PIETER BREEL
s2348373
7th of July 2023

The correlation between the precipitation and the representative high groundwater level (RHG) in urbanized Hengelo

UT SUPERVISOR
Dr.Ir. A.M. Łoboda

COMPANY SUPERVISORS
Ir. M. van Rees
Ir. J. Bouma

BZ Ingenieurs & Managers

UNIVERSITY OF TWENTE.

Preface

Before you lies the bachelor thesis about the 'the correlation between the precipitation and the representative high groundwater level (RHG)in urbanized Hengelo'. This thesis is the last step in completing my Bachelor of Science at the University of Twente. I worked on this thesis from the 1st of May to the 7th of July 2023. During this period I learned many new skills, especially setting up and performing research, as well as report writing skills. Additionally, this is the first time that I have worked in a company setting on a problem, which was also very insightful.
In the past years at the University of Twente, I have learned many skills and aspects regarding civil engineering. Although this project was different to the ones in the modules of the study, the subjects and projects have given me tools to tackle this research.
I want to thank Anna Łoboda as my supervisor from the University of Twente for her aid and feedback in creating this report and I want to thank Marnix van Rees and Johan Bouma for their help inside the company regarding understanding the data and background of the problem. Their constant feedback and critical thinking have helped shape this report to what you see before you.

I hope you enjoy reading this thesis.

Pieter Breel
Enschede, 07-07-2023

Abstract

This thesis researches the relationship between precipitation and its impact on groundwater levels, specifically focusing on the representative high groundwater level (RHG) in Hengelo. The RHG is defined as the height that is exceeded or met on only 10% of days in a year, often leading to a rise in water levels and resulting in damage. The aim of this study is to determine this relationship, using historical data, where groundwater levels were measured through the monitoring well network in Hengelo, and precipitation was measured by the weather station installed by the Dutch weather institute.

To address the main research question, the first step involves determining the amount of data necessary to accurately describe the RHG in Hengelo. Once this is established, a minimum threshold for data can be set, which should be followed when utilizing historical data. The found threshold stands at 6 years, however, the spread in earlier years are very small and thus in this area also can be used.

Subsequently, the study examines the extent to which the groundwater level rises before reaching an RHG situation. The investigation focuses on understanding the increase and exploring potential relationships with the number of days preceding an RHG event, aiming to identify any predictive capabilities. It was found that there is a linear relationship between the groundwater level rise and the number of days before a RHG, with a very high R^2 value.

Finally, the thesis delves into the precipitation patterns preceding an RHG occurrence. Similar to the groundwater level rise analysis, the study evaluates the relationship between precipitation and the number of days leading up to an RHG event, with the intention of determining predictive capabilities. These findings are then compared to the STOWA values for extreme precipitation, which are utilized in the absence of sufficient data. The study reflects on how the observed precipitation relates to the specified STOWA values and investigates the factors contributing to potential differences, while also discussing the appropriate and feasible usage of these values. The results showed that the precipitation and number of days before a RHG also had a linear relationship also with a very high R^2 value and that the STOWA value often overestimate the RHG levels.

As a result of both a linear relationship between the precipitation and groundwater level rise, the relation between these is also seen to be linear. The conclusion of this research shows the predictive formula that can be used to describe an RHG situation with precipitations in certain time frames. Although this correlation has a very high accuracy it should be noted that only two variables are taken into account. It is therefore advised in following research to build a model that looks at more factors such as topography and geology.

Abbreviations and terminology

Table 1: Abbreviations and terminology

IQR	Inter quartile range
NAP	Normaal Amsterdamse Peil
QQ-plot	Quantile against Quantile plot
RHG	Representative high groundwater level
RLG	Representative low groundwater level
STOWA (En)	Foundation applied research water management network aqua-thermics
STOWA (NL)	Stichting Toegepast Onderzoek Waterbeheer Netwerk Aquathermie

Contents

1 Introduction

In a world with an ever-changing climate, extremes events are becoming commonplace, including longer and more droughts as well as shorter and more intense rain showers. The climate is undergoing changes, assuring a future vastly different to the past or present [1]. The primary cause of these changes is the continuous and escalating emission of greenhouse gases [2].

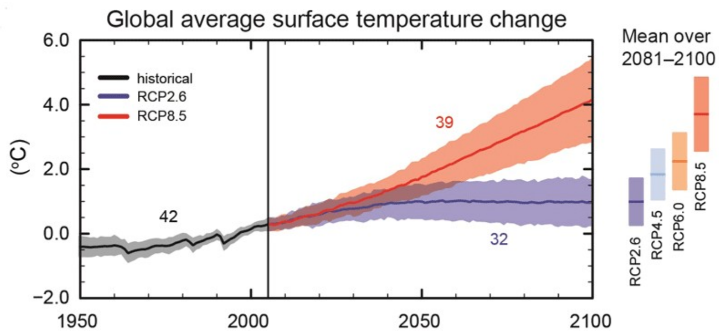

Figure 1: Temperature graph over the coming century [3]

Figure 1, shows that by 2050 the global temperature will increase by two to three degrees Celsius. These rising temperatures will also cause the sea level to rise as well as generally heating the surface of the earth. This causes far more (intense) heat-waves, which evaporate more water and thus, lead to more droughts [1]. Simultaneously, the changing climate also causes far more intense but shorter storms and precipitation. This increase in extreme weather situations now has to be taken into account when planning policy [3][4]. The increase in droughts also have an impact on something that is not directly visible, but affects the water-management in the Netherlands drastically [5]. These are the groundwater reserves in the Netherlands.

Groundwater is all the water stored below the surface of soils. The layers that retain this water are called aquifers. However, before the water reaches the aquifer, the precipitation has to travel through several layers of soil. Precipitation is defined as all water that comes from the atmosphere and falls to the ground where it becomes surface water [6]. This also means that the groundwater levels are mainly dependent on the precipitation. The layers of soil where the surface water has to seep through, or else infiltrate, to reach the aquifer is called the unsaturated zone. In the unsaturated zone, as its name suggests, the water does not completely fill all the empty spaces, contrary to the saturated zone or aquifer. If the precipitation does not infiltrate the soil it is called surface runoff. After infiltration the water flows due to pore pressure and gravity. This can either be directly down (gravity drainage) or sideways (interflow or throughflow), depending on the slope. The flow from the unsaturated zone into the saturated zone is called the recharge of the groundwater. The place where the aquifer starts is called the water-table and is often denoted in metres. If, however, the water table is located near the surface such that the roots of plants can reach it, then water is evapotranspirated through the plants back into the atmosphere. This is similar to evaporation except that an additional medium is used [7]. In Figure 2 this process is shown.

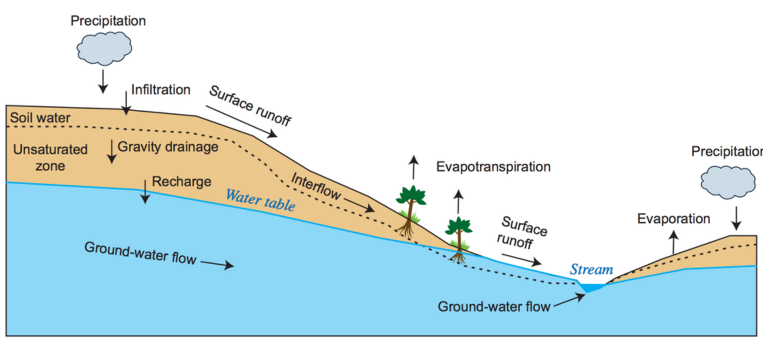

Figure 2: Groundwater movement diagram [8]

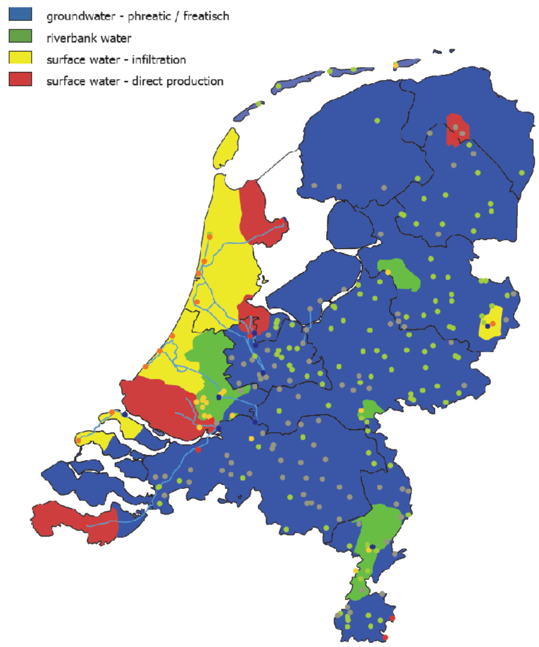

Figure 3: Prevalent drink water sources Netherlands [9]

Groundwater in the Netherlands is used heavily for daily consumption. Around 60% of drinking water is drawn from groundwater [4], as seen by the blue area in Figure 3. This makes it an extremely important resource for current inhabitants. Therefore, the ground-water level needs to be managed precisely, to be used as a water source as well as to prevent water damage. Unfortunately, droughts cause ground-water levels to fluctuate, as well as making infiltration into the hard and dry ground more difficult [5]. These fluctuations then give rise to water shortages, which was the case in the summer of 2018 [10]. Additionally groundwater is also used for growing crops and industrial uses, further emphasizing the importance of this resource.

Groundwater is mainly positively influenced by precipitation ([11] [12]), and negatively influenced by evaporation and evapotranspiration [13]. In the summer there is, in the Netherlands, a deficit of rainwater. Accompanied by higher temperatures the amount of water that is evaporated and evapotranspirated is higher. These two factors cause the groundwater level (or water table) to drop significantly. On the other hand, the winter is characterized by the opposite, with (usually) an excess of precipitation and a lack of evaporation and transpiration.. The result is a seasonal variation in groundwater level and precipitation. This seasonality can be seen in Figure 4, where data from a groundwater logger in the area of Hengelo is taken.

Figure 4: Groundwater level and precipitation fluctuation throughout 3 years [14]

The seasonality is important to understand as it influences the policy for the surrounding area. When more groundwater is available, the quality also becomes better and thus usage for drinking and irrigation are encouraged. Studies performed in India [15] and China [16], showed this relation and also encouraged the readers to work with this seasonality to adapt usage, so as to improve the quality and abundance of the groundwater. . Thus, in wetter seasons the water quality is improved so that more water can be used in the drier seasons. Although this research was performed in India and China, the main idea is applicable to the Netherlands, with respect to the salinity of the groundwater. To accurately describe this seasonality, a certain time frame has to be found. A study performed in the United States [11] showed that with 60 months of data (or approximately 5 years) the groundwater level could accurately be described.

This seasonality has changed over recent years due to climate change. Studies (such as [17], [18] and [19]) have found that the aforementioned relationships, together with seasonality, are changing and that adaptations should be made. Additionally there other factors that also should be looked that influence groundwater and that are influenced by the changing climate.

These factors make it difficult to manage groundwater accurately to provide enough and qualitative amounts of water for people. A study in California ([20]) describes the sustainable groundwater management act (SGMA) in California. Here, it was discussed that in dry periods, groundwater usage should be lessened by irrigating crops with flood flows as well as recharging groundwater basins. Another study (such as [21]) suggested to use smart technologies, such as artificial intelligence, big data or cloud computing, to manage groundwater resources in real time. These pave the way for strategic planning and development.

To conclude, to accurately describe the groundwater level, insights have to be made into the factors that influence groundwater. This, in the positive sense, is mainly the precipitation and in the negative sense the evaporation and evapotranspiration. As these factors have a degree of seasonality to them, it is important to take be able to enough data to accurately describe this variety and potentially make predictions and adaptations for each season.

1.1 Research aim and questions

From the introduction, some knowledge gaps can be identified and, based on them, research questions. From these knowledge gaps research questions and sub-questions can be defined. The main research question will be:

- What is the correlation between the precipitation in the winter and representative high (phreatic) groundwater-level (RHG) in urbanized Hengelo?

In order to the above mentioned question, several sub-questions have to be formulated and answered.

- How much of the available data for the given area counting from most recent to oldest,is needed to be able to accurately define the RHG level?
- How much does the groundwater level rise before a RHG situation?
- How much precipitation falls before a RHG situation and how does it compare to the normative precipitation?

2 Methodology

2.1 Research area

This study is an observational, historical (retrospective) data study of groundwater loggers and one weather station. The municipality of Hengelo has a network of 104 groundwater monitoring wells. In each well, the groundwater level is measured using a logger. The data of the groundwater loggers were collected in the period from August 2011 (earliest validated measurements) to April 2023 (latest validated measurements). The groundwater loggers document the groundwater level, measured in Normaal Amsterdamse Peil (NAP), each hour for in the given time period. The NAP is a standard measurement level in the Netherlands for all height differences and topograhpy. A NAP of zero metres is approximately equal to the mean sea level of the North Sea. The loggers are placed and maintained by a company called 'Eijkelkamp'. The location of the loggers is distributed over the area of Hengelo (located in the east of the Netherlands near the border of Germany) and the respective locations of each logger can be seen in Figure 5, represented by a blue dot.

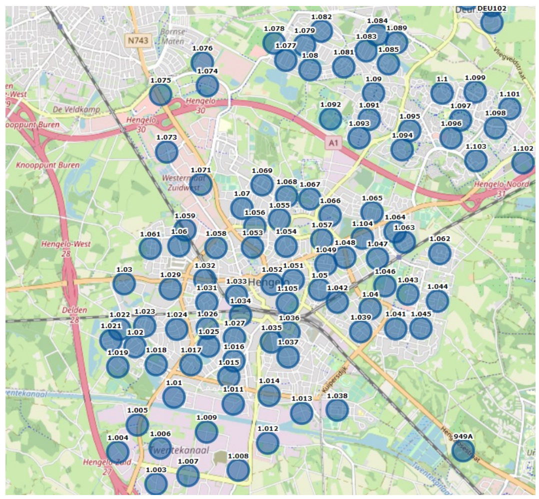

Figure 5: Location groundwater loggers Hengelo (Latitude: 52.2523195, Longitude: 6.79552559080801) [14]

A ground water monitoring well is essentially an open-ended tube, where the underground water pressure pushes water into the tube. From here, the water height in the tube is measured in NAP using a logger. In this study the focus is on the phreatic groundwater. Figure 6 shows a schematic example of a logger that is located in studied area with the corresponding soil at certain depths.

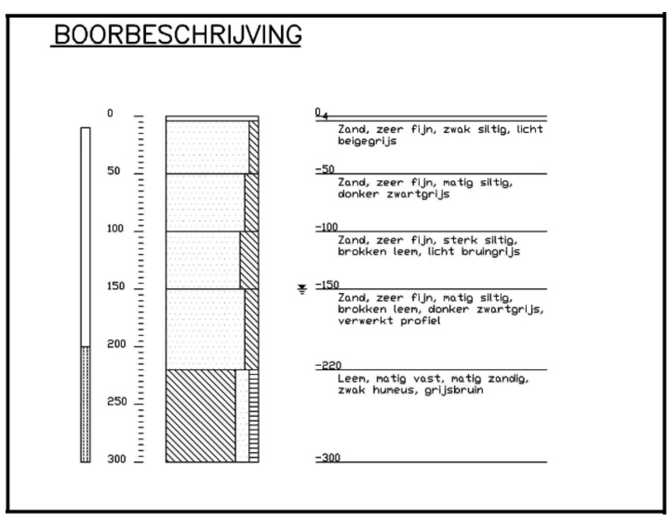

Figure 6: Schematic cross section of a groundwater logger [14]

The KNMI precipitation station, similar to the groundwater loggers, measures hourly the precipitation (in mm) that has fallen. Under precipitation is meant, any liquid (regardless of state) that is formed in the atmosphere and falls back to earth. The logger is located at the 'Twente Airport' and was placed by the Dutch weather institution or KNMI. The reason for this location is that this logger is the only one in the area that has the sum of precipitation per hour and also the closest. The KNMI has data from 1950 to recent times for this weather station. The data gets updated every few weeks after they have been checked and verified. As the groundwater logger data only starts in August 2011, the hourly rainwater measurements used in this study were also taken from August 2011 to April 2023 [22].

The location of the study area also has a certain precipitation regime (characteristic pattern) in the winter. The regimes are described in the 'Stichting Toegepast Onderzoek Waterbeheer Netwerk Aquathermie (STOWA) neerslagstatistieken en reeksen voor het waterbeheer'. Hengelo is located in the 'LL' regime, which as can be seen in Figures 7 and 8, shows a specific number of precipitations with a respective recurrence time, over a number of days. The current regime, without any additional data, shows a recurrence time of once or twice per year, over 9 days.

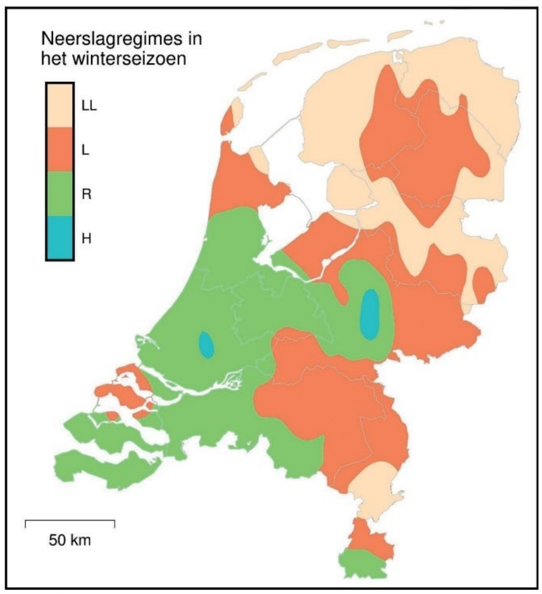

Figure 7: STOWA map with precipitation regions [23]

LL

Nov – feb	Dagen 1	2	4	8	9
2× per jaar	17.8	23.8	32.5	45.1	47.7
1× per jaar	21.1	28.3	38.7	53.9	57.2
1× per 2 jaar	24.6	33.0	45.0	62.6	66.3
1× per 5 jaar	29.5	39.5	53.5	73.5	77.7
1× per 10 jaar	33.5	44.6	60.1	81.5	85.9
1× per 20 jaar	37.6	49.9	66.7	89.3	93.7
1× per 25 jaar	39.0	51.7	68.9	91.7	96.2
1× per 50 jaar	43.5	57.3	75.7	99.1	103.6
1× per 100 jaar	48.1	63.1	82.6	106.3	110.7
1× per 200 jaar	53.0	69.2	89.6	113.3	117.5
1× per 250 jaar	54.7	71.2	91.9	115.5	119.6
1× per 500 jaar	59.9	77.6	99.0	122.2	126.0
1× per 1000 jaar	65.5	84.2	106.2	128.6	132.2

L

Nov – feb	Dagen 1	2	4	8	9
2× per jaar	19.5	26.1	35.6	49.4	52.3
1× per jaar	23.1	31.1	42.5	59.2	62.8
1× per 2 jaar	27.0	36.2	49.4	68.6	72.7
1× per 5 jaar	32.4	43.3	58.7	80.7	85.2
1× per 10 jaar	36.7	48.9	65.9	89.4	94.2
1× per 20 jaar	41.3	54.8	73.2	97.9	102.8
1× per 25 jaar	42.8	56.7	75.6	100.6	105.5
1× per 50 jaar	47.7	62.9	83.0	108.8	113.7
1× per 100 jaar	52.8	69.3	90.6	116.6	121.4
1× per 200 jaar	58.2	75.9	98.3	124.3	128.9
1× per 250 jaar	60.0	78.1	100.8	126.7	131.2
1× per 500 jaar	65.7	85.1	108.6	134.0	138.3
1× per 1000 jaar	71.8	92.4	116.6	141.1	145.0

R

Nov – feb	Dagen 1	2	4	8	9
2× per jaar	21.6	29.0	39.5	54.8	58.0
1× per jaar	25.7	34.4	47.0	65.6	69.5
1× per 2 jaar	29.9	40.1	54.7	76.0	80.6
1× per 5 jaar	35.9	48.0	65.1	89.4	94.4
1× per 10 jaar	40.7	54.2	73.0	99.1	104.4
1× per 20 jaar	45.8	60.7	81.1	108.5	114.0
1× per 25 jaar	47.4	62.9	83.8	111.5	116.9
1× per 50 jaar	52.8	69.7	92.0	120.5	126.0
1× per 100 jaar	58.5	76.8	100.4	129.3	134.6
1× per 200 jaar	64.5	84.1	108.9	137.7	142.8
1× per 250 jaar	66.5	86.6	111.7	140.4	145.4
1× per 500 jaar	72.9	94.3	120.4	148.5	153.2
1× per 1000 jaar	79.6	102.4	129.2	156.4	160.7

H

Nov – feb	Dagen 1	2	4	8	9
2× per jaar	23.2	31.2	42.5	59.0	62.4
1× per jaar	27.6	37.1	50.6	70.6	74.8
1× per 2 jaar	32.2	43.2	58.9	81.9	86.7
1× per 5 jaar	38.6	51.7	70.1	96.2	101.7
1× per 10 jaar	43.8	58.4	78.6	106.7	112.4
1× per 20 jaar	49.2	65.3	87.3	116.8	122.7
1× per 25 jaar	51.1	67.7	90.2	120.0	125.9
1× per 50 jaar	56.9	75.0	99.0	129.7	135.6
1× per 100 jaar	63.0	82.6	108.1	139.1	144.9
1× per 200 jaar	69.4	90.6	117.2	148.3	153.7
1× per 250 jaar	71.5	93.2	120.2	151.1	156.5
1× per 500 jaar	78.4	101.5	129.5	159.9	164.9
1× per 1000 jaar	85.7	110.2	139.0	168.3	172.9

Figure 8: STOWA table with precipitation recurrence times [23]

2.2 Data Cleaning

Before performing any statistical analysis or normality check, the data needs to be cleaned. Four processes are required to clean data:

1. Accuracy – checking for erroneous data, typographical errors, that all data are on the same scale and have the same coding

2. Missing data – this involves checking if the data is missing (completely) at random (MCAR or MAR), missing not at random (MNAR) or structurally missing. This will determine whether data may be imputed or not during statistical analysis. Data that is missing not at random or structurally missing should not be imputed as the missingness is related to the value of the variable itself.

3. Outliers – Checking for normality outliers are values within a dataset that fall outside of three standard deviations, i.e. they fall outside of the boundaries of the dataset. Outliers will not be removed from the dataset if they have correct values (this is checked in Step 1, Accuracy), they represent natural variation. Instead, a sensitivity analysis will be performed (analysis with and without the outliers) and both will be reported.

4. Homoscedasticity – the variances of the variables are checked for equality (i.e the residual term across variables is constant). As the value of the dependent variable changes, the error term does not vary very much. This is an important assumption in performing a linear regression.

2.3 Normality test

Once the data have been cleaned (incorrect or faulty data removed), they have to be tested for normality of distribution in order to determine whether parametric or non-parametric testing can be performed. In a normally distributed dataset, the mean and median values of the set are (almost) the same and form the centre point, with the data spread equally to each side, forming a symmetrical histogram [24]. Testing of normality can be done by two techniques- visualisation, utilising graphical representation such as histograms, boxplots or a quantile-quantile plots (QQ plot) [25], or formal testing utilising one of several statistical tests such as the Shapiro-Wilk -, Kolmogorov-Smirnov - or Anderson-Darling tests [26]. For these tests, the null hypothesis is that the data is normally distributed, hence this hypothesis is only rejected when a p-value larger than the significance value is found. However, the drawback of the usage of numerical methods is that the p-value is dependent on the sample size. For a large sample the p-value becomes very small even when a smaller sample of the same population has the same distribution. This produces false positive results. To avoid this, graphical methods will be used in this report [27]. Additionally, the (logger)data encompassed more than 100 000 data points. In mathematical probability, the central limit theorem states that if a dataset is large enough, the probability distribution will approximate a normal distribution when sampled. The cut-off point is generally accepted at 100 [28]. Hence, the dataset for this study was assumed to be (approximately) normally distributed.

2.4 Statistics

2.4.1 Boxplot

A boxplot is a visualisation of a few statistical characteristics of the tested data set, i.e. the first quartile, third quartile, minimum, maximum and median. Sometimes the mean can also be added or read however this is rarely done. The previously stated characteristics give insight into the distribution of the data set. Due to this the boxplot is often used as a precursor of a hypothesis test. Whether this is a (form of) t-test or analysis of variance (ANOVA), a boxplot is often used in understanding the distribution and variation of the data [29]. As stated earlier, the boxplot displays some characteristics of the dataset. In Figure 9, the anatomy of a boxplot is shown.

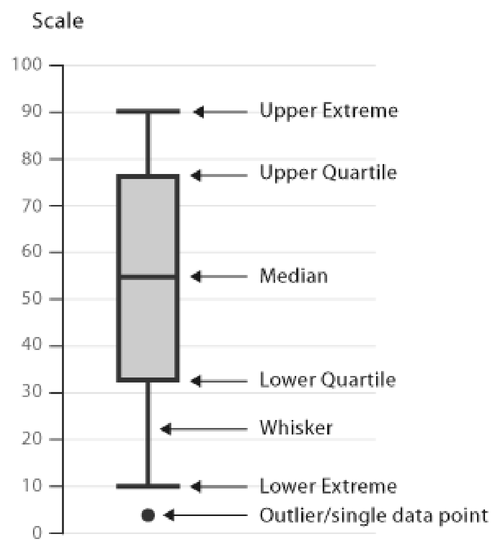

Figure 9: Anatomy of a boxplot [30]

In a normal distribution, the whiskers of a boxplot (the lines outside the middle box) are dependent on the interquartile range or IQR. The IQR is the height of the box which is the length of the first quartile to the third quartile and split by the median. The length of the whiskers are about 1.5 times the length of the IQR. If the data is skewed, either the position of the median or the length of the whiskers will reveal which way the data is skewed. Figure 10 shows the equal distribution of the data around the median, the relation to the distribution of the dataset. Additionally, some boxplots can have dots outside the range of the whiskers. These points are outliers as they have an abnormally large distance from the rest of the data (greater than the upper and lower boundaries) [29].

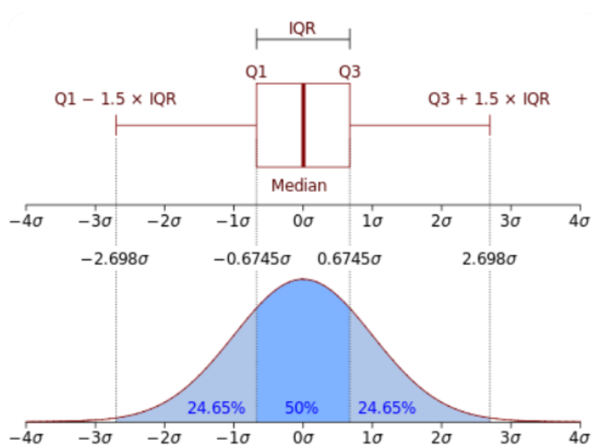

Figure 10: Boxplot and distribution relation [29]

A boxplot was used to show the difference in the point locator and spread of the RHG for all loggers. The boxplots for this dataset were created as followed: the initial plot contained all eleven years of data and the RHG value was calculated for this data. Subsequently, the data of one year was subtracted, one year at a time and each time a new RHG value was calculated. The RHG values were then subtracted from the total data and a point locator and spread calculated. The closer the point locator is to zero, the more is approximates the total dataset.

2.4.2 T-test

A t-test is a parametric statistical test that shows if there is a statistically significant difference between two means of data. To use the t-test there are a few assumptions that need to be. These are [31]:

- normally distributed data;
- data is collected from a representative randomly selected portion of the population;
- the data is continuous;
- the sample size is of adequate size;
- homogeneity of variance.

When the assumptions are met, the type of t-test can be identified based on the data. There are 3 types of t-test, these are:

- one sample t-test (test the sample mean against known mean);
- independent samples t-test (unpaired data);
- dependent samples t-test (paired data).

From these three tests only the dependent samples t-test was used. The reason for this is that the data is paired, i.e. the data tested will be on different time instances of the same logger (i.e. 1,2 or 3 years later). The t-statistic is the ratio of the spread between two groups as well as the spread within the respective groups. A larger value indicates more difference between groups, whereas a smaller score indicates more similarity. The t-value for a dependent, or paired t-test is:

$$t = \frac{\Sigma d}{\frac{n(\Sigma d^2)-(\Sigma d)^2}{n-1}} [32] \tag{1}$$

where 'd' is the difference per paired value and 'n' is the number of values. The t-statistic produced is the ratio of the difference of the estimated value of a variable from its hypothesised value to its standard error. The t-statistic is used to verify whether to accept or reject the null hypothesis.

2.4.3 Linear regression

A linear regression is a plot of a line that shows the (potential) relationship between two variables. This is done by plotting a line through data points that have a coupled x and y coordinate. For example, if on a certain day a certain groundwater level is measured and a certain precipitation is measured, then data-points can be made for each day. The principle on which a linear regression works is the so called ordinary least squares. This means that the smallest sum of squared errors is searched for and used. The error is the actual value subtracted by the predicted value. The squaring of variables is done to ensure that negative predictions aren't cancelled out [33]. The general formula for a line is:

$$y = ax + b \tag{2}$$

Here there are 2 unknowns, which are needed for the line to 'fit' the data. These are the slope 'a' and the intercept 'b'. The formulas to determine these coefficients are listed below:

$$a = \rho \frac{S_y}{S_x} \tag{3}$$

$$b = \bar{y} - a\bar{x} \tag{4}$$

where:
ρ is the Pearson correlation
S_x and S_y are the standard deviations of the values x and y
$\bar{x}$ and $\bar{y}$ are the means of x and y

In order to perform a linear regression, certain assumptions or prerequisites have to be met. These are partially the same as for a t-test. The prerequisites for a linear regression are[34]:

1. Linearity of data
   Which can be checked by looking at a scatter plot of the data.

2. Normality
   Which can be checked by looking at a QQ-plot of the data.

3. No or little multicollinearity
   Will not be assessed as only the correlation of 2 variables will be looked at.

4. No auto-correlation
   The error terms or residuals are correlated, decreasing accuracy of the model.

5. Homoscedascity
   The variances of the residual terms are constant.

When these assumptions are met, the linear regression can be performed. The amount of variance explained by the variables is represented by the R^2 value, which is represented as a percentage. An R^2 of 1 indicates a perfect correlation whereas an R^2 of 0 indicates absolutely no correlation [33].

2.4.4 Statistical significance

Statistical significance was assumed at $p<0.05$. This indicates that there is a 5% chance that the result found occurred by chance. The value ranges from 0 to 100%. The lower the value the lower indication that the result was not by chance. Once a p-value is chosen a table value for the t-statistic can be found if performing it manually. If this value is larger than the calculated value then the hypothesis is accepted and a similarity is proven, else the hypothesis is rejected and it can be said that the two datasets are different [35].

2.5 Outline action plan

The research mainly focuses on the RHG of 104 monitoring wells in the area around the city of Hengelo. The first step was to calculate this value, which is defined as the 90th percentile of the groundwater level. Hence, for each logger, a histogram was made and a line at the 90th percentile value was made. This resulted in a single value, called the RHG level. After this a QQ-plot of each logger was made to check the data for normality. Due to the large number of data points available, the normality check is a formality as there is relied on the central limit theorem. After this step, a choice was made to exclude certain data-points or loggers.

The next step was to determine the number of years needed for the model, before the RHG level stabilises. Ten years of hydrological data was available (measured from April to April of each year). The time-series starts with data from April 2022 to April 2023. Data was added per year, starting with the year of the most recent data and adding each year consecutively, backwards, until all ten years were in the model. For each of these intervals, the RHG was determined for each logger. A plot was then created to show the RHG fluctuation (or lack thereof) over time. The data from the different years was statistically tested against the total RHG (averaged over all the years). We then determined how many years were needed to find an RHG mean similar to the total dataset. Multiple boxplots were grouped to visualise the change in RHG means.

Consequently, the precipitation before a RHG occurrence was compared to the STOWA values for extreme precipitation occurring once per winter (the hydrological winter is described within the STOWA manual as the months from the first of October to the first of March). Winters are characterised by an excess of precipitation and a lack of evaporation, this results in a surplus of (rain/ground)water. These values are currently used as the reference data when there is a lack of data.

An RHG occurrence is a point in time where the groundwater-level is equal to or higher than the 90th percentile (RHG) value. As a RHG occurrence can more easily happen when the groundwater is high (due to a previous RHG or some precipitation), the first RHG occurrence in the hydrological winter (after summer) is the primary data point against which all values are compared. The situation immediately before this occurrence was examined and compared to the known data. The STOWA manual describes the sum of precipitation for 1, 2, 4, 8 and 9 days prior for extreme precipitations, hence the same days will be used within the model for comparison. The sum of precipitation for the aforementioned time instances was averaged per day, over all loggers, for example, if 100 loggers were used, 100 points for 1 day precipitation were averaged.

In addition to the precipitation values before an RHG occurrence, the change in groundwater-level was also calculated for the same time periods. This was done by taking the groundwater level on the day before the RHG occurrence (baseline value), and calculating the incremental (with respect to the previous time instance) increase or decrease of the groundwater-level. Similarly, the change in groundwater level for each day was averaged over all loggers.

Finally, the values of precipitation and change in groundwater-level before an RHG situation were plotted on a scatter plot. Within the scatterplot, the linear regression coefficients was determined and the line was plotted. This linear regression was tested by applying a R^2 value to see the degree of variance explained. The formula derived from this regression and degree of correlation enabled us to make a statement regarding the predictive capability of the model, for example, whether this amount of precipitation fell within a certain time frame and whether an RHG situation happened or not, for the given area.

3 Results

3.1 Correlation between precipitation and groundwater loggers

In Figures 11-12, the results are shown for the precipitation plotted on the x axis against the groundwater level rise on the y axis. The points are plotted for the consecutive days, thus, for day 1 there is a precipitation and groundwater- level rise, then for day 2, etc. In Figure 11, the correlation between the regression line and 1, 2, 4, 8, and 9 days are depicted (similar to the days that are used in the STOWA manual for precipitation recurrence). In Figure 12a and Figure 12b the regression lines for days 15, and 30 days, respectively, are depicted.

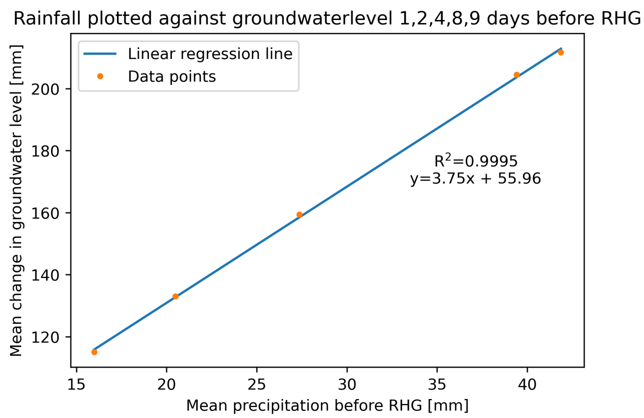

Figure 11: Precipitation against rise groundwater level for 1,2,4,8,9 days before RHG

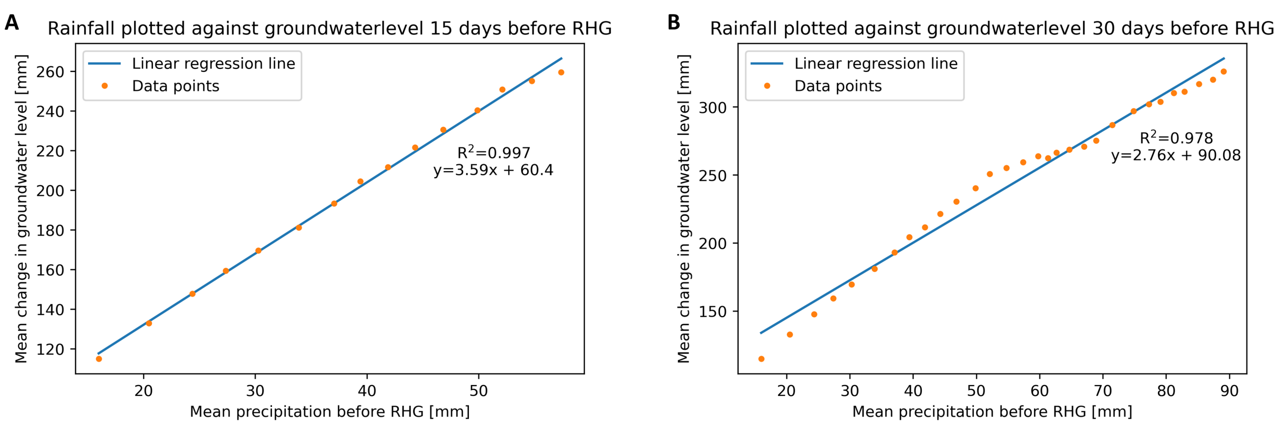

Figure 12: A: Precipitation against rise groundwater level 15 days behore RHG. B: Precipitation against rise groundwater level 30 days before RHG

All linear regressions show very high R^2 value. Additionally, it can be seen that in all figures except the 30 days a similar slope and intercept are depicted. This means that the found values are in accordance with the data point. In Table 2, the values of precipitation versus rise in groundwater level for days 1, 2, 4, 8 and 9 days, 15, are shown, correlating to Figure 11.

Table 2: Precipitation and groundwater level rise for 1, 2, 4, 8 and 9 days

Days before RHG	1	2	4	8	9
Mean rise GW [mm]	115.05	133.02	159.41	204.50	211.71
Mean precipitation [mm]	15.97	20.48	27.35	39.40	41.85

3.2 Test for normality

In Figures 13 and 14, we see examples of histograms and QQ plots of two of the 104 loggers. For each logger a histogram and QQ plot were generated. From these figures we can see that the data is approximately normally distributed, with the mean (red line) generally centred in the middle and that the data broadly follows a bell distribution. The QQ plots show that the majority of the data follows the line with a deviation at both ends, where the extreme values are found.

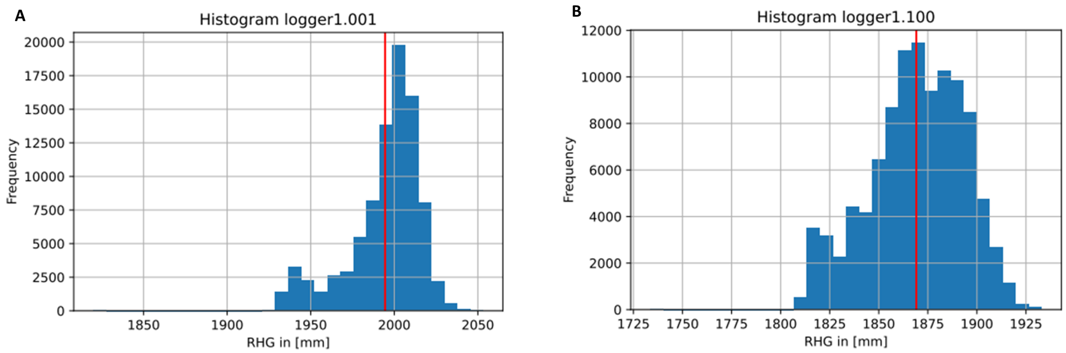

Figure 13: A: Example histogram of logger 1.001. B: Example histogram of logger and 1.100

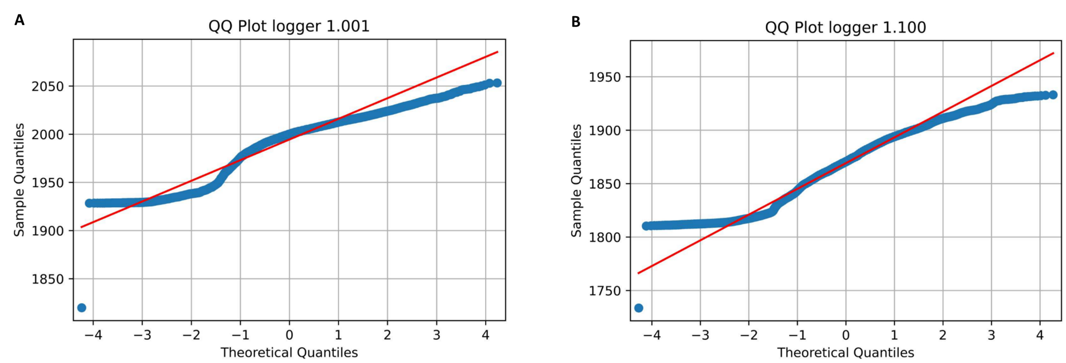

Figure 14: A: Example QQ plot of logger 1.001. B: Example QQ plot of logger and 1.100

Most of the data from the loggers follows an approximately normal distribution. However, there are a few that deviate from this, these are loggers 1.107, 1.106 and 1.076. They do not resemble the bell curve and the mean and median of these loggers are not equal to each other. Additionally, the QQ plot has a deviating shape as the tails or line entirely diverge from the theoretical line. The histograms are either a singular peak value or a lot of bars next to each other without a clear median of median. These results can be seen in Figures 15, 16 and 17. These loggers were excluded from further analysis.

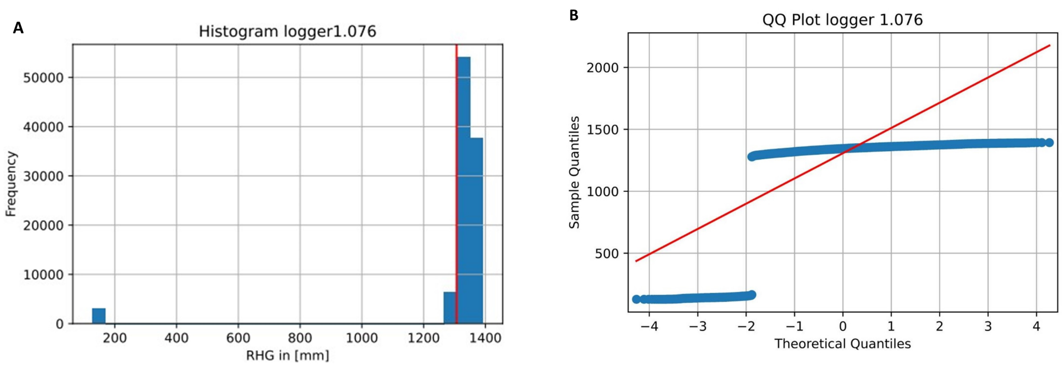

Figure 15: A: Histogram of logger 1.076. B: QQ plot of logger 1.076

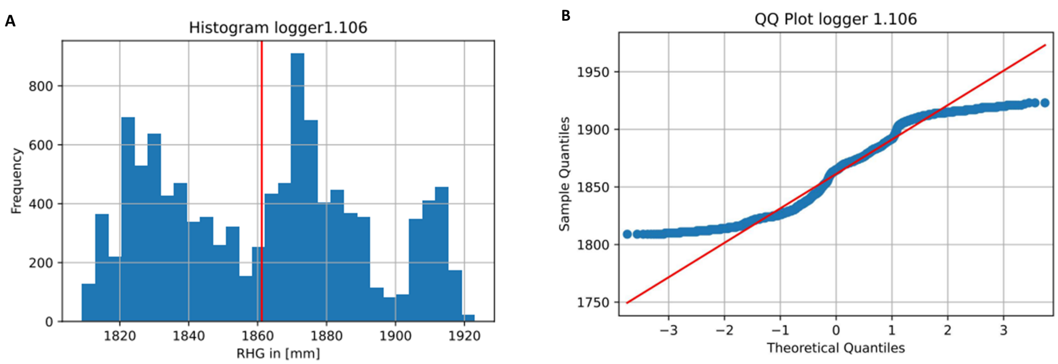

Figure 16: A: Histogram of logger 1.106. B: QQ plot of logger 1.106

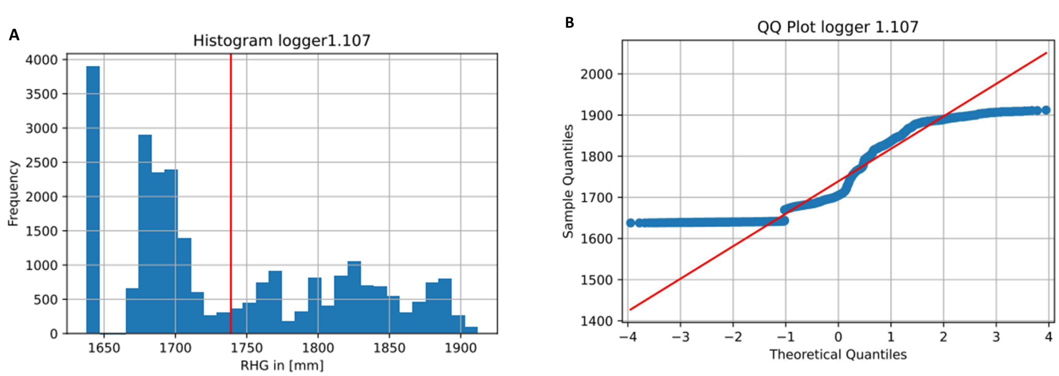

Figure 17: A: Histogram of logger 1.107. B: QQ plot of logger 1.107

When looking further into these loggers it can be seen that 1.106 and 1.107 were only placed very recently (last 2 years). Thus, these loggers do not have many data points and for this reason are very skewed. It can also be seen that from the years that they were measuring a lot of data is missing. Regarding logger 1.076 this many identical values, which could be due to the fact that it is located close to a pipe which could be leaking. This would explain the singular, but very often reoccurring, value. As stated earlier in the report these loggers were removed as they are not representative of the rest of the data and may skew the outcomes.

3.3 Number of years needed for RHG determination

Figure 18 shows a plot of the t-statistic and the p-value for the means of each of the years of data, correlated to the total dataset. In the graph with the p-values, we see that all p-values are statistically significant except for the six-year cumulative data, as this coincidentally has the exact same mean as the total dataset. As there is no difference in the means, the null hypothesis is not rejected and the p-value is not statistically significant.

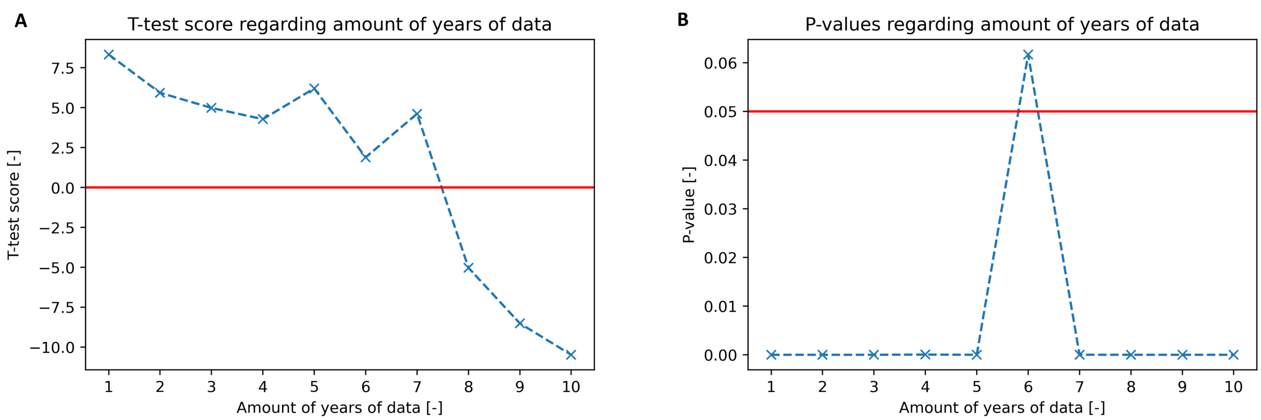

Figure 18: A: Result T-test t-score over amount of years. B: Result T-test p-value over amount of years

In Figure 19, an example of the line plots for the RHG can be found for two loggers. The RHG value fluctuates until the six-year cumulative dataset, thereafter the curve flattens as there is further little or no variance in the mean up to eleven years (red line). When more data is used in each successive point until the total dataset is used then the value should also converge to the total data RHG value.

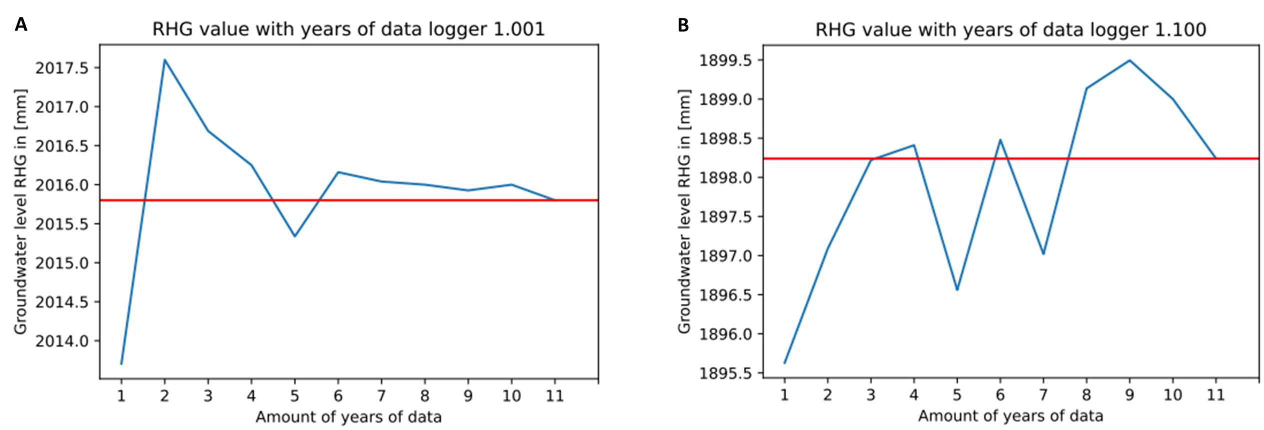

Figure 19: A: Example line plot of logger 1.001. B: Example line plot of logger 1.100

Finally, in Figure 20, the differences with respect to eleven years are represented. This means that the dataset for one year, two years and onwards, are subtracted from the dataset of eleven years, resulting in the boxplots of the differences with eleven years of data. Again, it is important to note that the boxplot of six years has its median in the middle of the box hence it is also the mean. Additionally, the median is similar to eleven years on the zero point (or extremely close to it).

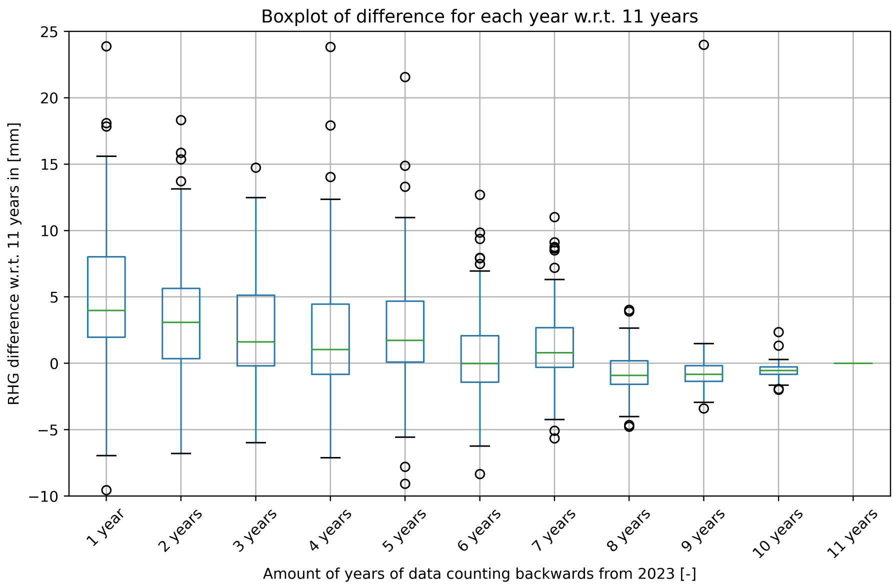

Figure 20: Boxplot of RHG differences

3.4 Precipitation and groundwater-level before RHG

The precipitation before an RHG situation is depicted in Figure 21a. as the light blue line. It shows the same shape line as the STOWA value for extreme precipitation in the winter that happens once or twice per winter. The STOWA lines do not directly correlate to a RHG situation, however they are used as an estimate in the absence of data (substitute outcome). The rise in groundwater-level before a RHG situation is depicted in Figure 21b. This graph has the same shape as the precipitation graphs when plotted on the same scale.

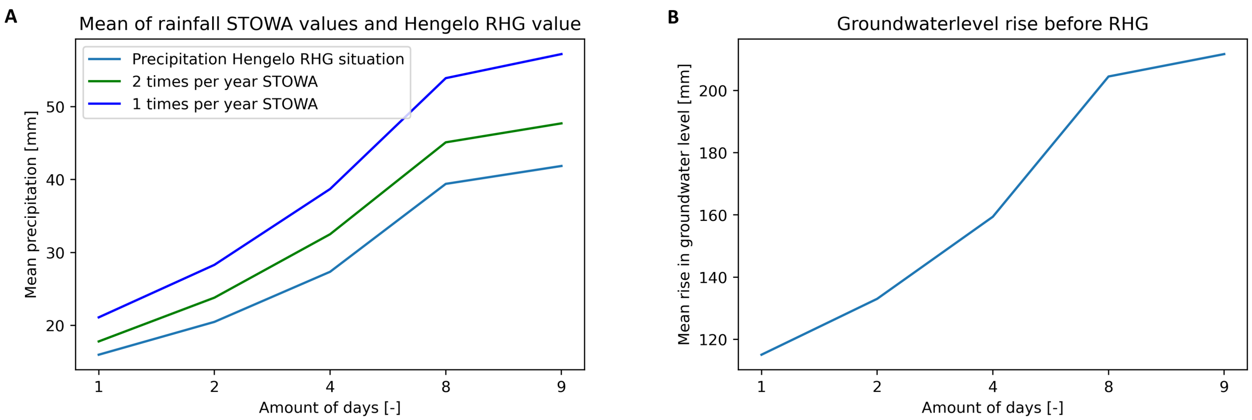

Figure 21: A: Precipitation before RHG situation. B: groundwater-level rise before RHG situation

In Figure 22, the points are plotted on a normal axis. The points are plotted for each day up until day 15. Then a regression line was taken through the points. The regression lines both have a very high R^2 value which indicated a high linear correlation. The graph should be read as follows: if a certain measured point is on or above the line the RHG value is met or exceeded, whereas a point under the line means the RHG value for a logger is not met.

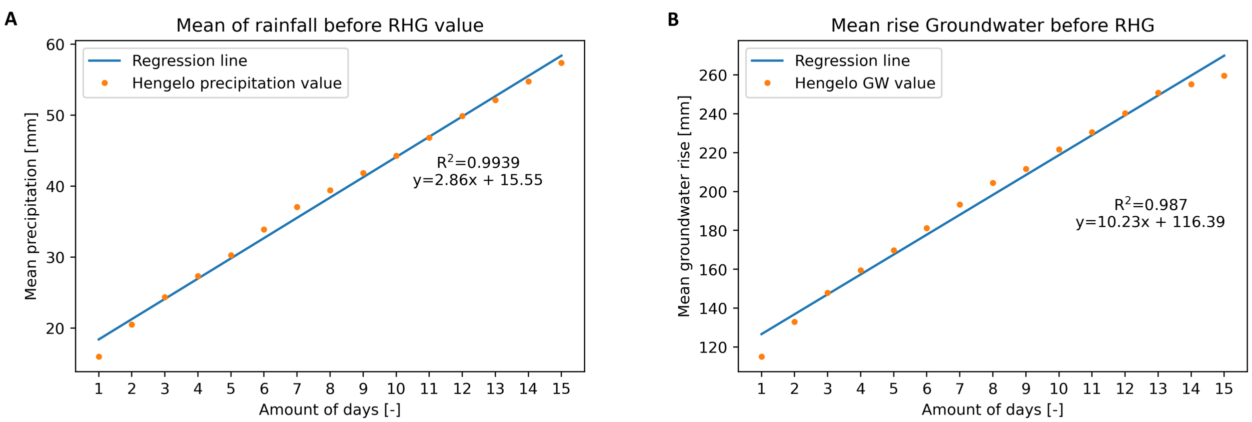

Figure 22: A: Precipitation regression before RHG situation. B: groundwater-level rise regression before RHG situation

4 Discussion

The aim of this study was to calculate how many years' worth of data is necessary to define an RHG occurrence, in order to make flood policy. A positive correlation between winter rainfall and phreatic groundwater levels was found. This is a logical conclusion as precipitation affects groundwater levels by recharging them. The more precipitation that falls, the higher the water levels in the wells are likely to be. This study is in line with that of the study of Smail et al. ([11]), who studied the correlation between monitoring wells and precipitation and which factors affected them. The main finding was between precipitation and the groundwater level. Their study also showed a positive correlation between groundwater levels and rainfall. Studies performed by Smail et al. and Damiba et al.([12] and [17]) also confirm this positive relationship.

The findings from the presented research showed a very strong, statistically significant correlation. However, the data are limited in that we only had two variables available and this is insufficient to produce a strong model. To date, most studies on the changes to water resources have been focused on surface water (such as [18] and [36]), and not on the factors affecting groundwater. Factors such as an increased demand for drinking water, agriculture and industry play an important role together with human activities and climate change on the quantity and quality of groundwater [19][37].

When looking at the graph depicted in Figure 12b, it can be seen that the slope and intercept are quite different when compared to the graphs depicted in Figures 11 and 12a. There, it can be seen that at around 60 mm of precipitation or the 15th point that the graph has a downwards nod. Additionally, it can be seen that the points are grouped much closer to each other. The question then arises if precipitation more than 15 days ago still has an influence on the phreatic groundwater level or that the water already has dissipated or through flowed to adjacent areas. Especially the latter reasons, as Hengelo is located between 17 and 27 meters above NAP on the side of a local hill that stretches across the neighbouring city of Enschede [38].

4.1 Number of years needed

The t-statistic is positive up to seven years' worth of data in the test model, with a local minimum at six years. After this the t-statistic becomes negative. An absolute value (either positive or negative) diverging away from the zero shows a weaker correlation than a value approaching zero. It would appear that in this dataset, six years' worth of data is needed to accurately describe the dataset for Hengelo. This is similar to the results seen in another study ([11]), which built an accurate model using five years' worth of data, in a similar study in the United States. However, it is noteworthy that the values after six years of data do not continue along the zero line. On the contrary they diverge, meaning that the values of the data from 2023-2017 are representative for the total dataset but the values of 2023-2016 and all years after that, are not representative. In our results, the mean of six years of data produces the same mean as the total dataset. This could be due to the fact that the earlier years have a deviating mean from the rest of the data due to climate change.

The boxplot of the differences (Figure 20) also shows, similar to the test results, that six years of data shows the same mean and median as the total dataset. In contrast, the rest of the data shows significant differences to the total dataset. Additionally, we see that the spread of the RHG data decreases with the number of additional years and that the mean and median converge to zero on the eleven year value. Nonetheless, when looking at the scale of the data, we see that the difference with one year of data compared to eleven years is only a minimum of -7 mm to a maximum of 16 mm, or a total range of 23 mm or 2.3 centimetres, with a point locator of 4mm.

It would seem from the results, that one year of data gives a good representation of the groundwater levels and that the differences per year in the additional data are so small that they hardly produce any variance, producing small t-statistics. The groundwater level in Hengelo can differ by 70 to 100 cm between winter and summer, so the 2.3 cm found is negligible. The earlier acceptance of only six years of data is thus justified. In addition, the points with less years of data are also acceptable and usable as the mean there is little deviation in the mean. However, logically, the variance of the years does decrease with an increase of data. To properly test if the variance across the means of the data is equal, it would be necessary to perform an ANOVA, or analysis of variance, statistical test.

4.2 Precipitation and groundwater-level before RHG

It can be seen that the precipitation before RHG has the same shape as the STOWA lines. However, there are some side notes to add to this. The STOWA lines are based on 1, 2, 4, 8 and 9 days and when plotted the first points initially double in axis value. Thus, when these are displayed, an exponential rise can be seen after which it smoothens out again on the 9th day. When this is plotted over an axis that ranges from 1 to 9 with all intermediate days, the line becomes linear. Thus, the depiction of this axis gives a skewed view of the otherwise linear data, but as the groundwater level rise is also depicted on a normal axis the result can still be compared. Additionally the precipitation and groundwater level rise are both linear, the only difference is in the quantity on the y-axis. The precipitation has value ranging between 20 and 60 millimetres, whereas the groundwater level rise ranges between 120 and 280 millimetres. Regrettably there further seems to be a paucity in literature to compare the precipitation before an RHG to. Many papers relate to the correlation between precipitation and rise, however, this is already discussed in an earlier section.

Next is that the precipitation before RHG does not have an intrinsic recurrence time like the STOWA lines and it is therefore difficult to compare them. The RHG, as it is defined, is met or exceeded 36 days a year or 10% of the year, but the RHG is dependent on precipitation that has a certain recurrence time. Currently, the first RHG situation coming from the summer to the (hydrological) winter is used, as it can then be said that the RHG value is not influenced by a previous RHG value. Yet, it in some situations it is possible, that the RHG could be influenced, for example, by the precipitation causing a first RHG coming from the summer - this only happens once per decade; or that (for example) the second RHG situation is caused by a precipitation once per year in the winter. As it stands, it is difficult to compare this with the STOWA values that are used in the absence of RHG data. Nevertheless, it can be seen that the line of the RHG value lies lower than the STOWA values. It was already expected that the values of the STOWA currently overestimate a RHG situation and this graph can show that this is indeed true. Furthermore, the graph below (Figure 22) can be read in a similar fashion to the correlation graphs, where all data shown (with the respective amount of days) on and above the line is a RHG situation.

When looking at the groundwater rise before an RHG situation over 1, 2, 4, 8 and 9 days, the same point regarding the scale of the axis can be made as the precipitation before an RHG situation. In this case, and in Figure 22, the same axis was chosen to be able to compare the two. From this it can be seen that they present a similar graph which means that the relationship between the groundwater rise and precipitation is likely to be linear. Unfortunately the groundwater level rise does not have standardized Dutch values to compare it to (next to the earlier mentioned points), thus another paucity in literature has been found.

5 Recommendation

As mentioned earlier, groundwater levels rely on many more factors than just precipitation. This research has focused only on the effect of precipitation on groundwater levels. For further research, it is advised to not use a regression but to build a predictive model. A model that has a moving average as well as an auto-regressive function (regression over time varying process), essentially a time series analysis, can predict the varying pattern of groundwater levels over time but has a certain repetition over time. Time series analysis uses models such as ARIMA (Autoregressive Integrated Moving Average) or SARIMA (Seasonal-ARIMA), are the most appropriate for the data. For a prediction model, many more variables are needed, the model is only as good as the values supplied [39]. To ensure this, research has to be done in all factors that can influence the groundwater. Factors that influence the groundwater but not limited to are:

- Porosity: this is the amount of water the soil can hold. The more porous the soil, the more water it can hold. [40]
- Permeability: this refers to the ability of water to flow through the soil or rock. The more permeable the soil or rock, the easier it is for water to flow through it. [41]
- Gravity: gravity is another factor that influences the flow of groundwater. Water flows downhill due to gravity, so the slope of the land can affect the movement of groundwater. [41]
- Climate: precipitation must exceed evaporation for groundwater to exist. The amount and timing of rainfall can affect the amount of groundwater that is available [42].
- Topography: the shape of the land can also influence the movement of groundwater. Water flows downhill due to gravity, so the slope and geology of the land can affect the movement of groundwater [42].
- Anthropogenic factors: human activities can also affect groundwater quality. For example roads, drainage and other man-made objects can influence groundwater [43].

It's important to note that these factors are interconnected, and changes in one factor can have cascading effects on others, ultimately impacting groundwater availability and quality.

In order to further enhance the research that was conducted, it is strongly advised to incorporate an ANOVA (Analysis of Variance) analysis to evaluate the impact of the number of years on the observed data. The inclusion of this analysis will provide statistical evidence of the variation and significance between different groups. Furthermore, employing ANOVA can yield outcomes that differ from those obtained using the t-test, adding a new perspective to the findings.

Hengelo's soil and landscape exhibit distinct characteristics, with a sand underground and silt layers covering a shallow hillside. However, it is intriguing to investigate whether variations exist when comparing other soil types or landscapes with differing slopes, which may be either larger or smaller in magnitude. Thus it is advised to perform the analysis on other places with these different soils.

6 Conclusion

The relation between the precipitation and the RHG groundwater level for Hengelo in the hydrological winter is positive and can be described within 15 days with the following formula :

$$y = 3.59x + 60.4 \tag{5}$$

where y is the average rise in groundwater (in [mm]) needed for an RHG to occur and x the amount of precipitation in mm. If the value of rise in groundwater with a given precipitation for a certain day is equal or higher than the line then the RHG value is met for Hengelo.

The number of years of data needed to accurately describe Hengelo is statistically proven to be 6. However, in this case the spread is very small and thus 1 is also sufficient.

Regarding the precipitation and rise in groundwater against the amount of days before a RHG occurrence, the following formulas were found:

$$y_2 = 2.86x_2 + 15.55 \tag{6}$$

$$y_3 = 10.23x_2 + 116.39 \tag{7}$$

where y_2 and y_3 (both in [mm]), respectively, are the precipitation and rise in groundwater level needed to reach the RHG level and x_2 ([-]) is the amount of days before the RHG occurence. When comparing the precipitation before a RHG occurence against the currently used STOWA values, the STOWA overestimates the RHG value. However the STOWA has a recurrence time of once or twice a year whereas the determined RHG does not have a recurrence time.

References

[1] European Climate Commission, "Consequences of climate change." [Online]. Available: https://climate.ec.europa.eu/climate-change/consequences-climate-change_en#:$\sim$:text=Climatechangeaffectsallregions,extremeheatwavesanddroughts.

[2] 350 Org, "CLIMATE SCIENCE BASICS." [Online]. Available: https://350.org/science/?gclid=Cj0KCQjwuLShBhC_ARIsAFod4fKOOz5pDHm_mivZjmIZ17bOmLqpZpqc_ZYhsh7ED9-G5fgI4UiuwJgaAka1EALw_wcB

[3] P. M. Berry, R. Betts, P. Harrison, and D. L. Sandars, "High-end climate change in Europe - impacts, vulnerability and adaptation," Tech. Rep., 2017. [Online]. Available: https://www.researchgate.net/figure/Changes-in-global-mean-temperature-relative-to-1985-2005-projected-for-the-CMIP5-climate_fig2_318013753

[4] "Drinkwater — Atlas Natuurlijk Kapitaal." [Online]. Available: https://www.atlasnatuurlijkkapitaal.nl/natuurlijk-kapitaal/drinkwater

[5] International Groundwater Resources Assessment Centre UN, "Groundwater Climate Change." [Online]. Available: https://www.un-igrac.org/areas-expertise/groundwater-climate-change

[6] "PRECIPITATION — English meaning - Cambridge Dictionary." [Online]. Available: https://dictionary.cambridge.org/dictionary/english/precipitation

[7] "Aquifers and Groundwater — U.S. Geological Survey." [Online]. Available: https://www.usgs.gov/special-topics/water-science-school/science/aquifers-and-groundwater

[8] "How relevant is evapotranspiration in the groundwater flow regime? — Hatari Labs." [Online]. Available: https://hatarilabs.com/ih-en/how-relevant-is-evapotranspiration-in-the-groundwater-flow-regime

[9] P. Smeets, G. Medema, and J. C. van Dijk, "The Dutch secret: How to provide safe drinking water without chlorine in the Netherlands," Tech. Rep., 2009. [Online]. Available: https://www.researchgate.net/figure/Source-water-types-in-the-Netherlands_fig1_40742083

[10] S. Y. Philip, S. F. Kew, K. V. D. Wiel, N. Wanders, G. J. V. Oldenborgh, and S. Y. Philip, "Regional differentiation in climate change induced drought trends in the netherlands," *Environmental Research Letters*, vol. 15, 9 2020.

[11] R. A. Smail, A. H. Pruitt, P. D. Mitchell, and J. B. Colquhoun, "Cumulative deviation from moving mean precipitation as a proxy for groundwater level variation in Wisconsin," *Journal of Hydrology X*, vol. 5, p. 100045, nov 2019.

[12] L. Damiba, A. Doumounia, V. Casey, A. Bounkoungou, F. Zougmoré, L. Damiba, A. Doumounia, V. Casey, A. Bounkoungou, and F. Zougmoré, "Analysis of Rainfall Variability on the Groundwater Levels of Wells in the Nouaho Basin in East-Central Burkina Faso," *Journal of Water Resource and Protection*, vol. 12, no. 11, pp. 964–974, nov 2020. [Online]. Available: http://www.scirp.org/journal/PaperInformation.aspx?PaperID=104546http://www.scirp.org/Journal/Paperabs.aspx?paperid=104546

[13] Y. K. Zhang and K. E. Schilling, "Effects of land cover on water table, soil moisture, evapotranspiration, and groundwater recharge: A Field observation and analysis," *Journal of Hydrology*, vol. 319, no. 1-4, pp. 328–338, mar 2006.

[14] "Twents Waternet - Actuele Grondwaterstanden." [Online]. Available: https://grondwater.webscada.nl/twentswaternet/

[15] "(PDF) Seasonal changes in groundwater quality and its suitability for drinking and irrigation uses." [Online]. Available: https://www.researchgate.net/publication/312023102_Seasonal_changes_in_groundwater_quality_and_its_suitability_for_drinking_and_irrigation_uses

[16] S. F. Yan, S. E. Yu, Y. B. Wu, D. F. Pan, D. L. She, and J. Ji, "Seasonal Variations in Groundwater Level and Salinity in Coastal Plain of Eastern China Influenced by Climate," *Journal of Chemistry*, vol. 2015, 2015. [Online]. Available: https://www.researchgate.net/publication/282515170_Seasonal_Variations_in_Groundwater_Level_and_Salinity_in_Coastal_Plain_of_Eastern_China_Influenced_by_Climate

[17] T. R. Green, "Linking climate change and groundwater," *Integrated Groundwater Management: Concepts, Approaches and Challenges*, pp. 97–141, jan 2016. [Online]. Available: https://link.springer.com/chapter/10.1007/978-3-319-23576-9_5

[18] R. M. M. Leith, P. H. Whitfield, R. M. M. Leithl, and P. H. Whitfield2, "EVIDENCE OF CLIMATE CHANGE EFFECTS ON THE HYDROLOGY OF STREAMS IN SOUTH-CENTRAL BC," *http://dx.doi.org/10.4296/cwrj2303219*, vol. 23, no. 3, pp. 219–230, 2013. [Online]. Available: https://www.tandfonline.com/doi/abs/10.4296/cwrj2303219

[19] I. S. Zektser and H. A. Loaiciga, "Groundwater fluxes in the global hydrologic cycle: past, present and future," *Journal of Hydrology*, vol. 144, no. 1-4, pp. 405–427, apr 1993.

[20] R. Langridge, "Implications of climate change to groundwater," *Encyclopedia of the World's Biomes*, vol. 4-5, pp. 438–453, jun 2020.

[21] A. S. Elshall, M. Ye, and Y. Wan, "Groundwater sustainability in a digital world," *Water and Climate Change: Sustainable Development, Environmental and Policy Issues*, pp. 215–240, jan 2022.

[22] "Uurwaarden van weerstations." [Online]. Available: https://daggegevens.knmi.nl/klimatologie/uurgegevens

[23] J. Beersma, H. Hakvoort, R. Jilderda, A. Overeem, and R. Versteeg, "Neerslagstatistiek en reeksen voor het waterbeheer 2019," *Stichting Toegepast Onderzoek Waterbeheer*, p. 205, 2019.

[24] "Normal Distribution — Examples, Formulas, Uses." [Online]. Available: https://www.scribbr.com/statistics/normal-distribution/

[25]

[26] "Normality Test: What is Normal Distribution? Methods of Assessing Normality — Editage." [Online]. Available: https://www.editage.com/blog/normality-test-methods-of-assessing-normality/

[27] "Test of Normality • Simply explained - DATAtab." [Online]. Available: https://datatab.net/tutorial/test-of-normality

[28] "Central Limit Theorem — Formula, Definition Examples." [Online]. Available: https://www.scribbr.com/statistics/central-limit-theorem/

[29] "Box Plot Explained with Examples - Statistics By Jim." [Online]. Available: https://statisticsbyjim.com/graphs/box-plot/

[30] "Box and Whisker Plots - Learn about this chart and its tools." [Online]. Available: https://datavizcatalogue.com/methods/box_plot.html

[31] "What Assumptions Are Made When Conducting a T-Test?" [Online]. Available: https://www.investopedia.com/ask/answers/073115/what-assumptions-are-made-when-conducting-ttest.asp

[32] "Paired Sample T-Test - Statistics Solutions." [Online]. Available: https://www.statisticssolutions.com/free-resources/directory-of-statistical-analyses/paired-sample-t-test/

[33] "Linear Regression Explained. A High Level Overview of Linear... — by Jason Wong — Towards Data Science." [Online]. Available: https://towardsdatascience.com/linear-regression-explained-1b36f97b7572

[34] "Assumptions of Linear Regression - Statistics Solutions." [Online]. Available: https://www.statisticssolutions.com/free-resources/directory-of-statistical-analyses/assumptions-of-linear-regression/

[35] "T Test (Student's T-Test): Definition and Examples - Statistics How To." [Online]. Available: https://www.statisticshowto.com/probability-and-statistics/t-test/

[36] T. C. Winter, "The interaction of lakes with variably saturated porous media," *Water Resources Research*, vol. 19, no. 5, pp. 1203–1218, oct 1983. [Online]. Available: https://onlinelibrary.wiley.com/doi/full/10.1029/WR019i005p01203https://onlinelibrary.wiley.com/doi/abs/10.1029/WR019i005p01203https://agupubs.onlinelibrary.wiley.com/doi/10.1029/WR019i005p01203

[37] S. S. D. O. I. H. J Bear, AHD Cheng, "Seawater Intrusion in Coastal Aquifers: Concepts, Methods and Practices - Google Boeken." [Online]. Available: https://books.google.nl/books?hl=nl&lr=&id=_e42XC6ekRoC&oi=fnd&pg=PR11&ots=d3-zLoIKf6&sig=NjbHX9Z0Tx960abPX_Tkn9GUsLc&redir_esc=y#v=onepage&q&f=false

[38] "Topografische kaart Hengelo, hoogte, relief." [Online]. Available: https://nl-nl.topographic-map.com/map-djxb3/Hengelo/?center=52.26774%2C6.78268&zoom=13

[39] J. Fattah, L. Ezzine, Z. Aman, H. El Moussami, and A. Lachhab, "Forecasting of demand using ARIMA model," *International Journal of Engineering Business Management*, vol. 10, oct 2018.

[40] "Factors that Influence Groundwater Movement - Video Lesson Transcript — Study.com." [Online]. Available: https://study.com/academy/lesson/factors-that-influence-groundwater-movement.html

[41] "Groundwater Movement Factors Influences — How Does Water Move Underground? - Video Lesson Transcript — Study.com." [Online]. Available: https://study.com/learn/lesson/groundwater-movement-factors-influences.html

[42] "7 FACTORS INFLUENCING EXISTENCE OF GROUND WATER." [Online]. Available: https://geographypoint.com/2022/08/7-factors-influencing-existence-of-ground-water/

[43] G. C. Johnson, T. M. Zimmerman, B. D. Lindsey, and E. L. Gross, "Factors affecting groundwater quality in the Valley and Ridge aquifers, eastern United States, 1993-2002," *Scientific Investigations Report*, 2011.

[44] H. Li, L. Yudong, Z. Ce, X. Zhang, W. Jing, and Z. Bao, "Seasonal and Inter-Annual Variability of Groundwater and Their Responses to Climate Change and Human Activities in Arid and Desert Areas :," 2020.

[45] American Society of Landscape Architects, "RESILIENT DESIGN: DROUGHT." [Online]. Available: https://www.asla.org/drought.aspx

[46] National Geographic, "Water Table." [Online]. Available: https://education.nationalgeographic.org/resource/water-table/

7 Appendix Normality table

Logger name	Mean	Mode	Median
1.001	1994.48	2003.0	2000.0
1.002	1976.65	1996.0	1983.79
1.003	1834.64	1833.0	1834.6
1.004	1781.66	1777.0	1780.0
1.005	1675.75	1691.0	1676.5
1.006	1721.32	1705.0	1719.5
1.007	1821.67	1818.0	1821.2
1.008	1840.74	1838.0	1841.0
1.009	1738.45	1729.0	1737.0
1.010	1644.90	1630.0	1642.0
1.011	1673.14	1666.0	1671.3
1.012	1751.02	1754.0	1752.0
1.013	1749.23	1742.0	1749.09
1.014	1735.80	1734.0	1737.4
1.015	1644.99	1641.0	1645.0
1.016	1622.15	1612.0	1621.7
1.017	1572.70	1564.0	1572.1
1.018	1576.03	1589.0	1582.0
1.019	1548.31	1567.0	1557.4
1.020	1568.61	1582.0	1573.16
1.021	1487.95	1510.0	1493.5
1.022	1480.98	1471.0	1480.9
1.023	1535.99	1532.0	1537.4
1.024	1539.53	1537.0	1540.34
1.025	1555.62	1543.0	1554.3
1.026	1571.65	1568.0	1572.08
1.027	1622.80	1623.0	1623.8
1.028	1518.74	1518.0	1519.88
1.029	1491.45	1486.0	1490.022
1.030	1488.50	1412.0	1496.42
1.031	1551.69	1550.0	1552.0
1.032	1525.93	1514.0	1526.0
1.033	1581.73	1594.0	1584.3
1.034	1615.22	1608.0	1615.0
1.035	1667.46	1666.0	1667.26
1.036	1685.52	1680.0	1686.23
1.037	1716.75	1711.0	1717.6
1.038	1883.52	1887.0	1882.19
1.039	1895.27	1882.0	1896.9
1.040	1873.64	1882.0	1872.43
1.041	1924.59	1933.0	1929.435
1.042	1771.70	1761.0	1773.0
1.043	1926.00	1939.0	1927.7
1.044	2069.78	2086.0	2074.57
1.045	2023.50	2027.0	2027.9
1.046	1872.11	1893.0	1870.5
1.047	1832.89	1832.0	1834.0
1.048	1783.95	1778.0	1785.4
1.049	1769.66	1762.0	1771.28
1.050	1750.47	1746.0	1751.4
1.051	1681.11	1681.0	1684.0
1.052	1675.15	1666.0	1673.9

Logger name	Mean	Mode	Median
1.053	1593.66	1576.0	1592.3
1.054	1657.93	1662.0	1662.7
1.055	1616.33	1612.0	1620.7
1.056	1583.17	1593.0	1586.5
1.057	1737.07	1738.0	1740.0
1.058	1517.35	1510.0	1517.48
1.059	1505.59	1492.0	1506.0
1.060	1514.20	1523.0	1513.0
1.061	1433.21	1422.0	1432.0
1.062	2025.76	2004.0	2029.5
1.063	1918.32	1921.0	1918.91
1.064	1876.62	1876.0	1877.73
1.065	1823.19	1844.0	1828.8
1.066	1702.22	1705.0	1704.61
1.067	1619.77	1623.0	1622.0
1.068	1543.96	1537.0	1541.0
1.069	1519.85	1511.0	1522.74
1.070	1556.99	1546.0	1558.0
1.071	1455.47	1454.0	1456.3
1.072	1482.38	1463.0	1481.0
1.073	1373.92	1378.0	1375.5
1.074	1374.06	1356.0	1374.0
1.075	1345.33	1340.0	1345.16
1.076	1306.68	1352.0	1344.37
1.077	1434.76	1428.0	1432.3
1.078	1419.79	1444.0	1425.0
1.079	1468.93	1474.0	1473.18
1.080	1522.19	1534.0	1532.0
1.081	1592.66	1595.0	1597.0
1.082	1481.44	1486.0	1484.0
1.083	1630.43	1625.0	1631.77
1.084	1651.37	1650.0	1654.41
1.085	1681.17	1676.0	1683.8
1.089	1717.58	1711.0	1719.0
1.090	1735.98	1758.0	1732.0
1.091	1704.09	1693.0	1705.4
1.092	1584.49	1560.0	1584.25
1.093	1681.04	1668.0	1683.61
1.094	1824.01	1833.0	1823.0
1.095	1841.30	1830.0	1842.0
1.096	1959.09	1986.0	1959.2
1.097	1951.95	1958.0	1951.7
1.098	2045.39	2041.0	2047.0
1.099	1959.05	1942.0	1957.5
1.100	1869.13	1861.0	1870.55
1.101	2089.48	2085.0	2091.43
1.102	2222.70	2220.0	2224.9
1.103	2049.36	2038.0	2050.7
1.104	1816.31	1805.0	1819.31
1.105	1675.96	1671.0	1677.4
1.106	1861.20	1824.0	1865.0
1.107	1738.90	1640.7	1704.7