Front page

Exam information
NGEA06012E - Geology-Geoscience Thesis 60 ECTS, Department of Geosciences and Natural Resource Management - Contract:126674 (Rolf Kauffmann)

Handed in by
Rolf Kauffmann
fxs369@alumni.ku.dk

Exam administrators
Eksamensteam, tel 35 33 64 57
eksamen@science.ku.dk

Assessors
Majken Caroline Looms Zibar
Examiner
mcl@ign.ku.dk
+4535322444

Hand-in information

Title: Optimized model for evaluating pesticide leaching
Title, english: Optimized model for evaluating pesticide leaching
The sworn statement: Yes
Does the hand-in contain confidential material: No

UNIVERSITY OF COPENHAGEN

FACULTY OF SCIENCE

Optimized model for evaluating pesticide leaching

Master thesis (60 ECTS) by Rolf Kauffmann

Supervisors: Majken Looms Zibar (KU) and Sachin Karan (GEUS)

Submitted on: 17/01/2022

Abstract:

Pesticide leaching poses a threat to the environment and groundwater. Which is why the Danish Pesticide Leaching Assessment Programme (PLAP) was established in 1999. The data from this project contains a variety of variables for a 20-year period. These variables are used to try and estimate groundwater table which acts as a proxy for the groundwater recharge, which is a key factor in pesticide leaching. This is done using linear regression techniques and data such as precipitation, potential evapotranspiration, and soil water content. The resulting models manage to recreate a groundwater table with a minimum of calibration time and variables.

Table of Contents

Introduction

Objective

The purpose of this thesis is to improve the modeling of the groundwater recharge that goes through the vadose zone to the saturated zone. This is done with the help of machine learning more specifically regression analysis. As there are no measurements for groundwater recharge to compare the results of these models against the groundwater table is used as a proxy for this measurement. Models for estimating groundwater recharge from groundwater table data are available see Cuthbert (2019) that uses the water table fluctuation for estimating the recharge. This estimation will not be done in this thesis.

The data used for the modeling is from the Danish Pesticide Leaching Assessment Programme (PLAP). Until now the modeling for PLAP has been done using the MACRO model, which is a physically based one-dimensional numerical model based on Richards' equation and the convection-dispersion equation (Rosenbom, 2021).

These models are more representative for larger domains. The problem with physically based modeling in a hydrologic context is that numerically implementing geological heterogeneities can be hard and they are inhibited by the understanding of the physics controlling the transport (Gumiere et al 2020).

The objective would be to be able to make a better model from the available data, that more accurately predicts the pesticide leaching. More specifically getting a better determination of the recharge, a flow of volume over time, through the vadose zone, a zone not fully saturated by water. As recharge is hard to measure and is not available for the sites a proxy is needed to approximate this. The proxy used is the groundwater table. All the modeling done here is in one dimension.

Linear regression is an algorithm used both within statistics and machine learning (Maulud, 2020). Similar use of linear regression can be seen in works such as Huang (2019) where linear regression as well as multi-layer perception and long short-term memory models are used for predicting groundwater recharge estimated using water table fluctuation for a region in Australia based on information from 465 boreholes for a timeframe of 42 years. In this thesis the data is from one site only and on a shorter timescale.

Pesticides and groundwater

Pesticides are a broad group of products that includes fungicides, herbicides, and insecticides. The use of pesticides is especially common in agriculture, with it being a main contributor of pesticide use in Denmark (Miljø- og Fødevareministeriet (2017)).

Leaching of pesticides and their degradation products poses a risk to groundwater resources and surface water systems. In agricultural practice pesticides are widely used and agricultural land takes up close to two-thirds of the area of Denmark, while drinking water is almost exclusively extracted from the groundwater. This makes monitoring of where leaching is likely to occur very relevant. The limit value for pesticides in drinking water and groundwater is set at 0.1 µg/L on an EU wide level. Recharge affects the leaching of pesticides into the groundwater as the pesticides and their breakdown products will flow into the groundwater. Numerous factors affect leaching of pesticides including agricultural practices, climate, hydrogeology, and soil type.

PLAP

In 1999 the Danish Pesticide Leaching Assessment Programme (PLAP) was established. The PLAP fields were established to test the impact of pesticides under field conditions and to not just rely on modeling and laboratory experiments, as the field conditions will influence the leaching. The project evaluates whether use of pesticides within the limits of given regulation will result in concentrations in the groundwater that exceed the permitted limits, as well as informing the scientific basis on which regulatory decisions are made. The focus is on the agricultural use of pesticides and their impact on the groundwater. The compounds that are to be evaluated for are selected based on information about sales of the compounds as well as knowledge about their mobility and how used it is per area combined with other factors. (Miljø- og Fødevareministeriet, 2018)

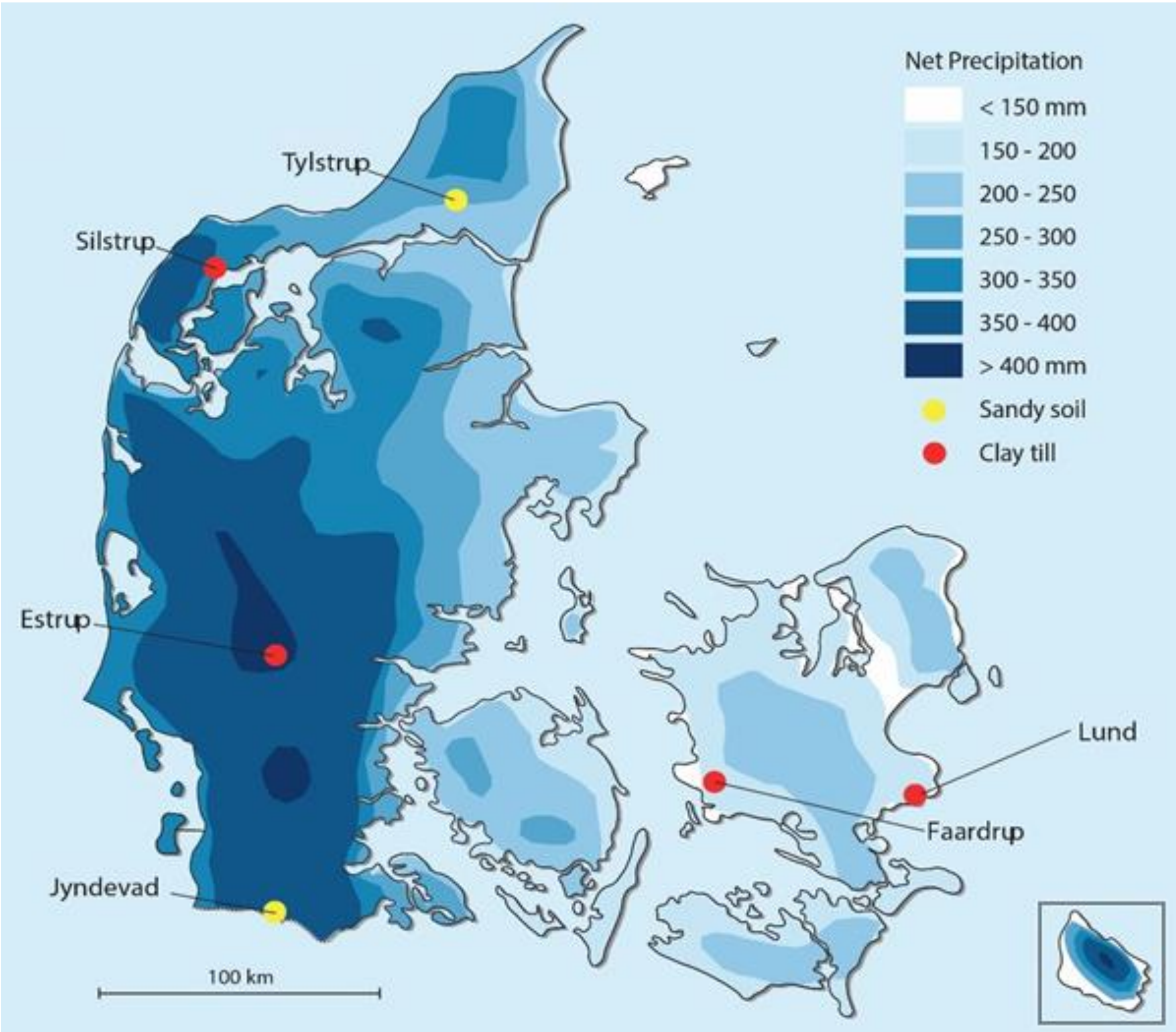

Figure 1: The locations of the six monitoring fields are marked throughout Denmark. With the soil type and the span of the annual net precipitation. (Modified figure from Rosenbom (2021))

PLAP consists of six monitoring fields as shown in figure 1. These sites have been chosen to represent different mean annual precipitations and soil types found throughout Denmark. The fields at Jyndevad and Tylstrup have sandy soils, while the fields at Estrup, Faardrup, Lund, and Silstrup have clayey till soils. All the fields were established in 1999 or 2000, except for Lund which was established in 2017. The monitoring for all fields is ongoing, except for the field at Tylstrup which was put on standby in 2019. The fields are monitored for different pesticides and their breakdown products. In figure 1 the yearly net precipitation is shown, which is the part of the precipitation that is not evaporated or transpired. Although the yearly normal precipitation is for 1931-1960, the tendencies are assumed to be the same although the numbers would be different (Miljøstyrelsen, 1992).

The data that has been collected spans from the beginning of the project until today. Although the main purpose is to measure the concentrations of pesticides, other parameters that affect leaching are also measured such as precipitation, soil water content and groundwater table.

The climate affects pesticide leaching, factors such as temperature, evaporation, and precipitation. As the temperature and evaporation does not differ much in Denmark, the sites have been chosen based exclusively on net precipitation. (Lindhardt, 2001)

Area of interest

In this report it was chosen to work on a sandy soil location. This choice was made as clayey tills are more unpredictable because fractures and biopores in the till can cause preferential flow that is harder to predict (Rosenbom, 2021). As Tylstrup one of the two sandy sites is on standby the other site which is Jyndevad was the site chosen to work on for this thesis.

In figure 2 the PLAP field at Jyndevad is shown. The inner white part is where crops are grown on the field, while the gray outer part is a buffer zone where grass is grown. The crops that are on the field are rotated for each year. All information regarding planting and sowing as well as the development of the plants as well as pesticide use, and irrigation are also documented in Rosenbom (2021) as well as earlier reports. The inner white field is 2.4 ha and measures 135 m x 180 m. The width of the buffer zone is 3 m in the east, 14 m in the south, 16 m in the west and 24 m in the north. The field is flat and is bordered on the eastern side by a windbreak. Monitoring for the field started in September 1999 and the sediment type is coarse sand.

As can be seen in figure 2 the measurement stations are placed in and close to the buffer zone. The general direction of the groundwater flow is in a western direction. Each year a new crop is planted, the crops most often planted on the Jyndevad field from 1999-2020 are spring barley, potatoes and winter wheat, and peas in that order.

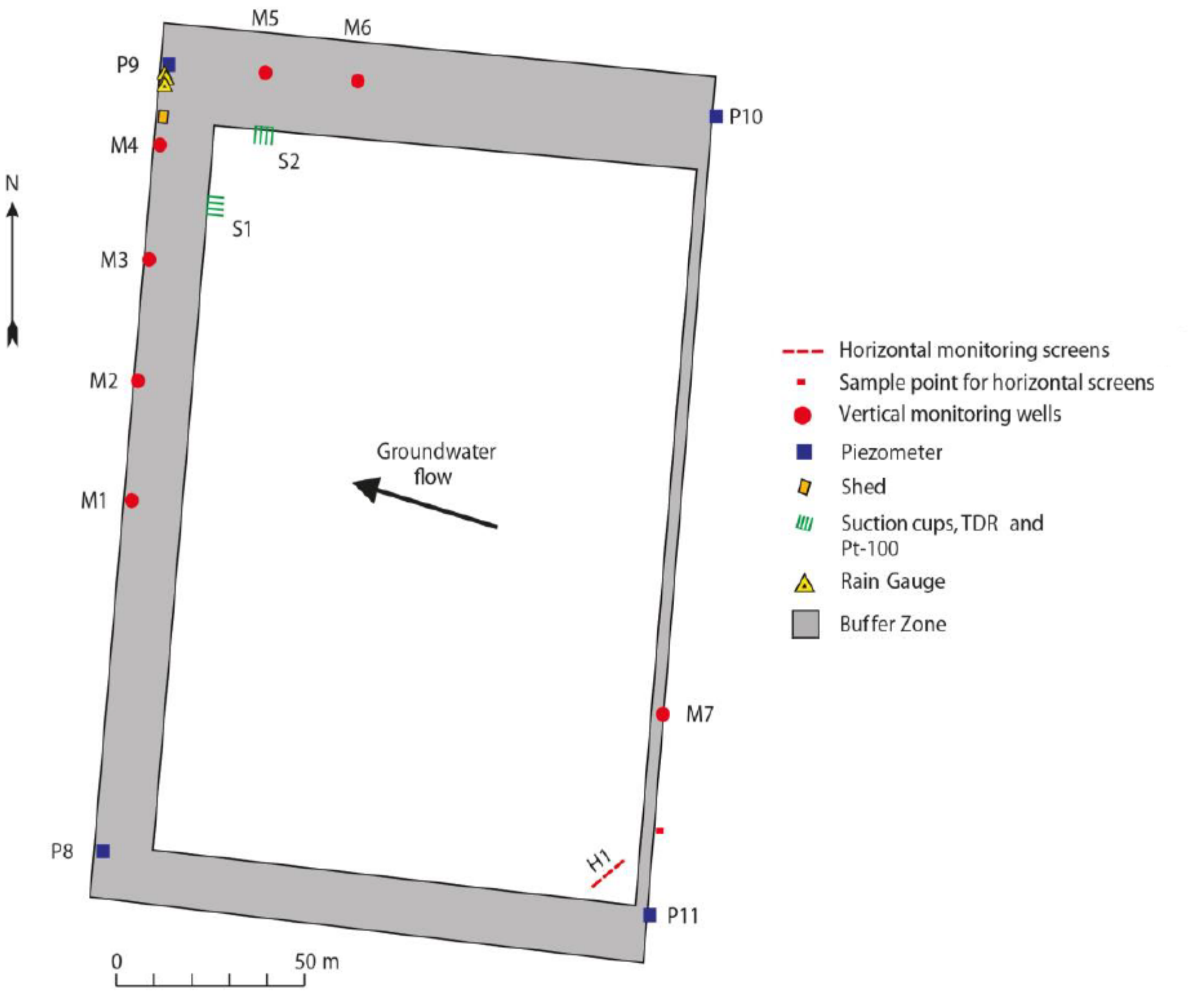

Figure 2: Outline of the Jyndevad field with the inner white part being the test field and the outer gray area being the buffer zone. The arrow denotes the flow of the groundwater and the different measurement stations are shown with colored markers with the name next to it. (Modified figure from Rosenbom (2021))

Theory and methods

Trend analysis

A first step was to investigate whether there was a trend in the time series data that is present for the field. The first step in this process is to plot the time series data in a graph and visually inspect it.

A trend analysis is done on the time series data, for this analysis a Mann-Kendall test is used. The Mann-Kendall test evaluates if there is a monotonic trend, which is a consistent upward or downward trend, for the data. The test is nonparametric, which means a normal distribution of the data is not assumed. The null hypothesis for the Mann-Kendall test is that there is no monotonic trend for the time series, while the alternate hypothesis is that there is a monotonic trend.

In equation 1 $f(t)$ is the monotonic trend as a function of time, while Q is a Sen's slope estimate for the slope of the trend and B is the estimate of the intercept. As we assume the trend is linear, we can use nonparametric Sen's method to estimate the slope and the intercept.

$$f(t) = Qt + B \tag{1}$$

Q is calculated using the values from the time series and the time as seen in equation 2.

$$Qi = \frac{x_j - x_k}{j - k} \tag{2}$$

Where x would represent the yearly or monthly values from the time series where j is more recent than k.

The Mann-Kendall test and the Sen's method of slope estimation is done using the Excel template from the Finnish Meteorological Institute. The data input into the model are time series data. Both annual and monthly values can be used in the template to evaluate if there are trends (Salmi, 2002).

To do the trend analysis, the S statistic is calculated. The S statistic indicates which way the trend goes. It is calculated by comparing earlier observations against later observations and categorizing them as either 1, 0, or -1 according to whether the difference between the observations is positive, equal, or negative. These values are then summed to get the value, which indicates if there is a trend.

The S statistic along with the variance of the statistic is then used for calculating the test statistic Z. The Z statistic is used to evaluate if there is a significant upward or downward trend, with a positive value of Z indicating a rising trend and vice versa. The equation is seen in equation 3, with the concept being like the S statistic.

$$Z = \begin{cases} \dfrac{S-1}{\sqrt{Var(S)}} & \text{if } S > 0 \\ 0 & \text{if } S = 0 \\ \dfrac{S+1}{\sqrt{Var(S)}} & \text{if } S < 0 \end{cases} \tag{3}$$

The Z test statistic can be used for evaluating if there is a significant trend.

In the Excel sheet a two-tailed test is performed for four different significance levels, ɑ. The significance levels are 0.1, 0.05, 0.01 and 0.001. These levels correspond to a percentage, for the chosen levels they go from 10% to 0.1% and this percentage describes the likelihood that the null hypothesis is being discarded even though it is true.

For a significance level of 0.001 this means that there is a probability of 0.1% that the null hypothesis is rejected even though it is true.

Another way of understanding the significance levels are the corresponding confidence levels, which can be calculated by subtracting ɑ from one. Using the above-mentioned significance levels the corresponding confidence levels are 90%, 95%, 99% and 99.9%. For each confidence level there is a corresponding value of the Z test, for 90% the value is ±1.645, for 95% it is ±1.960, for 99% it is ±2.576 and for 99.9% it is ±3.291. With a positive indicating an upward trend and a negative indicating a downward trend. This means that for a Z value of 3.291 or above there is 99.9% that the alternate hypothesis is true and that there is an upward trend.

Synthetic data

Synthetic data made to evaluate if it is possible to develop models and evaluate these models when the data is scarce (Mirus, 2011).

To set up a synthetic model Hydrus-1D is used. Hydrus-1D is a program that simulates water movement in one dimension. This is done by numerically solving Richards' equation seen in equation 4.

$$\frac{\partial \theta}{\partial t} = \frac{\partial}{\partial x}\left[K\left(\frac{\partial h}{\partial x} + \cos\alpha\right)\right] - S \qquad (4)$$

With θ being the volumetric water content, t being the time, x being the spatial coordinate, K being the unsaturated hydraulic conductivity, h being the water pressure head, and S being the sink term (Šimůnek, 2013).

In this case the synthetic model is used to recreate a groundwater table at the field in Jyndevad, with the input into the Hydrus-1D model being the precipitation and the potential evapotranspiration from Jyndevad.

To build up a groundwater table in Hydrus-1D two soil layers are used in the model. The upper layer must have a high hydraulic conductivity, while the lower layer must have a low hydraulic conductivity. The hydraulic model used is the van Genuchten-Mualem single porosity model. The upper boundary condition used is an atmospheric boundary condition with surface runoff, while the lower boundary condition is free drainage water boundary condition, which allows the water to flow through the lower boundary. The soil hydraulic parameters used for the two layers are shown in table 1. These values have been determined through trial and error trying to recreate the observed groundwater table as best as possible. In table 1 θ_r is the residual soil water content and is unitless, θ_s is the saturated soil water content and is unitless, ɑ and n are parameters in the soil water retention function and ɑ has a unit of 1/cm while n is unitless, K_s is the saturated hydraulic conductivity in cm/day, and I is a tortuosity parameter in the conductivity function and is unitless. The values of θ_r and θ_s influence the value of the soil water content, as these values limit the range within the synthetic soil water content can be.

Table 1: The settings used for the different layers in the Hydrus model. Material 1 is the top layer and material 2 is the bottom layer.

	θ_r	θ_s	ɑ	n	K_s	I
Mat. 1	0.13	0.30	0.04	2.68	2925	0.50
Mat. 2	0.10	0.34	0.01	1.09	0.19	0.50

The height of the soil column is set to 3 meters, with the approximate span in the data being around two meters, which allows space for the 0.3 m bottom layer. Observation points are put at a depth of 25 cm, 60 cm, as well as at the bottom of the profile. The observation point at the bottom of the profile outputs the value of the groundwater table, while the two other observation points are placed to detect the soil water content at those depths, as they are measured at these depths for the field in Jyndevad. The height of the soil profile is subtracted from the observations outputted from the observation node at the bottom of the profile to get the depth of the groundwater table.

The Feddes water uptake model is used and the parameters that are used for the root water uptake parameters are the ones from the build-in database in Hydrus for potatoes. A constant root depth is used, and it is set at 0.6 m. The time discretization for output results for the soil water contents as well as the groundwater table are daily.

Regression

Regression analysis is a method both within the fields of statistics and machine learning. The type of regression model that is used in this thesis is linear regression. Linear regression is a method of estimating the relationship between an independent or explanatory variable and a dependent or response variable.

A linear regression line in its simplest form follows the equation seen in equation 5:

$$y = a + bx \tag{5}$$

Where y is the dependent variable, a is a constant term, b is a coefficient and x is the independent variable. The inputs into would be both the dependent and the independent variables and the output of the model would be the constant and the coefficient.

The type of regression done here is linear regression, two types of linear regression are simple and multiple linear regression. Both simple and multiple linear regression have one dependent value, the difference is in the independent variables of which there is only one for simple regression while there is more than one for the multiple regression.

The tool used for regression analysis in this thesis is MATLAB. To solve the regressions, the MATLAB function fitlm, which is part of its Statistics and Machine Learning Toolbox, is used. This function uses QR decomposition as its main fitting algorithm (MathWorks, 2021).

QR decomposition can be used for estimating linear regressions using the ordinary least squares estimation method. The least squares regression line is the smallest distance between the actual and the predicted value. The ordinary least squares method is a way of minimizing the sum of squared residuals, to find the model that fits best (Taboga, 2021).

The fitlm function uses one or multiple input variables to find the best fit to the response variable. If five input variables were used with the response variable, the fitlm function would solve a multiple linear regression and the model output from the function would look like what is seen in equation 6:

$$y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \epsilon. \tag{6}$$

With y being the dependent variable and X being the independent variables, β are constants to the independent variables and β_0 specifically is the constant term and ϵ is an error term that is included in the equation to represent the uncertainty of the model and accounts for the lack of fit between the model and the dependent variable.

An example of the MATLAB script used for an example with five variables can be seen in the appendix.

For the regression models the coefficient of determination (R^2) and root mean square error (RMSE) are used to evaluate the results. Both statistics give information about the fit of the regression models.

The coefficient of determination or R^2 is given by the equation in equation 7.

$$R^2 = 1 - \frac{SSR}{SST} \tag{7}$$

With SSR being the sum squared regression, which is the difference between the predicted value and the mean of the observed value, while SST is the total sum of squares, which is the difference between the observed value and its average value.

For multiple regression analysis the adjusted coefficient of determination value makes more sense, as it incorporates that each additional independent variable adds more uncertainty to the model. In equation 8 the calculation of the adjusted R^2 is shown.

$$R^2_{adj} = 1 - (1 - R^2)\frac{n-1}{n-p-1} \tag{8}$$

With n being the sample size and p being the number of independent regressors.

A higher value of R^2 shows how much of a linear relationship there is between variables. The range is from 0 to 1, with the number corresponding to the percentage that can be explained in the y-variable by the x-variables.

The equation used for calculating the root mean square error or RMSE is seen in equation 9.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n}} \tag{9}$$

With n again being the sample size, y_i is the observed value, and $\hat{y}_i$ is the predicted value. The residuals between the observed and predicted values are squared and summed and the square root of this is taken. The lower limit for the RMSE is zero and the bigger the number the bigger the error. The RMSE gives an idea of how big the error of the model is compared to the data in the same unit as the unit of the value that is being evaluated.

Results

Data

The data that are used are from Jyndevad, the location of which can be seen in the lower left of fig. 1, and it consists of time series of differing lengths and measurement frequency.

The longest and most complete time series data are from above the ground surface, these being the time series for precipitation (P), daily minimum and maximum temperature (T), and potential evapotranspiration (PET). These time series are present for each day from 03/01/1984 until 30/06/2020, except for one missing measurement of temperature. The number of data points in these series are 13329 data points.

The precipitation data is uncorrected and irrigation data is registered for the field. For the temperature time series there were 49 cases where in the raw data for the maximum and the minimum temperatures had been accidentally transposed. These values were reversed back to their proper order.

The potential evapotranspiration has been calculated for Jyndevad. The calculation has been done using the Makkink method. This method is based on temperature and radiation (Kraalingen, 1997). The potential evapotranspiration is the possible amount of evapotranspiration that can take place if enough water is present.

From below the ground surface the length of the time series and the frequency of measurements varies for each of the variables. The measurements that are used from below the ground surface are soil water content (SWC) and groundwater table (GWT).

The soil water content is measured at six depths for stations S1 and S2 which are seen in fig. 2 in the northwestern corner of the test field. These are measured with time domain reflectometer (TDR) probes. They are measured at depths of 25, 60, 93, 110, 190 and 210 cm. The measurements are available from 02/09/1999 until 30/06/2020 with holes of various lengths in the time series. The mentioned time span equals 7181 days.

The series with the least available data are at 25 cm depth, with 6598 measurements at S1 and 6552 measurements at S2 and at 210 cm depth, with 6937 measurements at S1 and 6315 measurements at S2. The four other stations have data available for close to the total duration, with the range of available data for those stations being 7062 to 7171 measurements.

The tendencies that the soil water content for S1 and S2 are similar, while the absolute values differ, therefore an average is taken of the two time series. The soil water content is reported in % and describes how saturated the soil is with water.

The groundwater table is measured at piezometers P8-P11 that can be seen in the corners of the buffer zone around the field shown in figure 2. For each of these stations there are filters at three depths.

The groundwater table data from Jyndevad comes in two series. The first has a larger number of data points for a shorter time span and the data points are a continuous daily series measured using a data logger. It consists of three continuous series, the first is a four-year series from mid-2012 to mid-2016, while the two others are one-year series from mid-2017 to mid-2018 and mid-2019 to mid-2020. These series have been measured at the second filter depth at station P11, which is at the south-eastern edge of the field seen in figure 2.

The second series has fewer data points but has been measured over a longer time span, these are manually measured on a more irregular basis, with a total of 354 measurements dates. The series runs from mid-1999 until mid-2020. There are twelve time series, one for each filter depth at each of the four stations, with gaps of different lengths in the series. As these series mostly follow the same tendency an average is taken of all these twelve series.

For station P10 there were values that were positive, meaning that they would be above the ground surface; these were excluded from the series.

For the manually measured values an average was taken of all the four stations and all the filter depths, this was done as there is a variance even within the same piezometers, but the overall tendency is the same and lessens the impact of outliers, making the series more robust.

The manually measured groundwater table is taken as a monthly mean for these measurements as the sampling frequency varies quite a bit, with the shortest timespan between measurements being on the same day and the longest being 92 days apart. The average span between measurement values is 22.3 days and the median value of the spans is 21 days. This means that there is not a mean for each month, which can also be seen in figure 3 where there are missing values especially before 2006.

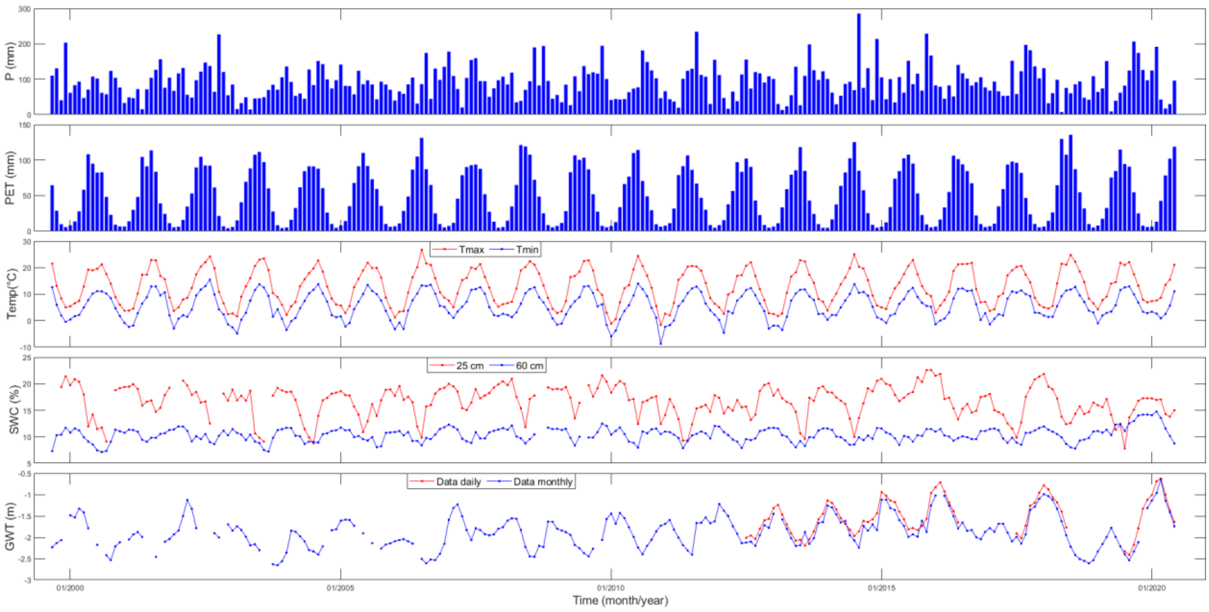

Figure 3: The monthly sums precipitation and potential evapotranspiration as well as the averages of the minimum and maximum temperature, the soil water content at depths 25 and 60 cm, and the groundwater table for the short and long series. The plotted time span is from September 1999 to June 2020.

The precipitation, as seen in fig. 3, varies quite a lot over the timespan with the lowest values being close to zero and the highest close to 300 mm per month.

The potential evapotranspiration and temperature seen in fig. 3 are quite similar in their trends with both being high in the summer months and low in the winter months, as the potential evapotranspiration is calculated based on the temperature and the radiation, which also affects the temperature, this is to be expected.

The soil water content has a period of higher soil water content in the winter month and lower in the summer months. The soil water content at the shallower depth has bigger fluctuations, while at the lower depth it is more stable. The groundwater table follows the same trend as the soil water content, being closer to the surface during the winter months and lower during the summer months.

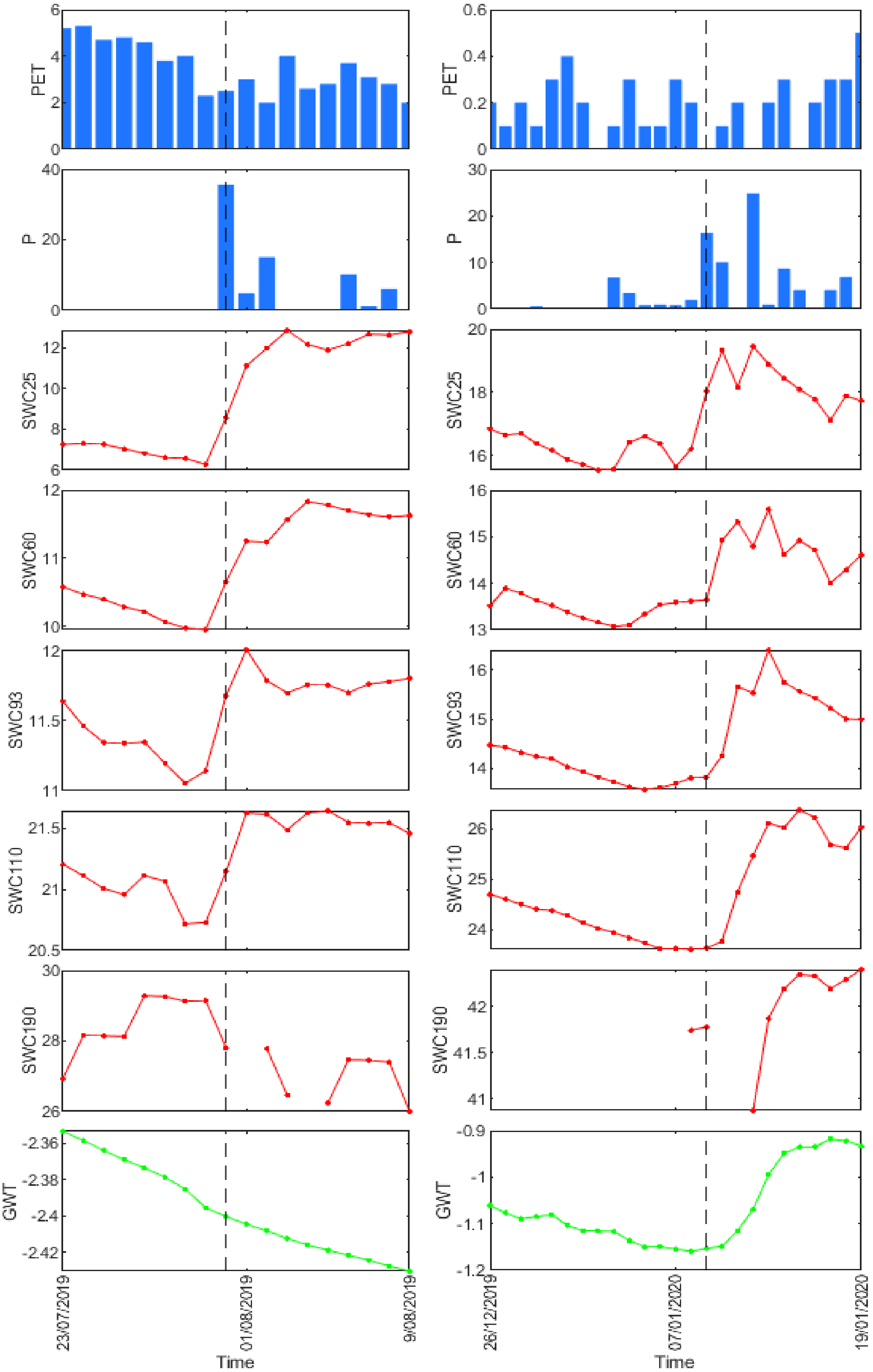

Figure 4: The flow of water in the summertime, on the left side, and the wintertime, on the right side.

In figure 4 the flow of the water through the soil is shown for summertime and wintertime. The dates have been chosen for a period with no or little rain before a bigger precipitation event. This is not as easily found for the winter months especially as the daily measurements for the groundwater table are only present for six years. The units for the axis are quite different, this reflects that there is a substantial difference in the water table and the soil moisture between the summer months and the winter months, the same is the case for the potential evapotranspiration.

The impact of the precipitation on the groundwater table is easy to see in the wintertime, while it is not apparent for the summertime. A reason for this is the difference in the depth of the groundwater table, with the difference between them being around 1.3 to 1.4 m. The groundwater table in the wintertime is at the same level as the soil water content at depth 110 cm. Also, the potential evapotranspiration is a factor of ten higher during the summertime. In the summertime the effect of the precipitation is seen within the day on the soil moisture down to 110 cm, while this is not the case in the wintertime. The effect on the groundwater table is apparent for the wintertime, while in the summertime no effect is seen.

Synthetic data

As there is scarcity for both cases of the groundwater table data that has been measured at the Jyndevad, synthetic data was created using the precipitation and potential evapotranspiration from the Jyndevad site.

With the help of the synthetic model groundwater table and soil water content at depth 25 and 60 cm is recreated from the 2nd of September 1999 to the 30th of June 2020. All the data used here is synthetic data, with only the precipitation and potential evapotranspiration being used to create the synthetic data as well as the groundwater table from the Jyndevad dataset being used for comparison.

Synthetic data for groundwater table and the soil water content at depths 25 and 60 cm is shown in figure 5 along with the measured values for comparison.

For the soil water content in the figure where the saturated soil water content, which is 30% as seen in table 1, that has been input in the Hydrus model limits how high the soil water content can go. The residual soil water content, the value of which is 13% and is seen in table 1 for the first material, gives a lower limit to the soil water content.

The synthesized soil water content at 25 cm is much more stable than the measured data as seen in figure 5. The measured values are between 5% and 25% and fluctuates very much, the synthetic data on the other hand is quite stable around 13% to 17% with a few peaks which go up to 30%. The synthetic data does not reflect the tendencies of the measured data very well. The limitations set on the model in form of the residual and saturated soil water contents does not help in this matter.

The synthetic soil water content at depth 60 cm and the measured data for the same depth both of which can be seen in figure 5. The measured data is more stable, with the highest values being around 15% and the lowest around 7%. The synthetic data is quite stable around 13% to 17%, but it differs by having multiple high peaks that do not appear in the data; these peaks hit the upper limit set by the saturated soil water content.

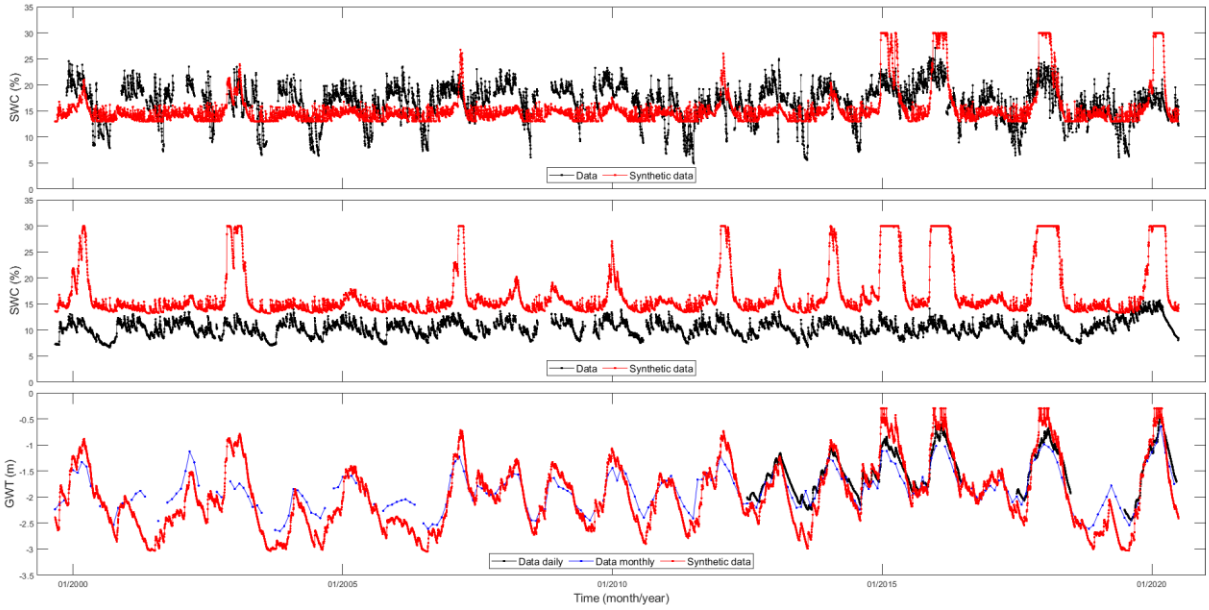

Figure 5: The synthetic data is shown in red, with the measured data shown in black and blue. The upper figure shows soil water content at a depth of 25 cm, the middle figure shows the same at depth 60 cm and the bottom figure shows the groundwater table.

The synthetic groundwater table and the measured values are more similar, with the same trends of rising and falling values being mostly the same and the biggest difference being the absolute values, especially for the period 2001 to 2007 where the values are different. From 2015 until 2020 the synthetic groundwater table hits an upper limit at around 0.3 m below the surface in four events, which correlates with the highest peaks seen in the measured data.

When setting up the synthetic data in Hydrus the resulting synthetic data was either better at modeling the groundwater table or the soil water content and not both, and as the groundwater table was the more important in this context the model was optimized for the groundwater table data.

Models and weighting

For the regression analysis a total of five models are used, each with a different number of variables. The 1-variable model is based solely on the weighted precipitation. Each subsequent model includes the variables from the previous models as well as a new variable. The 2-variable model uses the weighted potential evapotranspiration as well as the weighted precipitation. The 1- and 2-variable models are based on variables which are present for the total period between the 2nd of September 1999 to the 30th of June 2020.

The 3-variable model is based on the weighted variables as well as the soil water content at depth 25 cm, while the 4-variable model also includes the soil water content at depth 60 cm. The 5-variable model includes the groundwater table from one year before. The variables used in the 3-, 4- and 5-variable models each have missing data within the timeframe mentioned above. This means that with an increasing number of variables the available data for the model decreases, with the 5-variable model having the least available data and the 1- and 2-variable models having the most.

As the 1- and 2-variable models are based on variables that are present for the entirety of the series, it is only limited by the presence of the groundwater table data when trying to make a regression of the variables against that data. While the 3-, 4-, and 5-variable models all have further limitations, with all needing the soil water content from 25 cm, the two latter models needing the same from a depth of 60 cm and the last model needing the groundwater table from last year. This means that the more variables the more missing model data points.

For the 1- and 2-variable regressions a weighting has been done for the precipitation and the potential evapotranspiration. This was done as the raw values of precipitation and evapotranspiration do not give good regression results.

The weighting used is a linear weighting of value based on the values of the previous year. The weighting puts more weight on recent values and less on the older values. An example for a daily weighting of precipitation is shown in equation 10.

$$wP = \frac{P_n * \left(1 - \frac{0}{364}\right) + P_{n+1} * \left(1 - \frac{1}{364}\right) + \cdots + P_{n+363} * \left(1 - \frac{363}{364}\right) + P_{n+364} * \left(1 - \frac{364}{364}\right)}{182.5} \quad (10)$$

With wP being the weighted precipitation, P_n is the precipitation of the day that is being weighted and P_{n-364} is the value of the precipitation 364 days before that. The weight of the first value is one and the last value is zero and the weighting falls linearly between those numbers. The weights add up to 182.5, which is half of a year.

This weighting means that the effect of recent data is much greater than the data from a year ago, with the last value shown in equation 10 being totally discounted. A total of 365 precipitation values will be added together in a weighted average.

The same type of weighting is done for the monthly summed values of precipitation and potential evapotranspiration. This can be seen in equation 11.

$$wP = \frac{P_n * \left(1 - \frac{0}{11}\right) + P_{n+1} * \left(1 - \frac{1}{11}\right) + \cdots + P_{n+10} * \left(1 - \frac{10}{11}\right) + P_{n+11} * \left(1 - \frac{11}{11}\right)}{6} \quad (11)$$

Again, the data is divided by half of a year which in this case is 6 months.

This type of weight is used as it must be assumed that if these variables do affect the groundwater table that this effect is not dependent only on the variable on the day that it is expected to predict the value, but that there is an accumulated effect on the groundwater table.

When using this type of weighting the first year of data for the variables cannot be weighted as there is insufficient data. But as there is data for both precipitation and potential evapotranspiration from before the span when we are going to model for this is not a concern in those cases.

The groundwater table used as the data in the monthly calculations is an average of all the twelve measured time series that surround the field, while for the daily series it is for station P11 at the middle filter depth, as this is the only longer series.

From now on the shorter series will be referred to as the daily series, while the longer series will be referred to as the monthly series. For the monthly data, a monthly average has been taken of the groundwater table values.

For the variables used for the monthly regression models the monthly average is used for the soil water content while for the precipitation and potential evapotranspiration the monthly sum is used.

The weakness of the daily data is that so little of it is present. This is where the longer groundwater series comes in. Although the measurement frequency is not as high it has been consistently measured from the establishment of the test field and thus allows for comparison over a much longer time span.

The regression constants for both the synthetic and the regular models can be found in the appendix in tables 27 to 30.

Data trends

The results of the trend analysis can be seen in table 2 to table 6. In the tables the columns contain the first year and the last year as well as the total number of years (n) tested. Then comes the Z test statistic and the significance levels going from + for a significance level of 0.1, to * for a level of 0.05, ** for a level of 0.01, and *** for a level of 0.001. If no sign is present the significance level is above 0.1. Also shown are the slope and constant estimates as well as the change over the years calculated using the slope and the number of years tested.

Table 2 and 3 show the results for precipitation, potential evapotranspiration as well as the minimum and maximum temperatures. The precipitation is only significant at a level of 0.1 for August. As can be seen in the test Z values for the precipitation, except for August, these are quite far from reaching even a significance level of 0.1, with the closest being the yearly and for the month of March, which has a negative test value, which indicates a falling trend.

Table 2: The trend analysis for the precipitation, potential evapotranspiration, and the minimum and maximum temperatures. This is both for the monthly and yearly data.

Time series	First year	Last Year	n	Test Z				Significance level			
				P	PET	Tmax	Tmin	P	PET	Tmax	Tmin
January	1984	2020	37	-0.27	0.31	0.46	1.01				
February	1984	2020	37	0.05	1.99	1.37	1.90		*		+
March	1984	2020	37	-1.27	2.73	1.86	1.88		**	+	+
April	1984	2020	37	0.01	3.06	2.69	2.17		**	**	*
May	1984	2020	37	0.46	0.73	0.61	1.24				
June	1984	2020	37	0.54	2.26	2.68	2.68		*	**	**
July	1984	2019	36	0.53	1.77	2.00	1.23		+	*	
August	1984	2019	36	1.87	1.27	1.40	2.30	+			*
September	1984	2019	36	0.75	2.34	2.87	3.23		*	**	**
October	1984	2019	36	0.48	1.76	1.27	1.51		+		
November	1984	2019	36	0.00	0.59	1.76	2.38			+	*
December	1984	2019	36	0.99	2.26	2.08	2.14		*	*	*
Yearly	*1984*	*2019*	*36*	*1.43*	*3.56*	*3.56*	*3.88*		***	***	***

Table 3: The trend analysis for the precipitation, potential evapotranspiration, and the minimum and maximum temperatures. This is both for the monthly and yearly data.

Time series	Sen's slope estimate				Constant				Change over time period			
	P	PET	Tmax	Tmin	P	PET	Tmax	Tmin	P	PET	Tmax	Tmin
January	-0.25	0.00	0.02	0.05	108.1	6.7	3.4	-1.0	-9.1	0.2	0.9	1.7
February	0.05	0.07	0.06	0.08	64.2	11.7	2.7	-2.2	1.8	2.7	2.4	2.9
March	-0.73	0.22	0.07	0.05	72.3	27.8	6.2	-0.2	-27.1	8.1	2.5	2.0
April	0.01	0.52	0.06	0.05	44.1	51.8	11.2	2.4	0.2	19.3	2.4	1.9
May	0.22	0.15	0.02	0.03	48.0	88.2	16.2	6.6	8.3	5.4	0.7	1.0
June	0.44	0.54	0.07	0.05	80.1	86.8	17.7	9.1	16.2	20.2	2.7	2.0
July	0.39	0.38	0.07	0.02	76.5	93.6	20.5	11.6	13.9	13.5	2.5	0.9
August	1.40	0.17	0.03	0.04	76.1	81.0	20.9	11.2	50.4	6.2	1.2	1.5
September	0.65	0.33	0.07	0.06	94.1	43.4	16.1	8.6	23.5	12.0	2.5	2.1
October	0.31	0.10	0.04	0.04	106.3	23.3	12.3	5.6	11.1	3.5	1.3	1.4
November	-0.01	0.01	0.04	0.07	98.4	9.3	7.1	1.2	-0.4	0.3	1.5	2.5
December	0.58	0.03	0.05	0.07	85.7	4.4	4.3	-0.2	21.0	1.2	1.8	2.4
Yearly	*3.52*	*2.49*	*0.05*	*0.05*	*1008.5*	*529.6*	*11.7*	*4.3*	*126.8*	*89.7*	*1.8*	*1.9*

For the yearly data, the change is significant for the potential evapotranspiration as well as the temperatures. They are all significant at a level of 0.001, which is quite strong evidence that an upwards trend is present for the annual data. For a number of the months, such as April, June, and September there is also good evidence for there being an upward trend. Looking at the changes over the years the annual change in temperature over the 36 years is close to 2 °C and close to 90 mm of more potential evapotranspiration per year. All the slopes for these three variables are positive and so are all the trends that have any form of significance.

Table 4 shows the results for the soil water content and is based on data from 1999 to 2020. On an annual basis no trend is seen but for the months there are two months for depth 25 cm where there is weak evidence for a downward trend and one month for depth 60 cm where there is evidence for an upward trend. The slopes are mostly negative at the shallower depth and mostly positive at the deeper depth.

Table 4: The trend analysis for soil water content at depths 25 and 60 cm. With the monthly and yearly trends.

Time series	**First year**	**Last Year**	**n**		**Test Z**		**Significance**		**Slope**		**Constant**		**Change**	
			S25	**S60**	**S25**	**S60**	**S25**	**S60**	**S25**	**S60**	**S25**	**S60**	**S25**	**S60**
January	2000	2020	20	21	-0.88	0.82			-0.08	0.01	20.5	11.1	-1.5	0.3
February	2000	2020	21	21	-1.24	-0.45			-0.11	-0.01	21.7	11.5	-2.3	-0.2
March	2000	2020	21	21	-1.84	0.21	+		-0.15	0.00	21.9	11.0	-3.1	0.1
April	2000	2020	21	21	-1.72	0.94	+		-0.10	0.03	19.5	9.3	-2.0	0.6
May	2000	2020	21	21	-0.94	-0.09			-0.13	0.00	18.3	9.3	-2.6	-0.1
June	2000	2020	21	21	-0.09	0.33			-0.02	0.01	13.9	8.8	-0.3	0.3
July	2000	2019	19	19	-0.77	0.49			-0.10	0.03	15.1	8.7	-2.0	0.6
August	2000	2019	20	20	0.55	0.94			0.07	0.04	13.1	8.7	1.3	0.7
September	1999	2019	17	20	0.62	2.17		*	0.06	0.13	14.5	6.2	1.0	2.6
October	1999	2019	17	20	-0.21	0.62			-0.02	0.03	18.2	9.9	-0.4	0.5
November	1999	2019	21	21	-0.69	-0.03			-0.03	0.00	19.7	11.2	-0.7	0.0
December	1999	2019	20	21	0.10	1.18			0.03	0.02	18.5	10.9	0.5	0.4
Yearly	*2000*	*2019*	*20*	*20*	*-0.36*	*0.81*			*-0.03*	*0.01*	*17.3*	*10.0*	-0.6	*0.2*

In table 5 the trend analysis for the groundwater table for the monthly series is shown. The yearly average has evidence for an upward trend, indicating that the groundwater table is rising, as does January and with less significance May. The slope for the yearly average is 0.02 m, which means the total change for the period would be 0.3 m.

Table 5: The trends for the monthly groundwater table, both monthly and yearly.

GWT	**First year**	**Last Year**	**n**	**Test Z**	**Signific.**	**Slope**	**Const.**	**Change**
January	2000	2020	20	2.43	*	0.04	-2.7	0.9
February	2000	2020	19	1.47		0.03	-2.3	0.5
March	2000	2020	21	1.48		0.02	-2.1	0.5
April	2000	2020	21	1.51		0.02	-2.2	0.4
May	2000	2020	20	1.85	+	0.02	-2.3	0.4
June	2003	2020	17	1.61		0.01	-2.5	0.3
July	1999	2020	19	0.98		0.01	-2.5	0.2
August	1999	2020	19	0.63		0.01	-2.3	0.1
September	1999	2020	19	0.42		0.00	-2.3	0.1
October	1999	2020	20	1.33		0.01	-2.4	0.3
November	1999	2020	20	1.20		0.02	-2.4	0.4
December	2000	2020	21	1.30		0.02	-2.4	0.5
Yearly	*1999*	*2020*	*22*	*2.20*	*	*0.02*	*-2.2*	0.3

Table 6: The trends for the daily groundwater table, both monthly and yearly.

GWT	First year	Last Year	n	Test S	Signific.	Slope	Const.	Change
January	2013	2020	6	-9		0.09	-3.8	0.5
February	2013	2020	6	-13	*	0.08	-3.5	0.5
March	2013	2020	6	-13	*	0.12	-5.0	0.7
April	2013	2020	6	-11	*	0.09	-4.3	0.6
May	2013	2020	6	-13	*	0.07	-3.8	0.4
June	2013	2020	6	-7		0.06	-3.8	0.4
July	2012	2019	6	-1		0.02	-2.6	0.1
August	2012	2019	6	3		-0.04	-0.9	-0.2
September	2012	2019	6	1		0.00	-1.8	0.0
October	2012	2019	6	-5		0.02	-2.4	0.1
November	2012	2019	6	-11	*	0.04	-2.7	0.2
December	2012	2019	6	-11	*	0.06	-3.3	0.4
Yearly	*2012*	*2019*	*6*	*9*		*0.06*	*-3.5*	0.4

In table 6 the daily groundwater table series is seen, as there is only a limited amount of data, as can also be seen on the number of years of data that are present. To increase the number of full years for the yearly trend analysis, the yearly average was taken from July to June, as otherwise there would only be two years of complete annual data. This is also the only trend analysis where the S-test was done as the series was quite short. Although there are six months where there are significance levels of 0.05, this is also an extremely short dataset where even the smallest differences are amplified. This can also be seen when the slopes are compared to those in table 5, which should be quite similar as they are measured on the same field, but the slopes are multiple times higher for a shorter time span.

In figure 6 the yearly trends are plotted. The data is more spread out for the precipitation than for the other variables, where it is easier to see an upward trend.

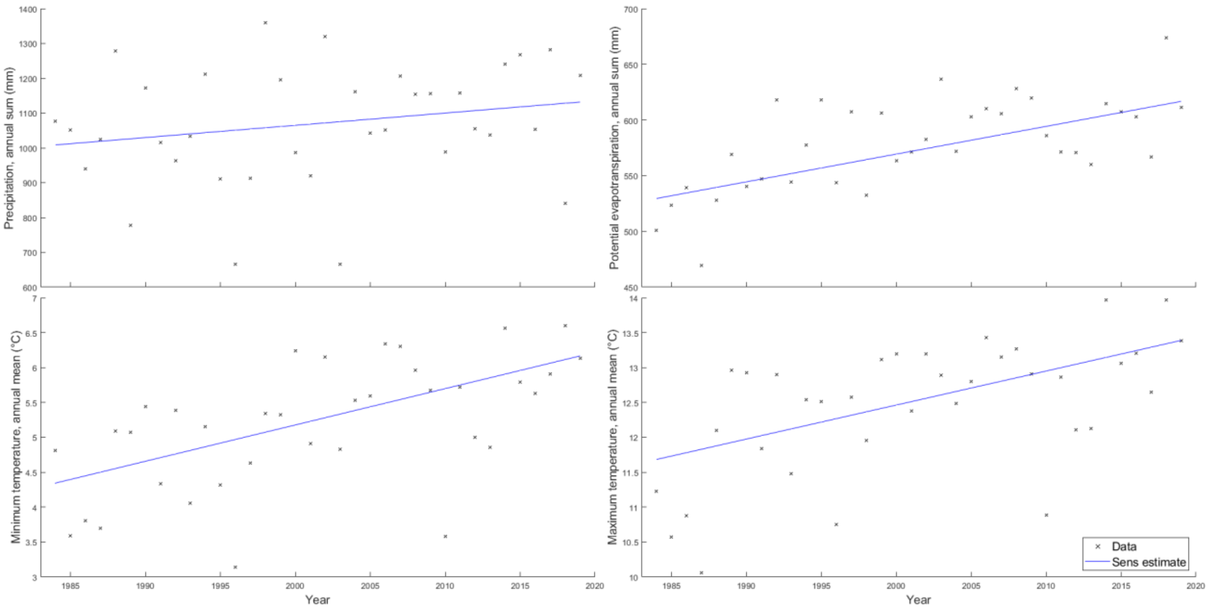

Figure 6: The trends for the precipitation, potential evapotranspiration as well as the minimum and maximum temperature plotted using the slope and the constant, shown in blue and the yearly data shown in red.

Synthetic data trends

The trends for the synthetic data are seen in table 7. Between September and February there are low significance trends for all the series except for two instances, while for the remaining months no significant trends can be seen. In all instances the trends are positive, which would indicate a rising groundwater table as well as a rising soil water content. The peaks seen in figure 5 for both the soil water contents has an effect on the trends and as these peaks are in the fall and winter, these are also the months where the significant rises are seen, without the peaks it is likely that no significant trends would be seen for the soil water contents.

Compared to the trends for the measured data in table 4 and 5 there are numerous differences. There are far more significant trends for the synthetic data and the trends are all positive, while for the measured data most of the significant trends for the soil water content are negative. For the groundwater table there is an overlap between the synthetic and the measured data in that there is a slightly significant upward trend in January, but for the synthetic data there is a more general trend for this upward trend in all the months that relate to the peaks that are seen in figure 5. This means that although the synthetic and measured data have similar trends, the difference in absolute values does affect the significance of the trends.

Table 7: The trends for the synthetic groundwater table, as well as the synthetic soil water content at depths 25 and 60 cm.

Time series	First year	Last Year	n	Test Z			Significance			Slope			Constant			Change		
				GWT	S25	S60	GWT	S25	S60	GWT	S25	S60	GWT	S25	S60	GWT	S25	S60
January	2000	2020	21	2.08	1.96	2.12	*	*	*	0.06	0.19	0.26	-3.0	10.9	10.6	1.36	4.07	5.51
February	2000	2020	21	1.96	1.90	1.97	*	+	*	0.05	0.16	0.34	-2.8	11.6	9.7	1.13	3.38	7.06
March	2000	2020	21	1.54	1.42	1.60				0.03	0.12	0.15	-2.2	12.4	13.9	0.67	2.58	3.13
April	2000	2020	21	1.30	0.57	0.88				0.03	0.02	0.07	-2.6	13.6	13.6	0.68	0.49	1.45
May	2000	2020	21	0.88	0.21	0.75				0.02	0.01	0.03	-2.5	13.6	13.5	0.34	0.14	0.61
June	2000	2020	21	0.94	0.57	1.36				0.02	0.01	0.02	-2.8	13.6	13.6	0.35	0.14	0.45
July	1999	2019	21	0.03	-0.63	-0.69				0.00	-0.01	-0.01	-2.5	13.9	14.5	0.07	-0.14	-0.29
August	1999	2019	21	0.45	1.24	1.60				0.00	0.02	0.02	-2.5	13.5	13.8	0.07	0.35	0.52
September	1999	2019	21	1.78	1.78	2.39	+	+	*	0.02	0.03	0.04	-2.9	13.4	13.4	0.49	0.57	0.88
October	1999	2019	21	1.96	0.94	1.90	*		+	0.03	0.01	0.04	-3.1	14.0	13.6	0.73	0.27	0.85
November	1999	2019	21	2.08	1.54	2.08	*		*	0.05	0.03	0.10	-3.1	14.0	13.0	1.06	0.63	2.07
December	1999	2019	21	2.20	2.33	2.24	*	*	*	0.06	0.06	0.18	-3.2	13.4	11.7	1.26	1.22	3.78
Yearly	*2000*	*2019*	*20*	*1.65*	*2.24*	*2.24*	+	*	*	*0.02*	*0.07*	*0.13*	*-2.6*	*13.2*	*13.0*	0.49	1.32	2.66

Synthetic data regression

To make a proof of concept of the regression models that are going to be made on the measured data from Jyndevad a synthetic model is set up. Synthetic data for the groundwater table and soil water content at depths 25 and 60 cm was recreated using the Hydrus-1D model as well as the precipitation and potential evapotranspiration data from the site at Jyndevad. In figure 7 the results for the synthetic groundwater table are seen compared with the monthly series of data from Jyndevad.

Synthetic data compared to measured data

In figure 7 the synthetically created data is compared to the data from Jyndevad. This was the best approximation found, and the same tendencies can be seen for the measured and synthetic groundwater table. The synthetic model is more extreme in the peak and trough values, good examples of this can be seen for peaks in early 2004, early 2015, early 2016, and early 2018 and the troughs late 2010, late 2011, as well as late 2013 where there also seems to be a slight discordance in between synthesized and measured data as to whether the groundwater is rising or falling.

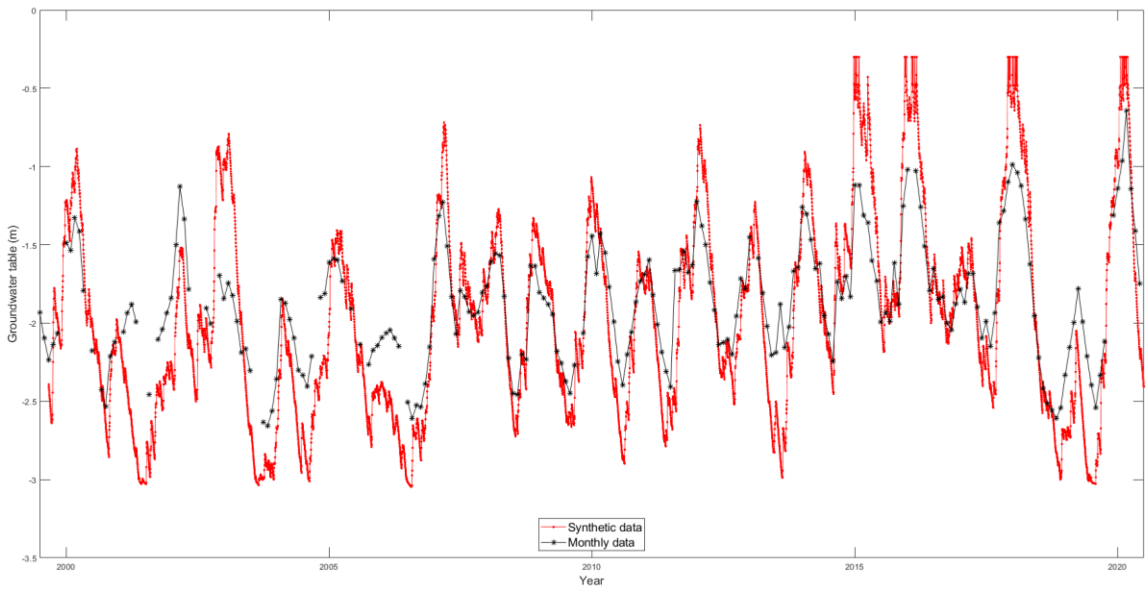

Figure 7: The synthetic groundwater table based on the Hydrus-1D model is in red, with the monthly groundwater table in black.

Simple regression

The next step is to use simple linear regression to predict the synthetic groundwater table using the weighted precipitation. This is done on both the daily synthetic groundwater table and the monthly averaged synthetic groundwater table.

The daily model is seen in figure 8, it is calibrated on the total amount of data. The general tendencies are reflected, with rising and falling groundwater tables being captured quite well. Around 2017 is a case where the model does not capture very well what happens in the synthetic data, with the model fluctuating a lot, while there are clear peaks and troughs in the data. From around the year 2000 to 2008 the model captures the highest peaks well, while mostly overestimating the groundwater table between those peaks. From 2012 onwards the model seems to have trouble capturing the height of the peaks seen in the data.

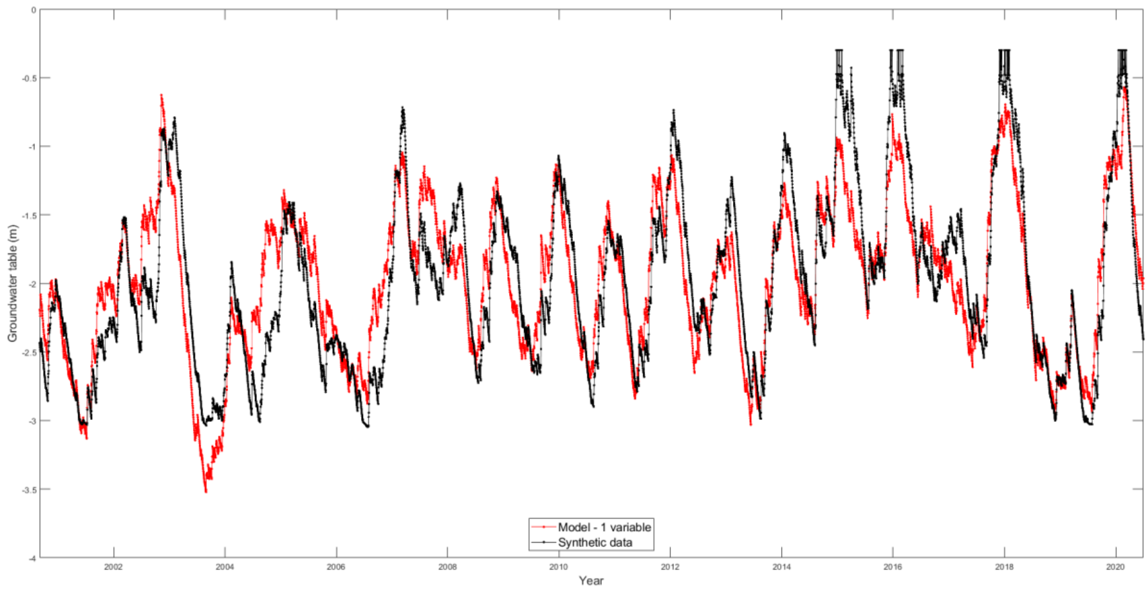

Figure 8: Simple regression on the total amount of the synthetic daily groundwater table.

The monthly synthetic data was made to resemble the mean groundwater table that is used in the monthly series for the measured data. In figure 9 the monthly data are shown, the same trends that are present in figure 8 are also present here, overall, the model fits well although there are with the model hitting the peaks of the synthetic data and a period from around 2004 to 2008 where the model barely captures the troughs and differs from the data.

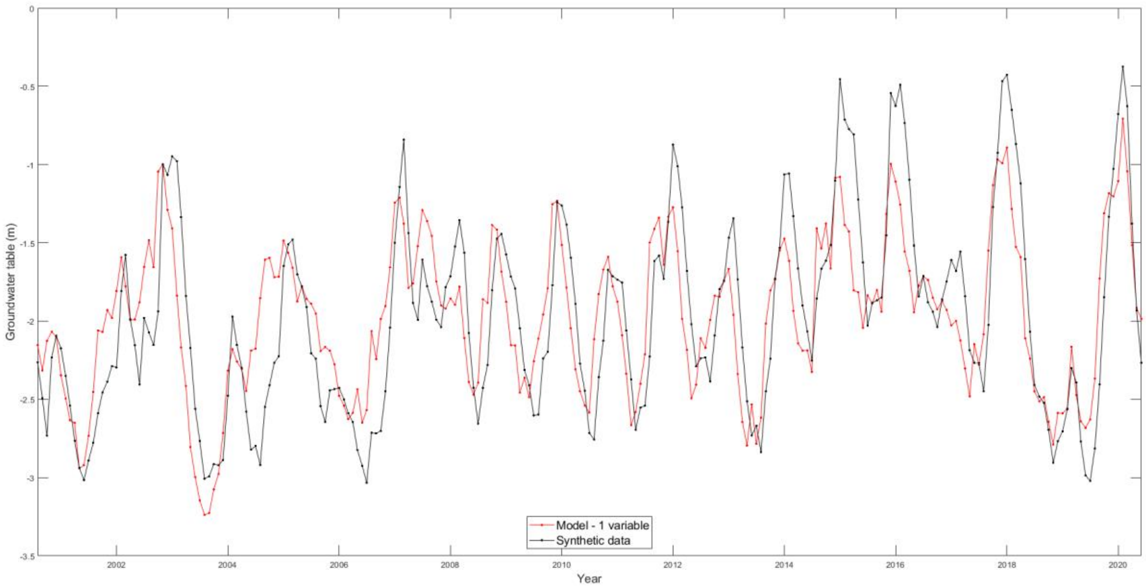

Figure 9: A simple regression for the total synthetic monthly groundwater table.

In table 8 the RMSE and R^2 values for both the daily and monthly models are shown. The daily model can explain 76% of the observed variation in the synthetic data, while the monthly model can explain 62%, which means that the weighted precipitation is able to explain much of the synthetic groundwater table. The RMSE is smaller for the daily model on daily data than for the monthly data.

Table 8: The R^2 and RMSE for the daily 1-variable model for the regression against the daily synthetic groundwater table and monthly model against the monthly synthetic groundwater table.

Synthetic models	**Total period**			
	R^2		**RMSE**	
Model	**Daily**	**Monthly**	**Daily**	**Monthly**
1 variable	0.76	0.62	0.31	0.38

Multiple regression

The daily model seen in figure 10 shows agreement between the models and the data for large parts of the years. Some examples where there is a discordance between the data and the models is in late 2003, where there is a trough where all the models are off by around 0.5 m, with the worst performing being the two-variable model. Periods with great fluctuation in the data, such as late 2001 to late 2002, early 2006, and late 2015 to early 2016 being examples. When the data rises or falls for a longer period the models do quite well. For the two-variable model it falls short for some of the peaks especially in 2015, 2016, 2018.

There is variation between the models when comparing the statistics in table 9, where the five-variable model best predicts the data, and the two-variable model predicts 7% less than that model and has an error that is 0.07 m higher.

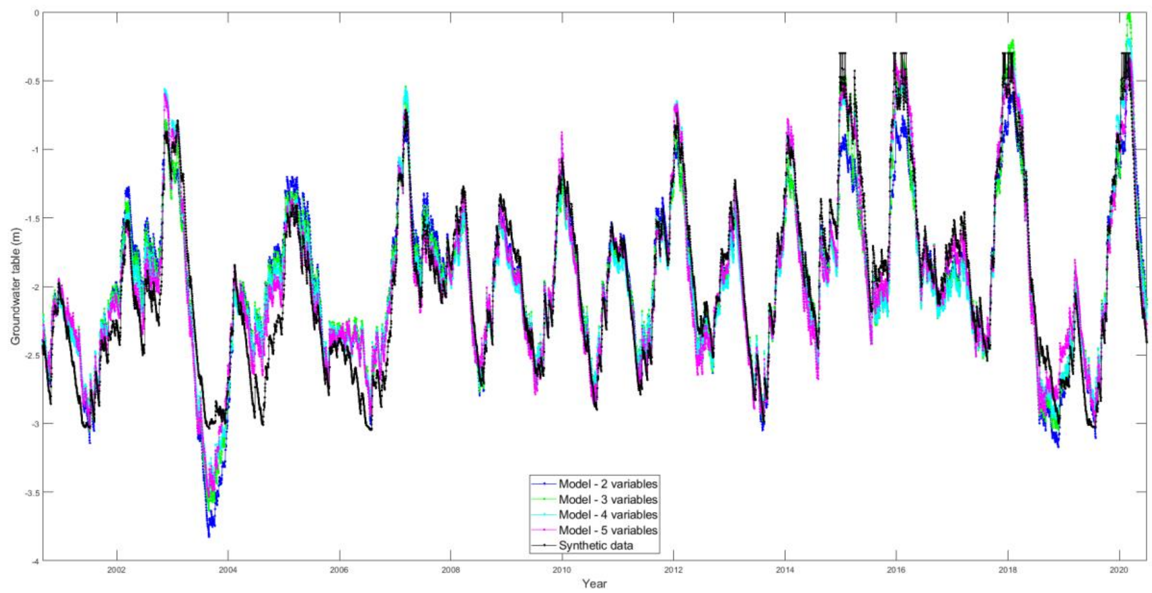

Figure 10: Four daily multiple regression models trained on the total daily amount of synthetic groundwater table.

In the monthly models in figure 11, it is easier to differentiate the models. The two-variable model is the worst of the models at predicting the synthetic data, with this being especially noticeable from the beginning of the plot in 2001 until early 2006 when it is either much higher or much lower than the data. This period seems to be the hardest for all the models to predict, as they all have the same tendency, with the five-variable model being closest to the data. The models simulate the rise and fall of the groundwater table well, while reaching the peaks and troughs is not as well simulated and the smaller fluctuations also cause problems.

Looking at the R^2 and RMSE for the monthly models in table 9, the models seem to improve slightly per added variable, with the span being an increase of 0.13 for R^2 and a decrease of 0.11 for the RMSE going from two to five variables. The biggest increase is going from two to three, within the multiple regression models although the biggest increase by far is going from one variable to two, which increases the model from being able to explain 62% to 80% of the variance of the data and a decrease of 0.1 m in the error.

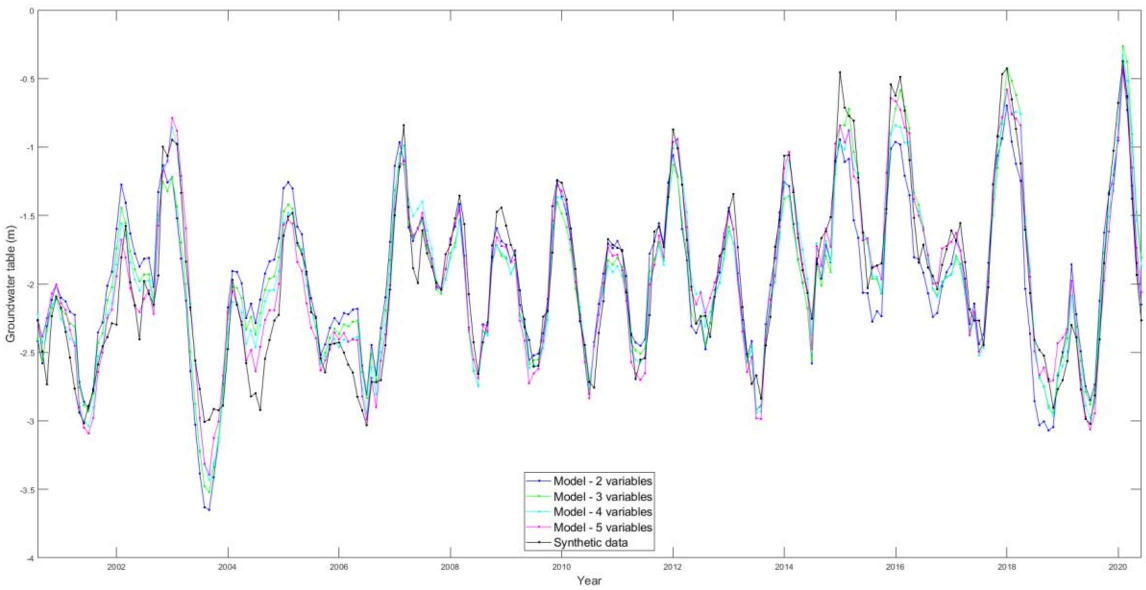

Figure 11: The monthly multiple regression models performed on the total amount of synthetic groundwater table, which is shown in black.

Comparing the statistics between the daily and monthly models in table 9 the five-variable models are identical in these statistics, while for the rest of the models the daily models are slightly better, with the differences between them getting bigger the fewer variables used. When comparing with the simple regression results in table 8 this is when the biggest difference can be seen between the daily and monthly models, with the daily model having a R^2-value that is 0.14 higher than the monthly and the RMSE being 0.7 lower. The biggest gain for both the daily and the monthly model is obtained when going from the one-variable model to the two-variable model.

Table 9: The R^2 and RMSE of the daily and monthly multiple linear regression models calibrated for the total period of synthetic groundwater table.

Synthetic models	Total period			
	R^2		RMSE	
Model	**Daily**	**Monthly**	**Daily**	**Monthly**
2 variables	0.86	0.80	0.24	0.28
3 variables	0.89	0.87	0.21	0.23
4 variables	0.91	0.90	0.19	0.20
5 variables	0.93	0.93	0.17	0.17

In table 10 the synthetic models are held up against the measured groundwater table. The models that perform worst for both the daily and the monthly models for the R^2 are the 1-variable models while for the RMSE the same is the case for the monthly model but for the daily models the 3-, 4- and 5-variable models perform slightly worse. Compared to the fit of the synthetic models to the synthetic data seen in tables 8 and 9 the only part where the fit is better in table 10 is for the R^2 values for the daily model except for the 5-variable model, for the rest of the values the fit is worse. And while in tables 8 and 9 the models get gradually better with more variables this is not the case in table 10, with the best daily model being the 2-variable model and the best monthly model being the 3-variable model. In table 10 there is a big improvement in the metrics when going from one variable to two variables, while the addition of variables beyond that makes the daily model slightly worse as is the case for the 5-variable model for the monthly model while the 4-variable model sees a slight increase in R^2 and a slightly worse RMSE while the 3-variable model sees small improvement in both R^2 and RMSE.

Table 10: The daily synthetic models are compared used on the daily groundwater table, while the monthly synthetic models are compared to the monthly groundwater table.

Synthetic compared to data	Total period			
	R^2		RMSE	
Model	**Daily**	**Monthly**	**Daily**	**Monthly**
1 variable	0.78	0.42	0.34	0.39
2 variables	0.94	0.72	0.30	0.31
3 variables	0.92	0.77	0.35	0.30
4 variables	0.92	0.75	0.35	0.32
5 variables	0.92	0.72	0.35	0.34

Regression

The first step for the regression analysis for the measured data from the site is to do a simple linear regression using the different measured variables to evaluate which ones could best describe the measured groundwater table.

These tests can be seen in figure 12 where the R^2 and the RMSE are plotted in blue in the figure, with the two left side plots showing the daily data and the two plots on the right side showing the monthly plots. As can be seen in the figure the best fit for the raw data is the soil water content at depth 190 cm which has a R^2 value of around 0.8 for both the daily and monthly groundwater table, while also having RMSE values of below 0.2. The reason for this is that the groundwater table is quite close to this depth, with the mean value of the groundwater table being close to this depth, which means that this value is acting as a proxy value for the groundwater table. This also means that using this variable or those around the depth at which the groundwater table fluctuates within is close to being the same as using the groundwater table to predict itself. Therefore, the soil water contents from 93 cm and down are excluded, as they are within the depths at which the measured groundwater table fluctuates within.

Notably the worst fit of all the raw variables is the precipitation, with it having the lowest R^2 values and the highest RMSE values, except for the soil water content at depth 60 cm which is worse. This is noteworthy as this variable would be expected to have a significant impact on the groundwater table. What may be missing is the temporal aspect, as it would be expected that there is a lag between precipitation events and the impact it would have on the groundwater table.

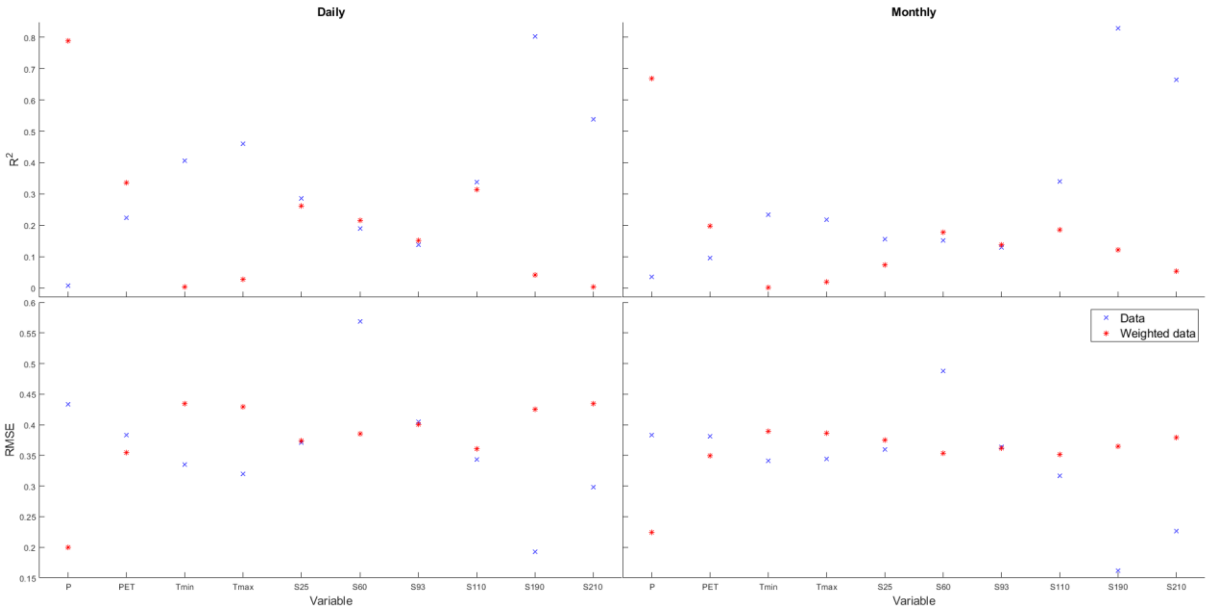

Figure 12: The R^2 and RMSE values for the daily and monthly series. The values for the raw data are shown with a blue cross, while the weighted data are shown with a red asterisk.

That is why a weighting is put on the variables to evaluate if this kind of weighting will impact the fit of the variables. The specific weighting is the above-described weighting where the measurements from the past year are used. The red points in the plot show the fits of these weighted variables. The weighting improves the how well the precipitation describes the groundwater table, improving it to the level that is quite like what the soil water depth at 190 cm had.

Simple regression

The following regressions give the total length of the groundwater table for calibrating the models. The model in figure 13 is calibrated against the groundwater table seen in black in the figure. This model uses the weighted precipitation as the explanatory variable and the daily groundwater table as the variable to be predicted. The monthly data has also been added to the figure to get an idea for how well the periods with no daily data fit.

Even with only one variable there is quite a high correlation between the lines, which can also be seen by the values for R^2 and RMSE reported in table 11, with the R^2 being 0.82 and the RMSE 0.19, which means that the weighted precipitation can explain around 82% of the variation the groundwater table and that the model has an error of around 0.19 m. The drawback is that the time series is short compared to the other available daily time series. Looking at the figure it never follows the data completely, but the tendencies seen in the data are mostly followed by the model seen in figure 13.

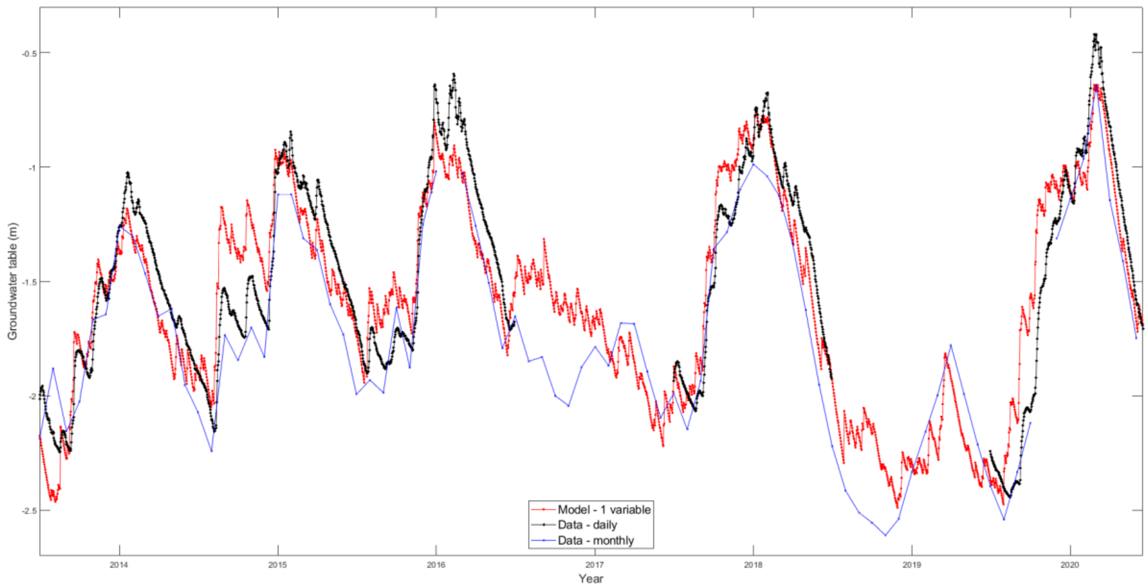

Figure 13: Simple regression, with the model in red, the daily groundwater table values in black, and the monthly groundwater table in blue.

In figure 14 a simple regression using only the weighted precipitation has been done for the total amount of monthly groundwater table data. The model fits worse than the daily model did on the daily data, there is still an overall fit in the tendencies, except for a few cases like in late 2005 to early 2006, where the groundwater table for the data and the model have opposite movements. Still the model mostly follows the tendencies that the data has although it does have problems reaching all the peak values of the data in the beginning of the years.

The values for the fit of models in table 11 show that the R^2 and RMSE of the monthly model is worse than that of the daily model. With the model being able to explain 67% of the variation of the data and the RMSE being slightly worse. This still means that two-thirds of the variation can be explained by only the precipitation for the monthly and more than 80% for the daily model. But we would expect other variables to also influence the groundwater table, which is why the next step will be multiple regression.

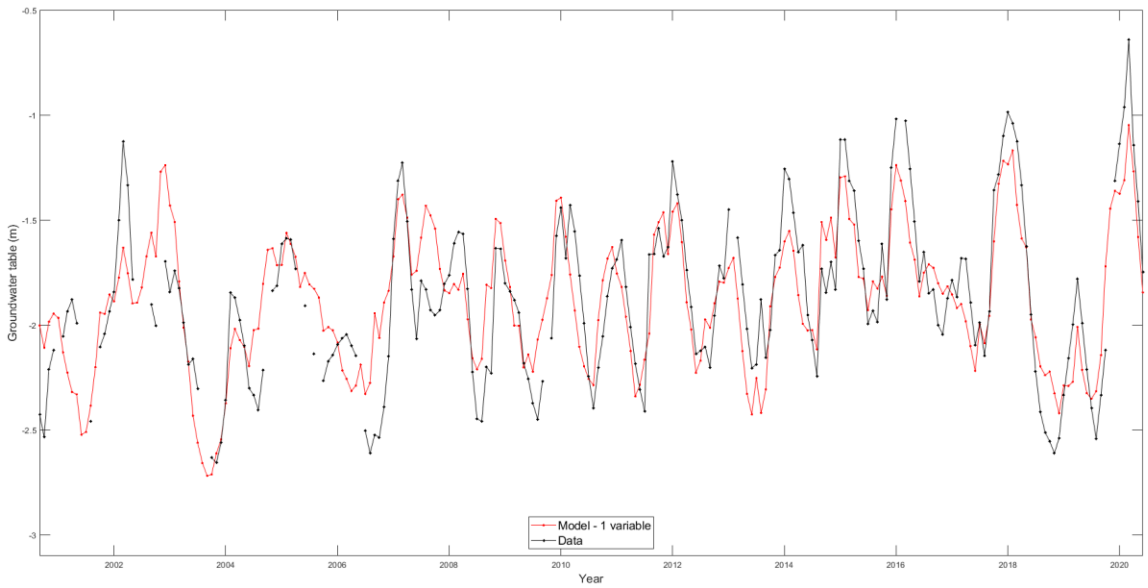

Figure 14: The fit of the model based on the weighted precipitation for the total period of monthly data, with the simple regression model in red and the groundwater table data in black.

Table 11: The coefficient of determination and the root mean square error for the regressions using weighted precipitation as the dependent variable.

Regression for the total period	R^2	RMSE
Daily	0.82	0.19
Monthly	0.67	0.22

Multiple regression

In these multiple regression models the models with 2, 3, 4 and 5 variables are used and evaluated. They are done both for the daily and the monthly data.

In figure 15 the models are compared to the daily data in black and the monthly data in blue. The models fit quite closely to the daily data and follow each other quite closely.

The monthly data has been included to give an idea of where the groundwater table would be in the two years when the daily data is missing, to evaluate if the models have the same tendencies as the data and looking at it the models have the same trend as the data with the 5-variable model performing worst. The model that differs the most is the 5-variable model, which differs for these two periods, and it is also absent from the period where daily data is present in 2017 to 2018 and 2019 to 2020, as it relies on the groundwater table from last year. This also only gives it three years where it is evaluated in comparison to the other models which are compared for five years. The difference between the models and the data is minor in comparison to the 1-variable model in figure 13, but there are instances where the models do not catch up to the peaks and troughs of the data, as seen in the trough in late 2013 and the peak in early 2016, or it lags behind the data when the groundwater table is falling, as seen in early 2014 and early 2018, also it seems that for some periods the models fluctuate more than the data, as seen in the trough in late 2015 and when the groundwater table is falling in mid-2014.

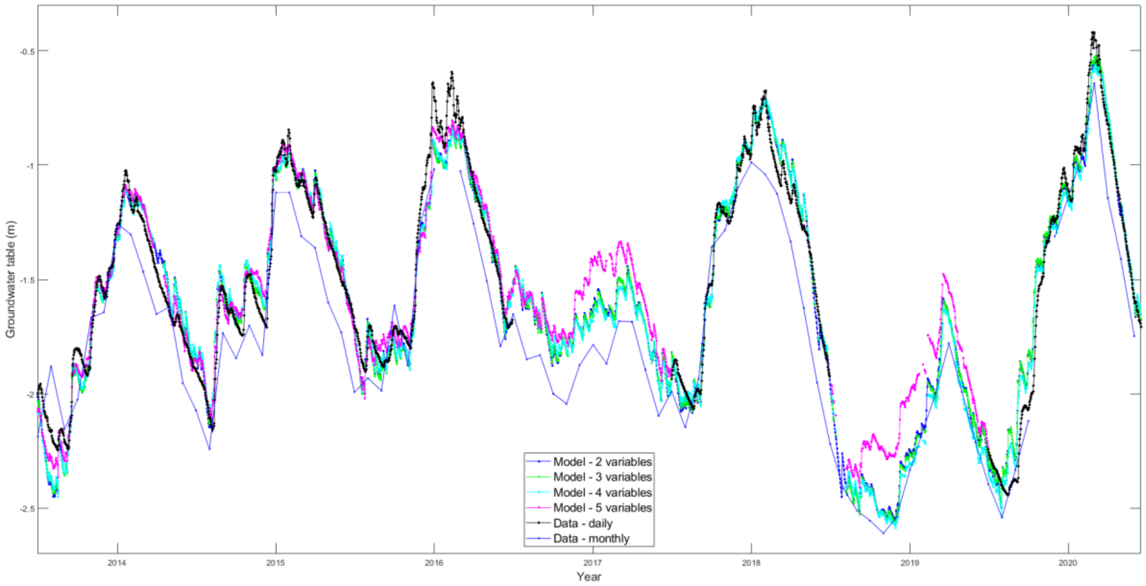

Figure 15: Multiple regression for the total duration of daily data. The models containing 2 to 5 variables are plotted against the daily groundwater table.

When comparing the fits of the models in table 12 they are almost identical and if a third decimal place were not included the only difference would be that the RMSE for the five-variable model is slightly lower. Compared to the one-variable model there is an improvement for both the R^2 and the RMSE, while the multiple regression models themselves are almost identical in this regard.

Table 12: The R^2 and the RMSE for the models with 2, 3, 4 and 5 variables that have been regressed against the total period of daily groundwater table data.

Regression for the total period - daily	R^2	RMSE
2 variables	0.958	0.092
3 variables	0.959	0.092
4 variables	0.962	0.089
5 variables	0.960	0.077

In figure 16 the fit for the multiple regression models against the monthly data is shown. When comparing the multiple regression models with the simple regression model in figure 14, there are several improvements that can be seen, among these being that the models follow the data much more closely and they are better at reaching the peaks that was a problem for the 1-variable model. Also, the discordance for the simple model and the data seen in early 2006 is not seen in these models. But there are points where the simple model was better at modeling the data, such as in late 2013 where the data has a small peak in between the two bigger peaks, the simple model was able to approximate this while the multiple regression models have a trough that goes quite low. In late 2004 the simple model better approximates the trough that is seen there, while the multiple models go to a much lower depth of around 3 m. There are still a few places where the models do not capture the highs and lows of the model, such as the peaks in start 2002, 2003, and 2005, where the models fall short and the troughs in mid to late 2008, 2009, and 2011, where the models are a little too shallow.

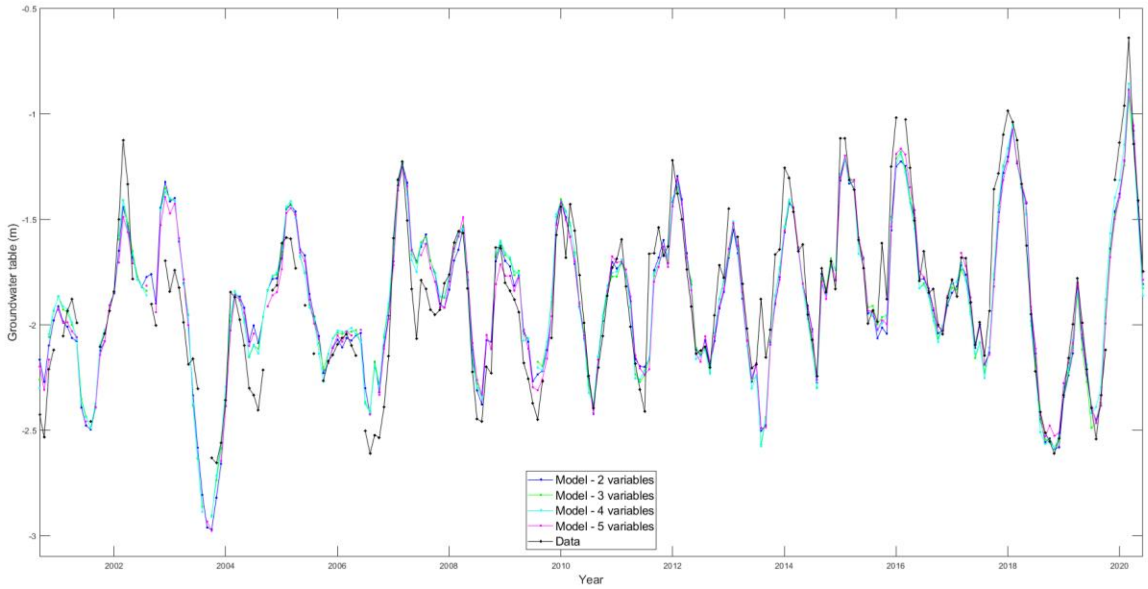

Figure 16: The multiple regression models compared to the monthly groundwater table data, shown in black.

In table 13 the statistics for the fit of the multiple regression models are shown, again as for the daily models they are quite similar again, although the difference is slightly bigger for the R^2 and smaller for the RMSE. The best fit of the models is for the 5-variable model.

Table 13: The R^2 and the RMSE for the multiple regression models that have been compared to the monthly groundwater table.

Regression for the total period - monthly	R^2	RMSE
2 variables	0.833	0.159
3 variables	0.842	0.156
4 variables	0.846	0.153
5 variables	0.848	0.151

Calibration and validation

In this part different calibration lengths are evaluated for the different models. The part of the data that is not used for calibration is used for the validation. The fit of the models is reported for the calibration data, the validation data, as well as the total available amount of data. The models are calibrated from the beginning of the time series as well as from the end of the time series.

Daily models

For the daily models only the models with one to four variables are used, as the amount of groundwater table data is quite limited and will not be sufficient to make a calibration and validation model using the old groundwater table data that the five-variable model uses.

There are two calibration lengths evaluated for the daily models. For the first scenario the models are fed one year of data, and for the second scenario they are fed three years of data. The exact number of days where both model results and groundwater table measurements are present for both scenarios is also reported. The figures and tables for an additional calibration length of two years are in the appendix in figures 26 and 27 and in tables 31 to 33.

One-year calibration

In figure 17 the models calibrated from the beginning are seen, while in figure 18 the models calibrated from the end are seen both for the first scenario. The corresponding R^2-values can be seen in table 14 and the values for RMSE are seen in table 15, while in table 16 the number days used in the various phases are shown.

For all the models seen in figure 17, except for the one-variable model, it is hard to distinguish between the models containing 2, 3, and 4 variables, as mostly these models overlap. For the calibration period the models have good fit with the model, except for the one-variable model which is quite limited in the highs and lows. For the total length of the time series the model with one variable never seems to hit the peaks or troughs the groundwater table data does, and barely goes below -2 m and above -1 m, keeping that model quite constricted. The other models do seem to follow the groundwater table data very well when it rises sharply towards the surface but seems to lag when it falls again; this can be seen in the beginning to middle of the years, especially in 2014 and 2018.

In figure 18 the fit of the models is best for the calibration period. The 1-variable model differs most from the other models. As opposed to the forward calibrated model it is better at fitting the extreme values. It fits best for the calibrated period and the multiple regression models again follow the same tendencies although the 2-variable model has a slightly higher groundwater table than the 3- and 4-variable models.

There are noticeable differences between the models in figure 17 and figure 18, while for the model calibrated from the beginning there is barely a difference between the 2-, 3-, or 4-variable models, while in the backward calibrated model the two-variable model is slightly higher than those two other models. For both calibration methods of the models the worst fit is for the one-variable model, with the forward run model being much more limited in the range compared to the backward run model which spans wider.

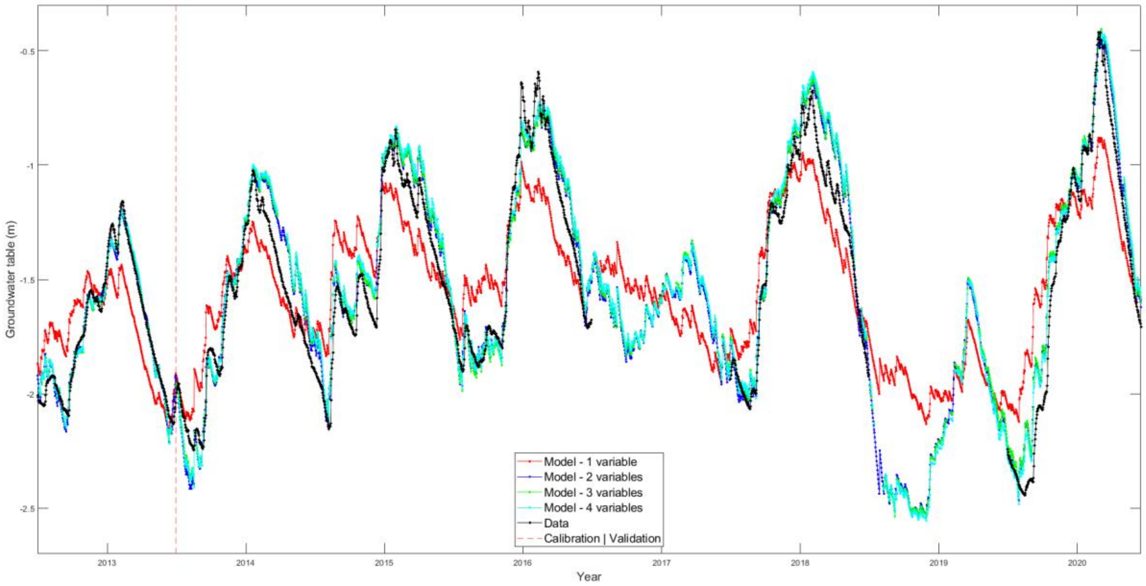

Figure 17: The forward calibration for the first scenario, with the dashed line separating the calibration period on the left side and the validation period on the right side.

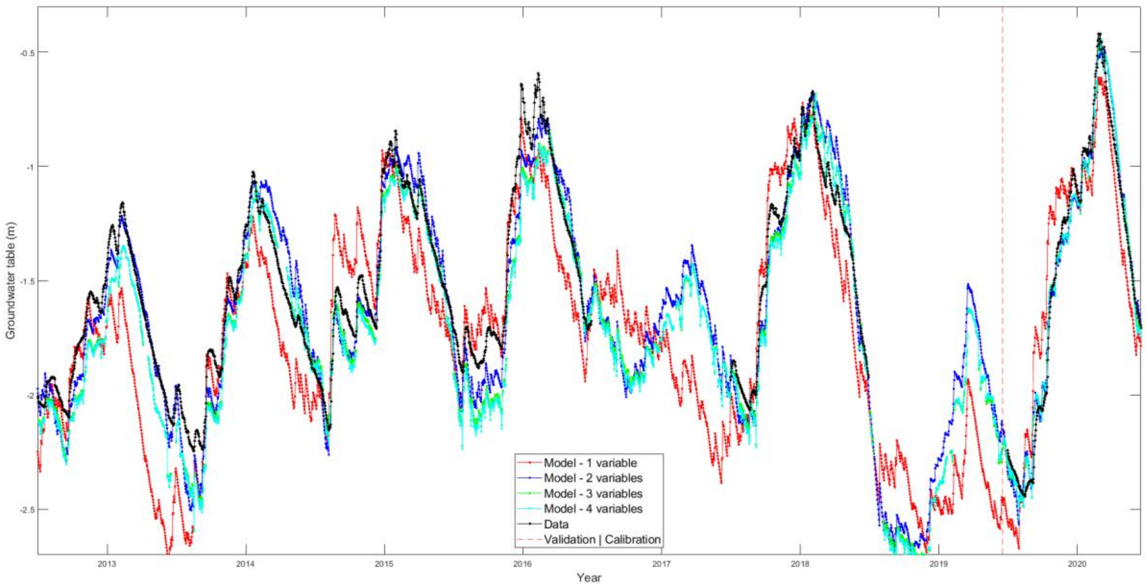

Figure 18: The reverse calibrated model, with the right side of the dashed line showing the calibration period and the left side of the line showing the validation period.

The R^2-values in table 14 the values for the total run of the models are quite similar for both the forward and backward calibrated models, the model calibrated from the beginning has slightly better values, with the models with two or more variables being almost the same and the 1-variable model being the worst for both. The model calibrated from the end performs better for the calibration period but worse for the validation period, which results in a slightly worse total period, which makes sense as the validation period is longer.

The RMSE values in table 15 are worst for both models for the 1-variable model. The RMSE is lower for the model calibrated from the beginning for all three phases and the error is almost the same for the 2-, 3- and 4-variable models within each phase. For the model calibrated from the end there is a slightly bigger difference, 0.03 to 0.04 m, between the errors for the 2-variable model and the 3- and 4-variable model for the validation and total length, with the 2-variable model having the lowest error for those two phases.

When comparing the R^2 and RMSE for the models calibrated on one year of data and the statistics of those models calibrated on the total length of the daily time series in table 11 and 12 to the results in tables 14 and 15, the values are not that far apart. For the forward run models, the R^2 is almost identical, while for the backward run model it is around 0.01 to 0.03 smaller, while the RMSE is noticeably worse, 0.05 to 0.06 m, for the 3- and 4-variable models that are calibrated from the end and slightly less so for the rest of the runs.

In table 16 the number of days for each period is seen. The lengths are almost identical for the differently calibrated models and the data that has been validated on is around five times as much data as it was trained on. The available days for the 1- and 2-variable models equals the available groundwater table data, as these series are unbroken. The missing days for the 3- and 4-variable models is caused by the time series used in these models, the soil water content at different depths, having gaps in the time series.

Table 14: The results for R^2 for the model calibrated for one year of daily data from the beginning (B) and the end (E) of the time series.

Calibration and validation - daily	R^2					
	Calibration		Validation		Total	
Scenario	**B**	**E**	**B**	**E**	**B**	**E**
1 variable	0.501	0.869	0.822	0.776	0.789	0.789
2 variables	0.966	0.982	0.956	0.934	0.957	0.945
3 variables	0.969	0.984	0.955	0.930	0.956	0.939
4 variables	0.969	0.984	0.958	0.928	0.958	0.937

Table 15: The RMSE for the models calibrated from the beginning, B, and the end, E, for the daily data.

Calibration and validation - daily	RMSE					
	Calibration		Validation		Total	
Scenario	**B**	**E**	**B**	**E**	**B**	**E**
1 variable	0.197	0.221	0.230	0.247	0.225	0.243
2 variables	0.052	0.082	0.113	0.121	0.105	0.115
3 variables	0.049	0.078	0.120	0.151	0.111	0.142
4 variables	0.049	0.078	0.125	0.161	0.116	0.150

Table 16: The days used in for the calibration, validation, and the total number of days for the four models for both the calibration models.

Days available for calibration and validation	Days				
	Calibration		Validation		Total
Scenario	**B**	**E**	**B**	**E**	**B,E**
1 variable	352	365	1825	1812	2177
2 variables					
3 variables	335	337	1723	1721	2058
4 variables			1718	1716	2053

Three-year calibration

In figures 19 and 20 the models are calibrated on three years of data, and they are calibrated from the beginning and the end of the time series, respectively. The biggest differences between them are for the 1-variable model that is the worst fitting of the models in both cases, with it being more constricted in the forward run model. The 1-variable model only very occasionally fits well with the data compared to the other models, mostly being above or below the measured value. Looking at the other models there are cases where one model fits better, but this is also usually in the calibration phase, such as late 2013 to early 2014, which fits better in figure 19, and the peak in 2020, which fits better for figure 20, where the models calibrated on the data mentioned perform better.

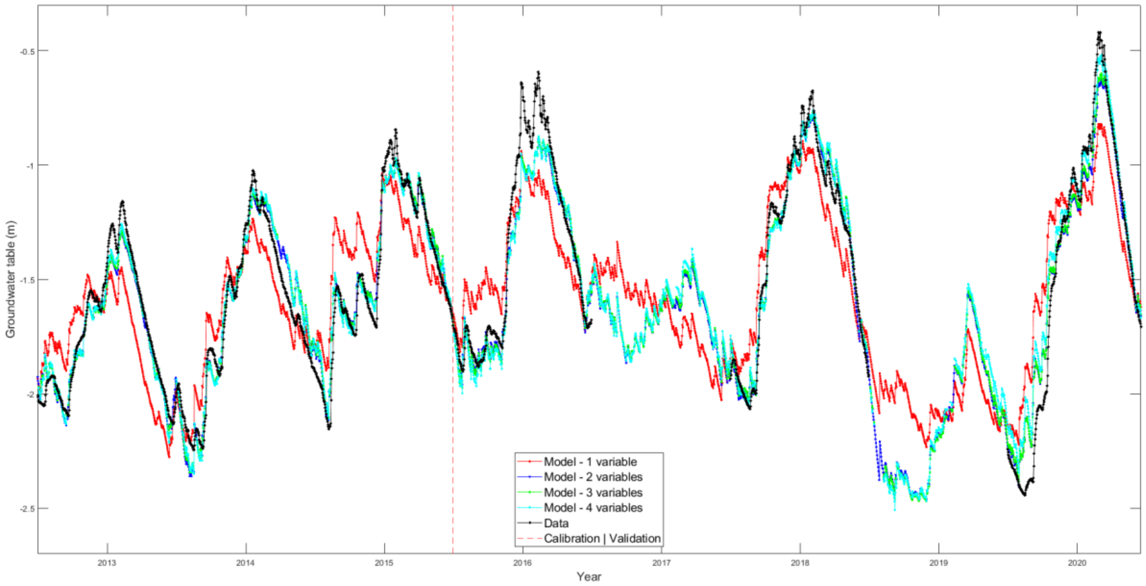

Figure 19: Calibration from the beginning of the time series, with the dashed line showing the boundary between the calibration period, on the left, and the validation period, on the right.

The values for the R^2 and the RMSE in tables 17 and 18 are quite similar, what is worth noticing is that for the model calibrated from the beginning the R^2 values are slightly worse for most of the calibration period, than the validation period, while the opposite is the case for the RMSE. Usually, it would be expected that the calibration period is when the data has the best fit, which is also the case for the calibrated from the end. Evaluating the models on the total groundwater time series the models with multiple variables are comparable except for the 4-variable model which is slightly worse in both fits. The best fits overall are for the 2-variable model, although by a very slim margin. As the calibration and validation periods have more or less just been flipped as each period is around three years of data it makes sense that the results for each are similar. The 1-variable model is the same for both calibration methods in the R^2 while the model calibrated from the beginning is slightly better in the RMSE.

As seen in table 19 the calibration and validation values for both the models are almost equal and the amount of data used for calibration and validation is close to being the same.

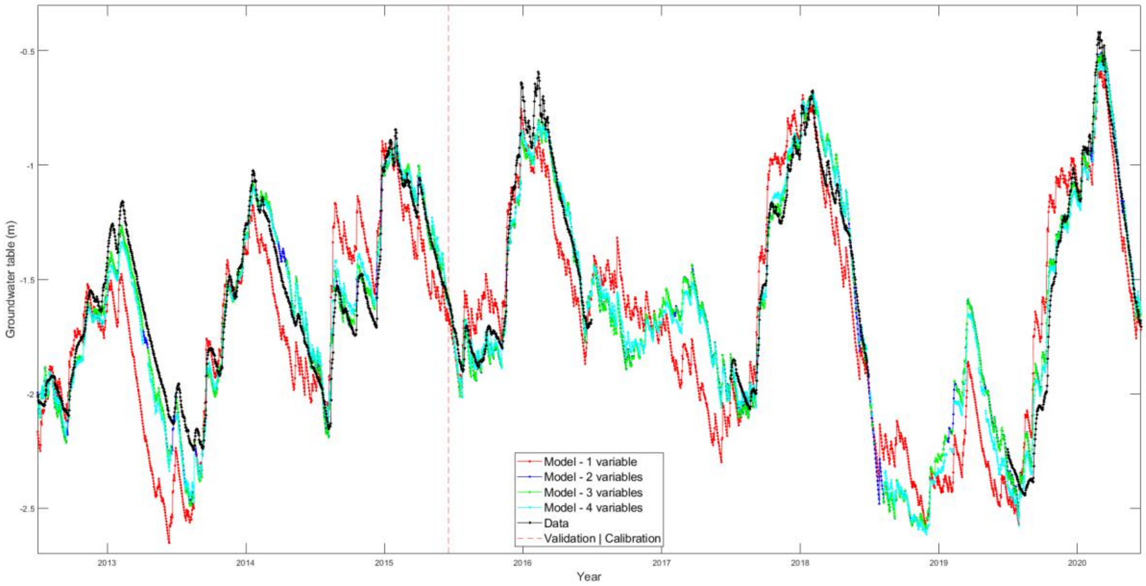

Figure 20: The calibration from the end of the time series, the calibration period is on the right of the dashed line, with the validation period on the left of the line.

Looking at both the calibration lengths the differences are barely noticeable between the models, with all the models being quite similar for both lengths of calibration and with the 2-, 3-, 4-variable models performing quite equally, though the 2-variable model performs the best on the total time series for all the scenarios.

Table 17: The R^2 for the daily models with three years of calibration going from the beginning of the time series as well as the end of the time series.

Calibration and validation - daily	R^2					
	Calibration		Validation		Total	
Scenario	**B**	**E**	**B**	**E**	**B**	**E**
1 variable	0.692	0.852	0.852	0.692	0.789	0.789
2 variables	0.951	0.962	0.961	0.949	0.958	0.957
3 variables	0.953	0.962	0.959	0.950	0.957	0.957
4 variables	0.955	0.967	0.944	0.934	0.948	0.956

Table 18: The RMSE for the three-year calibration from the beginning and end for the daily models.

Calibration and validation - daily	RMSE					
	Calibration		Validation		Total	
Scenario	**B**	**E**	**B**	**E**	**B**	**E**
1 variable	0.188	0.187	0.230	0.244	0.210	0.217
2 variables	0.075	0.095	0.111	0.090	0.095	0.093
3 variables	0.074	0.096	0.112	0.091	0.095	0.094
4 variables	0.072	0.090	0.125	0.101	0.103	0.095

Table 19: The days available for each model for both the three-year calibration period running from the beginning and the end.

Days available for calibration and validation	Days				
	Calibration		Validation		Total
Scenario	**B**	**E**	**B**	**E**	**B,E**
1 variable	1094	1095	1083	1082	2177
2 variables					
3 variables	1047	1052	1011	1006	2058
4 variables	1045	1047	1008		2053

The differences between the results for the models calibrated on three years of groundwater table seen in tables 17 and 18 for the R^2 and the RMSE and the values for the models calibrated on the total length of groundwater table in tables 11 and 12 is minor with the biggest differences being that the 1-variable model is worse, and the 4-variable model calibrated from the beginning performs slightly worse.

Monthly models

The same calibration and validation models are also done for the monthly data, and they are also done with calibration from the beginning of the time series and from the end. This is done for a total of five models and for two different calibration lengths. Additional figures and tables for a calibration length of 90 months are in the appendix in figures 28 and 29 and in tables 34 to 36.

60 months of calibration

In figure 21 the models calibrated on 60 months of data from the beginning are shown. The models with 2-, 3-, and 4-variables are quite like one another, with the two latter being almost indistinguishable and the first differing slightly from these. The 1- and 5-variable models differ more from the other models and each other. In the calibration phase both the 1- and 5-variable models fit quite poorly. All the models barely hit the peaks that the data does, which can be seen in large parts of the validation phase, especially in the beginning of the years from 2014 onwards, but also during the calibration, particularly in early 2002.

When looking at the statistics in tables 20 and 21 the worst performing model is the 5-variable model, which is worse than the 1-variable model, except for the R^2 in the calibration phase, in both RMSE and R^2, while the best performing model is the 4-variable model, with the 3-variable model being quite similar in its performance.

For the models calibrated from the end in figure 22, the models fit very well with the peaks and mostly fit well with the data. The models sometimes overshoot as seen in the trough in late 2003 and the peak in late 2007. The 1-variable model has the worst fit with the data, going the opposite direction to the data in late 2005 to early 2006 and being higher than the other models at multiple occasions such as late 2007 and late 2015, where the other models are also higher than the data. A noticeable disagreement between all the models and the data can be seen in late 2013. Overall, the models seem to perform best when there is a sharp rise or fall into a clear peak or trough, and smaller fluctuations, such as the period during the calibration phase from early 2016 to later 2017.

The R^2 and RMSE for the backward run model are shown in tables 20 and 21. The 1-variable model performs the worst in all metrics for all the phases, while in the calibration phase the 3- and 4-variable model perform best, in the validation phase those same as well as the 5-variable model perform well and for the total length the best performing model is the 5-variable model closely followed by the 4- and 3-variable models and the 2-variable model is slightly behind that.

When comparing the two figures 21 and 22 the beginning period from 2001 to late 2006 seems to be the period of the data that is hardest for both calibration methods to fit. The models that have been calibrated from the beginning have been calibrated on that period of data. Therefore, those models have a worse fit overall compared to the model calibrated from the end. The R^2 in table 20 for the different calibrations are noticeably different, with the calibration period from the forward run model performing much worse than the other model and the calibration period performing worse than the validation results, which is the opposite of the expected result, which is that the calibration period has a better fit than the validation period. When looking at the RMSE in table 21 the model calibrated from the beginning is worse in all phases when compared to the model run from the end. When comparing their fit for the total time series the end-calibrated model has an overall better fit, with the 1-variable models being most alike for the total period and the biggest difference for the total period being for the 5-variable models.

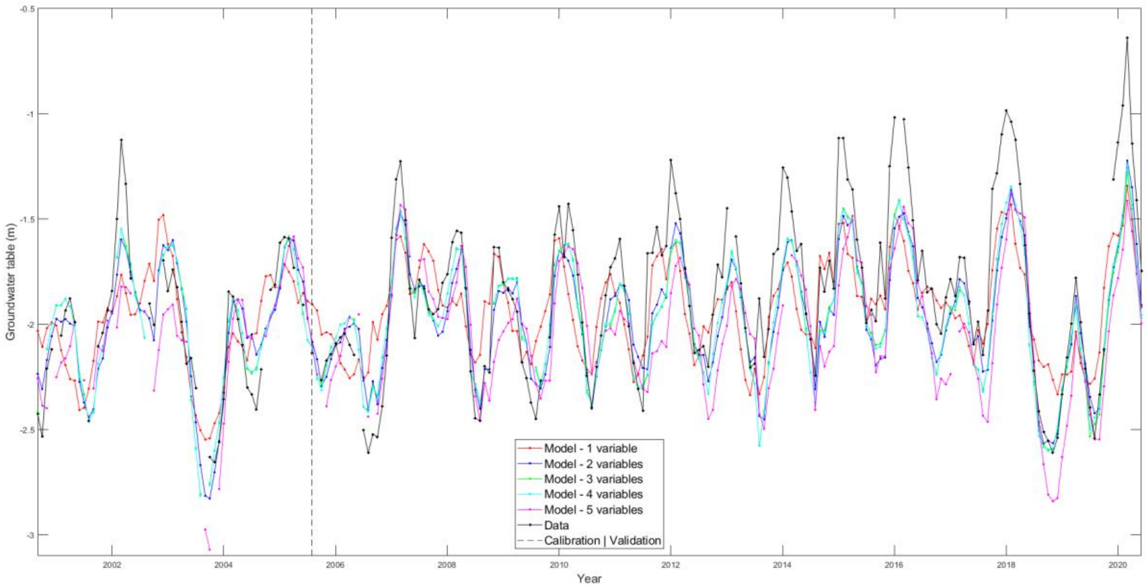

Figure 21: The calibration length used is 60 months and is indicated with the dashed line, with the data on the left of it being used for calibrating the models and the data on the right being used for validation.

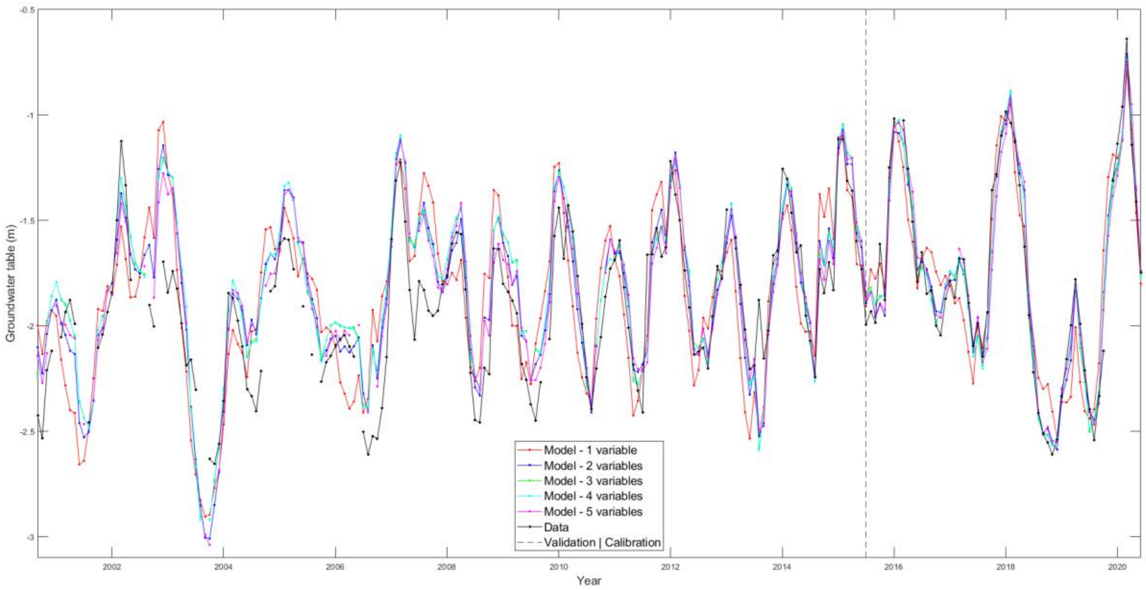

Figure 22: The models are calibrated from the end of the time series, which is on the right side of the dashed line and validated on the data on the left of the dashed line.

Table 20: The R^2 values for the calibration period of 60 months with the values for both the calibration from the beginning and the calibration from the end of the time series.

Calibration and validation - monthly	R^2					
	Calibration		Validation		Total	
Scenario	**B**	**E**	**B**	**E**	**B**	**E**
1 variable	0.511	0.878	0.705	0.558	0.668	0.668
2 variables	0.767	0.951	0.823	0.751	0.798	0.822
3 variables	0.819	0.958	0.824	0.780	0.807	0.839
4 variables	0.820	0.958	0.830	0.781	0.812	0.841
5 variables	0.575	0.947	0.622	0.777	0.612	0.845

Table 21: The RMSE for the 60-month calibration period, with the results for both the calibration from the beginning and the end of the time series.

Calibration and validation - monthly	RMSE					
	Calibration		Validation		Total	
Scenario	B	E	B	E	B	E
1 variable	0.229	0.167	0.255	0.265	0.250	0.242
2 variables	0.158	0.105	0.223	0.200	0.211	0.180
3 variables	0.143	0.098	0.225	0.199	0.211	0.177
4 variables	0.143	0.098	0.223	0.198	0.209	0.176
5 variables	0.239	0.112	0.311	0.178	0.299	0.163

When looking at table 22 where the available months are seen for the models, the forward run model has about a year less of data compared to the calibration from the end. As the 1- and 2-variable models are based on precipitation and evapotranspiration and those series are unbroken, the months used by those models equal the available groundwater table data. So, comparing those models to the ones with 3-, 4- and 5-variables the calibration from the end has more available data for the calibration period as the values are almost the same for all models when the models are calibrated from the end. Both the gaps in the groundwater table data from the beginning of the time series as well as the gaps in the soil water contents have affected that the results of the forward run models are slightly worse than the model run from the end.

Table 22: The months of data available for the models based on the groundwater table data and the presence of the variables used for the models.

Months available for calibration and validation	Months				
	Calibration		Validation		Total
Scenario	B	E	B	E	B,E
1 variable	47	59	172	160	219
2 variables					
3 variables	42	59	169	152	211
4 variables					
5 variables	38	56	164	146	202

120 months of calibration

In this section the models calibrated on a period of 120 months are shown, the exact number can be seen in table 25. The calibration from the beginning is shown in figure 23 and the calibration from the end is shown in figure 24.

Many of the same faults found in the models with 60 months of calibration in figures 21 and 22 are also found in figures 23 and 24. Overall, the figures are quite similar for the models calibrated the same way. The biggest differences in the models calibrated from the beginning seem to be that the models are slightly closer to the data, such as the peaks in 2014, 2015, 2016 and 2018. For the model calibrated from the end a slight improvement can be seen for the trough in late 2003.

Comparing the R^2 and the RMSE for the 120-month calibration models, which can be seen in tables 23 and 24, the calibration from the end gives the better results in all metrics except for the R^2 for the validation and for the total time for the 1-variable model where both models are equal. This is the same that was seen for the models calibrated on 60 months of data and compared to those the biggest improvement is seen for the model calibrated from the start of the time series, which is better in almost every metric, while the improvement for the other models is negligible in comparison. The best models for the total length of the 120 months of calibration from the end are the 4-variable model, which performs best for the R^2-values and the 5-variable model which performs best in the RMSE, closely followed by the 3-variable model and the 2-variable model. For the model calibrated from the start the RMSE is lowest for the 2-variable model and the R^2 highest for the 4-variable model and those along with the 3-variable model are performing well. The 5-variable model is improved compared to the 60-month calibration, but still performs worse than the other models. The worst performing model in both cases is the 1-variable model which performs the same for the 60 and the 120 months of calibration time with only small variations in the RMSE differentiating these.

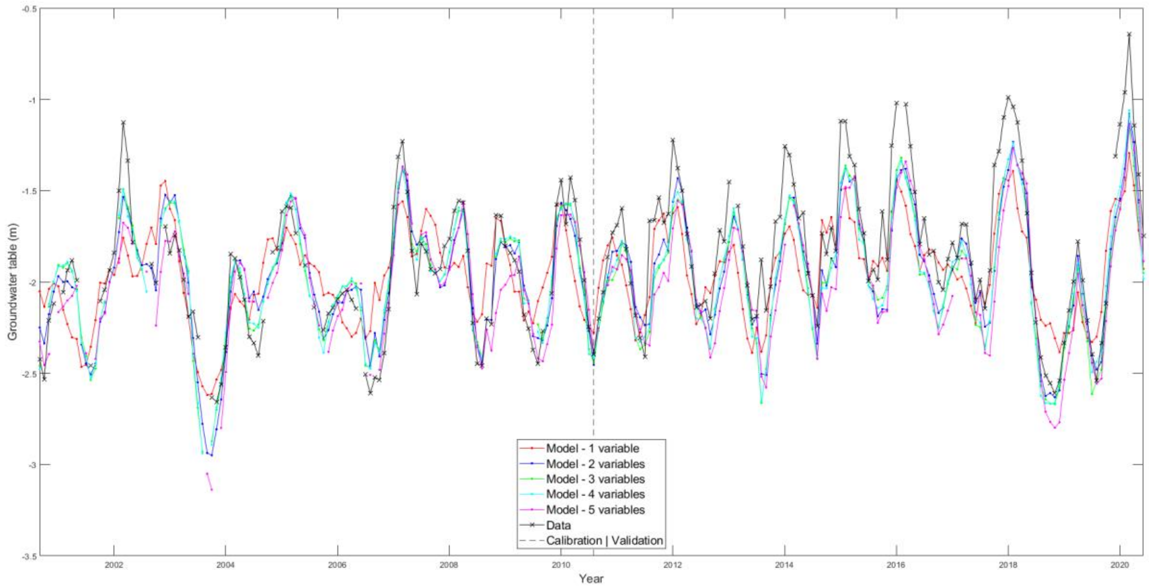

Figure 23: The models are calibrated with 120 months of data from the beginning of the time series with the calibration period on the left side of the dashed line and the validation period on the right side.

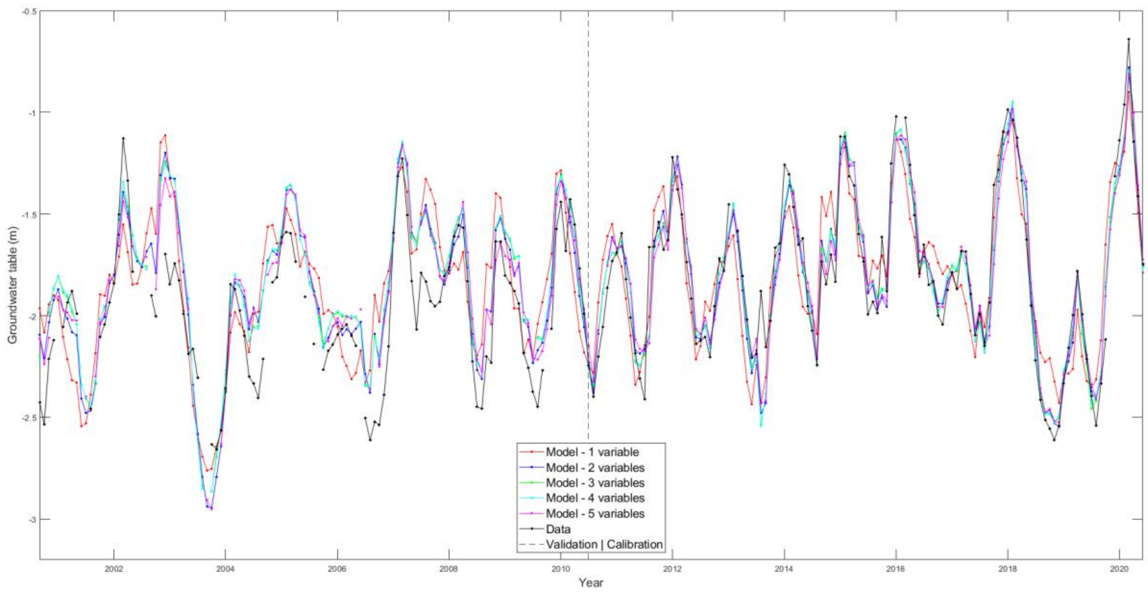

Figure 24: The models are calibrated with 120 months of data from the end of the time series with the dashed line dividing the calibration period on the right and the validation period on the left.

Table 23: The R^2 for the models calibrated on 120 months of data, both from the beginning and the end of the time series.

Calibration and validation - monthly	R^2					
	Calibration		Validation		Total	
Scenario	**B**	**E**	**B**	**E**	**B**	**E**
1 variable	0.516	0.796	0.790	0.507	0.668	0.668
2 variables	0.812	0.902	0.857	0.757	0.816	0.826
3 variables	0.849	0.907	0.879	0.804	0.819	0.839
4 variables	0.853	0.908	0.879	0.810	0.828	0.842
5 variables	0.752	0.895	0.787	0.791	0.748	0.841

Table 24: The RMSE of the models using 120 months for the calibration from both the beginning and the end of the time series.

Calibration and validation - monthly	RMSE					
	Calibration		Validation		Total	
Scenario	**B**	**E**	**B**	**E**	**B**	**E**
1 variable	0.233	0.185	0.257	0.279	0.246	0.233
2 variables	0.145	0.128	0.218	0.214	0.187	0.173
3 variables	0.132	0.124	0.237	0.212	0.196	0.170
4 variables	0.130	0.123	0.228	0.211	0.190	0.169
5 variables	0.177	0.134	0.285	0.194	0.243	0.163

Again, as in the models calibrated on 60 months of data there is more data used for the calibration from the end and the variable data is more intact from this end. As can be seen in the total the models with 1 and 2 variables have 219 points of groundwater table data to compare to, while the 3- and 4-variable model only has 211, as the three extra variables limit the number of months that overlap where both the data and variables are present. The difference between the total amount of months is eight and the difference for the months of calibration for the forward run models with 1- and 2-variables and models with 3- and 4-variables is eight, meaning that all of months where soil water content is missing are in the calibration months of the model calibrated from the beginning.

Table 25: The months of data available where both groundwater data and model variables are present for both calibration period and validation period as well as the total length of the time series.

Months available for calibration and validation	Months				
	Calibration		Validation		Total
Scenario	**B**	**E**	**B**	**E**	**B,E**
1 variable	104	117	115	102	219
2 variables					
3 variables	96	114	115	97	211
4 variables					
5 variables	90	114	112	88	202

Daily model for the monthly data

In figure 25 the results are shown for the daily models calibrated on the total amount of daily data and are used to predict the monthly groundwater data. Compared to the monthly models calibrated on the total length of groundwater data, which are seen in figure 14 and 16, the models are not as good at getting the exact right values of the groundwater table but are good at capturing the trends of the rising and falling groundwater table quite well, the 1-variable model being the exception. The other three models align with each other most of the time and are mostly hard to distinguish from each other. For the timespan when daily data is present the models perform well, which makes sense as the models have been calibrated on this data, while for the monthly data there is a bigger offset between the models and the data.

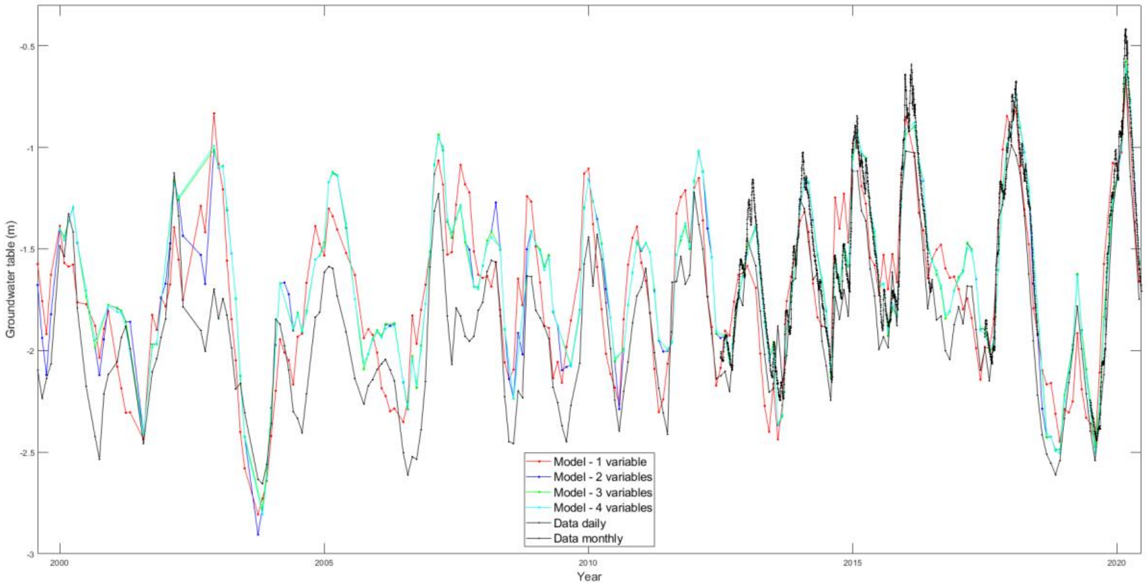

Figure 25: The daily model used on the monthly data, with the daily data also shown. The presence of data is shown with a dot on the lines, which in this figure are all linked with lines even when there are missing months between the dots.

The R^2 and RMSE for the daily model fitted against the monthly data is shown in table 26. When comparing the R^2-values against the values for the monthly model in tables 11 and 13 the values for the 2-, 3- and 4-variable models are very similar to the results of the monthly model, while the 1-variable model is worse. When comparing the RMSE against the monthly model in tables 11 and 13 the values are noticeably worse for all the models, being around 0.11 m higher. This is quite a high number and is also noticeable in figure 25, where the models do not align as well with the data as when compared to the monthly models in figures 14 and 16.

Table 26: The R^2 and RMSE for the daily model used on the monthly data.

Daily model on monthly data	R2	RMSE
1 variable	0.577	0.329
2 variables	0.848	0.269
3 variables	0.846	0.267
4 variables	0.843	0.265

Discussion and conclusion

In the trends shown in the results, the trends were increasing at a significant rate for both the temperature and the potential evapotranspiration, while the same is not the case for the precipitation even for a timespan of 36 years. If more data were present, it might be that a trend with some form of significance level could be found for the precipitation. It would be expected that if both the temperature and the potential evapotranspiration have a significant upward trend and the precipitation stays the same, that the groundwater table would see a fall. But this is not the case for the trends shown, with the yearly groundwater table seeing a significant upward trend. It could be speculated that the precipitation also has an upward trend but that more data would be needed to find it at a significant level, as the precipitation seems to fall more randomly over the years. It would also be interesting looking at the other PLAP fields to see if the same trends are seen for all the places or if a trend can be seen for the same variables at those sites.

The linear weighting used is quite simple to implement and is based on the assumption that the data that lies before has an accumulative effect on the present-day groundwater table, in this case a year before and favors the more recent data. The weighting used is not necessarily the most descriptive for what is taking place but as can be seen in the results it is quite robust and for the variables it is used on it can help explain most of the change seen in the groundwater table. The weight is linear and uses a year of data. Both of these factors could be changed using e.g., a Poisson distribution and playing with the amount of data that is used for the weighting and finding the optimal for the chosen variables. Further work that could be done is to further explore other types of weighting that may better reflect the effect that the precipitation and potential evapotranspiration has on the groundwater table.

Looking at the results of the daily regression models in tables 11 and 12, which use the total amount of data as calibration, the worst performing daily model is the 1-variable model with the other models performing similarly and better in both R^2 and RMSE.

When comparing the results for the calibration on the total amount of data to the models validated on the total amount of data using one year of data as calibration either from the beginning or the end of the time series seen in tables 14 and 15 the best performing model is the 2-variable model. The R^2 is almost identical in the forward run model and off by 0.013 in the backward run model when compared to the values in table 11 and 12, while for the RMSE the values are off by 0.013 and 0.023 for the forward run and backward run 2-variable model. This is the best performing of the models with the 3- and 4-variable models being more off in both the R^2 but also in RMSE.

When comparing the R^2 and RMSE in tables 11 and 12 with the ones for the total length of time series in tables 17 and 18, which have been calibrated on three years of data, the best performing models are the 2- and 3-variable models, which are almost identical in R^2 and RMSE compared to the values in table 11 and 12, while the 4-variable model is not far behind although the forward run model is slightly worse in both R^2 and RMSE being off by 0.01 in both.

Overall, for the daily models the 2-variable model seems to perform the best overall and being the most robust both in the amount of calibration data needed and being a quite simple model. The 2-variable model is much improved compared to the 1-variable model and adding more variables does not add much to the models.

In the calibration validation for the daily data the amount of calibration data does not make a substantial difference for the quantity of data present here. It is a short time series for the groundwater table compared to the length of some of the other daily time series. The weighting of the precipitation and the evapotranspiration do give the models a lot of information about the previous year, so even though the calibration is only for a year the weighted variables contain information relating to the previous year.

For the monthly models in tables 11 and 13 the 1-variable model is the worst performing model in both RMSE and R^2. With the 2, 3, 4 and 5 variable models increasing slightly in R^2 and decreasing slightly with RMSE with each variable.

When comparing with the model calibrated on 60 months of data and validated on the total amount of data which can be seen in tables 20 and 21 again the 1-variable model performs bad, but the 5-variable model actually performs worse than all the other models in the data calibrated from the beginning and better than all the models when calibrated from the end the difference in R^2 for the good and the bad model is 0.23 and for the RMSE it is 0.13 m which is a big difference. The backward and forward calibrated 2-, 3- and 4-variable models differ 0.02 for the 2-variable model and 0.03 for the other two models in the R^2, while for the RMSE the difference is 0.03 m. This is low in comparison to the 5-variable model.

For the models calibrated on 120 months of data which are seen in tables 23 and 24 the R^2 and RMSE validated on the entire period do improve compared to the values for the 60 months of calibration, although the models calibrated from the beginning still perform worse than those calibrated from the end, with the 5-variable model still performing very badly.

Although the 2-, 3- and 4-variable models perform quite similarly, the addition of variables beyond the two does not improve the performance much, sometimes even making it worse. The improvement when going from the 1-variable to the 2-variable model is more notable.

The daily models used on the monthly data in table 26 perform quite a lot like the monthly models in table 11 and 13 when it comes to R^2, except for the 1-variable model which performs worse but is much worse when it comes to the RMSE where it is 0.11 m higher. Although the daily model has been trained on more data points, it has a shorter time horizon than the total length of the model but is similar in length to the 60-month calibration period, which has a better RMSE.

As can be seen for the monthly models in particular, the period chosen for calibration can have a big effect on how well the regression model performs. With the 5-variable model being especially vulnerable, the 1-variable model being the least vulnerable, and with the 2-, 3- and 4-variable models being affected, but still close to the model modeled on the total amount of data. So, while those three models do perform worse when calibrated on uncharacteristic data, the results are still quite robust.

The 1-variable model performs quite similarly for all the calibration methods, while the 5-variable model is the most vulnerable either being slightly better than the other models or much worse than all but the 1-variable model. The middle ground between those models are the 2-, 3- and 4-variable models which all perform better than the 1-variable model in all cases.

Other factors may also affect the groundwater table, factors that are outside the field at Jyndevad, such as a change in the number of trees. A change in groundwater table is not the same as a change in recharge

Although for the monthly calibration and validation there some of the models with more variables perform slightly better, the 2-variable model is quite robust, while being quite a simple model. Precipitation is easily available for most areas, while potential evapotranspiration can be calculated based on temperature and radiation. The variables used in the 2-variable model make sense in a hydrological perspective as it would be assumed that the main input is the precipitation and evapotranspiration would be a limiter on how much of the water reaches the groundwater table. The same would be the case for the recharge.

Although the 3- and 4-variable models do improve the models in some cases it is usually a minor improvement and the soil water content at those depths do not help much in the model and are also not necessarily readily available at most sites.

The worst performing model is the 1-variable model, which uses only the weighted precipitation. But although it is only worse when compared to the other models. It can explain quite a lot with a minimum of input.

The inclusion of the groundwater table from the past year was done to get a model that would be the best, as this model performed quite well for some tests, e.g., the synthetic models and in the total regressions in table 12 and 13. The model is not very feasible in a real-world scenario as the groundwater table is used to predict itself. The model that includes the variable is very variable and is not particularly useful.

In the ideal scenario which is used when making models for the synthetic data each added variable increases the R^2 and decreases the RMSE. This happens for both the daily and monthly synthetic data. But when using the measured data the same is not the case. It is interesting to note that the synthetic models in tables 8 and 9 perform worse in both R^2 and RMSE than the regression models based on the measured data in tables 11 to 13.

The time discretization that would make most sense for making a model is the daily measurements and a longer daily series would be an interesting case to work with and the actual evapotranspiration could be of interest.

Further work would be to investigate other methods of machine learning such as long short-term memory (LSTM), which is a type of recurrent neural network typically used for time series forecasting. This approach could be used to try improving upon the model, although it would require more time and not necessarily improve much on the results.

The models found here could be evaluated for the other sites within PLAP. Testing the model found for this sandy site on the other sandy site in Tylstrup could be further work. Also, whether the 2-variable model with the same type of weighting would perform similarly on the other sites within PLAP, both the sandy and clayey till sites.

Overall, the regression models can simulate the groundwater table quite well using only the 2-variable model which consist of data that is easily available for most areas in Denmark.

References

Cuthbert, M.O., Taylor, R.G., Favreau, G., Todd, M.C., Shamsudduha, M., Villholth, K.G., MacDonald, A.M., Scanlon, B.R., Kotchoni, D.O.V., Vouillamoz, J.-M., Lawson, F.M.A., Adjomayi, P.A., Kashaigili, J., Seddon, D., Sorensen, J.P.R., Ebrahim, G.Y., Owor, M., Nyenje, P.M., Nazoumou, Y., … Kukuric, N. (2019). Observed controls on resilience of groundwater to climate variability in sub-Saharan Africa. Nature (London), *572*(7768), 230–234. https://doi.org/10.1038/s41586-019-1441-7

Gumiere, S.J., Camporese, M., Botto, A., Lafond, J.A., Paniconi, C., Gallichand. J., Rousseau, A.N. (2020). *Machine Learning vs. Physics-Based Modeling for Real-Time Irrigation Management*. Frontiers in Water, 2. https://doi.org/10.3389/frwa.2020.00008

Huang, X., Gao, L., Crosbie, R.S., Zhang, N., Fu, G., & Doble, R. (2019). Groundwater Recharge Prediction Using Linear Regression, Multi-Layer Perception Network, and Deep Learning. Water (Basel), 11(9), 1879–. https://doi.org/10.3390/w11091879

Kraalingen, D.W.G. van and Stol, W. (1997). Evapotranspiration Modules for Crop Growth Simulation - Penman, Makkink and Priestley-Taylor. AB-DLO.

Lindhardt, B., Abildtrup, C., Vosgerau, H., Olsen, P., Torp, S., Iversen, B.V., Jørgensen, J.O, Plauborg, F., Rasmussen, P., and Gravesen, P. (2001). The Danish Pesticide Leaching Assessment Programme: Site characterization and monitoring design, Geological Survey of Denmark and Greenland, Copenhagen.

http://pesticidvarsling.dk/wp-content/uploads/Etableringsrapport/plap1_sept-2001.pdf

Maulud, D., & Abdulazeez, A.M. (2020). A Review on Linear Regression Comprehensive in Machine Learning. *Journal of Applied Science and Technology Trends*, *1*(4), 140-147. https://doi.org/10.38094/jastt1457

MathWorks. (2021). Fit linear regression model.
https://www.mathworks.com/help/stats/fitlm.html

Miljøstyrelsen (1992). Danmarks fremtidige vandforsyning. ISBN: 87-503-9581-5.

https://www2.mst.dk/Udgiv/publikationer/1992/87-503-9581-5/pdf/87-503-9581-5.pdf

Miljø- og Fødevareministeriet (2017). Pesticidstrategi 2017 - 2021. ISBN: 978-87-7120-898-6. https://mst.dk/media/141516/pesticidstrategi2017-2021.pdf

Miljø- og Fødevareministeriet (2018). Procedure for udvælgelse af stoffer til grundvandsovervågningen (GRUMO) og pesticidlisten.

https://mst.dk/media/169559/bilag-til-afrapportering-fra-vandpanelet.pdf

Mirus, B.B., Loague, K., Cristea, N.C., Burges, S.J., and Kampf, S.K. (2011). A synthetic hydrologic-response dataset. *Hydrological Processes*, *25*(23), 3688–3692. https://doi.org/10.1002/hyp.8185

Rosenbom, A.E., Karan, S., Badawi, N., Gudmundsson, L., Hansen, C.H., Nielsen, C.B., Plauborg, F. and Olsen, P. (2021). The Danish Pesticide Leaching Assessment Programme: Monitoring results May 1999-June 2019. Geological Survey of Denmark and Greenland. Copenhagen. ISBN: 978-87-7871-532-6

pesticidvarsling.dk/wp-content/uploads/2021/01/The-Danish-Pesticide-Leaching-Assessment-Programme-2019-.pdf

Salmi, T., Määttä A., Anttila, P., Ruoho-Airola, T. and Amnell, T. (2002). Detecting trends of annual values of atmospheric pollutants by the Mann-Kendall test and Sen's slope estimates – the Excel template application MAKESENS. Finnish Meteorological Institute. Helsinki. ISBN: 951-697-563-1.

Šimůnek, J., Šejna, M., Saito, H., Sakai, M., and van Genuchten, M.T. (2013). The HYDRUS-1D Software Package for Simulating the One-Dimensional Movement of Water, Heat, and Multiple Solutes in Variably-Saturated Media, Version 4.17. Department of Environmental Sciences, University of California Riverside.

Taboga, M. (2021). QR decomposition, Lectures on matrix algebra.

https://www.statlect.com/matrix-algebra/QR-decomposition

Appendices

MATLAB code example

Example of MATLAB code (in italics):

```
%% Read data and convert to timetable
data = readtable('C:\...\Data.xlsx');
ttdata = table2timetable(data);
%% weighting of precipitation and evapotranspiration
step = 1/364;
w = 0:step:1;
w = w';
P = ttdata.Precip
PET = ttdata.PEvap
start = aaa % define start
end = zzz % define end
wP=[];
wPET=[];
for i=start+365:end
   Pr(i) = sum(P(i-364:i).*w);
   wP(i-364) = Pr(i);
   Et(i) = sum(PET(i-364:i).*w);
   wPET(i-364) = Et(i);
end
wP = (wP'/182.5);
wPET = (wPET'/182.5);
%% Define other variables
GWT = ttdata.GWT(start+365:end)
GWTold = ttdata.GWT(start:end-365)
S25 = ttdata.S25(start+365:end)
S60 = ttdata.S60(start+365:end)
PETw = wPET
Pw = wP
%% 5 variables / daily
```

%% All these variables need to have the same length and fit with the dates

X = [S25, S60, GWTold, PETw, Pw];

mdl = fitlm(X,GWT)

%% For plotting the model results

x1 = S25;

x2 = S60;

x3 = GWTold;

x4 = ET;

x5 = Pw;

estim = mdl.Coefficients.Estimate;

*y1 = estim(2)*x1 + estim(3)*x2 + estim(4)*x3 + estim(5)*x4 + estim(6)*x5 + estim(1);*

Constants for the regression models

When doing linear regression analysis each independent variable results in a constant being multiplied to this variable as well as a constant term that is added to the equation. In the tables 26 to 29 these constants that are part of the regression equation and denoted in the equation in the method section as β are shown with each column showing the values that the constants would have for the different models. Interc. is the intercept with the y-axis while the remaining values are the constants that are used on the variables to get the best linear regression fit.

Table 27: The best estimates for the constants in the regression models for the daily measured groundwater table.

Regression models for the total period - daily	Constants					
	Interc.	Pw	PETw	S25	S60	GWTold
1 variable	-4.11	0.80	-	-	-	-
2 variables	-3.01	0.68	-0.44	-	-	-
3 variables	-3.01	0.70	-0.44	0.00	-	-
4 variables	-2.97	0.73	-0.43	0.00	-0.02	-
5 variables	-2.48	0.63	-0.33	0.00	-0.01	0.22

Table 28: The best estimates for the five regression models for the total length of the daily synthetic time series.

Synthetic model - daily	Constants					
	Interc.	Pw	PETw	S25	S60	GWTold
1 variable	-5.05	1.03	-	-	-	-
2 variables	-4.17	1.01	-0.50	-	-	-
3 variables	-4.42	0.85	-0.43	0.04	-	-
4 variables	-4.34	0.75	-0.37	0.00	0.04	-
5 variables	-4.16	0.80	-0.32	0.00	0.04	0.14

Table 29: The best estimates for the constants in the regression models for the monthly measured groundwater table.

Regression models for the total period - monthly	Constants					
	Interc.	Pw	PETw	S25	S60	GWTold
1 variable	-3.61	0.02	-	-	-	-
2 variables	-2.98	0.02	-0.01	-	-	-
3 variables	-3.12	0.02	-0.01	0.01	-	-
4 variables	-3.26	0.02	-0.01	0.01	0.03	-
5 variables	-3.14	0.02	-0.01	0.01	0.02	0.07

Table 30: The best estimates for the constants in the regression models for the monthly synthetic groundwater table.

Synthetic model - monthly	Constants					
	Interc.	Pw	PETw	S25	S60	GWTold
1 variable	-4.47	0.03	-	-	-	-
2 variables	-3.39	0.03	-0.02	-	-	-
3 variables	-4.53	0.02	-0.02	0.09	-	-
4 variables	-3.95	0.02	-0.02	0.15	-0.12	-
5 variables	-3.69	0.02	-0.02	0.15	-0.13	0.14

Additional calibration and validation results

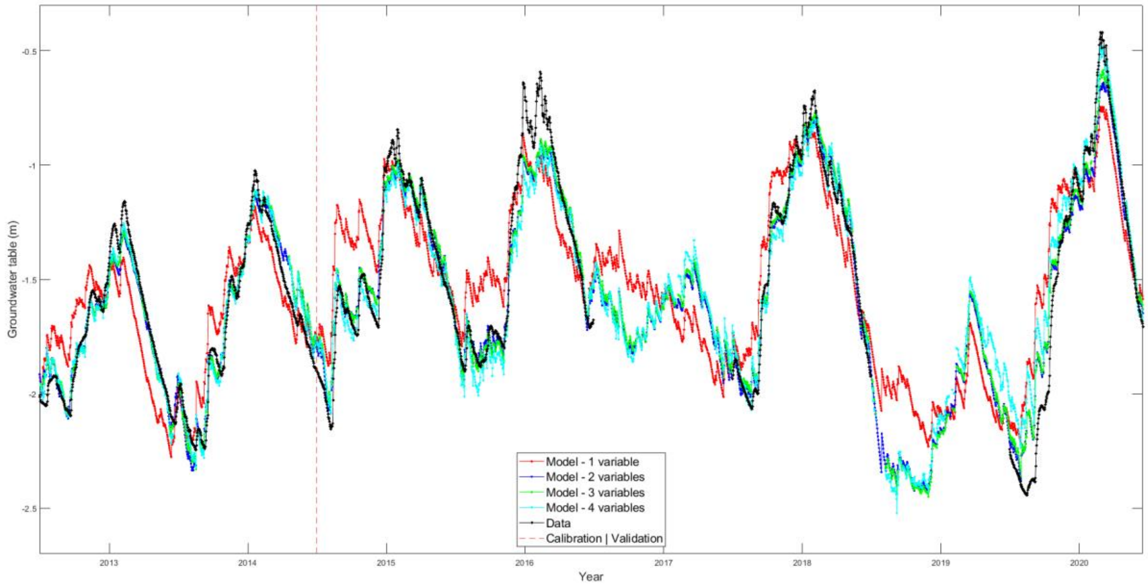

Figure 26: Daily models calibrated from the beginning on 730 days of data.

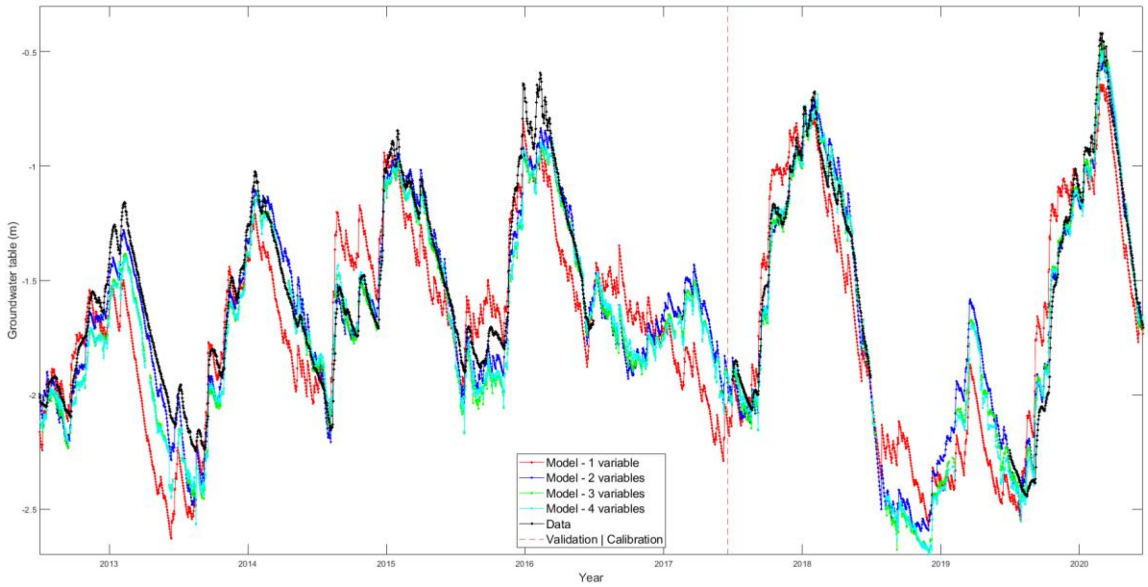

Figure 27: Daily models calibrated from the end on 730 days of data.

Table 31: The R^2 for models calibrated on 730 days of data.

Calibration and validation - daily	R^2					
	Calibration		Validation		Total	
Scenario	B	E	B	E	B	E
1 variable	0.704	0.864	0.822	0.732	0.789	0.789
2 variables	0.940	0.972	0.962	0.947	0.958	0.958
3 variables	0.945	0.978	0.960	0.936	0.957	0.947
4 variables	0.952	0.980	0.914	0.926	0.925	0.945

Table 32: The RMSE for models calibrated on 730 days of data.

Calibration and validation - daily	RMSE					
	Calibration		Validation		Total	
Scenario	B	E	B	E	B	E
1 variable	0.170	0.192	0.228	0.229	0.210	0.218
2 variables	0.076	0.088	0.108	0.100	0.099	0.096
3 variables	0.074	0.079	0.109	0.142	0.099	0.125
4 variables	0.069	0.076	0.145	0.140	0.125	0.123

Table 33: The exact number of days used for calibration and validation for the models.

Days available for calibration and validation	Days				
	Calibration		Validation		Total
Scenario	B	E	B	E	B,E
1 variable	717	730	1460	1447	2177
2 variables					
3 variables	674	691	1384	1367	2058
4 variables		688	1379	1365	2053

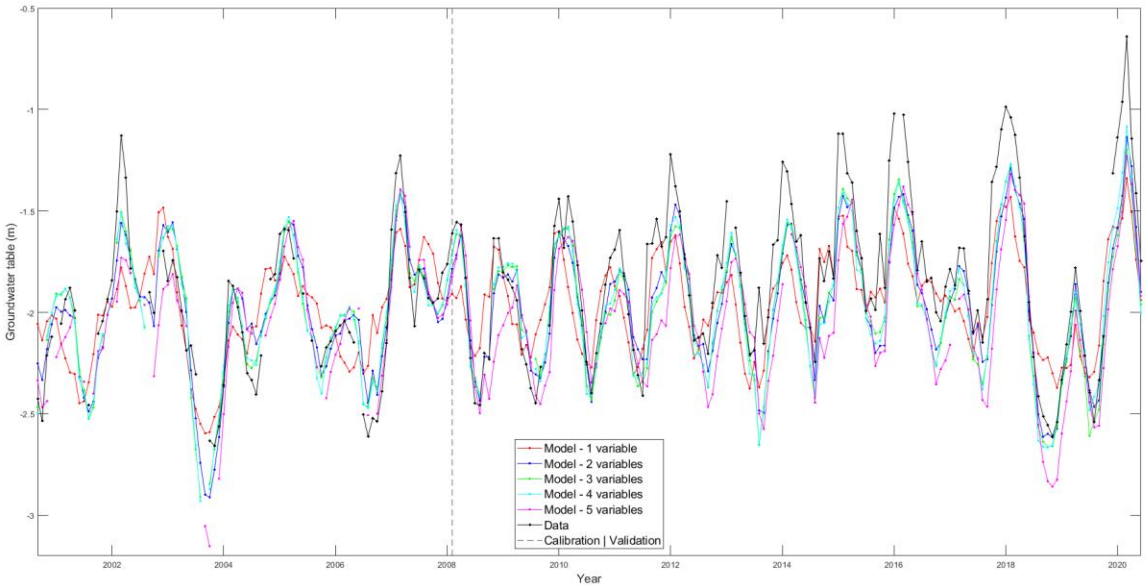

Figure 28: The monthly models calibrated on the first 90 months of data.

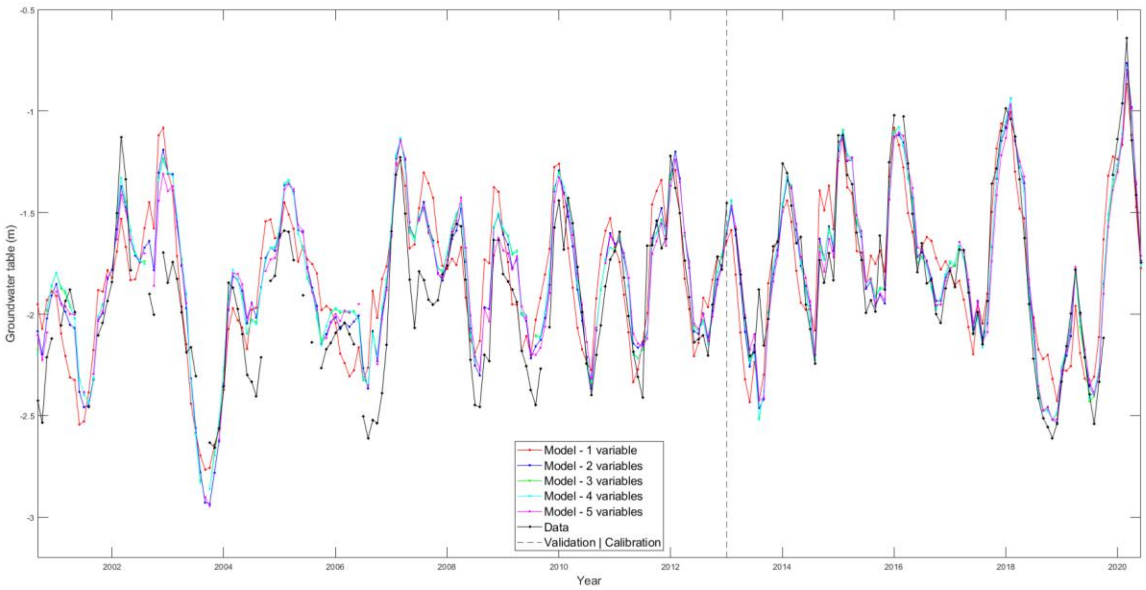

Figure 29: The monthly models calibrated on the last 90 months of data.

Table 34: The R^2 for the models calibrated on 90 months of data.

Calibration and validation - monthly	R^2					
	Calibration		Validation		Total	
Scenario	B	E	B	E	B	E
1 variable	0.507	0.809	0.753	0.546	0.668	0.668
2 variables	0.788	0.916	0.842	0.769	0.808	0.829
3 variables	0.838	0.921	0.843	0.794	0.812	0.841
4 variables	0.843	0.922	0.852	0.798	0.823	0.843
5 variables	0.719	0.909	0.696	0.784	0.683	0.840

Table 35: The RMSE for the models calibrated on 90 months of data.

Calibration and validation - monthly	RMSE					
	Calibration		Validation		Total	
Scenario	**B**	**E**	**B**	**E**	**B**	**E**
1 variable	0.235	0.189	0.262	0.267	0.253	0.239
2 variables	0.154	0.125	0.217	0.206	0.198	0.178
3 variables	0.137	0.121	0.230	0.205	0.204	0.175
4 variables	0.135	0.121	0.222	0.203	0.197	0.174
5 variables	0.196	0.132	0.307	0.190	0.278	0.168

Table 36: The exact number of months that the models have been calibrated and validated on.

Months available for calibration and validation	Months				
	Calibration		Validation		Total
Scenario	**B**	**E**	**B**	**E**	**B,E**
1 variable	75	88	144	131	219
2 variables					
3 variables	70	88	141	123	211
4 variables					
5 variables	61	85	141	117	202