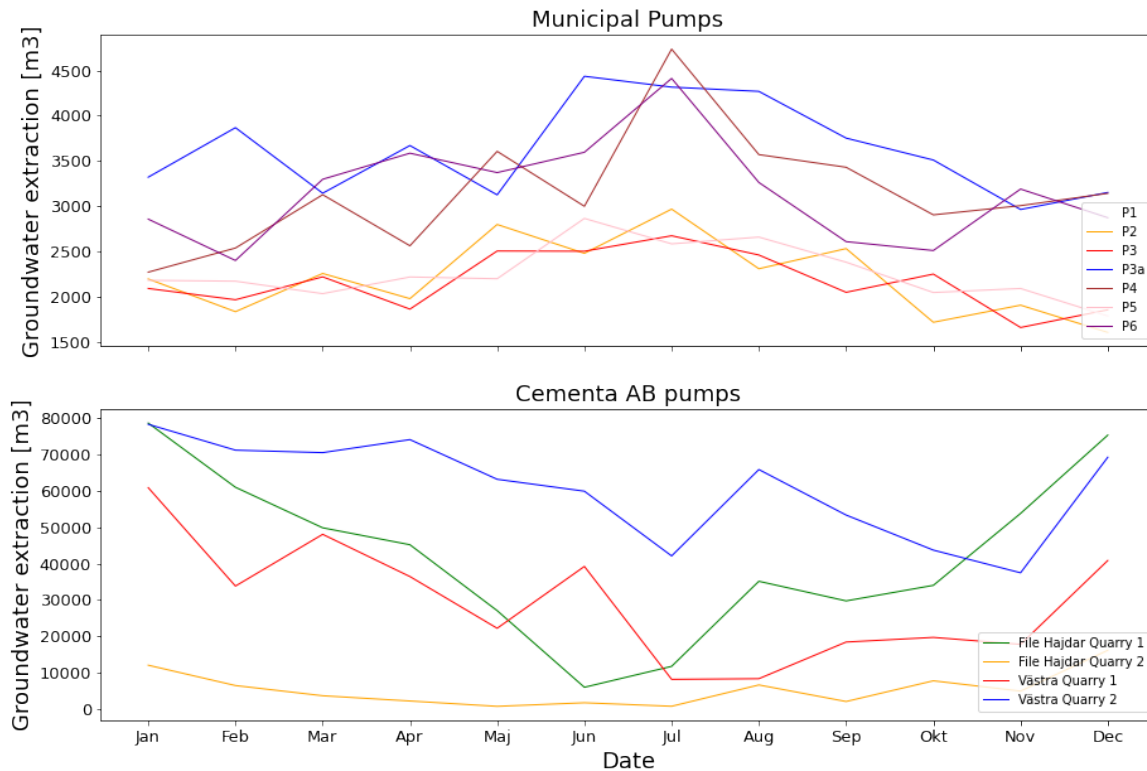


Figure 8.36: a) Average extraction per year for all municipal pumps b) Average extraction per year for Cementa AB pumps

Time-series related information regarding each bore

- **BH86:** The time-series for this bore has semi-consistent observation head data over the 40 year period, with the exception of a 10 year gap that is characteristic of the area.
- **BH96:** The dependent time-series for this bore contains negative values, therefore the entire time-series was raised by 10 meters.
- **BH80:** The dependent time-series for this bore contains negative values, therefore the entire time-series was raised by 20 meters.
- **Slite 9:** The dependent time-series for this bore contains an interval of 20 years with no recorded groundwater levels.
- **BH2006:** A rough estimate was made for when the tin can could have blocked this borehole, for the years where obvious abnormal fluctuations could be observed head data was removed.

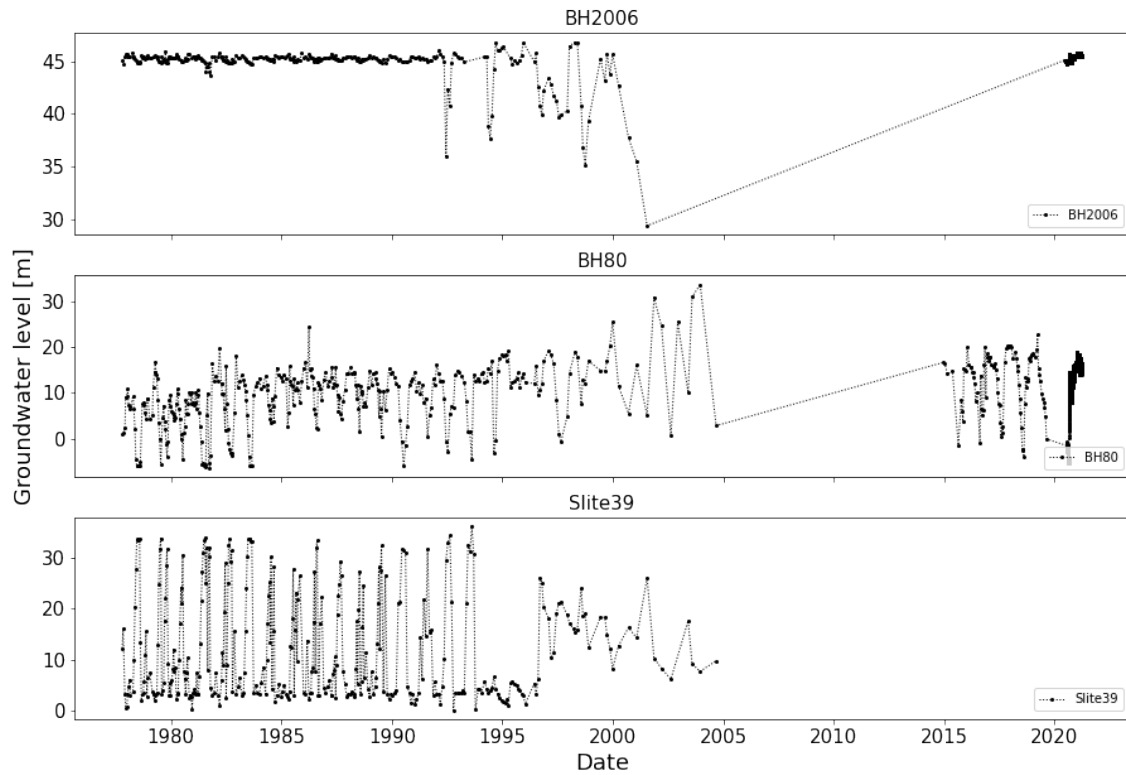


Figure 8.37: Recorded time-series containing a) BH2006 b) BH80 c) Slite 39

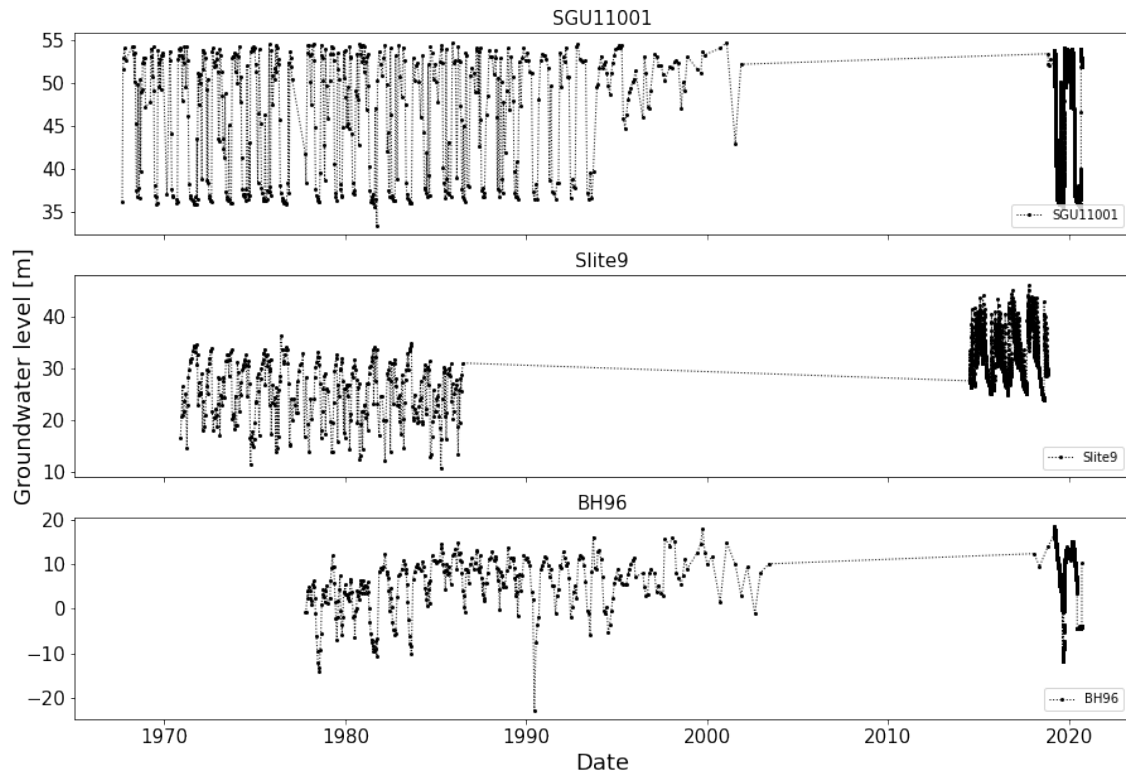


Figure 8.38: Recorded time-series containing a) SGU11001 b) Slite 9 c) BH96

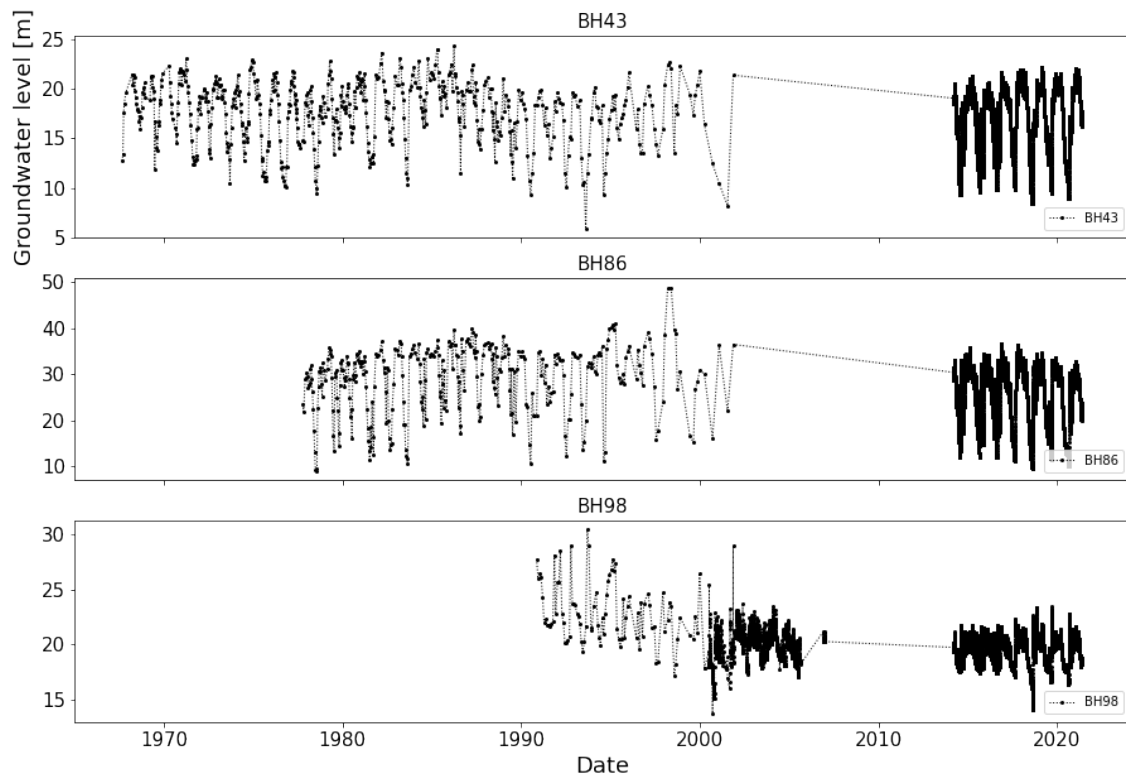


Figure 8.39: Recorded time-series containing a) BH43 b) BH86 c) BH98

E Appendix

The average wind speed required has to be measured at 2 meters height from the ground, u_2 , as SMHI measures the average wind speed at 10 meters from the ground Eq. (8.1) is used to adjust it.

$$u_2 = u_h \frac{4.87}{\ln(67.8h - 5.42)} \quad (8.1)$$

u_2 is the calculated average wind speed [m/s] and u_h is the measured average wind speed [m/s] at elevation, h [m]. The height of the actual measurement elevation is denoted with h , and is 10 meters for the wind speed data from SMHI.

Calculating the slope of saturation vapor pressure curve, Δ [$\frac{kPA}{^\circ C}$], is the next step in the process. This slope describes the relationship between the temperature and the saturation vapor pressure and is calculated according to Eq. (8.2).

$$\Delta = \frac{4098 * (0.6108 * \exp(\frac{17.27 * T_{mean}}{T_{mean} + 237.3}))}{(T_{mean} + 237.3)^2} \quad (8.2)$$

where T_{mean} is the mean daily temperature [$^\circ C$].

Thereafter, the psychrometric constant, γ [$\frac{kPA}{^\circ C}$], is calculated. This constant represents the ratio of energy needed to increase the temperature one degree for one unit mass of moist air under constant pressure, also known as specific heat at constant pressure, to latent heat of vaporization (Zotarelli et al., n.d.), it's calculated according to Eq. (8.3).

$$\gamma = \frac{C_p * P}{\epsilon * \lambda} = 0.000665P \quad (8.3)$$

where λ is the latent heat of vaporizing, which is estimated to be 2.45 [$\frac{MJ}{kg}$], ϵ is a ratio of the molecular weight of water and dry air; 0.622 [-]. C_p is the specific heat at constant pressure, for average atmospheric conditions a value of $1.013 * 10^{-3}$ [$\frac{MJ}{kg^\circ C}$] is used. Finally, P is the average atmospheric pressure at 20 $^\circ C$ and is 100.76 [kPa] (Zotarelli et al., n.d.). Both atmospheric pressure, P , and specific heat at constant pressure, C_p , are based on temperature and humidity, respectively, and change over time. Instead of calculating both of the parameters for every single day, averaged values for the entire 40 year period was used for both. This is partially due to the time-series for humidity data from Visby had an irregular measurement frequency and contained large intervals with no recorded data. Moreover, the difference between the calculated values and averaged values was determined to be small enough to be considered negligible.

FAO-56 uses three terms as auxiliary calculations for the radiation (delta term; DT [-]) in Eq. (8.4), wind (Psi term; PT [-]) in Eq. (8.5) and temperature (temperature

term; TT [-]) parameters in Eq. (8.6). These terms are calculated individually and are used to facilitate the calculation process.

$$DT = \frac{\Delta}{\Delta + \gamma(1 + 0.34u_2)} \quad (8.4)$$

$$PT = \frac{\gamma}{\Delta + \gamma(1 + 0.34u_2)} \quad (8.5)$$

$$TT = \frac{900}{T_{min} + 273} \quad (8.6)$$

where T_{min} is the daily minimum temperature [$^{\circ}\text{C}$].

After the auxiliary terms have been calculated, the mean saturation vapor pressures, e_s [kPA], and the actual vapor pressures, e_a [kPA], are derived. The mean saturation vapor pressure, e_s , as seen in Eq.(8.9), is based on the daily air temperature and acquired using $e_{sT_{min}}$ [kPA] and $e_{sT_{max}}$ [kPA], see Equations (8.7) and (8.8), respectively.

$$e_{sT_{min}} = 0.6108 * \exp\frac{17.27T_{min}}{T_{min} + 237.3} \quad (8.7)$$

$$e_{sT_{max}} = 0.6108 * \exp\frac{17.27T_{max}}{T_{max} + 237.3} \quad (8.8)$$

$$e_s = \frac{(e_{sT_{min}} + e_{sT_{max}})}{2} \quad (8.9)$$

where, T_{max} represents the daily maximum temperature [$^{\circ}\text{C}$].

As mentioned the relative humidity data from SMHI contained missing data, and could not be considered adequate enough to be used. For such instances, it is suggested by the guide to use $e_{sT_{min}}$ in place of e_a . This assumes that the air is saturated with water vapor at T_{min} and thus the relative humidity is considered to be nearly 100 % (Zotarelli et al., n.d.). Therefore, e_a is replaced by $e_{sT_{min}}$.

In order for the solar radiation to be accurately represented in the ET_o calculations, the inverse distance between northern Gotland's relative position on earth and the sun needs to be derived, d_r [-], this is done using Eq. (8.10). The solar declination, δ [-], over a year is derived through Eq. (8.11), this parameter is then used to calculate the changing sunset hour angle, ω_s [rad] according to Eq. (8.12), over a year.

$$d_r = 1 + 0.033\cos\frac{2\pi}{365}J \quad (8.10)$$

$$\delta = 0.409\sin\left[\frac{2\pi}{365}J - 1.39\right] \quad (8.11)$$

$$\omega_s = \arccos[-\tan(\varphi)\tan(\delta)] \quad (8.12)$$

where J represents the day of the year, relative to the date [-] and φ is the latitudinal coordinate of northern Gotland [rad].

Equations (8.10), (8.11) and (8.12) are used to calculate the extraterrestrial radiation, R_a [$\frac{MJ}{m^2 day}$] as presented in Eq. (8.13). Which in turn is used to calculate the clear sky solar radiation, R_{so} [$\frac{MJ}{m^2 day}$], using Eq. (8.14).

$$R_a = \frac{24 * 60}{\pi} G_{sc} * d_r [(\omega_s \sin(\varphi) \sin(\delta)) + (\cos(\varphi) \cos(\delta) \sin(\omega_s))] \quad (8.13)$$

$$R_{so} = (0.75 + 2 * 10^{-5} z) R_a \quad (8.14)$$

where; G_{sc} is a solar constant, 0.0820 [$\frac{MJ}{m^2 min}$] and z is the elevation above sea level [m].

Net long-wave radiation, R_{nl} [$\frac{MJ}{m^2 day}$], is calculated using Eq. (8.7) and (8.14) as presented in (8.15).

$$R_{nl} = \sigma \left[\frac{(T_{max} + 273.16)^4 + (T_{min} + 273.16)^4}{2} (0.34 - 0.14 \sqrt{e_a}) * [1.35 \frac{R_s}{R_{so}} - 0.35] \right] \quad (8.15)$$

where, σ , is the Stefan-Boltzmann's constant, $4.903 * 10^{-9}$ [$\frac{MJ}{K^4 m^2 day}$] and R_s [$\frac{MJ}{m^2 day}$] is the incoming solar radiation retrieved from Copernicus database. Net shortwave radiation R_{ns} [$\frac{MJ}{m^2 day}$] is calculated using Eq. (8.16);

$$R_{ns} = (1 - \alpha) R_s \quad (8.16)$$

where, α is a canopy reflection coefficient; 0.23 [-] for grass. Net shortwave radiation, R_{ns} [Eq. (8.16)] and net long-wave radiation, R_{nl} [Eq. (8.15)] in combination provide net radiation, R_n [$\frac{MJ}{m^2 day}$] [Eq. (8.17)];

$$R_n = R_{ns} + R_{nl} \quad (8.17)$$

In order to convey the net radiation, R_n , in terms of evaporation, it is revised using Eq. (8.18) to calculate R_{ng} [$\frac{MJ}{m^2 day}$];

$$R_{ng} = 0.408 R_n \quad (8.18)$$

ET_0 , seen in Eq.(8.21), can be calculated by the addition of ET_{rad} [$\frac{mm}{day}$] and ET_{wind} [$\frac{mm}{day}$], which are derived using Eq. (8.19) and (8.20).

$$ET_{rad} = DT * R_{ng} \quad (8.19)$$

$$ET_{wind} = PT * TT(e_s - e_a) \quad (8.20)$$

$$ET_o = ET_{wind} + ET_{rad} \quad (8.21)$$

F Appendix

Table 6: Time-interval for all linear trends applied

Borehole	From	To
BH43	2000-01-01	2005-01-01
BH86	2000-01-01	2020-01-01
BH98	1993-01-01	2001-01-01
BH96	1980-01-01	1986-01-01
BH80	1980-01-01	1987-01-01
Slite39	-	-
Slite 9	-	-
SGU11001	-	-
BH2006	-	-

Table 7: R_{adj}^2 for simulated results before and after using linear trends

Borehole	Before Trend	Linear	After Linear Trend
BH43	0.73		0.75
BH86	0.74		0.77
BH98	0.43		0.57
BH96	0.48		0.61
BH80	0.46		0.61

G Appendix

Table 8: R_{adj}^2 for simulated results using two different StressModels with the AR(1) noise model

Borehole	StressModel 1		StressModel 2	
	Gamma	Exponential	Gamma	Exponential
BH43	0.60	0.51	0.60	0.61
BH86	0.66	0.58	0.64	0.65
BH98	0.32	0.53	0.29	0.50
BH96	0.50	0.54	0.5	0.54
BH80	0.58	0.59	0.58	0.59
Slite 39	0.48	0.52	0.48	0.52
Slite 9	0.53	0.63	0.39	0.63
SGU11001	0.46	0.39	0.46	0.39
BH2006	0.53	0.52	0.53	0.52

Table 9: R_{adj}^2 for simulated results using both StressModels but without the AR(1) noise model

Borehole	StressModel1		StressModel2	
	Gamma	Exponential	Gamma	Exponential
BH43	0.75	0.75	0.75	0.75
BH86	0.77	0.76	0.77	0.76
BH98	0.56	0.56	0.57	0.57
BH96	0.61	0.60	0.61	0.60
BH80	0.61	0.60	0.61	0.60
Slite 39	0.53	0.53	0.53	0.53
Slite 9	0.71	0.70	0.71	0.70
SGU11001	0.47	0.46	0.47	0.46
BH2006	0.53	0.52	0.53	0.53

Table 10: R_{adj}^2 for all simulated scenarios for LH model without the AR(1) noise model

	StressModel1		StressModel2	
BH43	0.75	0.75	0.75	0.75
BH86	0.78	0.77	0.78	0.65
BH98	0.62	0.62	0.63	0.62
BH96	0.61	0.68	0.66	0.67
BH80	0.64	0.63	0.61	0.60

Table 11: R_{adj}^2 for all simulated scenarios for WellModel without the AR(1) noise model

	StressModel1		StressModel2	
BH43	0.74	0.74	0.74	0.74
BH86	0.78	0.77	0.78	0.77
BH98	0.63	0.60	0.63	0.60
BH96	0.60	0.63	0.60	0.59
BH80	0.61	0.60	0.61	0.60

H Appendix

[7]:

	Checks	Statistic	P-value	Reject H0
Shapiro	Normality	0.52	0.00	True
D'Agostino	Normality	2079.92	0.00	True
Runs test	Autocorr.	-18.39	0.00	True
Stoffer-Toloi	Autocorr.	852.80	0.00	True

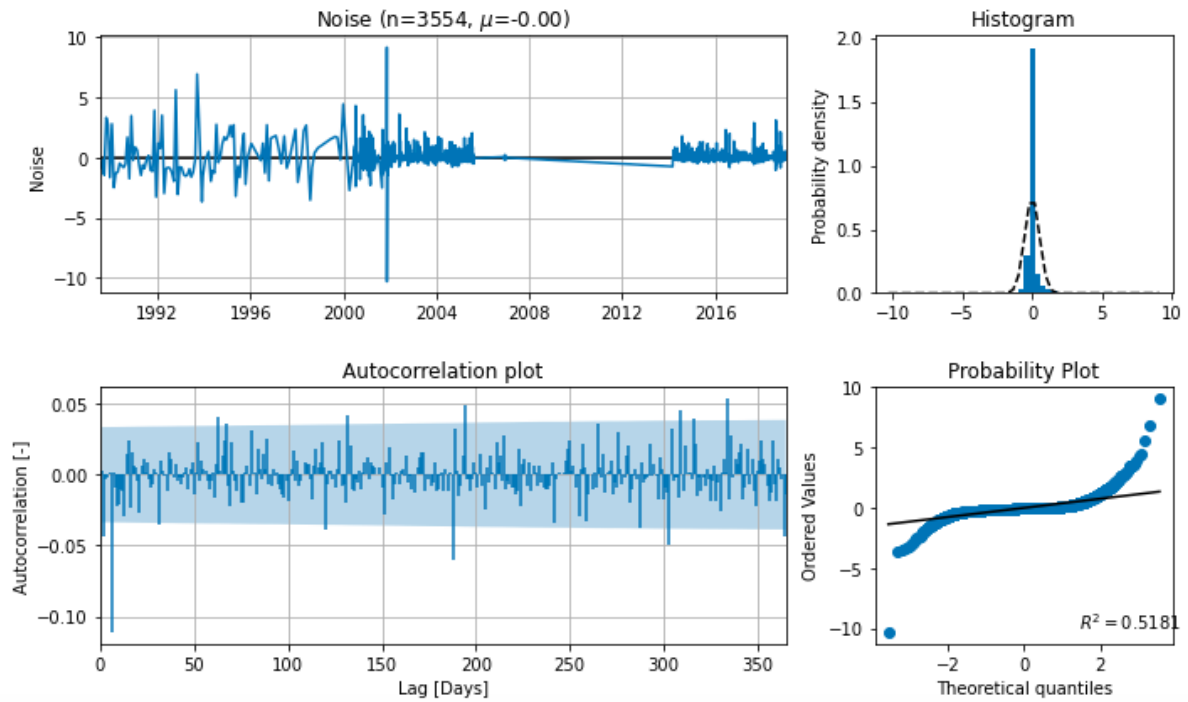


Figure 8.40: Diagnostic hypothesis tests with AR(1) Noise Model - BH98

[5]:

	Checks	Statistic	P-value	Reject H0
Shapiro	Normality	0.96	0.00	True
D'Agostino	Normality	418.70	0.00	True
Runs test	Autocorr.	-49.42	0.00	True
Stoffer-Toloi	Autocorr.	15660.64	0.00	True

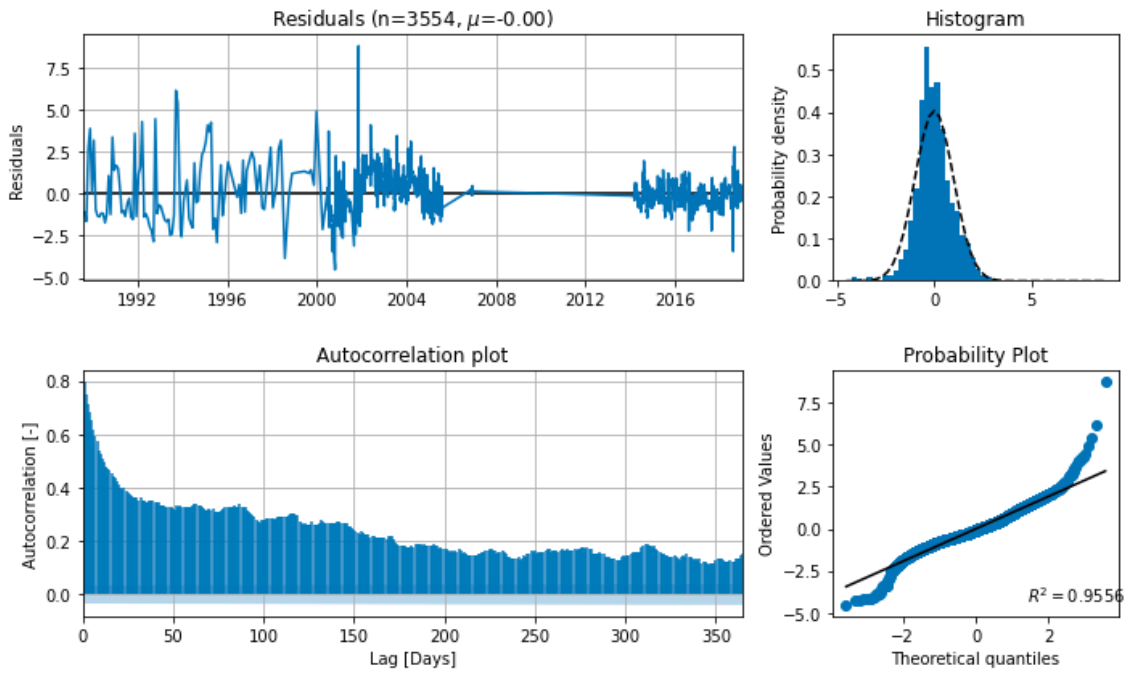


Figure 8.41: Diagnostic hypothesis tests without AR(1) Noise Model - BH98

I Appendix

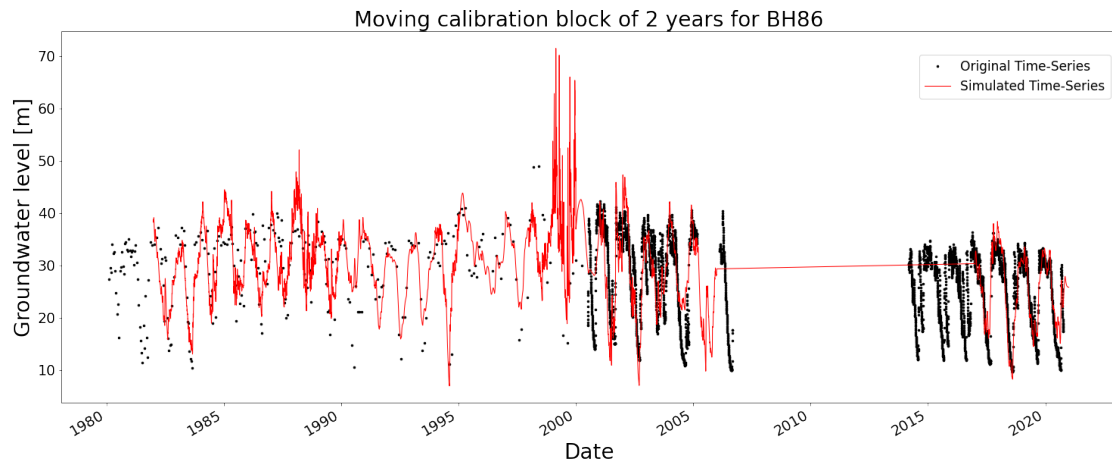


Figure 8.42: A moving calibration block of 2 years through the time-series of BH86

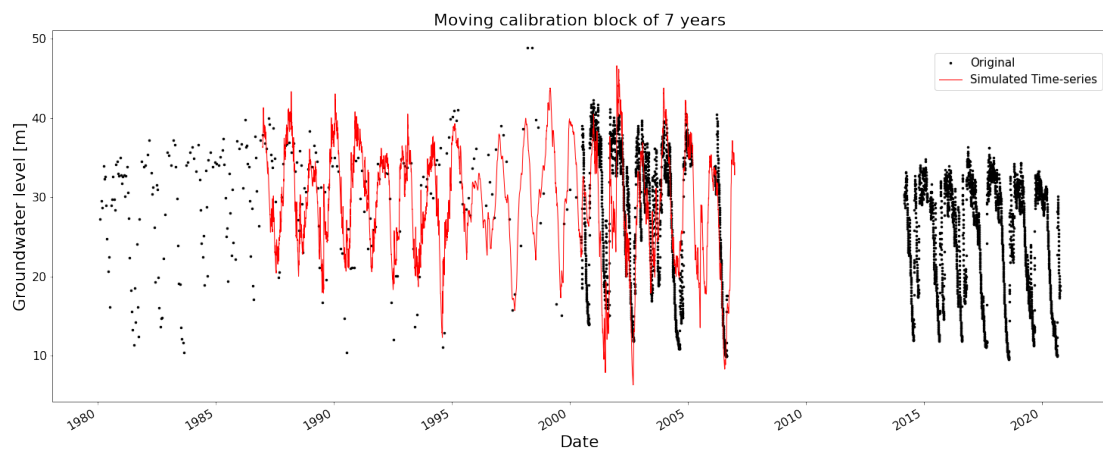


Figure 8.43: A moving calibration block of 7 years through the time-series of BH86

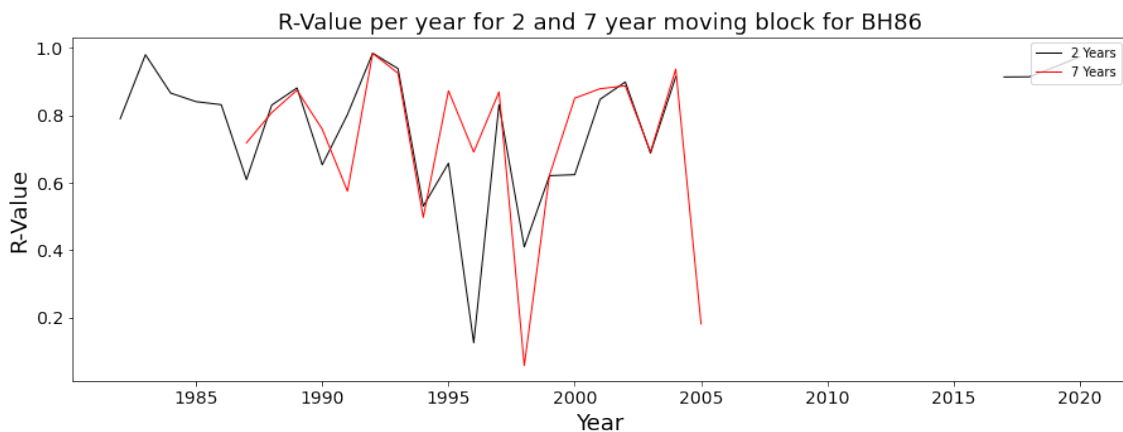


Figure 8.44: The year-to-year Pearson coefficient of correlation for two and seven year moving block simulations BH86

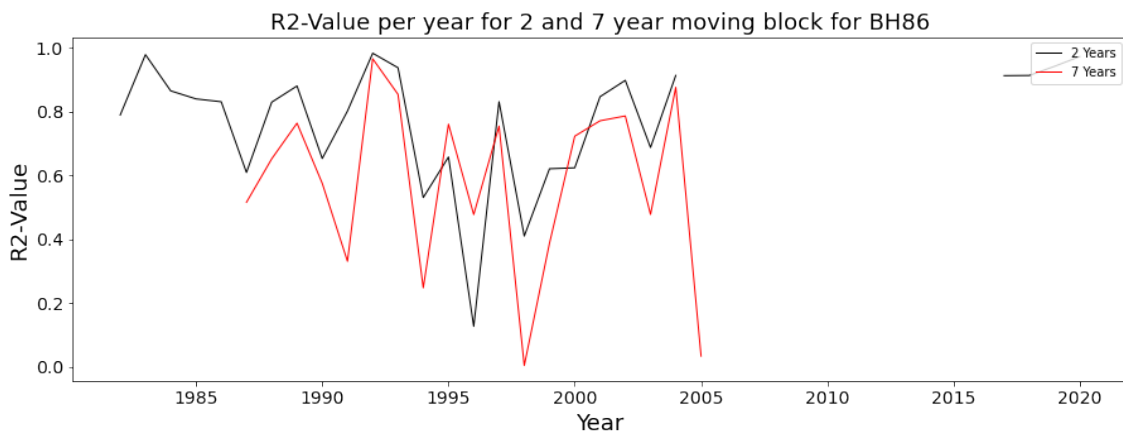


Figure 8.45: The coefficient of correlation for all simulated years with a calibration period of 2 and 7 years for BH86

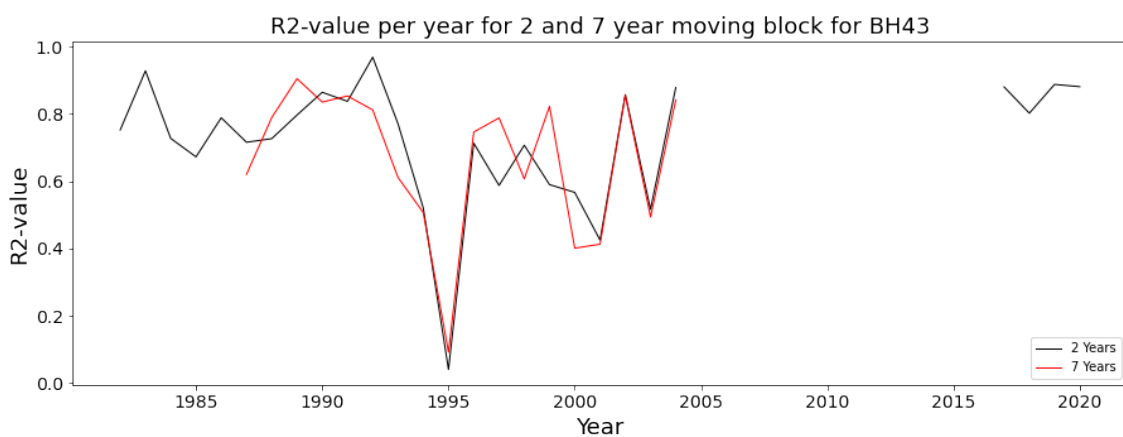


Figure 8.46: The coefficient of correlation for all simulated years with a calibration period of 2 and 7 years for BH43

J Appendix

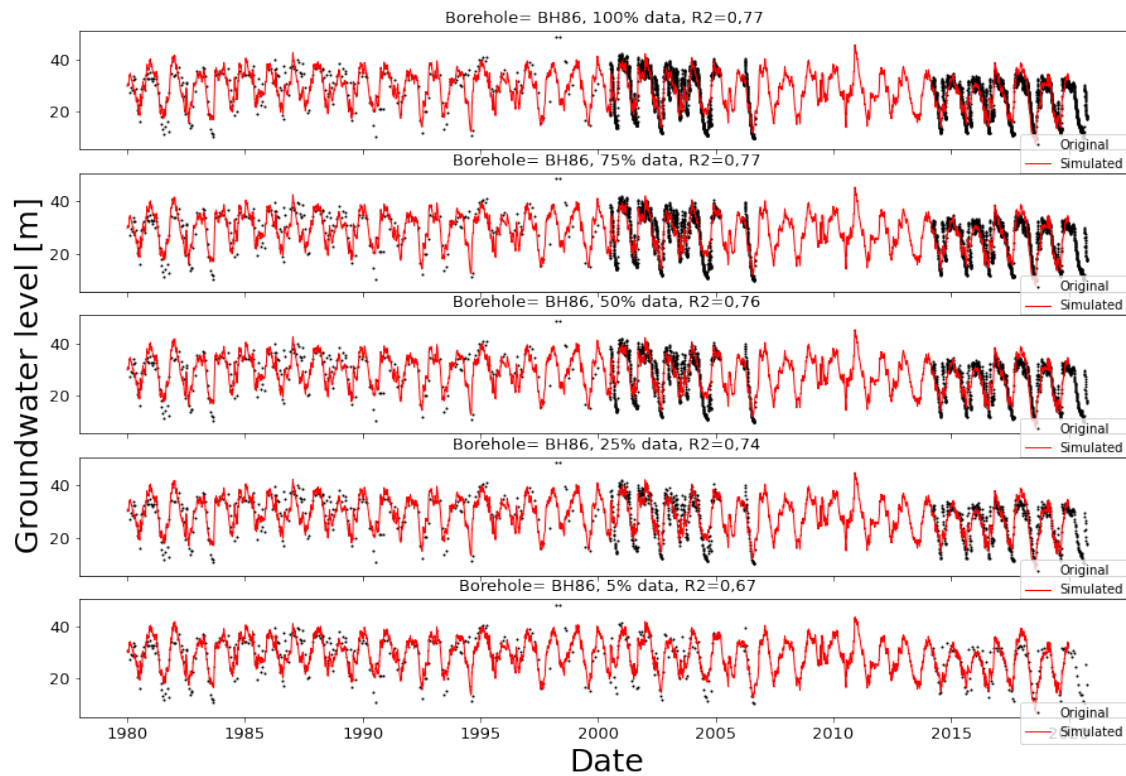


Figure 8.47: Systematic percentile removal of automatic measurements for BH86 of a) 100% b) 75% c) 50% d) 25% and e) 5%

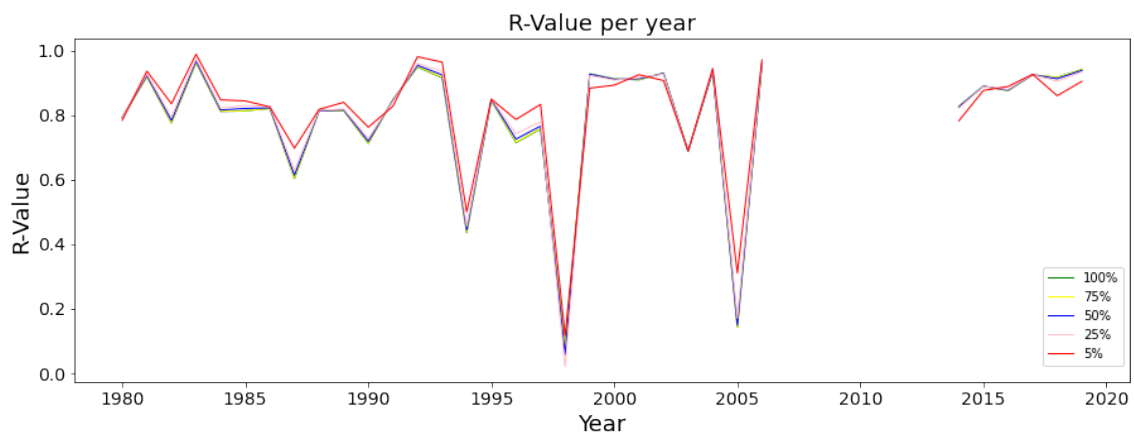


Figure 8.48: BH86: Correlation between simulated and recorded groundwater levels for each individual year

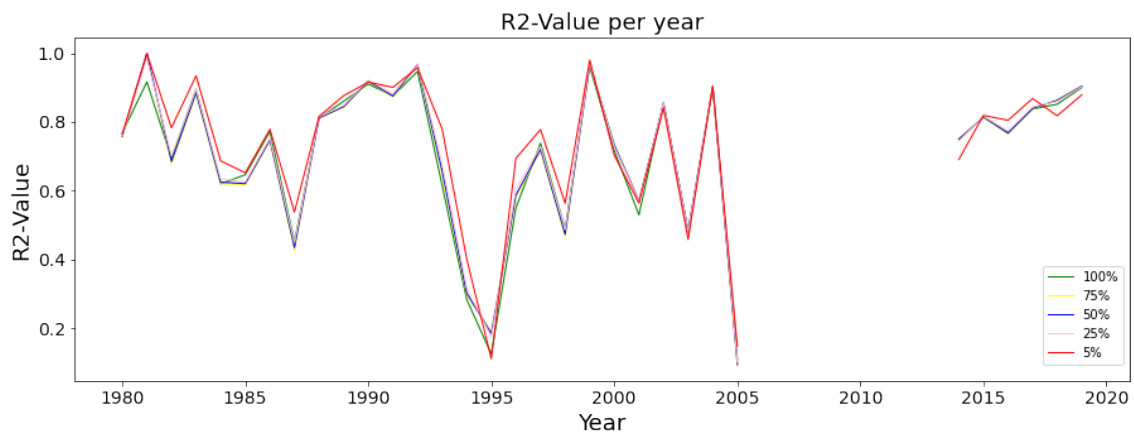


Figure 8.49: BH43: Coefficient of determination between simulated and recorded groundwater levels for each individual year

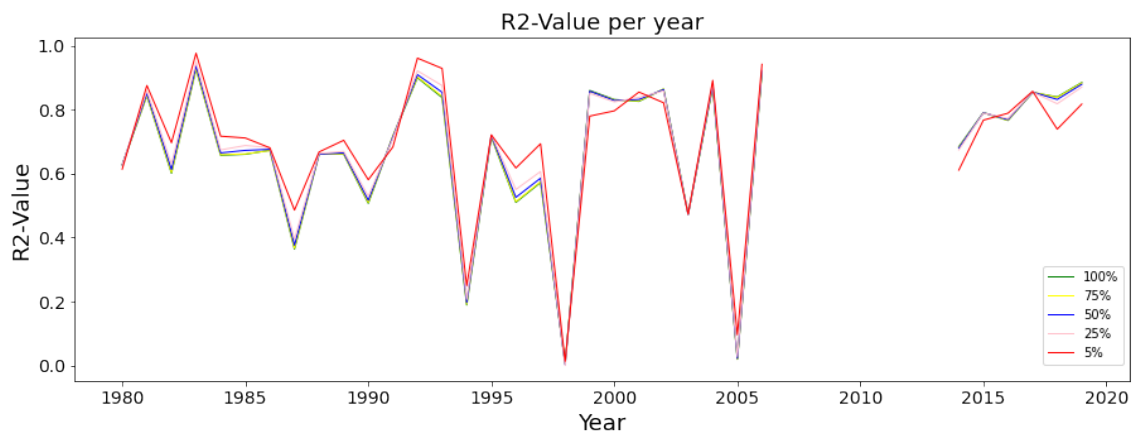


Figure 8.50: BH86: Coefficient of determination between simulated and recorded groundwater levels for each individual year

K Appendix

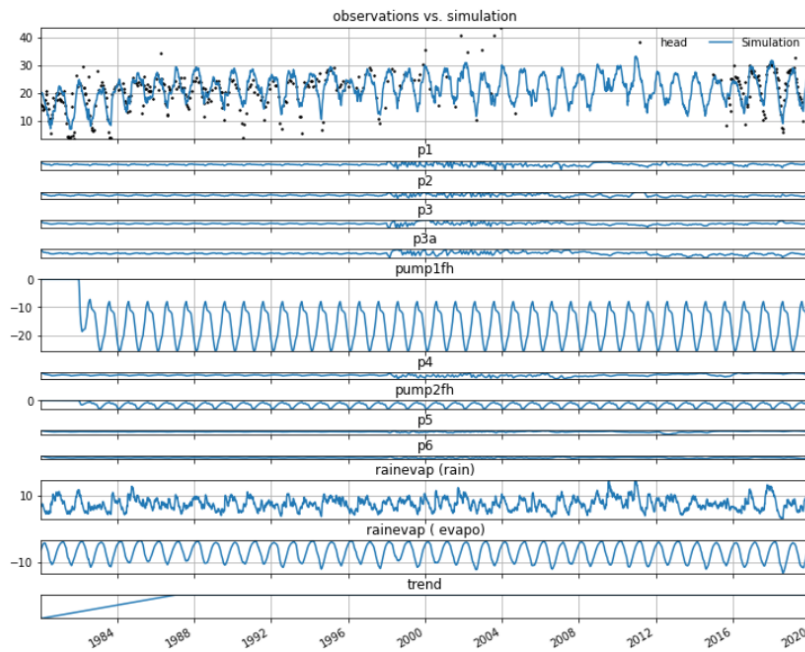


Figure 8.51: Diagnostics of the influences of pumps for BH80 in WellModel

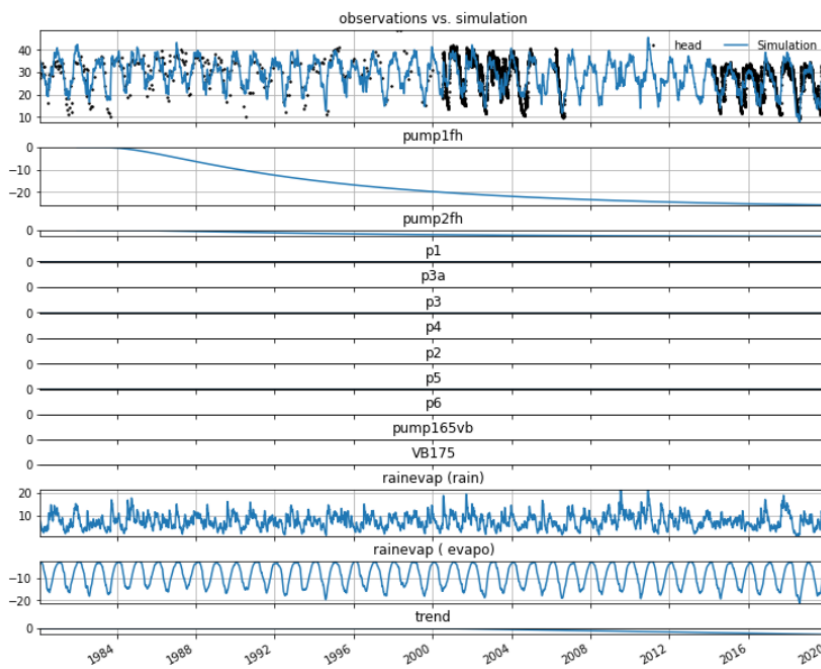


Figure 8.52: Diagnostics of the influences of pumps for BH86 in WellModel

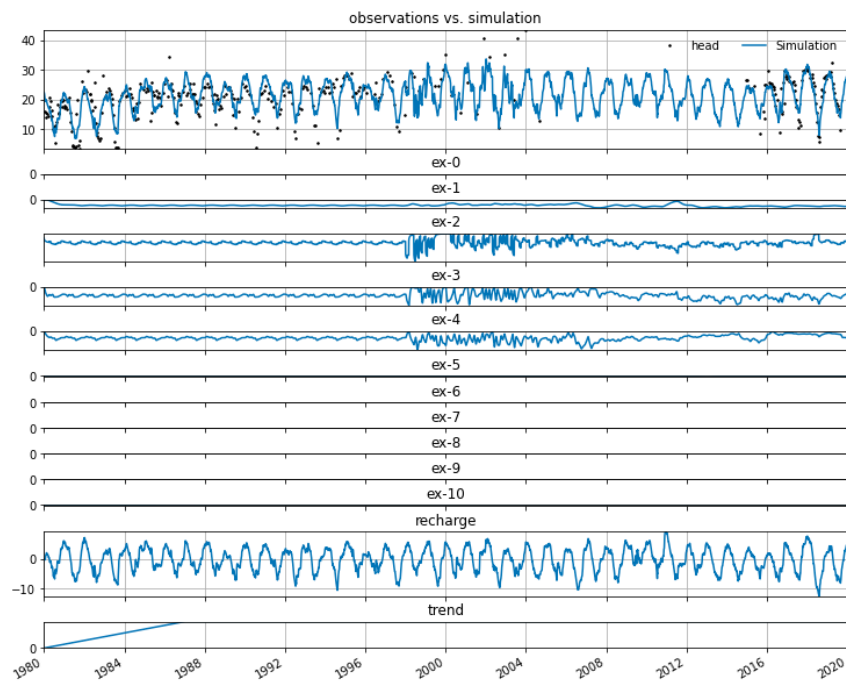


Figure 8.53: Diagnostics of the influences of pumps for BH80 in LH Model

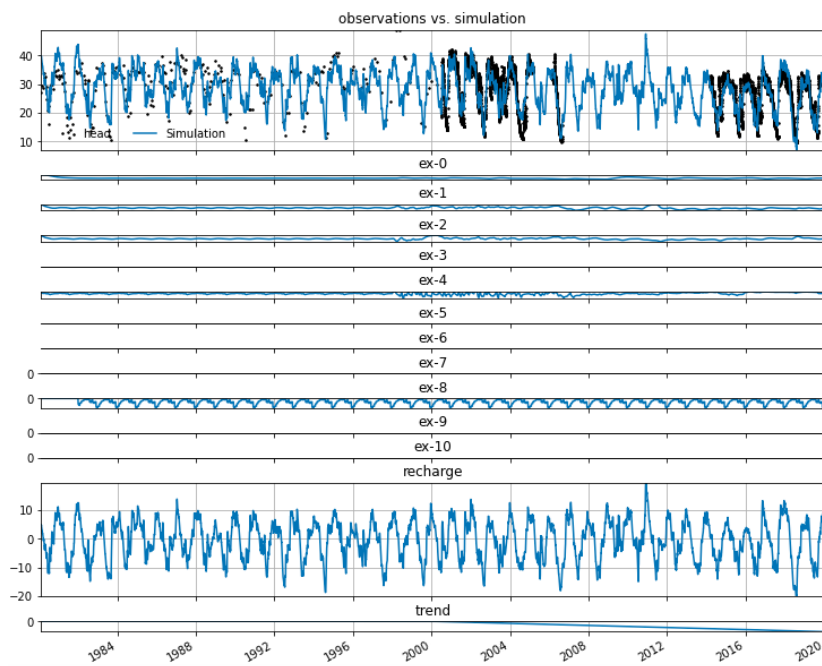


Figure 8.54: Diagnostics of the influences of pumps for BH86 in LH Model

L Appendix

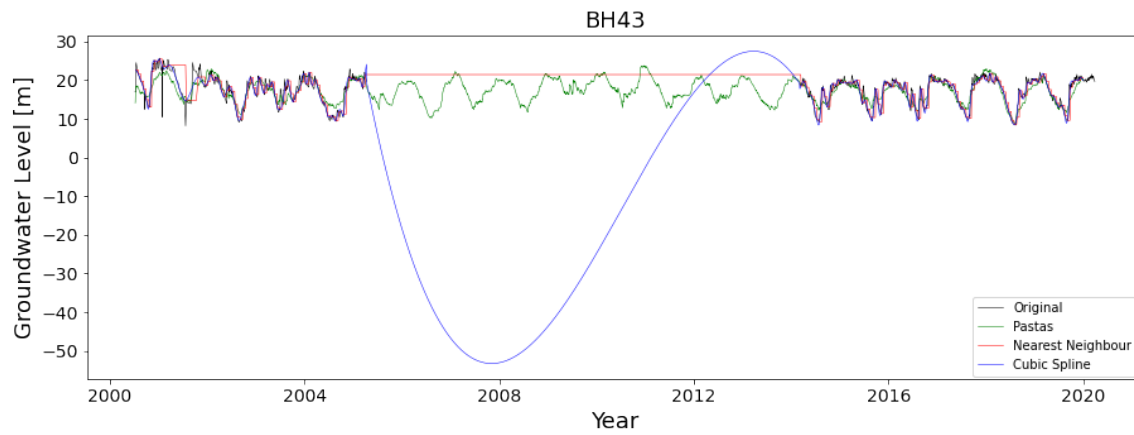


Figure 8.55: A comparison between the original data, NN, CS & TFN simulation for BH43

Table 12: R^2 -values for the TFN model, NN and CS

Method	BH43	BH86
Pastas	0.65	0.67
Nearest Neighbour	0.76	0.72
Cubic Spline	0.93	0.91



Figure 8.56: a) Every two weeks of the time-series is calibrated or simulated for b) Every six weeks of the time-series is calibrated or simulated for

Table 13: R^2 -values for the TFN model, NN and CS

Method	2W	4W	6W	8W
Pastas	0.77	0.77	0.76	0.76
Nearest Neighbour	0.90	0.74	0.61	0.63
Cubic Spline	0.93	0.71	0.68	0.12

DEPARTMENT OF ARCHITECTURE AND CIVIL
ENGINEERING
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2021
www.chalmers.se



CHALMERS
UNIVERSITY OF TECHNOLOGY