GÖTEBORGS UNIVERSITET

INSTITUTIONEN FÖR
GEOVETENSKAPER

# FROM SPACE TO THE SUBSURFACE

## Examining Relations Between Vegetation Indices and Local Groundwater Storage.

**Gustav Antonsson**

Uppsats för avläggande av filosofie masterexamen med huvudområdet geografi
2023, 120 hp
Avancerad nivå
B1279

# Abstract

This study aimed to examine the relationship between vegetation indices, NDVI and NDWI, and groundwater levels in the county of Kalmar, utilizing correlation and regression analysis. Further, by examining related geospatial features the study aimed to interpret the statistical outcomes to identify significant temporal and spatial patterns. Data used involved NDVI, NDWI, derived from Sentinel 2 level-1C imagery, as well as groundwater measurement. The data was presented in time series showing bi-weekly maximum values, extending over different ranges between 2015 and 2022. Additionally, data related to land cover, soil type, topographic location, distance between groundwater and ground surface have been observed and compared between measurement stations to create a framework for interpretation.

While few definite patterns have emerged, results of the study provided notable observations from performed analysis. Results showed varying strengths of correlations over the measurement stations studied, for both indices in relation to groundwater levels, with strongest correlation generally found after three-month time-lag of vegetation indices. NDVI showed positive correlation, indicating high NDVI values correlating with low groundwater levels and vice versa. NDWI over most stations showed negative correlation, indicating high NDWI values indicating high groundwater levels. Also, while soil type and median groundwater depth were features that provided notable findings when analyzed separately in relation to found correlations, combining several features make patterns less certain. Results of this study show variation in correlation being due to variations in local geospatial features.

Future studies should more extensively and separately examine the effect of geospatial features related to vegetation and groundwater correlation. The use of different remote sensing data sources such as Sentinel 2 level-2A and Radar could present a more informative result. Also, addressing more careful data collection, seasonal focus and use of visualization. The addition of precipitation data could further be used to provide more detail in explaining the relationship between vegetation and groundwater.

# Acknowledgement

Completing this thesis marks the culmination of my five-year journey in geography studies. It has been an enriching experience that has broadened my understanding of the world around us and deepened my appreciation for the impact that geography has on our daily lives.

I would like to express gratitude to my supervisor, Roland Barthel, for providing me with guidance throughout this thesis work. Your understanding has been instrumental in helping me navigate the challenges of this research and bring it to a conclusion. I would also like to thank my fellow students, particularly those I shared an office with during this last term, for creating a fun and supportive learning environment. You have made these years very enjoyable and for that I'm thankful.

Lastly, I am excited to apply what I have learned during my university studies in working life. I look forward to exploring the applicability of geography and doing good work within the field. To my fellow students, I wish you all the best in your future endeavors and hope that we can continue to learn from each other as we embark on this new chapter in our lives.

All the best!
Gustav Antonsson

# Table of content

Abstract ..... 2

Acknowledgement ..... 3

1. Introduction ..... 6

1.1 Aim ..... 7

2. Background and key themes ..... 7

2.1 Groundwater ..... 7

2.2 Remote sensing and groundwater ..... 10

2.3 Vegetation indices ..... 12

2.4 Statistical analysis ..... 14

2.5 Previous research ..... 15

3. Study area ..... 18

4. Data ..... 22

4.1 Remote sensing data ..... 22

4.2 Groundwater data ..... 22

4.3 Additional data ..... 22

5. Method ..... 23

5.1 Collecting and processing remote sensing data ..... 24

5.2 Spearman's correlation coefficient and linear regression analysis ..... 25

5.4 Comparison of time series and found correlations to geographical features ..... 27

6. Results ..... 29

6.1 Correlation and regression outcomes ..... 29

6.2 Interpretation of time series ..... 38

6.3 Further interpretation of geospatial features and found correlations ..... 42

7. Discussion ..... 49

7.1 Comparison of $r_s$ based on visual interpretation and statistical analysis. ..... 49

7.2 Variation over time and geospatial features ..... 50
7.3 Methodology and input parameters ..... 55
8. Conclusions ..... 58
9. References ..... 59
10. Appendix ..... 63

# 1. Introduction

Groundwater is one of the most important natural resources to societies worldwide. It provides us with drinking water and the access to adequate groundwater is of utmost importance for a society to function (Eveborn, et al. 2017). During recent years, groundwater levels in Sweden have varied a lot and longer periods of low groundwater levels, both for large and small aquifers, have occurred. The reasons for this can be found both in nature-based processes and human activities. With climate change, changing weather patterns are predicted, including drought, scarce snowmelt, changing vegetation seasons, and varying precipitation levels, which all in their specific way affect groundwater recharge (Lång, et al. 2022). Especially precipitation, playing a determining role in groundwater recharge (Eveborn, et al. 2017). Adding to this, increased human groundwater usage affects the local groundwater availability. Other regional to local variations that affect groundwater recharge are variations in geology, topography and vegetation, creating differences between several localities within Sweden as a whole (Lång, et al. 2022).

The county of Kalmar, which consists of twelve municipalities in the southeastern parts of Sweden, has been estimated to be an area at risk of increasingly low groundwater levels. The reason for this has many potential answers. Kalmar County as part of the southern Baltic Sea area is to a larger extent than most other parts of Sweden occupied by agricultural areas which tend to be quite demanding on groundwater resources, especially during peak growing season (Vattenmyndigheten Södra Östersjön, 2016). Also, lower yearly mean values for effective precipitation as well as higher yearly mean evapotranspiration[1] have been observed for this district in relation to the rest of Sweden (Eveborn, et al. 2017; SGU, 1994). The mapped soil depth, presented by Geological Survey of Sweden (SGU) (SGU, 2023, a), indicates that large parts of the county of Kalmar have relatively shallow soil depths, which too can affect the groundwater recharge reducing the amount of soil water able to be maintained in the soil layer.

In order to keep up with the changing groundwater conditions continued measurements and research is required. The complex character of groundwater and phenomenon surrounding and affecting it requires studies to be done at different scales and with methodologies and data from

[1] Evapotranspiration – The water which departs to the atmosphere as fumes, post evaporation or transpiration from ground, water, or vegetation surfaces (SGU, n.d, c).

different scientific backgrounds. An example of this is the growing use of remote sensing[2] to collect and analyze information about groundwater quality and quantity, as well as conditions affecting these. An increase in regional groundwater measurements is available thanks to the use of automated loggers provided at most observation wells (Lång, et al. 2022). Combining frequent groundwater level measurements with frequent remote sensing data can present some insight in local groundwater recharge over a longer time period (Meijerink, et al. 2007).

## 1.1 Aim

The main aim of this study is to examine the relationship between vegetation features and groundwater levels, further analyzing groundwater situations within the county of Kalmar. This, by using statistical analysis on time series with Sentinel 2 level-1C data, displaying NDVI as well as NDWI, and groundwater level measurements. Additionally, using geospatial features, trying to explain variations within the statistical outcomes. Geospatial features used involves land cover mainly vegetation type, soil type, topography and depth to aquifer.

### 1.1.1 Research questions

- What correlations within time and space exist between groundwater levels and vegetation indices, NDVI and NDWI, Sentinel 2 level-1C imagery, in the case of Kalmar County?
- How do studied geospatial features affect the correlation between groundwater levels and vegetation indices, NDVI and NDWI, within the county of Kalmar?

# 2. Background and key themes

## 2.1 Groundwater

Groundwater is water that exists below the ground surface within pores in the fractures of different soil types. Precipitation can infiltrate the ground and percolate[3] towards the groundwater table which is the upper boundary of the saturated zone. Aquifers are the geological features within which the groundwater is contained. Groundwater aquifers can take

[2] Remote sensing – The process of measuring physical characteristics of an area from a distance, using its reflected and emitted radiation (Mason, 2016).

[3] Percolation – Infiltrated water which makes it way downwards below ground surface, towards the groundwater table (SGU, n.d, c).

many different shapes and sizes and are often classified as large or small aquifers in Sweden. Larger aquifers tend to have less groundwater level variations than the small aquifers where groundwater levels tend to fluctuate more. This additionally varies depending on the aquifer's depth below ground surface and its soil type (Bergström, 2001).

### 2.1.1 Groundwater and vegetation

According to Eveborn et al. (2017) precipitation minus evapotranspiration is a main determinant for groundwater recharge. Vegetation plays a crucial role in this process through both interception and transpiration. Interception occurs when precipitation lands on vegetation and evaporates before reaching the ground, while transpiration describes the process where vegetation collects water from the ground before releasing it through plant fumes. During seasons of healthy vegetation growth, smaller amounts of groundwater can be restored compared to when the ground surface is more accessible. The vegetation's ability to create evaporation through interception and transpiration further depends on its capacity to collect water from the groundwater surface level and its leaf area index[4] (LAI). Also, vegetation can act as a barrier, slowing down surface runoff and thus potential infiltration. While groundwater levels do not typically increase during the growing season, there are exceptions, such as during cloudbursts and heavy rainfall that can cause groundwater recharge infiltration and overcome interception and transpiration (Eveborn, et al. 2017).

LAI being a determinant for interception indicates that dense healthy vegetation covering surface areas result in low amounts of effective precipitation, i.e the difference between total precipitation and total evapotranspiration. This in turn suggests that forested areas generally would have higher interception than non-forested areas (Mason, 2016). Research by Johansson (1998) has further shown that deforestation can lead to groundwater recharge in areas where smaller vegetation patches exist. This is due to the reduced evaporation through interception and transpiration resulting from the lack of vegetation. Without this barrier, precipitation can reach the ground surface more easily and be absorbed by unsaturated soil (Johansson 1998). Building on this, soil water, which is the water contained in the soil above the groundwater aquifer, plays a crucial role in vegetation's ability to collect water from the ground surface. In Sweden, early summers are typically characterized by low precipitation, making vegetation more dependent on soil water during this time. The most soil water is usually found during

[4] Leaf area index – The total one-sided leaf area of a vegetation feature per unit ground area (Jones & Vaughan, 2010)

spring, after snowmelt has infiltrated the ground, and during autumn when precipitation amounts are generally higher and vegetation's ability to absorb water is lower (Grip et al. 1994).

### 2.1.2 Topography, soil type and groundwater

Three topographic locations related to groundwater are recharge areas, discharge areas and intermediate areas. Recharge areas are areas where groundwater flow is predominantly directed downwards, in discharge areas it is predominantly directed upwards. Recharged areas are often located in topographical highs, at the catchments boundaries. Discharge areas are in topographical lows, usually close to surface water bodies or wetlands. Intermediate areas are located between recharge and discharge areas and processes can differ depending on seasons and weather conditions (Condon & Maxwell, 2015). Watersheds are the boundaries that divide the water flow into different directions. Surface watersheds are determined by topography and runoff water following the terrain. Ground watersheds cannot be studied from above ground (Haraldsson, 2015).

If the ground surface is saturated, there is an increase in surface runoff due to the grounds inability to absorb more water. Surface runoff also occurs during intense precipitation when the ground absorption falls behind, and excess water continues its path above ground. During such conditions, the ground can only absorb the water when it enters a low and loses its momentum (Grip, et al. 1994).

While topography can affect groundwater recharge, soil type oftentimes is even more determinant. The permeability of different soil types affects how much and how quickly infiltrated precipitation can reach the groundwater aquifer and restore its surface levels. When soil permeability is high the topography is less determinant of the infiltration and travel of runoff precipitation than when permeability is low (SGU, 2023, b). Within sand or gravel-based soil types, such as glacial sediments, the permeability is generally high, while mixed particles such as within moraine the permeability is estimated to a medium rate. Dense soil types with fine particles such as clay result in low permeability (Lindström, et al. 2011). While permeability determines how quickly and effectively water can move through a porous medium, porosity determines how much water can be stored in rocks or soils. Porosity can be described as the pore volume in relation to the total volume and pore volume varies depending on the particle size within each soil type (Karlsson, et al. 2021).

Groundwater aquifers located deep below ground surface tend to see less fluctuation in groundwater level variation, while aquifers located just a few meters below ground surface (m.b.g.s) tend to have more fluctuation in groundwater levels. Shallow laying groundwater aquifers react faster to changes in precipitation and are at higher risk in being drained through vegetation transpiration and drought. Aquifer size and soil type are also determinants within the above-described pattern (Bergström, 2001).

Above mentioned aspects can create a delay in groundwater recharge. This means that even if precipitation has infiltrated the ground surface, there could be a wide range of time passing before the soil water reaches the groundwater table (Grip, et al. 1994).

## 2.2 Remote sensing and groundwater

Using remote sensing within groundwater studies can involve qualitative as well as quantitative approaches. The wide variety of approaches suitable for a specific study is determined by parameters such as data, instruments, temporal and spatial resolution, scale, land cover, climate and weather conditions, policies, etcetera. All these parameters are examples of what makes remote sensing applicability to groundwater studies extensive but also complex (Meijerink, et al. 2007).

At its core, remote sensing applicability to groundwater studies is defined by the different ways in which earth surface features react to different wavelengths on the electromagnetic spectrum (Meijerink, et al. 2007). As an example, healthy vegetation reflects a high level of incident light in the near-infrared part of the spectrum, while it absorbs much and reflects little of the incident light in the red part of the spectrum (Jones & Vaughan 2010). With the use of electro-optical sensors, which register the sunlight's reflectance from earth surface objects, spectral properties can be examined. Additional sensors commonly used today in remote sensing are Radar and Lidar sensors using microwaves and laser respectively to collect information about earth's surface (Olsson, et al. 2018). Whether it is reflecting or absorbing, different wavelengths within the bands of the electromagnetic spectrum react differently to different earth features making it possible to collect information about the feature based on its spectral properties (Meijerink, et al. 2007).

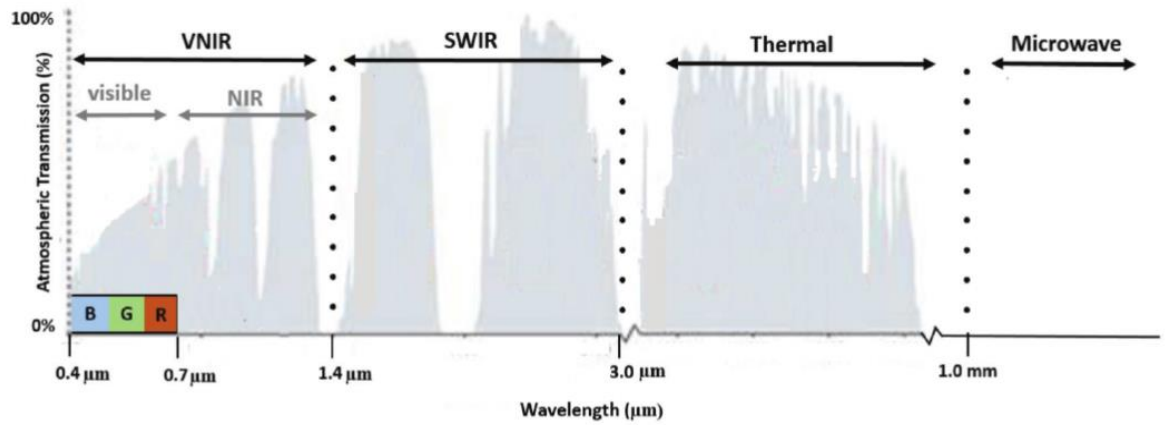


*Figure 1, The electromagnetic spectrum, x-axis displays wavelengths stretching from shorter to longer and y-axis showing percentage of atmospheric transmission, i.e., percentage of radiation passing through Earth's atmosphere (Campbell, et al. 2022).*

Different sections of the electromagnetic spectrum can further be targeted for different utilizations. The visible part of the spectrum in combination with the near infrared lights as well as short wave infrared lights can be used to interpret different kinds of land cover. Vegetation, water sources and urban areas, just to name a few, all have potential implications on the water balance[5] at a local scale. Moving up on the spectrum, into the microwave section, Radar systems use microwaves, also referred to as radio waves, to collect information on for example ground shifts and movements, soil moisture and vegetation, and lineament mapping.

In addition to this, elevation models can be created through height values in laser point clouds data, often collected from airborne LIDAR-scanners. Elevation models tell the user about runoff flow patterns as well as depressions where water can be collected. Using remote sensing in groundwater studies is preferably done using different datasets with different characteristics over different dates. As an example, multispectral imagery can be supplemented with topographic, geologic, hydrologic data in order to create a more extensive interpretation and analysis (Meijerink, et al. 2007).

Today there are several commercial as well as open-source data available for a wide range of users. There are several electro-optical satellites that provide multispectral imagery at different scales with different temporal and spatial resolutions (Holden, 2017). Some examples of satellite programs capturing multispectral imagery are Sentinel 2, Landsat 7 and 8, and Terra equipped with the MODIS instrument (ESA, n.d, a; Meijerink, et al. 2007). Examples of

[5] Water balance – The water balance for a period describes how the inflow of water to an area relates to the outflow and storage in the area (SMHI, 2022)

satellite missions collecting radar data are Sentinel 1 and SMAP (Adams, et al. 2022; Entekhabi, et al. 2010). Lastly an example of a satellite using a K-Band Ranging System measuring variation in earth's gravitation, the GRACE mission. Specifics on these missions, which data is available open source, as well as websites where data can be accessed are displayed in table 1.

| Mission | Agency | Spatial resolution | Temporal resolution | Instrument/Scanner | Bands | Status |
|---|---|---|---|---|---|---|
| *Sentinel-2* | *European Space Agency (ESA)* | *10, 20 and 60m* | *10 days single satellite, 5 days combined* | *MultiSpectral Instrument (MSI)* | *13 Spectral Bands* | *Active* |
| *Landsat 7* | *United States Geological Survey & NASA* | *30, 15, and 100m* | *16 days* | *Enhanced Thematic Mapper Plus (ETM+)* | *8 Spectral Bands* | *Active* |
| *Landsat 8* | *United States Geological Survey & NASA* | *30, 15, and 100m* | *16 days* | *Operational Land Imager (OLI)* | *11 Spectral Bands* | *Active* |
| *Terra* | *NASA* | *250, 500, 1000m* | *1-2 days* | *Moderate Resolution Imaging Spectroradiometer* | *36 Spectral Bands* | *Active* |
| *Sentinel-1* | *European Space Agency (ESA)* | *5m to 400km* | *1-3 days* | *C-band Synthetic Aperture Radar* | *C-Band* | *Active* |
| *SMAP* | *NASA* | *3, 9 and 36km* | *2-3 days* | *L-Band Radiometer & L-Band Radar* | *L-band* | *Active* |
| *GRACE* | *NASA & German space agencies (DLR)* | *ca 100km* | *-* | *Several instruments, i.e K-Band Ranging System (KBR)* | *K-band* | *Active* |

*Table 1, Remote sensing missions with groundwater applicability, addressing their responsible agency, spatial and temporal resolution, what instrument or scanner is used collecting data, what sections of the electromagnetic spectrum (bands) are targeted, and whether they are active or not.*

## 2.3 Vegetation indices

### 2.3.1 NDVI

Vegetation indices use the characteristics of certain wavelengths in relation to vegetation features on earth's surface creating a new variable. By using two spectral bands, it's possible to calculate the normalized difference between different spectral bands based on their reflective properties (Jones & Vaughan 2010). Also, the use of ratioed images, while not removing them completely, can mitigate atmospheric effects[6] on surface reflectance. The atmospheric effect can alter the spectral characteristics of observed images (Lillesand, et al. 2015).

One of the most common indices used for vegetation studies is the Normalized Difference Vegetation Index (NDVI) which uses the Near InfraRed (NIR) band and the Red band (Rouse, et al. 1973), see equation 1. In this case, thanks to the reflective properties of chlorophyll, the contrast between the low reflectance of the red band and the high reflectance of NIR can help

[6] Atmospheric effect – scattering and absorption of electromagnetic radiation by gas and particles within the atmosphere (Lillesand, et al. 2015).

the user identify pixels of dense healthy vegetation (Jones & Vaughan 2010). NDVI ranges from -1 to 1. In general terms high values indicate dense healthy vegetation while a zero or negative value indicates no vegetation at all. Adding to this, NDVI as a more accessible alternative gives a good estimate of the LAI (Jones & Vaughan 2010).

$$NDVI = (Nir - Red)/(Nir + Red) \quad \text{Eq. 1}$$

### 2.3.2 NDWI

Normalized Difference Water Index (NDWI), also known as Normalized Difference Moisture Index (NDMI), provides indication of vegetation and soil moisture. It has several iterations and usages depending on the band wavelengths used. Originally Gao (1996) presented a formula using the central wavelengths of 860 nm and 1240 nm, to study vegetation water content through remote sensing data.

$$NDWI = (\rho 860\,nm - \rho 1240\,nm)/(\rho 860\,nm + \rho 1240\,nm) \quad \text{Eq. 2}$$

Later, Chen et al. (2005) would modify the original formula and present an alternative available originally for Landsat data. The difference between the two is using the central wavelength 1640 nm of one of the SWIR bands. This band is also available through the Sentinel-2 mission.

$$NDWI = (NIR - SWIR)/(NIR + SWIR) \quad \text{Eq. 3}$$

According to Jones & Vaughan (2010), the latter NDWI formula shows less saturation than the original formula presented by Gao (1996). NDWI can be viewed as a complement to NDVI (Gao, 1996). It ranges from -1 to 1 (Jones & Vaughan 2010). High values indicate high presence of moisture while low values indicate drought. More specifically -1 to 0 indicate different degrees of dry- or wetness for no to low canopy cover. 0 to 0.4 can indicate high or low water stress for average canopy cover. 0.4 to 0.8 covers indication of water stress for mid-high canopy cover to very high canopy cover, no water stress. 0.8 to 1 indicates full canopy cover no water stress (EOS Data Analysis, n.d, a).

Lastly, a third iteration of NDWI is presented by McFeeters (1996) using the green band as well as the NIR band in order to identify surface water bodies. The chosen band combination aims to exclude influences from vegetation and soil, opposite to the previous two.

$$NDWI = (Green - NIR)/(Green + NIR) \quad \text{Eq. 4}$$

## 2.4 Statistical analysis

Applying remote sensing measurements in relation to groundwater measurements within statistical analysis can help researchers examine groundwater and earth features relationships. Examples of this found in previous research are presented in section 2.5, specifically regarding vegetation index and groundwater levels examined through correlation and regression analysis.

### 2.4.1 Spearman's correlation coefficient

Correlation coefficient is a statistical measure that describes the potential association between variables. There are several different ways to calculate and analyze the correlation coefficient. The method used for this study, Spearman's correlation coefficient ($r_s$), is similar to the more common Pearson's correlation coefficient (r) which calculates r based on the linear relationship of the input data variables. However, $r_s$ is based on ranks estimated from the values of the input data variables, applied to a linear relationship. Spearman's method allows for non-linear data, such as seasonally dependent time series data, to be used when calculating the correlation coefficient, keeping the calculation relatively simple while not neglecting the non-linear variations of the data set (Schober, et al. 2018). Output $r_s$ shows a value between -1 to 1 where 0 shows no correlation, 1 shows perfect positive correlation, and -1 shows perfect negative correlation. Positive correlation indicates similarly ranked values correlating and negative indicates the opposite, high rank values correlating with low rank values (Körner & Wahlgren, 2015).

It is important to disclose that a strong correlation value, whether negative or positive, doesn't necessarily mean that variable A triggers variable B, and their relationship is causal. Rather the case might be that both variables have the same separate factor to trigger causation, or the relationship between the two simply being a coincidence. Due to this, correlation coefficients can be supplemented with further analysis, such as regression analysis looking closer at the variance of used variables (Schober, et al. 2018).

### 2.4.2 Regression analysis

Regression analysis is a statistical method widely used within several scientific branches. The purpose of regression analysis is to further examine and analyze the relationship between a response and one or more predictor variables. The analysis creates a model based on the response variable and whether this can be predicted based on the values of the predictor

variables. Regression analysis can also be used to test your input predictor variables and further validate input data by estimating error values as well as other statistical measures. In this case measure of interest is Coefficient of Determination ($R^2$). $R^2$ describes to what extent the independent variables can explain the variance in the dependent variable and it stretches from zero to one. Here, 1 would indicate that the independent variable would explain 100% of the dependent variable variability. 0 would indicate that the independent variable would explain 0% of the dependent variable variability (Kuhn & Johnson, 2013). There are several different models to use when performing a regression analysis. Which model to use depends on several different aspects surrounding your subject and data. For this study, similar to Song, et al. (2021) (see section 2.5), a univariate linear regression analysis was used.

## 2.5 Previous research

There are several examples of previous studies where remote sensing data has been applied to groundwater studies. More specifically, there are several studies where vegetation indices are used to analyze groundwater phenomenon and occurrences. Bhanja, et al. (2019) used NDVI derived from MODIS satellite data, in order to study correlation with in-situ groundwater level measurements for large areas of India. The aim of the study was to use vegetation index data as an estimator for groundwater levels. They did this by collecting widely available MODIS data, precipitation measures and ground-based groundwater observations from the year 2005 to 2013, which then was applied to a machine learning process in order to create a prediction model describing potential groundwater levels. The groundwater observations were measured four times a year, once for early monsoon, once for late monsoon, once for pre-monsoon, once for mid-monsoon. Hence, they captured measurements occurring during extremely humid to semi-arid climates. The results show high spatial variability in correlation between groundwater levels and NDVI, acknowledging the wide variation of land use and land cover of the study area. Further, strong correlation values, $r > 0.6$ with a statistical significance of $p < 0.01$ according to Bhanja, et al. (2019), are shown from naturally vegetated areas, such as forests, in central India, as well as from the shrublands of western India. However weak to moderate correlation values shown in ($r = 0$ to 0.6) were found in most parts of southern and northern India and in the case of agricultural areas, which they argue are due to human impact. Further, the results vary over the seasons with highest positive correlation during early and late monsoon and low to moderate negative correlation during monsoon. Looking at precipitation

and NDVI, what Bhanja, et al. (2019) argues to be low to moderate correlation ($0 < r < 0.7$) was found.

Zhu, et al. (2015) also used in situ groundwater depth measurements and remote sensing data in order to examine the correlation between groundwater levels and vegetation health through NDVI. The study area in this case was six subareas in the Xiliao River Plain located within Inner Mongolia in northern China, an arid to semi-arid area on a regional scale. The main land cover within the study area varies between forests, grasslands and cultivated land. Data used for the study also involved precipitation data, and the time span for which all data covered was 1981-2010, with values extracted on a monthly basis. Due to the time span of 1981-2010, both the Advanced Very High-Resolution Radiometer (AVHRR) on National Oceanic and Atmospheric Administration satellites (NOAA) as well as MODIS was used for NDVI data collection. Further, both were analyzed separately due to potential calibration differences. The NOAA NDVI data from 1981-2000 and MODIS NDVI data from 2001-2010. Results show strong correlation ($r > -0.6$, similar to Bhanja, et al. (2019)) between max NDVI and groundwater depth for two of the six subareas, during dry periods between 2001-2010. One subarea showed moderate correlation $r = -0.54$, and the three sub areas remaining showed low correlation with values between $r = -0.33$ to $-0.22$. For the same period, the first two sub areas showed low correlation between precipitation and NDVI. Zhu, et al. (2015) argues that this strengthens the relation between NDVI and groundwater depth for these two areas. Strongest correlation between precipitation and NDVI for the years 2001-2010 shows values of $r = 0.7$ and $r = 0.6$. For the same areas, during the same time span, correlation between NDVI and groundwater levels show weak correlation. This further suggests vegetation being less dependent on groundwater than precipitation. Zhu, et al. (2015) argues that precipitation is a determinant for natural vegetation in the study area.

Song, et al. (2021) has performed a similar study as Zhu, et al. (2015) in the sense that the phenomenon of interest is groundwater level in relation to vegetation. This study also derives from China with the study area of Tarim Basin and Qaidam Basin, two arid areas in the west of China. Song, et al. (2015) used MODIS satellite data in order to capture NDVI from the year 2000 to 2016 and combined it with groundwater depth measurements in order to examine their correlation. In addition to this, Song, et al. (2015) use regression analysis to further explore the relationship between groundwater and vegetation cover. Different time spans were applied for different subareas of the total study area. While the results show what Song, et al. (2015) argues

to be a strong correlation r = -0.505 in the Kongque River Catchment, over the time span 2000-2004, the regression analysis shows a different indication. The coefficient of determination only showed a value of 0.255. This in turn would indicate that the chance of the response variables (in this case groundwater levels) variation being due to the independent variables (in this case NDVI) variation is only about 25 %.

Conclusions from above-described studies include the method of using vegetation indices to examine the relationship of surface cover, specifically vegetation, and groundwater surface depth. This, through time series data and statistical analysis in the form of correlation coefficients and regression analysis. Adding to this, there is value in supplementing correlation analysis with regression analysis, further aiming at validating your data and further exploring the relation between variables used.

The above-described studies have all taken place at a regional scale, while also considering differences in different subareas of the regions. Capturing the relationship between vegetation indices and groundwater depth at this scale could be argued to require a dense network of measurement stations for in-situ groundwater observations. Studying vegetation indices and groundwater depth correlation at a more local scale could potentially present a more distinct result. The remote sensing data used is also of relatively low spatial resolution due to the regional scale of the study areas. Looking at a local scale would require high resolution remote sensing data in order to properly display the vegetation indices. The time aspect arguably also plays a part in the results of the studies above. The input data used ranges from a monthly to a quarterly basis. Having more frequent measures within the input data could present a more revealing result.

The above studies took place within different climates, mainly occurring within arid to semi-arid regions while also acknowledging correlation between NDVI and groundwater levels during monsoon, creating a very humid climate. Arguing that vegetation more heavily relies on groundwater within arid to semiarid areas, due to lack of precipitation, applying this type of study to more humid climates over a longer time span might give varying results to the ones found in arid or semi-arid study areas, further exploring the potential impact of precipitation.

# 3. Study area

This study has examined 57 sub-samples of the larger study area of Kalmar county. The sub-samples were derived from 57 of SGU's groundwater measurement stations within Kalmar county. Additionally, out of the 57 sub-samples, ten focus samples have been further examined. The study area of Kalmar county as well as the sub- and focus samples are displayed in figure 2.

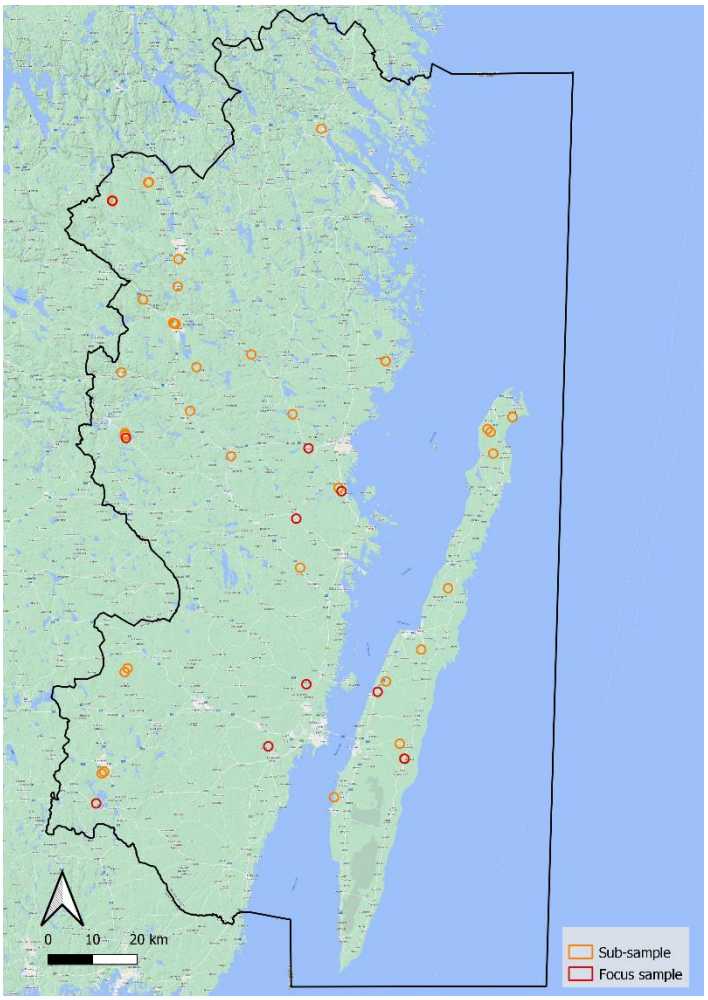


*Figure 2, The 57 total sub-samples within Kalmar County used for this study. Ten of 57 are further marked as focus samples, target for additional analysis.*

## 3.1 Groundwater

Kalmar county as part of the southern Baltic Sea area has varying prerequisites when looking at groundwater. As an example, different soil types and depths, topography, land cover and usages play vital roles in the county's groundwater situation. While there are occurrences of good groundwater resources supplying water of high quality, the groundwater recharge is one of the lowest in Sweden (Vattenmyndigheten Södra Östersjön, 2022).

## 3.2 Climate

Historically, the county of Kalmar has seen lower yearly mean precipitation than the rest of Sweden, with less than 200 mm precipitation yearly between 1961-1990. To compare, areas on similar longitudes but on the west coast of Sweden experience among the highest yearly mean precipitation in Sweden for the same time span. Building on this, the county also historically shows longer vegetation seasons ranging from 210-220 to 250-260 days per year, in different areas of the county, from the year 1961-2013 (ibid.). The county of Kalmar also saw higher mean evapotranspiration than most of Sweden, as well as lower mean runoff for 1961-1990 (SGU, 1994).

### 3.3 Land cover

Looking at land cover, generally the vegetation cover consists of crop land and a mix of coniferous and deciduous forests. The county has populated areas of different densities, varying from sparse populated areas in the countryside to dense urban areas within the twelve municipalities that constitute the county.

### 3.4 Geology

While the most occurring soil type is moraine there are also other soil types of different permeability present, as displayed in figure 3. Looking at soil depth in figure 4, this becomes relevant in groundwater recharge. Large parts of Kalmar County have relatively shallow soil depths, reducing the amount of water that can be stored in the soil.

The above-mentioned aspects are what make the county of Kalmar an area of interest when studying groundwater. Each aspect further has implications for groundwater conditions and recharge. Kalmar county as part of the southern Baltic Sea area has been recognized by Swedish water authorities as one of the more critical areas in Sweden, when discussing groundwater.

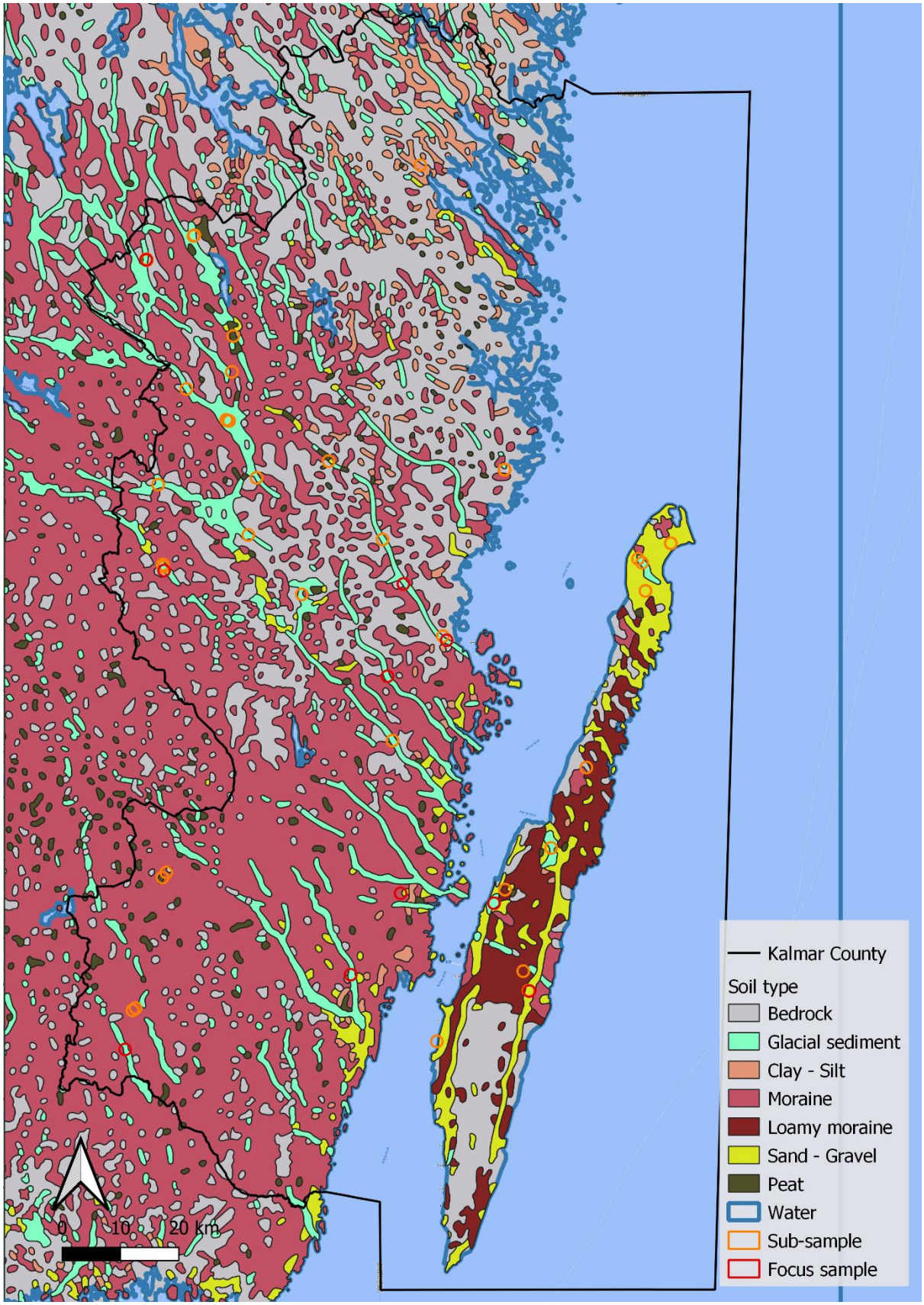


*Figure 3, Soil type map over the county of Kalmar. Additionally, the 57 sub-samples (including ten focus samples) are visualized above their respective soil type.*

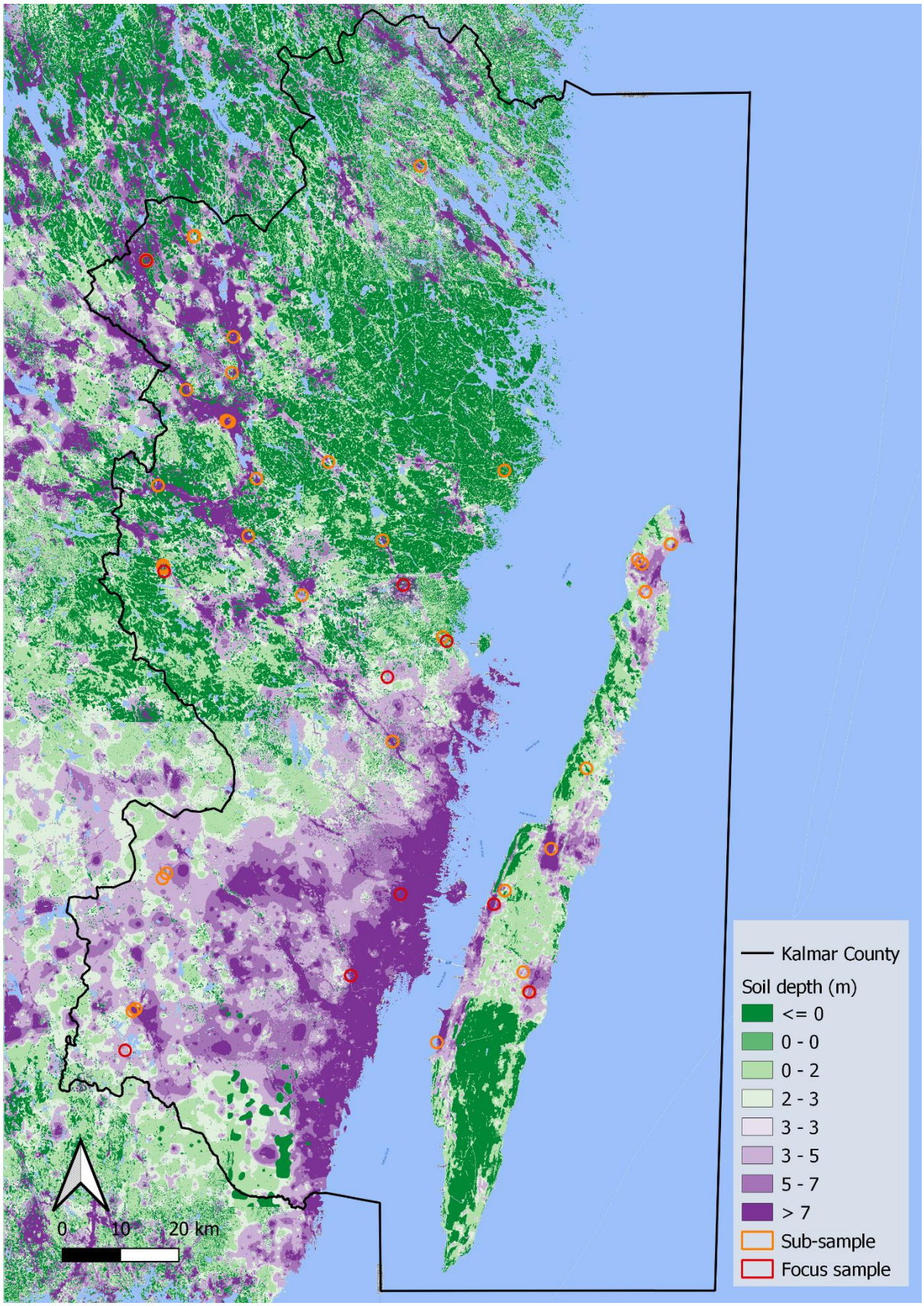


*Figure 4, Soil depth map over the county of Kalmar. Additionally, the 57 sub-samples (including ten focus samples) are visualized above their respective soil type.*

# 4. Data

## 4.1 Remote sensing data

For this study, Sentinel-2 Level-1C data has been used for the time period of 2015-06-23 to 2022-12-31. The Sentinel-2 satellites, Sentinel-2A and -2B were launched in 2015 and 2017 respectively and use multispectral instruments to collect reflected sunlight over thirteen spectral bands. Part of the Sentinel-2 mission is to contribute to land monitoring and climate change studies (ESA, n.d, a). The Sentinel-2 data was extracted from Google Earth Engine (GEE). GEE provides an effective way of extracting larger sets of data, rather than collecting image by image from open-source databases. The image collection used, *"Harmonized Sentinel-2 MSI: Multispectral Instrument, Level-1C"* contains images from 2015-06-23 to the present, with a temporal resolution of five days around the equator and two to three days around the latitude of Sweden, using images from both satellites involved in the European Space Agency (ESA) Sentinel-2 mission (Harrie, 2020). The spatial resolution varies between the thirteen spectral bands available, but for the bands used in this study the spatial resolution is 10x10 or 20x20 meters (ESA, n.d, a).

## 4.2 Groundwater data

The county specific groundwater data originates from SGU. The data collection came as an CSV.file containing measurements from all observation wells within the county of Kalmar, spanning over a variation of time periods and with varying temporal resolution. More specifically the data used for this study involves groundwater levels in m.b.g.s, the name of each measurement station, the time period during which data has been collected and the coordinates for each site. Further, the data collected also covers attributes such as aquifer type, soil type, whether it's an open or closed aquifer and its topographic location.

## 4.3Additional data

Within the subject of groundwater studies, key features to analyze are related to vegetation and other land cover types, soil type and depth as well as topography. Additional data used for comparison and further analysis regarding these features that affect correlation between groundwater levels and vegetation indices, NDVI and NDWI, are presented as follows. Land cover data from Nationella marktäckedata developed by Swedish Environmental Protection Agency (Naturvårdsverket, 2018) was used in order to display different land cover extent in

each focus area. Each focus sample with respective land cover can be found in figure 23, within the results. Additionally, elevation data provided by Lantmäteriet (Lantmäteriet, n.d) was used in order to capture topography and surface flow models, created by Scalgo (Scalgo, n.d). Lastly, data over mapped groundwater aquifers provided by SGU (2022) have been used.

# 5. Method

The study aimed at exploring potential correlation between vegetation indices and groundwater level measurements for 57 sub-samples derived from measurement stations in the county of Kalmar. To each measurement station a one-kilometer buffer was applied, creating a sub-sample area of 3,14 $km^2$.

The time series format used for the three main variables, NDVI, NDWI and groundwater levels, was chosen to examine the progression of each variable over time. Applying analysis of time series data over a set period demands certain qualities of the sample data. For this study, biweekly maximum values for targeted indices over a seven-and-a-half-year period have been used. This further demanded high temporal resolution data, which was targeted through the Sentinel-2 image collection, with five days between each image at the equator and two to three around Sweden's longitude, using both satellites (ESA, n.d, a; Harrie, 2020). However, the temporal resolution was not the only aspect deciding data availability. Image quality can vary depending on cloud cover and other weather conditions. Some images are rendered useless due to high cloud cover. To work around this, while not a perfect solution due to the modification of actual values, interpolation was used to enable collection of time series data. Interpolation is the process of creating estimated values based on adjacent data. Images containing empty pixels, as a result of the cloud mask removing values from the pixels visualizing clouds, can further be filled with data from measurements correlated to the incomplete image (Harrie, 2020).

Found correlation between vegetation indices and groundwater measurements was further examined in relation to geospatial features vegetation, land cover, soil type, topographic location, and median groundwater table depth as an indicator for depth of aquifer. This further examination was applied to ten focus samples.

## 5.1 Collecting and processing remote sensing data

For this study, Sentinel-2 level-1C data was extracted and processed in Google Earth Engine: Code Editor. Code Editor is an integrated development environment that uses JavaScript for geospatial data processing (Google Earth Engine, 2021). The workflow used is presented in figure 5. The original JavaScript used was created by Ujval Gandhi (2021).

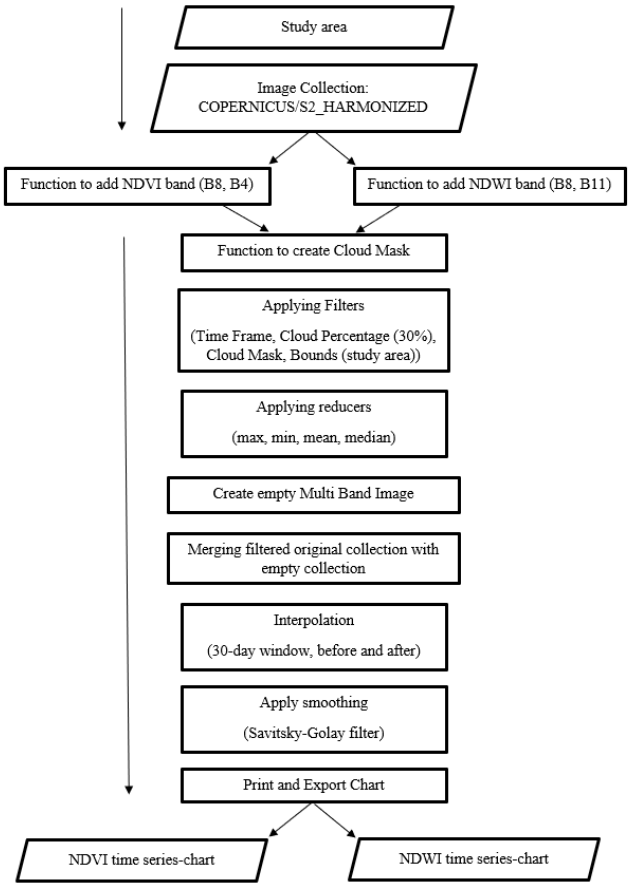


*Figure 5, Workflow describing data remote sensing data processing and collection. The process was performed using JavaScript and each square within the workflow describes a section of the script.*

The script merges a filtered original image collection containing Sentinel-2 level-1C images, with an empty collection. This merged image collection is then targeted for interpolation. Before that some presets were applied which targeted images within the original image collection. First a function that creates new separate bands for NDVI (Eq.1) and NDWI (Eq.3) as well as applying these to each image within the original images collection. Then a function that defines the cloud mask used. The cloud mask utilizes a separate cloud mask band (QA60) that was provided by the original image collection. The cloud mask band in question is in turn based on spectral values of specifically the blue band, band B10 and one of the SWIR bands which all three can be used to identify dense clouds and cirrus (ESA, n.d, b). Additionally, filters used within this script targeted images based on a specific time frame (2015-06-23 to 2022-12-31), the sub-samples area, certain cloud cover percentage (30%), as well as the pre-defined cloud mask. Lastly reducers were defined, these extract a single pixel value from all the pixel values found within each 3,14 $km^2$ sample, based on a certain criterion. Reducers for max, min, mean, and median values were added, however the study only used the max values. The reducers extracted values from the final interpolated images.

The interpolation required knowing the time of acquisition of each image, this in a millisecond timestamp format. Used interpolation method further acquired pixel values from the first adequate image both before and after the time of acquisition of the image with the missing pixels. If for example the before image would be closer adjacent in time to the timestamp of the original image than the after image is, the before image would have greater influence on the fill than the after image. This was achieved through a time ratio formula (Eq.5), where t = interpolation timestamp, t1 = before image timestamp, t2 = after image timestamp.

$$TimeRatio = (t - t1) / (t2 - t1) \quad \text{(Harrie, 2020)} \qquad \text{Eq. 5}$$

The complete interpolation formula used (Eq.6) describes a linear interpolation, where y = interpolated image, y1 = before image, y2 = after image.

$$y = y1 + (y2 - y1) * ((t - t1) / (t2 - t1)) \quad \text{(Harrie, 2020)} \qquad \text{Eq. 6}$$

Further, in this case the interpolation was restricted to 30 days before and after with the aim of maintaining seasonal relevance of the data, while still being able to find images that pass the filters set.

The last step of the time series processing involved smoothing. Smoothing was used to compute local averages to the data in order to reduce some of the noise that might occur within digitized data. In this case, the Savitzky-Golay filter was used, which performs moving window averaging over the time series data, with a three-point window. Using moving window averaging allows for the variation in a time series to be taken into consideration instead of using a constant averaging (Press, et al. 2007).

## 5.2 Spearman's correlation coefficient and linear regression analysis

In the case of this study, positive correlation indicates a high or low vegetation index value correlating with a low or high groundwater level. This is explained by the groundwater levels being presented as meters below ground, hence a high groundwater measurement value represents a lower actual groundwater level than a low groundwater measurement value does. Negative correlation indicates that a high vegetation index value correlates with a low groundwater measurement value, hence a high groundwater level. It could also indicate that a low vegetation index value correlates with a high groundwater measurement value, hence a low groundwater surface level.

### 5.2.1 Calculating Spearman's correlation coefficient

Spearman's Correlation Coefficient ($r_s$) was calculated for each separate indices in relation to groundwater levels for each measurement station, as part of a larger Python script. This script used the output CSV-files from the above-described JavaScript, containing time series with each index for each measurement station. It also used the site-specific groundwater data, separated to a CSV-file per measurement station. The script mainly used the pandas and NumPy plugins.

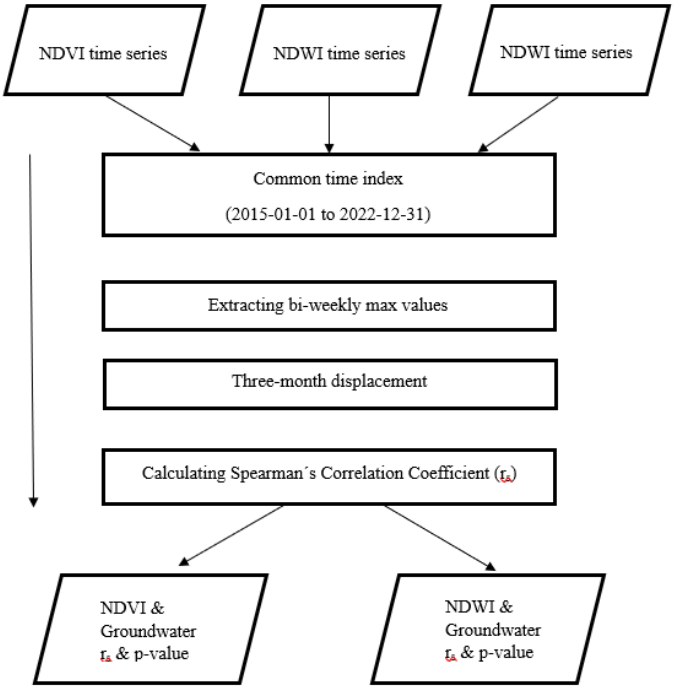


*Figure 6, Workflow describing process of calculating groundwater and index correlation through Python scripts*

Before reaching calculation of $r_s$ the input CSV-files needed further processing. First step involved changing the CSV format of each file to XLSX, the panda's plugin being more compatible with XLSX-files. Then values collected within each file were applied to a common timeline stretching over each day between 2015-01-01 to 2022-12-31. This was part of preparation for calculating the biweekly max values of each variable. Calculating biweekly max values helped narrow the data set down, making it more comprehensive as well as aligning the values of each variable at common dates, which was a necessity for calculating $r_s$. Using a biweekly basis also ensured that potential seasonal variations are maintained within the data set. Adding to this, the temporal resolution of Sentinel 2 data is as mentioned two to three days around Sweden's latitude, when using imagery from both satellites. A weekly basis would risk being dependent on single images while a biweekly basis ensures at least two often three images per two weeks stretch, depending on cloud cover.

Post calculating biweekly max values for each variable related to each measurement station, $r_s$ was calculated. The process of groundwater recharge is known to be delayed, as described in section 2.1. In order to account for the delay each index's time series was displaced within three

months before the groundwater levels timeline. Several time spans for time-lag were tested, ranging up to 6 months with a monthly difference for each try. Before three-month time-lag there was an increase in correlation strength over most measurement stations and their respective variables. After three months there was a decrease, hence the three-month time-lag output was most appropriate for further analysis.

From here ten focus samples were chosen for further analysis. The selection process was based on correlation values, aiming to select stations with a variation of correlation strength, as well as different time spans, geographical location and other geospatial features.

### 5.2.2 Performing linear regression analysis

The univariate linear regression analysis was performed in Excel, targeting each focus sample and their respective bi-weekly max variable values, structure in time series. Within a scatter plot, groundwater levels were placed as the dependent variable at the Y-axis and index values were placed as the independent variable at the X-axis. The full extent of the time series was used when performing the analysis. Output $R^2$-values were collected for each focus area. It is important to note that the statistical analysis, both correlation and regression analysis, only utilized data at dates that were common to both NDVI/NDWI and groundwater level measurements.

## 5.3 Comparison of time series and found correlations to geographical features

As the found correlations between groundwater levels and NDVI, NDWI varied a lot between stations (see results section), it was further explored which factors, such as vegetation, land cover, soil type, topography and groundwater level depth have an impact on the strength and significance of found correlations. The goal was also to determine whether the statistical outcomes follow specific spatial patterns that can be related to geographical features. The performed comparison involved ten focus samples.

| Station | Soil type | Topografic location | Levels observed (m.b.g.s) | Time span |
|---|---|---|---|---|
| *Dörby 38* | *Sand* | *Intermediate area* | *4.14 - 5.7* | *4/18/2019 - 12/22/2022* |
| *Emmaboda 21* | *Gravel or coarser* | *Intermediate area* | *1.3 - 2.25* | *5/2/2019 - 12/22/2022* |
| *Fliseryd 1* | *Gravel or coarser* | *Discharge area* | *1.32 - 2.41* | *5/2/2019 - 12/22/2022* |
| *Forshult 67* | *Sand* | *Intermediate area* | *0.83 - 1.7* | *10/29/2020 - 12/22/2022* |
| *Läckeby 1* | *Moraine* | *Recharge area* | *4.35 - 6.24* | *5/16/2019 - 9/17/2020* |
| *Påskallavik 2* | *Gravel or coarser* | *Recharge area* | *0.86 - 2.19* | *5/16/2019 -12/22/2022* |
| *Rälla 36* | *Sand* | *Intermediate area* | *11.89 - 14.18* | *1/1/2015 - 12/22/2022** |
| *Trekanten 63* | *Sand* | *Intermediate area* | *0.99 - 1.94* | *10/29/2020 - 12/22/2022* |
| *Vimmerby 55* | *Sand* | *Recharge area* | *4.44 - 4.94* | *1/1/2015 - 11/10/2022* |
| *Vimmerby 104* | *Moraine* | *Recharge area* | *7.06-8.09* | *1/1/2015 - 12/22/2022** |

*Table 2, Focus samples with respective geospatial features considered.*

### 5.3.1 Vegetation index maps

NDVI and NDWI images were collected for March, July and December for each year within the time series to examine where time series index values were derived from. Vegetation indices images were based on Sentinel 2 level 1C data extracted from GEE and visualized in QGIS before visual interpretation. Months were chosen to capture seasonal variety. Due to the number of images collected to provide each stations time series, this was not repeated for each stations respective images.

### 5.3.2 Land cover maps

Land cover data was mapped to create a geospatial context for each focus area. This aimed to provide an indication of the extent of certain land covers to see how well the collected index maximum values represent the overall study area. Nationella marktäckesdata, from Naturvårdsverket, was used. Certain forest land cover classes were merged during visualization in QGIS to make the maps more comprehensive. These cover classes described variations within coniferous and deciduous forests but were merged to create single deciduous and coniferous forest classes respectively.

### 5.3.3 Map exploring topography

One sub-sample (Forshult 67) was chosen for closer study to examine the effect if above surface topography on vegetation health as measured by NDVI from April to September 2021. The area was chosen based on its shallow groundwater levels and the above surface topography ranging from around 40 to 55 meters above sea level, directly above the aquifer. NDVI images were collected April, May, June, July, August and September 2021 in order to examine the effect of topography on vegetation health when groundwater levels are low, and potential dependence of groundwater higher.

### 5.4. Using averaged correlation strength

In order to further examine potential patterns found within the statistical outcome, averaged correlation strength between index and groundwater levels were compared to soil type, topographic location and median groundwater depth. In excel, soil type and topographic location were used as classes examining how the strength of correlation over all focus samples, for each index in relation to groundwater levels, was divided over different soil types and topographic locations. The classes were gravel or coarser, sand and moraine regarding soil type and intermediate, recharge and discharge areas regarding topographic location. These are the classes used in the collected data from SGU. The median groundwater depth was calculated and used as responsive variable in relation to averaged correlation strength between index and groundwater levels as dependent variable, to explore their linear regression.

# 6. Results

## 6.1 Correlation and regression outcomes

Results from Spearman's correlation coefficient show varying $r_s$-values between the two vegetation indices and groundwater levels, for all the 57 sub-samples. Results from correlation analysis without the time-lag is displayed in table 3.

| | | NDVI | NDWI | | | NDVI | NDWI |
|---|---|---|---|---|---|---|---|
| | **Station** | **$r_s$** | **$r_s$** | | **Station** | **$r_s$** | **$r_s$** |
| **1** | *Äleklinta 17* | *0,478* | *-0,569* | **30** | *Påskallavik 1* | *0,062* | *0,080* |
| **2** | *Björnhult 68* | *-0,060* | *0,070* | **31** | *Påskallavik 2* | *0,181* | *0,088* |
| **3** | *Blomstermåla 1* | *-0,027* | *0,227* | **32** | *Rälla 36* | *-0,345* | *0,313* |
| **4** | *Böda 7* | *-0,193* | *0,329* | **33** | *Silverdalen 57* | *0,209* | *-0,407* |
| **5** | *Böda 9* | *-0,117* | *0,387* | **34** | *Sjöstorp 32* | *-0,017* | *0,077* |
| **6** | *Böda 34* | *0,122* | *-0,092* | **35** | *Sjöstorp 33* | *-0,004* | *0,079* |
| **7** | *Böda 35* | *-0,248* | *0,062* | **36** | *Storebro 61* | *0,336* | *-0,465* |
| **8** | *Dörby 38* | *0,212* | *-0,217* | **37** | *Trekanten 63* | *0,627* | *-0,003* |
| **9** | *Dörby 39* | *-0,199* | *0,066* | **38** | *Vermagasinet 60* | *0,174* | *-0,263* |
| **10** | *Emmaboda 1* | *0,392* | *-0,152* | **39** | *Vimmerby 101* | *0,395* | *-0,137* |
| **11** | *Emmaboda 3* | *0,418* | *-0,291* | **40** | *Vimmerby 102* | *0,460* | *-0,241* |
| **12** | *Emmaboda 21* | *0,485* | *0,028* | **41** | *Vimmerby 103* | *0,536* | *-0,302* |
| **13** | *Fliseryd 1* | *0,306* | *-0,367* | **42** | *Vimmerby 104* | *0,244* | *-0,110* |
| **14** | *Forshult 67* | *0,382* | *-0,310* | **43** | *Vimmerby 105* | *0,541* | *-0,345* |
| **15** | *Gamleby 62* | *0,532* | *-0,146* | **44** | *Vimmerby 106* | *0,550* | *-0,295* |
| **16** | *Gosjökulla 64* | *0,055* | *-0,139* | **45** | *Vimmerby 12* | *-0,142* | *0,175* |
| **17** | *Högrum 19* | *-0,007* | *-0,033* | **46** | *Vimmerby 13* | *0,152* | *-0,061* |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| **18** | *Hultsfred 65* | *-0,016* | *-0,074* | **47** | *Vimmerby 14* | *0,006* | *0,177* |
| **19** | *Jernforsen 59* | *0,139* | *-0,212* | **48** | *Vimmerby 15* | *0,190* | *-0,080* |
| **20** | *Kleva 21* | *0,609* | *-0,424* | **49** | *Vimmerby 18* | *0,144* | *-0,001* |
| **21** | *Köpingsvik 31* | *-0,058* | *0,146* | **50** | *Vimmerby 41* | *0,425* | *-0,028* |
| **22** | *Läckeby 1* | *-0,195* | *-0,380* | **51** | *Vimmerby 42* | *0,382* | *-0,034* |
| **23** | *Långrälla* | *-0,407* | *0,225* | **52** | *Vimmerby 51* | *-0,172* | *0,453* |
| **24** | *Laxemar 1* | *0,497* | *-0,486* | **53** | *Vimmerby 52* | *0,077* | *-0,068* |
| **25** | *Laxemar 2* | *0,475* | *-0,461* | **54** | *Vimmerby 53* | *0,162* | *-0,117* |
| **26** | *Laxemar 3* | *0,495* | *-0,452* | **55** | *Vimmerby 54* | *-0,273* | *0,501* |
| **27** | *Mörlunda 58* | *-0,009* | *-0,180* | **56** | *Vimmerby 55* | *0,476* | *-0,177* |
| **28** | *Nybro 1* | *0,475* | *-0,381* | **57** | *Vimmerby Hamra 66* | *0,222* | *-0,282* |
| **29** | *Nybro 2* | *0,475* | *-0,340* | | | | |

*Table 3, Results from Spearman's correlation for the 57 sub-samples, without time-lag*

The strength of the $r_s$-values increased over most measurement stations, when using a time-lag of three months for the vegetation indices. Results from correlation analysis with the three-month time-lag is displayed in table 4. Remaining results presented regard data with three-month time-lag.

| | | **NDVI** | | **NDWI** | |
|---|---|---|---|---|---|
| | **Station** | **$r_s$** | **p-value** | **$r_s$** | **p-value** |
| **1** | *Äleklinta 17* | *0,459* | *0,000* | *0,200* | *0,057* |
| **2** | *Björnhult 68* | *0,667* | *0,000* | *-0,442* | *0,001* |
| **3** | *Blomstermåla 1* | *0,443* | *0,000* | *-0,291* | *0,005* |
| **4** | *Böda 7* | *0,522* | *0,000* | *-0,184* | *0,015* |
| **5** | *Böda 9* | *0,441* | *0,000* | *-0,107* | *0,158* |
| **6** | *Böda 34* | *0,688* | *0,000* | *-0,451* | *0,000* |
| **7** | *Böda 35* | *0,583* | *0,000* | *-0,338* | *0,001* |
| **8** | *Dörby 38* | *0,706* | *0,000* | *-0,297* | *0,003* |
| **9** | *Dörby 39* | *0,681* | *0,000* | *-0,393* | *0,000* |
| **10** | *Emmaboda 1* | *0,461* | *0,000* | *-0,567* | *0,000* |
| **11** | *Emmaboda 3* | *0,352* | *0,001* | *-0,569* | *0,000* |
| **12** | *Emmaboda 21* | *0,424* | *0,000* | *-0,528* | *0,000* |
| **13** | *Fliseryd 1* | *0,555* | *0,000* | *-0,554* | *0,000* |
| **14** | *Forshult 67* | *0,696* | *0,000* | *-0,656* | *0,000* |
| **15** | *Gamleby 62* | *0,233* | *0,090* | *-0,669* | *0,000* |
| **16** | *Gosjökulla 64* | *0,672* | *0,000* | *-0,519* | *0,000* |
| **17** | *Högrum 19* | *0,445* | *0,000* | *-0,355* | *0,000* |
| **18** | *Hultsfred 65* | *0,671* | *0,000* | *-0,515* | *0,000* |
| **19** | *Jernforsen 59* | *0,525* | *0,000* | *-0,419* | *0,002* |
| **20** | *Kleva 21* | *0,266* | *0,004* | *-0,532* | *0,000* |
| **21** | *Köpingsvik 31* | *0,520* | *0,000* | *0,004* | *0,966* |
| **22** | *Läckeby 1* | *0,778* | *0,000* | *-0,306* | *0,069* |

| | | | | | |
|---|---|---|---|---|---|
| **23** | *Långrälla* | *0,689* | *0,004* | *-0,421* | *0,118* |
| **24** | *Laxemar 1* | *0,207* | *0,133* | *-0,533* | *0,000* |
| **25** | *Laxemar 2* | *0,259* | *0,061* | *-0,585* | *0,000* |
| **26** | *Laxemar 3* | *0,230* | *0,097* | *-0,589* | *0,000* |
| **27** | *Mörlunda 58* | *0,632* | *0,000* | *-0,485* | *0,000* |
| **28** | *Nybro 1* | *0,276* | *0,040* | *-0,470* | *0,000* |
| **29** | *Nybro 2* | *0,150* | *0,271* | *-0,510* | *0,000* |
| **30** | *Påskallavik 1* | *0,626* | *0,000* | *-0,638* | *0,000* |
| **31** | *Påskallavik 2* | *0,664* | *0,000* | *-0,633* | *0,000* |
| **32** | *Rälla 36* | *0,292* | *0,001* | *0,119* | *0,175* |
| **33** | *Silverdalen 57* | *0,339* | *0,013* | *-0,611* | *0,000* |
| **34** | *Sjöstorp 32* | *0,700* | *0,000* | *-0,525* | *0,000* |
| **35** | *Sjöstorp 33* | *0,702* | *0,000* | *-0,527* | *0,000* |
| **36** | *Storebro 61* | *0,337* | *0,014* | *-0,473* | *0,000* |
| **37** | *Trekanten 63* | *0,386* | *0,003* | *-0,187* | *0,167* |
| **38** | *Vermagasinet 60* | *0,559* | *0,000* | *-0,559* | *0,000* |
| **39** | *Vimmerby 101* | *0,258* | *0,129* | *-0,276* | *0,104* |
| **40** | *Vimmerby 102* | *0,268* | *0,009* | *-0,390* | *0,000* |
| **41** | *Vimmerby 103* | *0,299* | *0,004* | *-0,420* | *0,000* |
| **42** | *Vimmerby 104* | *0,146* | *0,171* | *0,082* | *0,440* |
| **43** | *Vimmerby 105* | *0,117* | *0,263* | *-0,327* | *0,001* |
| **44** | *Vimmerby 106* | *0,118* | *0,262* | *-0,412* | *0,000* |
| **45** | *Vimmerby 12* | *0,510* | *0,000* | *-0,322* | *0,001* |
| **46** | *Vimmerby 13* | *0,552* | *0,000* | *-0,420* | *0,000* |
| **47** | *Vimmerby 14* | *0,507* | *0,000* | *-0,423* | *0,000* |
| **48** | *Vimmerby 15* | *0,593* | *0,000* | *-0,432* | *0,000* |
| **49** | *Vimmerby 18* | *0,423* | *0,000* | *-0,355* | *0,000* |
| **50** | *Vimmerby 41* | *0,607* | *0,000* | *-0,474* | *0,000* |
| **51** | *Vimmerby 42* | *0,645* | *0,000* | *-0,498* | *0,000* |
| **52** | *Vimmerby 51* | *0,353* | *0,117* | *0,149* | *0,520* |
| **53** | *Vimmerby 52* | *0,141* | *0,542* | *0,353* | *0,116* |
| **54** | *Vimmerby 53* | *0,592* | *0,000* | *-0,515* | *0,000* |
| **55** | *Vimmerby 54* | *0,179* | *0,437* | *-0,082* | *0,724* |
| **56** | *Vimmerby 55* | *0,492* | *0,000* | *-0,545* | *0,000* |
| **57** | *Vimmerby Hamra 66* | *0,316* | *0,024* | *-0,506* | *0,000* |

*Table 4, Results from Spearman's correlation for the 57 sub-samples, with three-month time-lag.*

Using the progression of five of the total ten focus samples respective NDVI $r_s$-values, it is noticeable that not all peak at the same time-lag (see figure 7). Emmaboda 21 and Vimmerby 55 peak only after one-month time-lag, while Rälla 36 peak post five-month time-lag. Påskallavik 2 and Dörby 38 peak post three-month time-lag and hence display the more

common pattern over the largest amounts of the original 57 measurement stations. Looking at the NDWI and groundwater level $r_s$-values (figure 8), Rälla 36 peak post four-month time-lag and Dörby 38 post one-month time-lag. Emmaboda 21, Vimmerby 55 and Påskallavik 2 all peak post three-month time-lag.

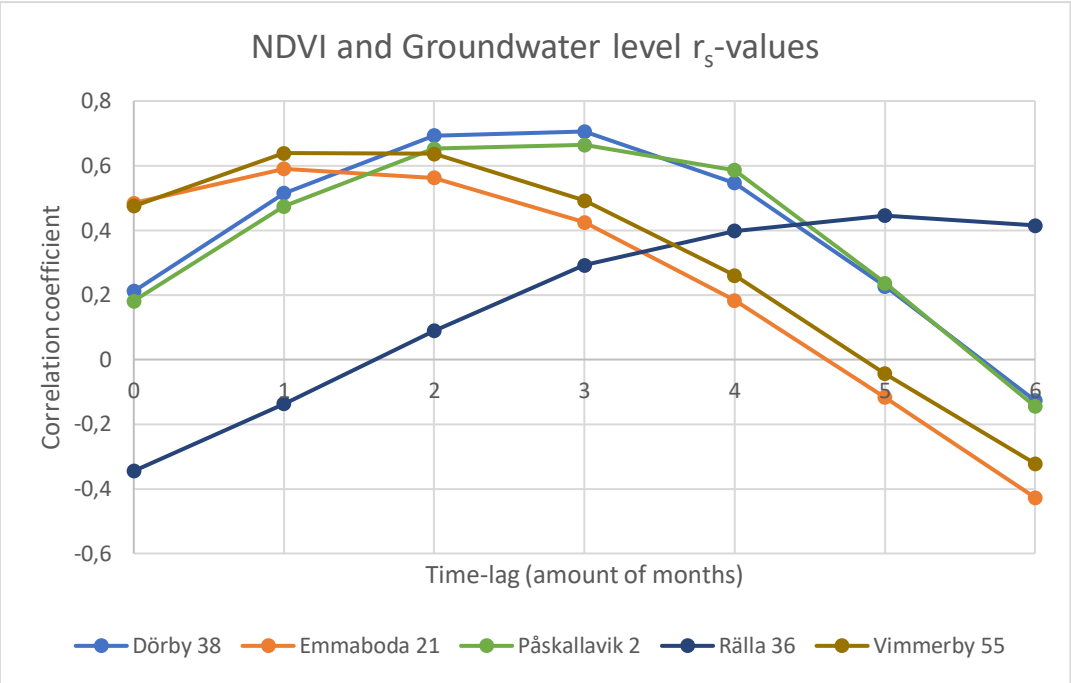


*Figure 7, Examples of NDVI and groundwater correlation coefficients post one to six months time-lag at five of the ten focus samples.*

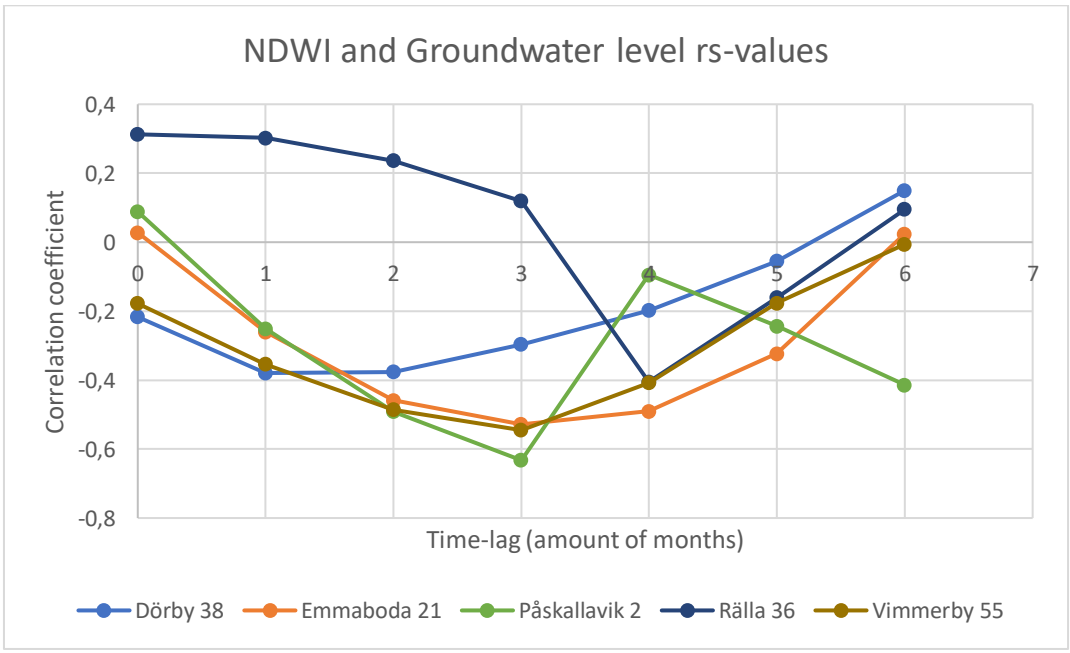


*Figure 8, Examples of NDWI and groundwater correlation coefficients post one to six months time-lag at five of the ten focus samples.*

Whether strong or weak, the $r_s$-values are always positive for NDVI and groundwater levels indicating that periods of higher NDVI values are followed by lower groundwater levels. The positive correlation values are due to groundwater measurements presented as m.b.g.s, hence higher values show lower groundwater levels. 16 measurement stations displayed in table 5, showed strong correlation ($r_s$ >= 0.6) between NDVI and groundwater levels, with p-values below 0.05 which indicates statistical significance. Most measurement stations show negative $r_s$ between NDWI and groundwater levels. Negative correlation indicates that periods of high vegetation moisture are often followed by increasing groundwater levels. The negative correlation values are due to higher groundwater levels being presented by lower values because of its unit. Five measurement stations, displayed in table 6, show strong correlation between NDWI and groundwater levels ($r_s$ <= -0.6) with p-values below 0.05, indicating statistical significance.

| | | NDVI | |
|---|---|---|---|
| | **Station** | **rs** | **p-value** |
| **2** | *Björnhult 68* | *0.667* | *0.000* |
| **6** | *Böda 34* | *0.688* | *0.000* |
| **8** | *Dörby 38* | *0.706* | *0.000* |
| **9** | *Dörby 39* | *0.681* | *0.000* |
| **14** | *Forshult 67* | *0.696* | *0.000* |
| **16** | *Gosjökulla 64* | *0.672* | *0.000* |
| **18** | *Hultsfred 65* | *0.671* | *0.000* |
| **22** | *Läckeby 1* | *0.778* | *0.000* |
| **23** | *Långrälla* | *0.689* | *0.004* |
| **27** | *Mörlunda 58* | *0.632* | *0.000* |
| **30** | *Påskallavik 1* | *0.626* | *0.000* |
| **31** | *Påskallavik 2* | *0.664* | *0.000* |
| **34** | *Sjöstorp 32* | *0.700* | *0.000* |
| **35** | *Sjöstorp 33* | *0.702* | *0.000* |
| **50** | *Vimmerby 41* | *0.607* | *0.000* |
| **51** | *Vimmerby 42* | *0.645* | *0.000* |

*Table 5, The 16 measurement stations showing $r_s$ >= 0.6 between NDVI and groundwater levels as well as their respective p-value.*

| | | NDWI | |
|---|---|---|---|
| | **Station** | **rs** | **p-value** |
| **14** | *Forshult 67* | *-0.656* | *0.000* |
| **15** | *Gamleby 62* | *-0.669* | *0.000* |
| **30** | *Påskallavik 1* | *-0.638* | *0.000* |
| **31** | *Påskallavik 2* | *-0.633* | *0.000* |
| **33** | *Silverdalen 57* | *-0.611* | *0.000* |

*Table 6, The five measurement stations showing $r_s$ <= -0.6 between NDWI and groundwater levels as well as their respective p-value.*

The spatial distribution of $r_s$-values over the 57 original measurement station are displayed in figure 9 and 10. Further, the spatial distribution of 16 measurement station showing $r_s$ >= 0.6 between NDVI and groundwater levels, is displayed in figure 11. The spatial distribution of the five measurement stations showing $r_s$ <= -0.6 between NDWI and groundwater levels are displayed in figure 12.

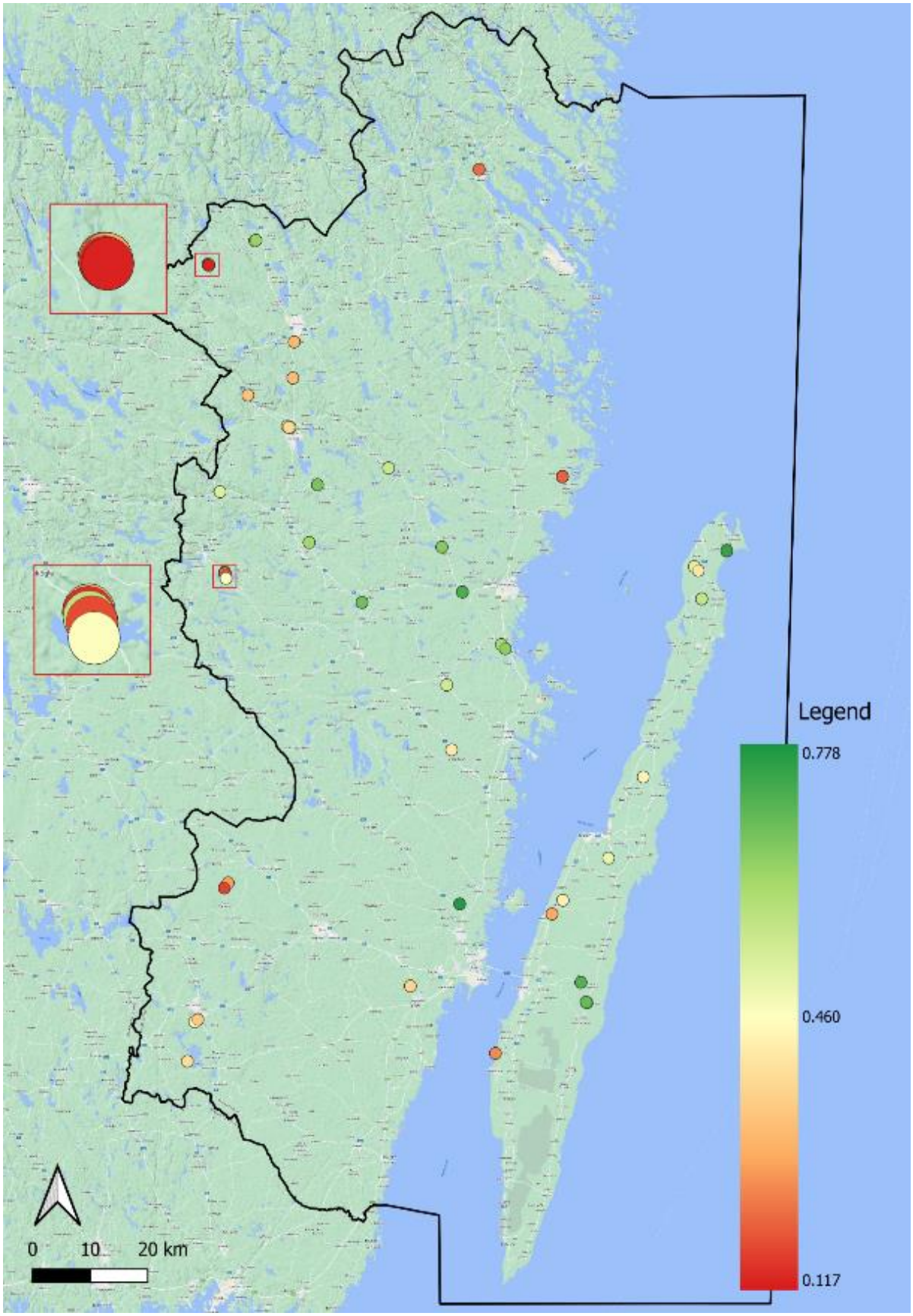


*Figure 9, The spatial distribution of rs-results for all 57 measurement stations, displaying NDVI and Groundwater level correlation. The legend describes NDVI and groundwater level $r_s$-values.*

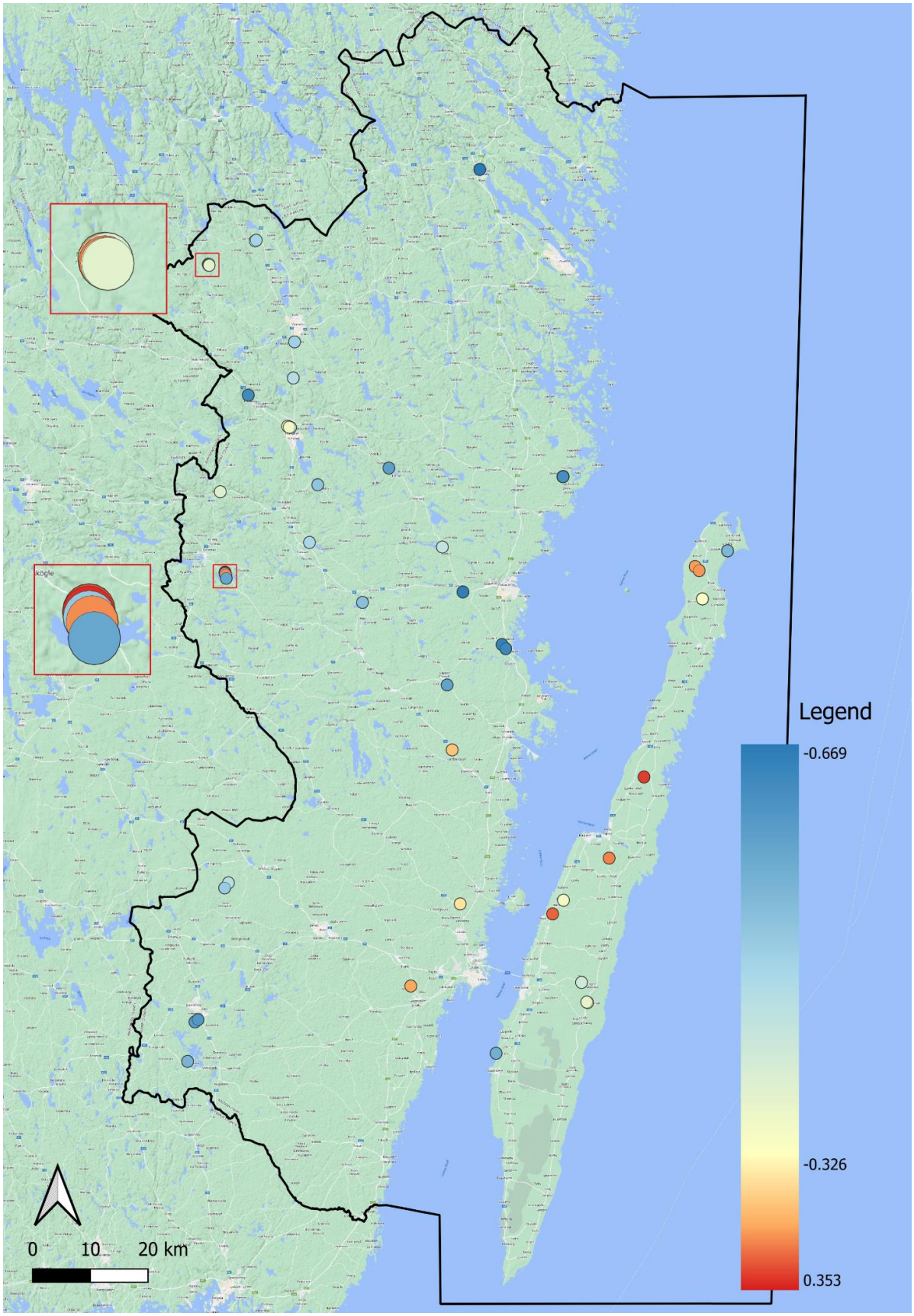


*Figure 10, The spatial distribution of rs-results for all 57 measurement stations, displaying NDVI and Groundwater level correlation. The legend describes NDWI and groundwater level $r_s$-values.*

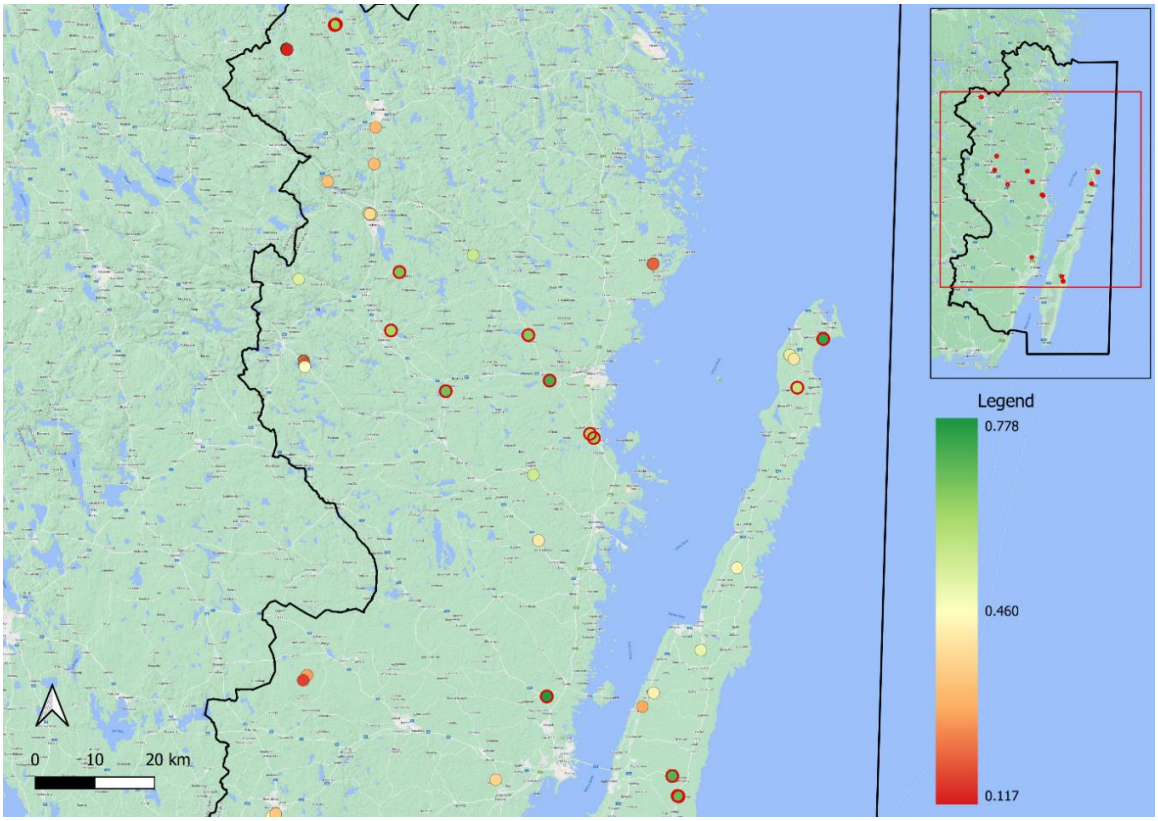


*Figure 11, The spatial distribution of the 16 of 57 measurement stations showing $r_s$ >= 0.6 between NDVI and groundwater levels. The legend describes NDVI and groundwater level $r_s$-values.*

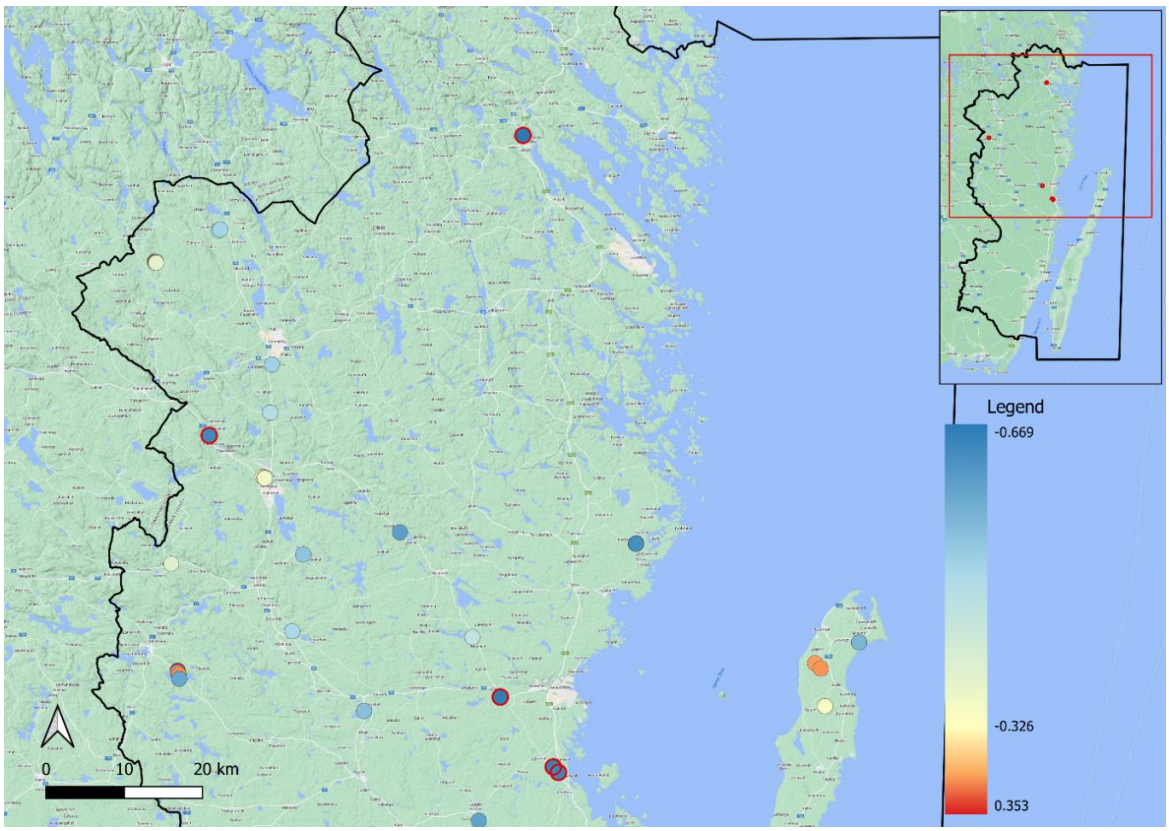


*Figure 12, The spatial distribution of the 5 of 57 measurement stations showing $r_s$ <= -0.6 between NDWI and groundwater levels. The legend describes NDWI and groundwater level $r_s$-values.*

As presented in section 5.3, ten focus samples were extracted from the original 57 measurement stations for further analysis. The overall strongest $r_s$ for these focus samples are found between NDVI and groundwater levels which show an average $r_s$ of 0.514 for the ten focus samples. NDWI and groundwater levels for the ten focus samples show an average $r_s$ of 0.391. All averaged values presented in the result section will be positive whether the correlation is positive or not. This is to represent the strength of correlation, averaging positive and negative values combined can lead to false indication of correlation strength. Calculating the average $r_s$ over the original 57 measurement stations the same pattern is found, NDVI provide strongest average $r_s$ (0.455), followed by NDWI (0.424).

| | NDVI | | NDWI | |
|---|---|---|---|---|
| **Station** | **$r_s$** | **p-value** | **$r_s$** | **p-value** |
| *Dörby 38* | *0,706* | *0,000* | *-0,297* | *0,003* |
| *Emmaboda 21* | *0,424* | *0,000* | *-0,528* | *0,000* |
| *Fliseryd 1* | *0,555* | *0,000* | *-0,554* | *0,000* |
| *Forshult 67* | *0,696* | *0,000* | *-0,656* | *0,000* |
| *Läckeby 1* | *0,778* | *0,000* | *-0,306* | *0,069* |
| *Påskallavik 2* | *0,664* | *0,000* | *-0,633* | *0,000* |
| *Rälla 36* | *0,292* | *0,001* | *0,119* | *0,175* |
| *Trekanten 63* | *0,386* | *0,003* | *-0,187* | *0,167* |
| *Vimmerby 55* | *0,492* | *0,000* | *-0,545* | *0,000* |
| *Vimmerby 104* | *0,146* | *0,171* | *0,082* | *0,440* |

*Table 7, Spearman's correlation coefficient results for each focus area, with related p-value.*

The results from the Spearman's correlation as well as linear regression analysis are presented in table 7 and 8. Presented $R^2$-values indicate that there is a low to moderate percent chance that variations in dependent variable groundwater levels are a result of variation in independent variable NDVI and NDWI. The exception being $R^2$ for NDVI at Läckeby 1. At this station a relatively strong linear relationship was seen for said vegetation index. This indicates that the chance of groundwater levels being affected by healthy vegetation cover at Läckeby 1 is relatively strong. Examples of regressions found are seen in figures 13 and 14, as well as to figures 35 to 42 presented in the appendix.

| | NDVI | NDWI |
|---|---|---|
| **Station** | **$R^2$** | **$R^2$** |
| *Dörby 38* | *0,473* | *0,132* |
| *Emmaboda 21* | *0,17* | *0,233* |
| *Fliseryd 1* | *0,257* | *0,271* |
| *Forshult 67* | *0,401* | *0,31* |
| *Läckeby 1* | *0,563* | *0,143* |
| *Påskallavik 2* | *0,448* | *0,367* |
| *Rälla 36* | *0,069* | *0,002* |
| *Trekanten 63* | *0,131* | *0,016* |
| *Vimmerby 55* | *0,174* | *0,358* |
| *Vimmerby 104* | *0,002* | *0,004* |

*Table 8, regression analysis results for each focus sample and respective $R^2$-values.*

Looking closer at Läckeby 1, while based on few measurements compared to the rest of the measurement stations, this station had the strongest linear relationship between groundwater levels and NDVI (0.563). The same linear strength does not apply for NDWI (0.143) however. Strongest linear relationship between groundwater levels and NDWI was found at Påskallavik 2. Påskallavik 2 also showed relatively strong, in relation to all focus samples, linear relationship between groundwater levels and NDVI.

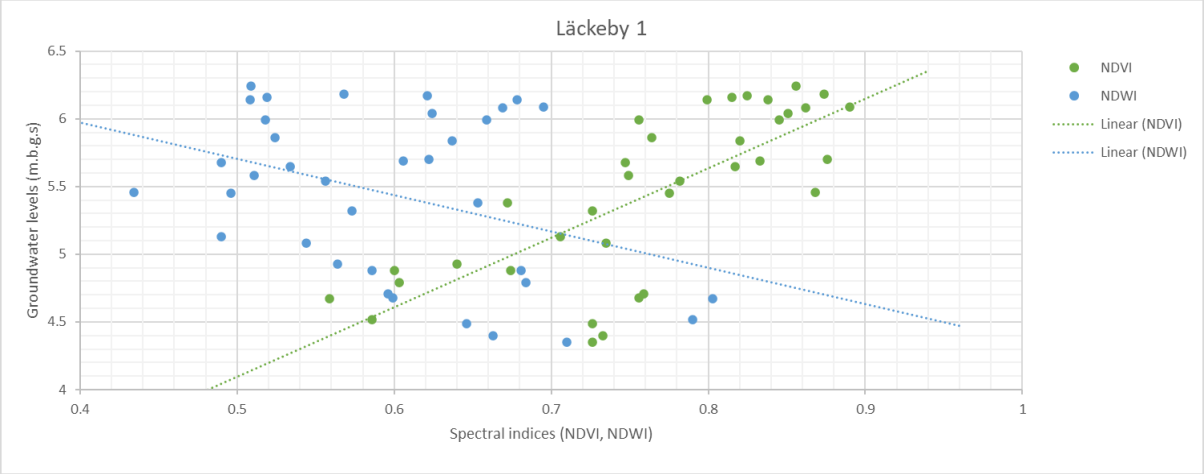


*Figure 13, Läckeby 1 regression analysis. $R^2$-values between Groundwater level and NDVI show 0.563 and NDWI 0.143.*

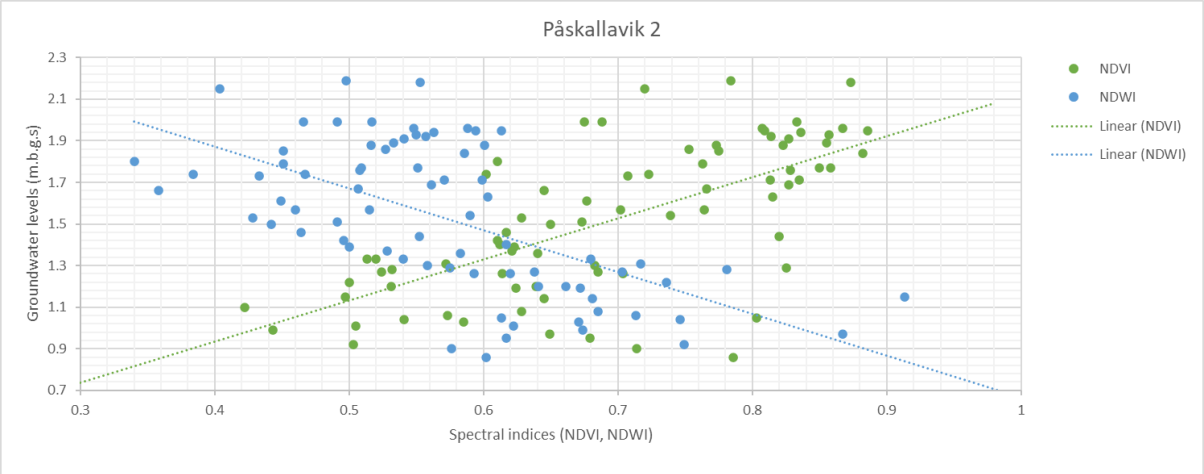


*Figure 14, Påskallavik 2 regression analysis. $R^2$-values between Groundwater level and NDVI show 0.448 and NDWI 0.367.*

## 6.2 Interpretation of time series

For observation, comparison and further interpretation, time series of NDVI, NDWI and groundwater levels in m.b.g.s have been assembled into shared graphs for each station within the ten focus samples (see figures 17 to 20 and 24, and figures 30 to 34 in the appendix).

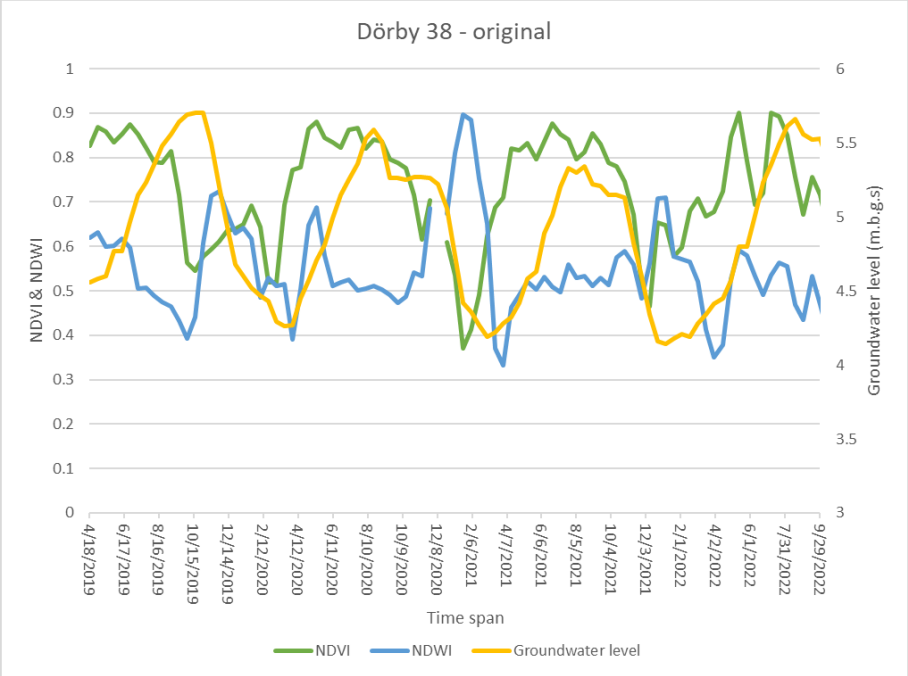


*Figure 15, Time series of NDVI, NDWI and groundwater level for Dörby 38*

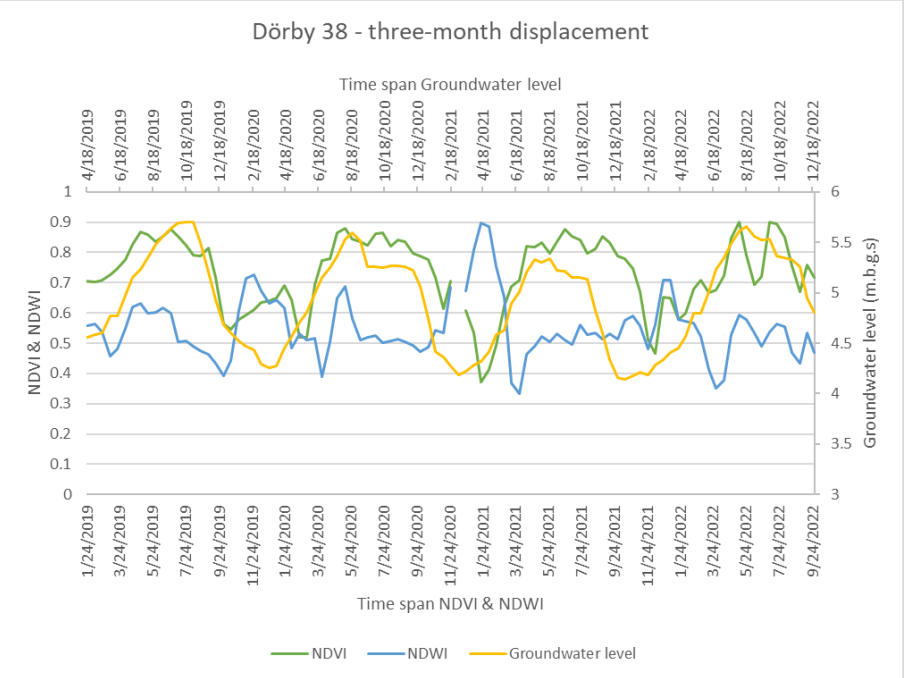


*Figure 16, Time series of NDVI, NDWI and groundwater level for Dörby 38, where groundwater level values are shifted by a period of three months relative to the measurements of NDVI, NDWI. The upper x-axis refers to groundwater levels.*

For most stations, the general pattern of the index values within their respective graph show that NDVI has started to increase around the month of March and is further maintained during the summer before declining around September and remaining low during fall and winter. NDWI time series shows an opposite pattern to NDVI with values starting to incline after

September when NDVI values decrease. The decline of NDWI values proceeds around February and March.

Looking at figure 17, 18 and 19, Dörby 38 presents an example where evident periods of high NDVI values are followed by low groundwater levels. The same general pattern is found at Vimmerby 55 but over a longer time span. Rälla 36, however, presents less evident alignment between periods of high NDVI values being followed by low groundwater level measurements.

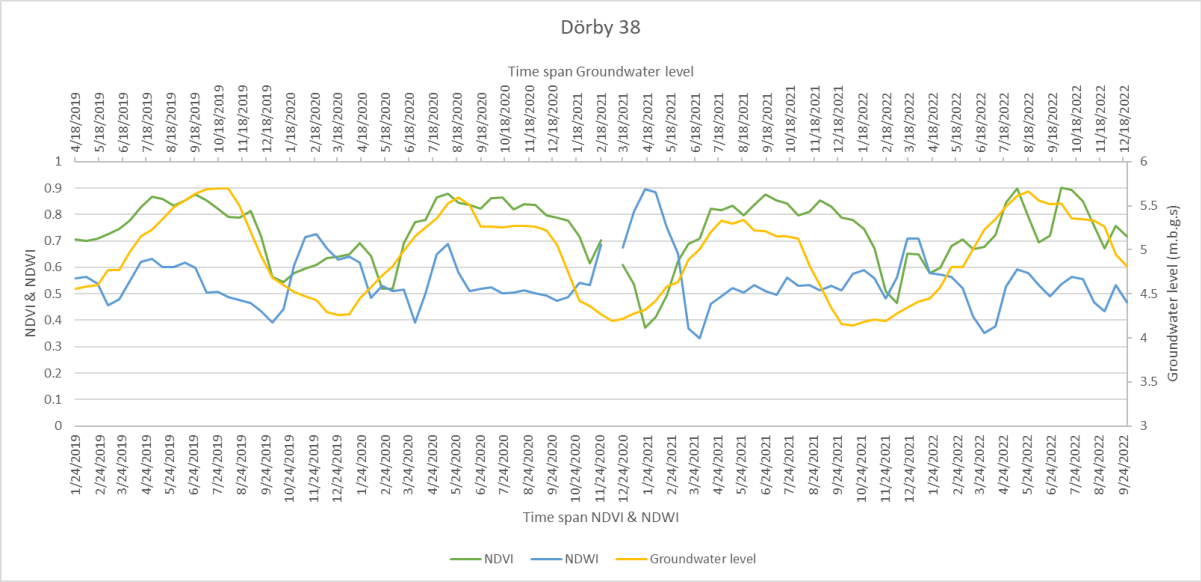


*Figure 17, Dörby 38 time serie, displaying NDVI and NDWI values related to the lower time span and groundwater level measurements related to the upper time span.*

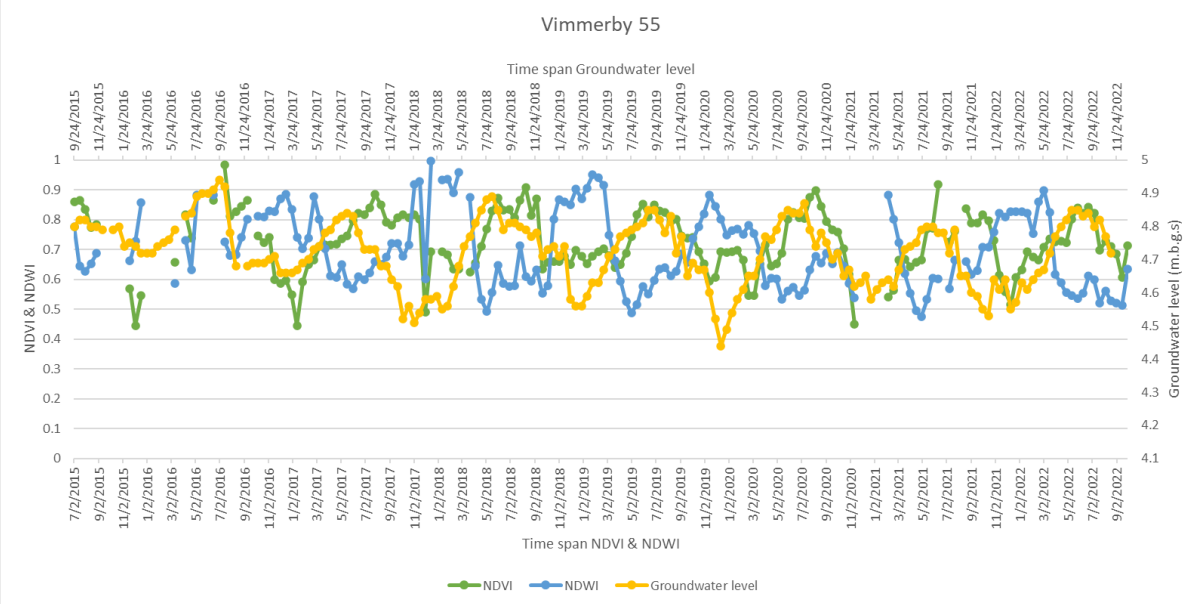


*Figure 18, Vimmerby 55 time serie, displaying NDVI and NDWI values related to the lower time span and groundwater level measurements related to the upper time span.*

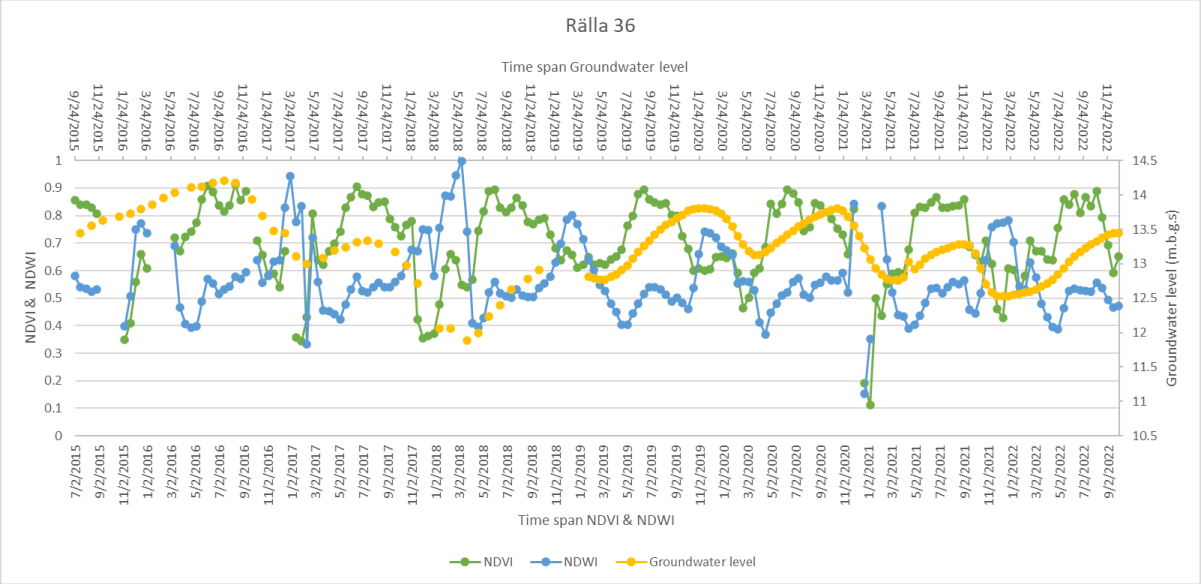


*Figure 19, Rälla 36 time serie, displaying NDVI and NDWI values related to the lower time span and groundwater level measurements related to the upper time span.*

Looking at figure 17 to 20, there are smaller peaks in NDWI during the summer months. These peaks, in relation to the overall lower NDWI values during summer season, indicate occurrences of intense precipitation and insufficient evapotranspiration resulting in higher vegetation moisture. An example of this is around June 2020 at Emmaboda 21. Also, seen at Dörby 38 a peak in NDWI around April and May 2020 is followed by a slight increase in groundwater level. This is one example of where slight groundwater recharge can be spotted during spring and summer months. However, this does not apply for all peaks in NDWI during summer seasons and doesn't occur at each station.

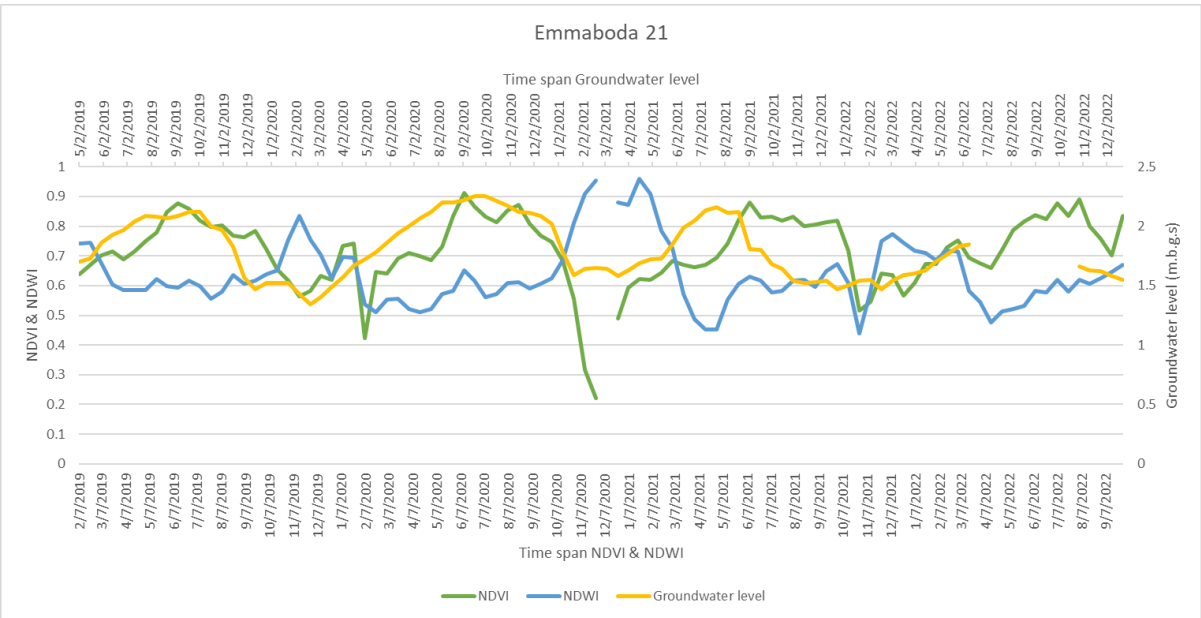


*Figure 20, Emmaboda 21 time serie, displaying NDVI and NDWI values related to the lower time span and groundwater level measurements related to the upper time span.*

## 6.3 Further interpretation of geospatial features and found correlations

### 6.3.1 Land cover

Each index value within the time series displays the maximum pixel value found within each sub-sample image on a two-week basis. Hence the index values displayed, deriving only from a single pixel, represent the sub-sample area to different extents depending on similarities between the maximum value pixel and remaining pixels within the sub-sample. This in turn is based on the land cover within the sub-sample as well as the season during which the image is taken.

High NDVI max values, such as those around 0.8, were found within pixels displaying coniferous and deciduous forest as well as healthy crops. When tapering off towards fall and winter, the NDVI max values remain relatively high, around 0.6. During this time max NDVI value pixels have been derived from patches of coniferous forests which maintain relatively high NDVI year-round. Examples of how the extent of different samples are represented by their index max value can be found in figure 21. During July 2021 at Dörby 38 NDVI values were more similar over a large land cover extent, involving forest as well as healthy crops. This, in comparison to March and December when fewer pixels align with the max NDVI values found within dense sections of coniferous forests. Påskallavik 2 shows an example of an area where the max NDVI values found align with remaining vegetation cover to a lesser extent even in July. Here, max NDVI values are found within pixels displaying mixed coniferous and deciduous forests. Pixels showing high NDVI values above 0.6 during March and December are very few but exist and hence disregards remaining vegetation cover to an extent (see figure 21).

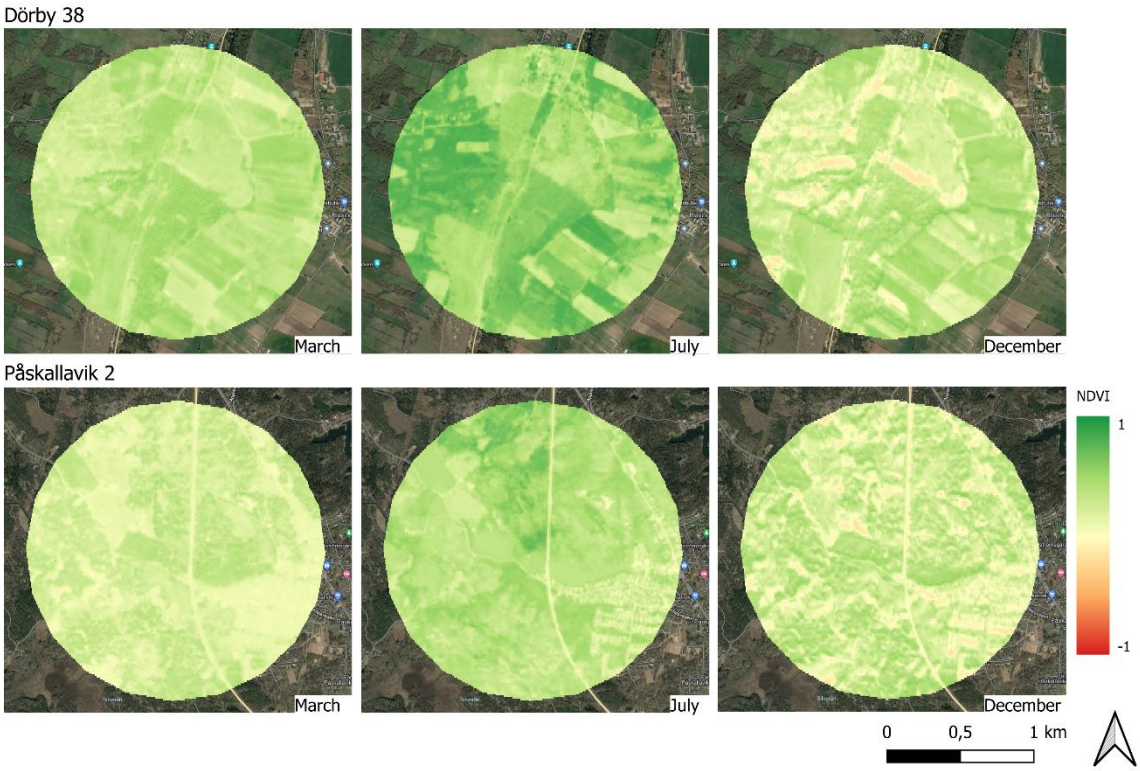


*Figure 21, NDVI maps over Dörby 38 and Påskallavik 2, presented to examine from what vegetation cover NDVI maxvalues have been derived. These examples show how different max values represent the full extent of the sub-samples to different extents.*

The high NDWI max values displayed in the time series can be explained by pixels with dense coniferous forests, maintaining high canopy cover during fall and winter, while also experiencing no water stress. Lower NDWI max values, found during the summer season are displayed by pixels with forest cover indicating low to no water stress. Max NDWI values during summer months can also be found at crop areas, indicating these being newly irrigated or not yet harvested, in comparison to other crop fields showing lower NDWI values. Of note is that while high vegetation shows low water stress during the summer season, lower vegetation might still suffer from moisture depletion (see figure 22).

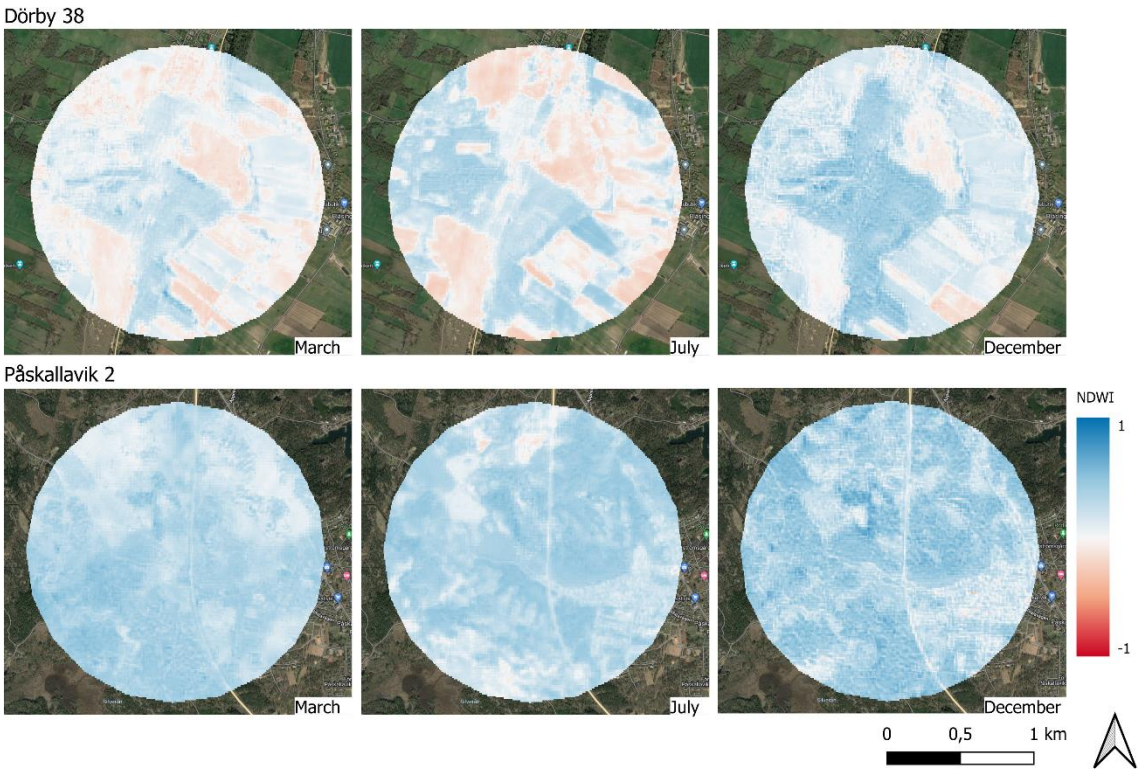


*Figure 22, NDWI maps over Dörby 38 and Påskallavik 2, presented to examine from what vegetation cover NDWI max values have been derived. These examples show how different max values represent the full extent of the sub-samples to different extents.*

Figure 23 shows that the land cover of the focus samples consists of a mix of forests, crop fields, populated areas, surface water bodies, which are all present to different degrees. Dörby 38, Forshult 67, Läckeby 1, and Påskallavik 2 being focus samples that show strong $r_s$ between NDVI and groundwater levels, differ in vegetation cover. The land cover of Påskallavik 2 and Forshult 67 mainly consist of coniferous forest of different density, Dörby 38 consist mainly of crop fields surrounding a section of coniferous forests, and Läckeby 1 while showing strongest $r_s$ as well as $R^2$ between groundwater levels and NDVI mainly consists of a populated area and crop fields with just a smaller section of coniferous forests and sections of deciduous trees present within the area.

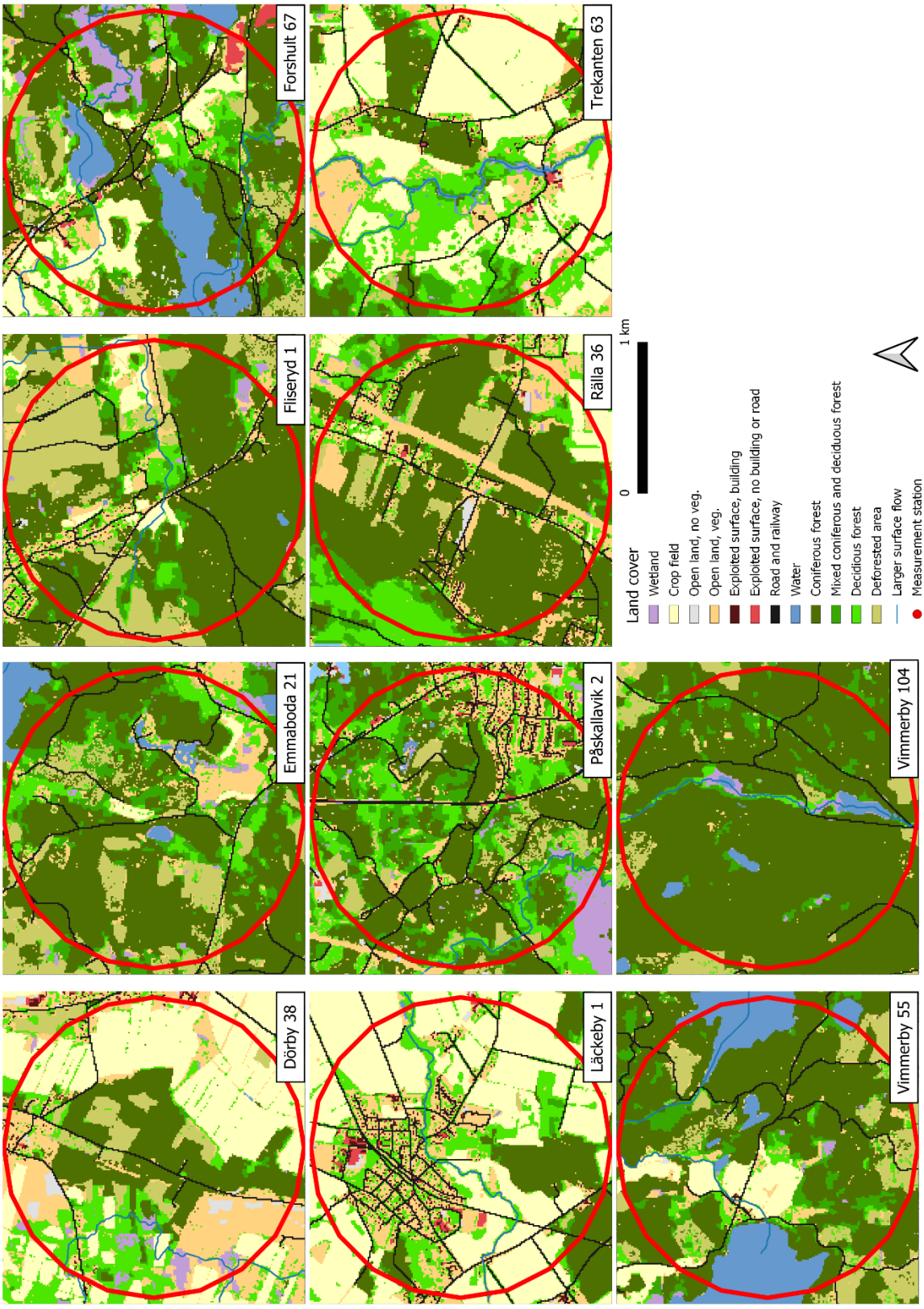


*Figure 23, Land cover maps for each focus sample based on Nationella Marktäckesdata (Naturvårdsverket, 2018). Each area is outlined by the red buffer, and the red center represents the measurement station.*

Strong $r_s$ between NDWI and groundwater levels are found within Forshult 67 and Påskallavik 2. Emmaboda 21, Fliseryd 1 and Vimmerby 55 show moderate $r_s$ being above 0.5. $R^2$ for the respective samples extend from 0.233-0.367. All these samples have the presence of surface water within the focus area. More specifically Emmaboda 21, Forshult 67 and Vimmerby 55 have larger water bodies near their measurement station. While Påskallvik 2 and Fliseryd 1 don't, both have larger surface flows percent within their area. However, Vimmerby 104 while having presence of several surface water bodies within the area, show both weak $r_s$ and $R^2$ between NDWI and groundwater levels.

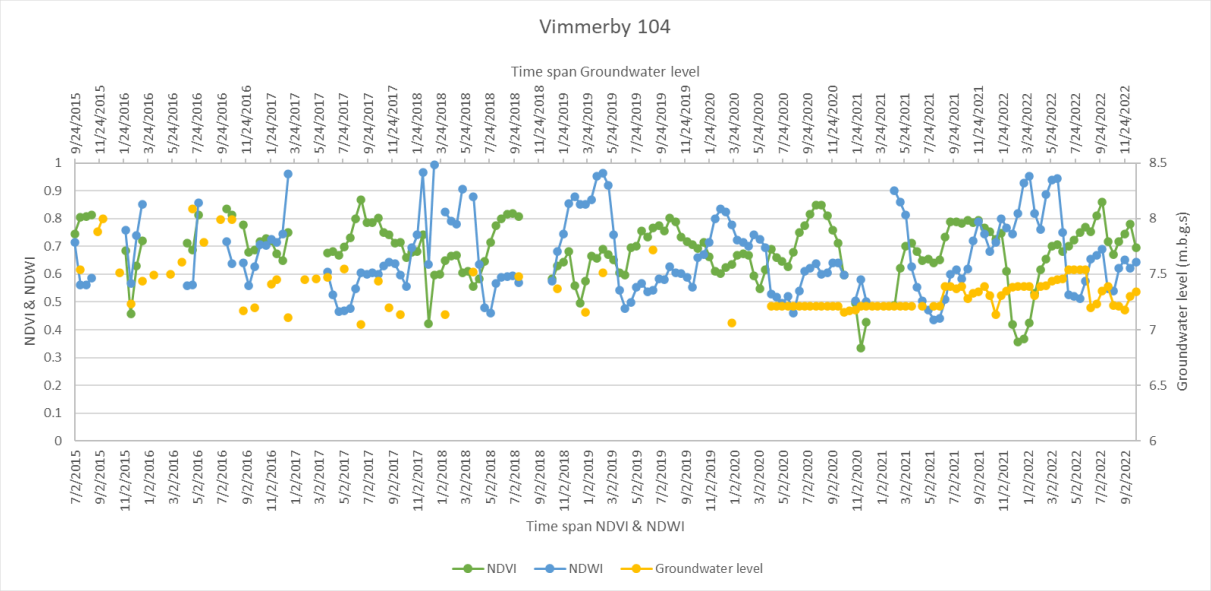


*Figure 24, Vimmerby 104 time serie, displaying NDVI and NDWI values related to the lower time span and groundwater level measurements related to the upper time span.*

### 6.3.2 Soil type and median groundwater levels

Looking at soil type in relation to correlation values, results are presented in figure 25. Strongest averaged $r_s$-values for NDVI and NDWI separately were found for aquifers within gravel or coarser soil type. Sand is placed thereafter and lastly moraine, which could indicate that soil type plays a large part in the correlation output.

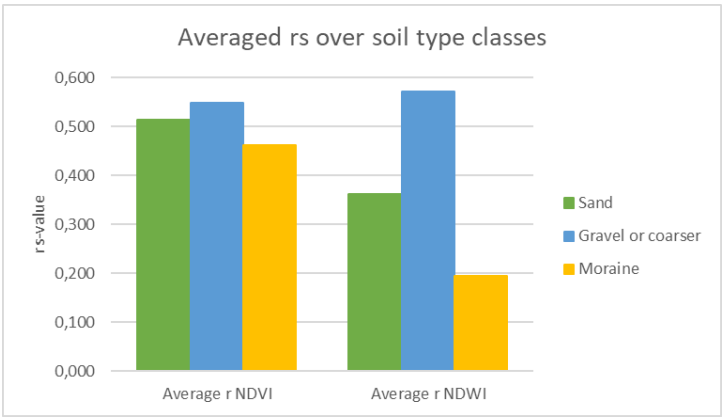


*Figure 25, Averaged correlation coefficients in relation to soil type. The Gravel or Coarser soil type class show strongest average rs-values. Then sand and lastly moraine.*

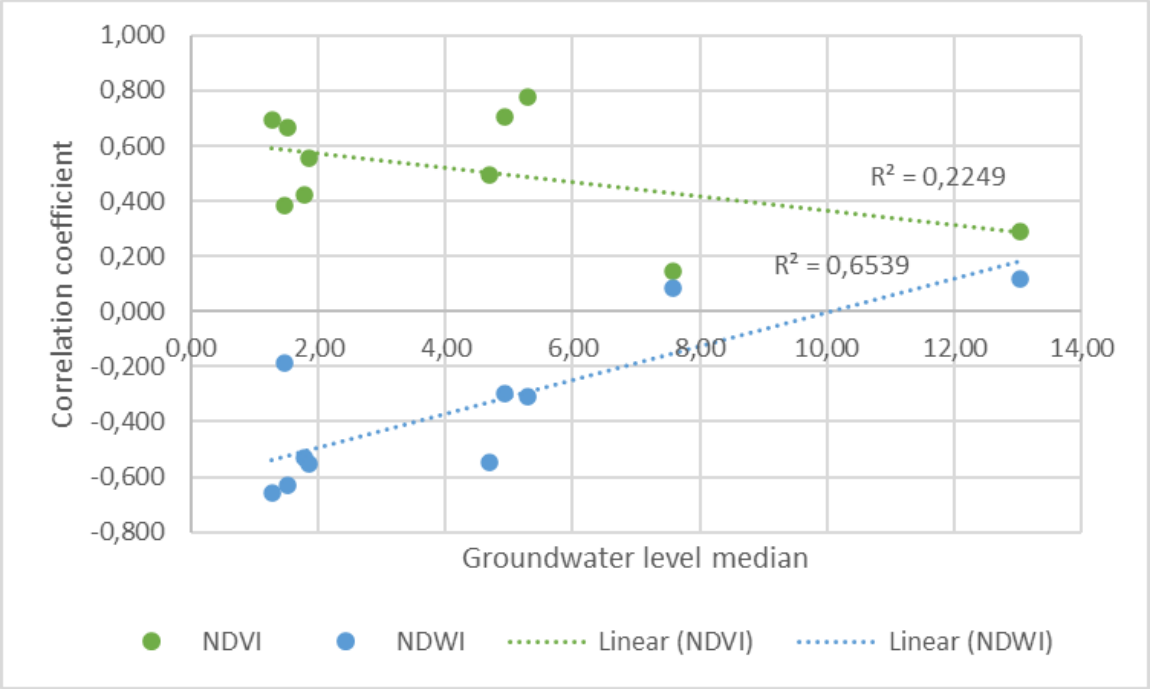


*Figure 26, Linear relationship between correlation coefficient as dependent variable in relation to median groundwater depth as response variable. Stronger linear relationship was found between NDWI correlation with groundwater and median groundwater depth.*

Comparing the median depth value for groundwater levels at each focus sample with the $r_s$-values for NDVI and NDWI, displayed in figure 26, reveals that no strong linear relationship between NDVI and groundwater level correlation coefficients and groundwater level median can be found. This indicates that in the case of these focus samples, median groundwater level depth did not play a vital part in the output correlation coefficient between NDVI and groundwater levels. Looking at NDWI, a much stronger linear relationship between correlation coefficient and groundwater median level is seen. This indicates that correlation between NDWI and groundwater levels are more dependent on median groundwater levels than NDVI.

| Station | GWL median | NDVI rs | NDWI rs |
|---|---|---|---|
| *Dörby 38* | *4.92* | *0.706* | *-0.297* |
| *Enmaboda 21* | *1.775* | *0.424* | *-0.528* |
| *Fliseryd 1* | *1.865* | *0.555* | *-0.554* |
| *Forshult 67* | *1.265* | *0.696* | *-0.656* |
| *Läckeby 1* | *5.295* | *0.778* | *-0.306* |
| *Påskallavik 2* | *1.525* | *0.664* | *-0.633* |
| *Rälla 36* | *13.035* | *0.292* | *0.119* |
| *Trekanten 63* | *1.465* | *0.386* | *-0.187* |
| *Vimmerby 13* | *5.315* | *0.552* | *-0.420* |
| *Vimmerby 55* | *4.69* | *0.492* | *-0.545* |
| *Vimmerby 104* | *7.575* | *0.146* | *0.082* |

*Table 9, Correlation coefficient in relation to median groundwater depth for each focus area.*

### 6.3.3 Topographic location

Figure 27 shows averaged $r_s$-values for each separate indices grouped based on topographic location. The discharge area class only contains $r_s$-values from one measurement station, Fliseryd 1, which show relatively strong correlation between NDWI and groundwater levels.

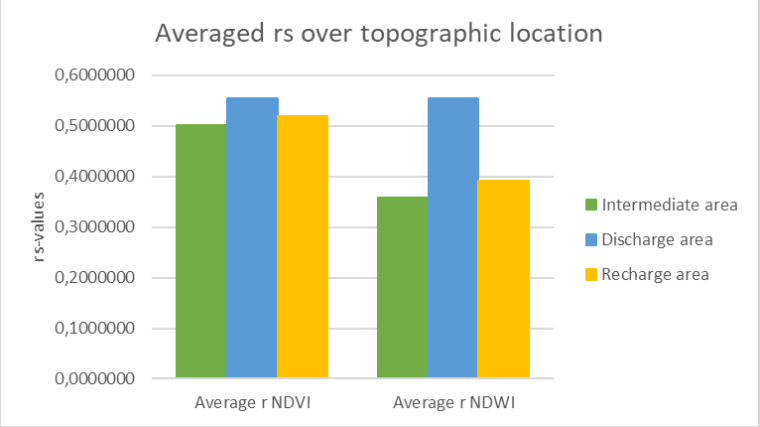


*Figure 27, Correlation coefficients in relation to soil depth. The discharge areas class show strongest average rs-values. Then recharge and lastly intermediate.*

The difference between $r_s$-values within the three topographic locations is larger in the case of NDWI correlating with groundwater. There is a pattern of stronger averaged rs-values deriving from recharge areas rather than intermediate areas. However, the strength of both is relatively similar.

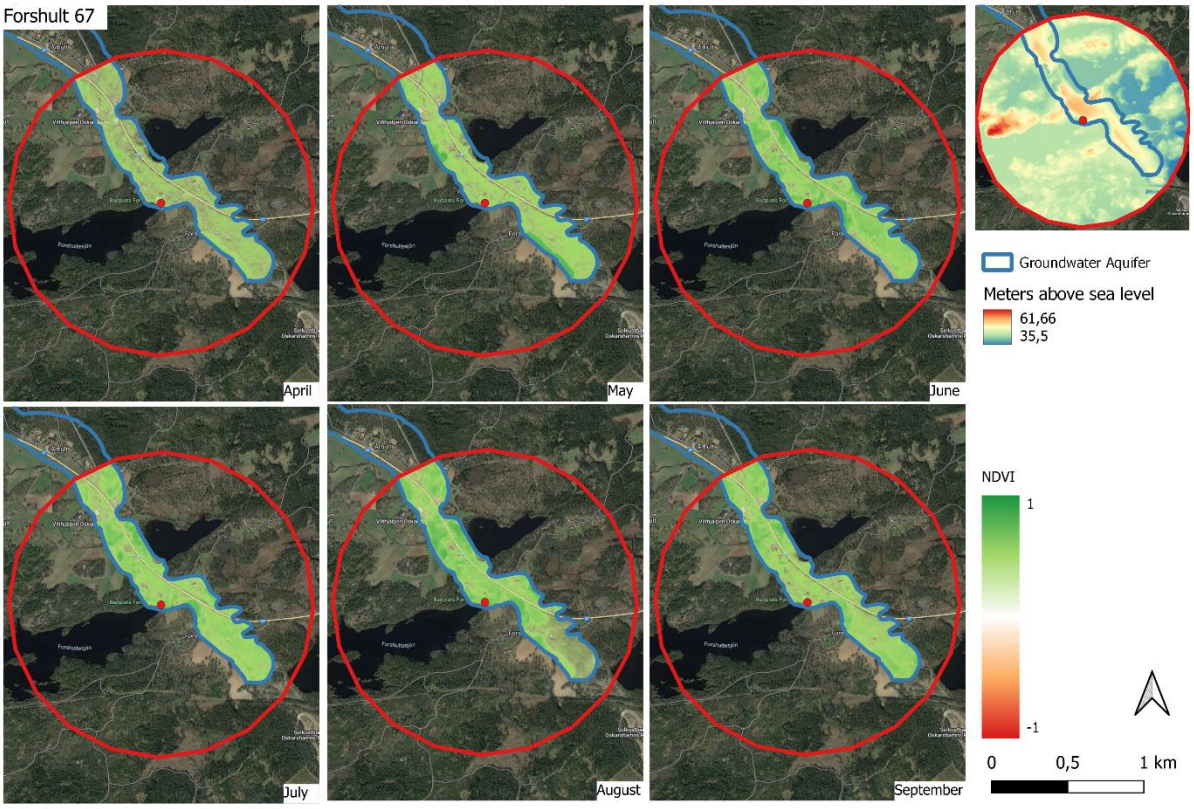


*Figure 28, NDVI over summer months at Forshult 67. The map aimed at examining NDVI variation as a result of above surface topography, during season when groundwater dependence seemingly increase.*

Results from topographic and NDVI maps of Forshult 67, aiming at examining variations in vegetation health due to above surface topography, during seasons where potential

groundwater dependence is higher, show no evident results related to groundwater uptake. While a lower laying section above the measurement station shows consistently high NDVI within the images from May and forth, this is likely due to the surface waterflow passing through that specific section, rather than the topographical low. Other variations are likely linked to vegetation types rather than topography looking at the crops displayed above the southern section of the aquifer showing as well as sections of deciduous forests showing high NDVI. These being vegetation types showing high NDVI during summer months.

# 7. Discussion

## 7.1 Comparison of $r_s$ based on visual interpretation and statistical analysis.

Presented result show that 16 of the original 57 measurements stations show $r_s >= 0.6$ between NDVI and groundwater levels, indicating strong correlation. In the case of NDWI and groundwater level, five of the original 57 measurements stations show $r_s <= -0.6$. These $r_s$-values follow the statistical threshold of strong correlation found in Bhanja, et al. (2019) and Zhu, et al. (2015), i.e., $r_s >= 0.6$ and $r_s <= -0.6$. However, when visually interpreting the time series found for example in figure 20, displaying sub-sample Emmaboda 21, strong correlation between vegetation indices and groundwater levels seems evident with extremes displayed for each variable following each other. This, even though the statistical $r_s$-values, 0.424 and -0.528 for NDVI and NDWI correlation with groundwater levels respectively, indicate moderate statistical correlation according to Bhanja, et al. (2019) and Zhu, et al. (2015). Combining statistical thresholds with visual interpretation could provide the user with revealing results otherwise potentially missed using only one of the two interpretation approaches.

Applying a lower threshold for strong correlation, $r_s >= -0.5$ and $r_s <= -0.5$, also supported by Song, et al. (2015) would in the case of this study result in 27 measurement stations showing strong correlation between NDVI and groundwater levels and 22 measurement stations showing strong correlation between NDWI and groundwater levels. Stretching the threshold even further down to +-0.4, would involve 35 measurement stations in the case of NDVI and groundwater levels and 35 measurement stations in the case of NDWI and groundwater levels. However, this suggestively requires additional interpretation of each time series that respective correlation is based on.

## 7.2 Variation over time and geospatial features

The positive correlation between NDVI and groundwater levels indicates that higher NDVI values, representing healthy dense vegetation, are followed by lower groundwater levels, aligning with the concept that healthy vegetation acts as a barrier to effective precipitation and groundwater recharge (Eveborn, et al. 2017). This pattern is further observed in the negative correlation values between NDWI and groundwater levels, aligning with the notion that the largest extents of groundwater recharge take place during winter and spring when vegetation health is reduced, and potential for effective precipitation increased (Grip et al. 1994).

The regression analysis results also demonstrate the same pattern, with higher NDVI values corresponding to lower groundwater levels and vice versa, along with opposite patterns for NDWI and groundwater levels. However, variations in groundwater levels for similar index values result in relatively weak $R^2$-values and suggest potential influence from geospatial features examined in this study.

Further, the strongest correlation between vegetation indices and groundwater levels generally occurred after a three-month time-lag. While most stations displayed this pattern, there were exceptions. Emmaboda 21 exhibited the strongest correlation after one month for NDVI, while Rälla 36 showed it after five months. For NDWI, Påskallavik 2 showed peak correlation after three months, while Dörby 38 displayed it after one month. However, the peaks in NDWI correlation did not always align with the same time-lag as NDVI correlation for the same stations. The delay in groundwater recharge, which the time-lag of vegetation indices aimed at capturing, is influenced by the time it takes for groundwater recharge to occur, indicating the importance of groundwater aquifer depth and soil type permeability. However, since NDVI $r_s$-values and the NDWI $r_s$-values for shared stations don't always peak at the same time-lag vegetation index variations found at different stations to some degree also affect the correlation. While a broader seasonal variation within the vegetation indices might seem evident, the more detailed variation in vegetation index values causes enough effect to potentially affect the temporal relation between index and groundwater levels.

### 7.2.1 Discussing geospatial features

As mentioned in the previous section, variation in correlation values is attributable to variations in geospatial features. These will be discussed further below.

In cases where there is weak correlation between vegetation indices and groundwater levels, it appears that groundwater measurements play a more significant role than the values of the indices themselves. On a general basis, the seasonal variation in vegetation indices is more consistent across the focus samples compared to the groundwater levels. For instance, while the vegetation indices exhibit seasonal highs and lows, the groundwater levels at Rälla 36 (figure 19) do not follow the same pattern as those at Dörby 38 (figure 17) and Vimmerby 55 (figure 18). This discrepancy could conceivably be attributed to the deep location of the groundwater table at Rälla 36, which ranges from approximately 12 to 14 meters below the ground surface. The greater distance that infiltrated water needs to travel to reach the groundwater table seemingly results in less alignment with above-surface patterns such as seasonal vegetation health and moisture represented by NDVI and NDWI.

Building on this, while high permeability can diminish the influence of topography, the variations in groundwater level measurements are influenced by a combination of geological and topographical characteristics, rather than single factors alone (SGU, 2023, b). The results of the soil type comparison in section 6.3 indicate that soil types with higher permeability and quicker response times may exhibit stronger correlation to the vegetation indices, likely due to responsive groundwater levels, following the suggestion in previous paragraph.

Another example where the potential influence of groundwater table depth and soil type can be seen and discussed is found around late summer the year 2021. For most stations the groundwater levels are restored a few months earlier this year indicating an increase in precipitation at this time 2021. Involving Emmaboda 21 (figure 20) it is noticeable that groundwater recharge occurs around September compared to October previous years. The groundwater table depth ranges from around 1.4 to 2.25 m.b.g.s at Emmaboda 21. At Dörby 38, the groundwater table ranges from around 4.25 to 5.75 m.b.g.s. Here, the groundwater level drastically increases around October 2021 compared to around November and December previous years. Of note is the one-month difference between the two stations the year 2021, indicating that Emmaboda 21 with shallower groundwater table depth saw recharge sooner than Dörby 38 the year of 2021. However, Emmaboda 21 is located within the soil type sand, and Dörby 38 within gravel or coarser, which could initially suggest faster recharge at Dörby 38 due to the higher permeability of said soil type. This further supports the notion that single factors alone don’t affect groundwater recharge, rather it is a question of combinations (SGU, 2023, b).

It is important to consider vegetation type as well, as different types of vegetation can produce similar vegetation index values. For example, forests and crops have both been found to produce high NDVI values at certain stages. However, there is an uncertainty introduced by the maximum values parameter, as it does not encompass the full range of vegetation types and their correlation to groundwater levels. This means that not all values within the time series can be definitively associated with a particular vegetation type. Hence the role of particular vegetation types remains relatively unknown within this study, while they likely play a part in the varying index values on a detailed level.

The weak correlation observed at Vimmerby 104 (figure 24) between both NDVI and NDWI and groundwater levels can potentially be attributed to various geospatial features. As illustrated in Figure 25 in section 6.3, the seasonal variation observed at other measurement stations is absent in the groundwater measurements at Vimmerby 104. This could be attributed to a combination of factors, including the deep groundwater table and the moraine soil type, which both require more time for infiltrated water to reach and replenish groundwater levels (Bergström, 2001; Lindström, et al., 2011). Also, the extensive vegetation cover at this location, which is relatively accurately represented by the maximum vegetation pixel values, can also affect groundwater recharge (Eveborn, et al., 2017). The year-round cover of a coniferous forest acts as a barrier, potentially reducing groundwater recharge and resulting in less seasonal variation in groundwater levels, while the vegetation index values remain variable.

Lastly, looking at topographic location, the upward flow of discharge areas, typically found in low-lying topographical regions (Condon & Maxwell, 2015), may support the correlation between NDWI and shallow aquifers, such as the one at Fliseryd 1. This is potentially due to the close proximity between above-surface and below-surface moisture. When examining the averaged NDVI and NDWI rs-values within intermediate and recharge areas, both show a relatively similar strength, indicating limited confidence in any determining role of these topographic locations regarding the correlation between vegetation indices and groundwater levels. The increased difference between NDWI and groundwater levels $r_s$-values within the three topographic locations potentially indicates that the topographic location plays a more significant role in NDWI correlating with groundwater levels than in the case of NDVI.

### 7.2.2 Discussing combinations in geospatial features

| Station | Soil type | Topografic location | Levels observed (m.b.g.s) | Time span | NDVI $r_s$ | NDWI $r_s$ |
|---|---|---|---|---|---|---|
| *Dörby 38* | *Sand* | *Intermediate area* | *4.14 - 5.7* | *4/18/2019 - 12/22/2022* | *0,706* | *-0,297* |
| *Emmaboda 21* | *Gravel or coarser* | *Intermediate area* | *1.3 - 2.25* | *5/2/2019 - 12/22/2022* | *0,424* | *-0,528* |
| *Fliseryd 1* | *Gravel or coarser* | *Discharge area* | *1.32 - 2.41* | *5/2/2019 - 12/22/2022* | *0,555* | *-0,554* |
| *Forshult 67* | *Sand* | *Intermediate area* | *0.83 - 1.7* | *10/29/2020 - 12/22/2022* | *0,696* | *-0,656* |
| *Läckeby 1* | *Moraine* | *Recharge area* | *4.35 - 6.24* | *5/16/2019 - 9/17/2020* | *0,778* | *-0,306* |
| *Påskallavik 2* | *Gravel or coarser* | *Recharge area* | *0.86 - 2.19* | *5/16/2019 -12/22/2022* | *0,664* | *-0,633* |
| *Rälla 36* | *Sand* | *Intermediate area* | *11.89 - 14.18* | *1/1/2015 - 12/22/2022** | *0,292* | *0,119* |
| *Trekanten 63* | *Sand* | *Intermediate area* | *0.99 - 1.94* | *10/29/2020 - 12/22/2022* | *0,386* | *-0,187* |
| *Vimmerby 55* | *Sand* | *Recharge area* | *4.44 - 4.94* | *1/1/2015 - 11/10/2022* | *0,492* | *-0,545* |
| *Vimmerby 104* | *Moraine* | *Recharge area* | *7.06-8.09* | *1/1/2015 - 12/22/2022** | *0,146* | *0,082* |

*Table 10, Focus samples, respective geospatial features considered and correlation coefficients.*

Reviewing table 10, it is difficult to establish evident combinations of geospatial features within the focus samples, that provide similar correlation values. While results from analyzing variables separately have shown patterns which conceivably align with key themes within the subject, combining these make patterns less confident. For example, Forshult 67 and Trekanten 63 are both measurement stations mapped within the soil type sand, both are located within an intermediate area, both show similar groundwater levels observed, and data has been collected over the same time span. Still, Trekanten 63 shows weak correlations for both indices and groundwater levels while Forshult 67 shows strong correlation for both indices and groundwater levels. Påskallavik 2 and Dörby 38 however, located within different soil types, topographic locations, time spans, levels observed and, time spans used both show strong correlation in terms of NDVI correlation to groundwater levels. Different combinations of geospatial features such as soil type, topographic location, aquifer depth, etcetera seemingly can provide similar correlations as seen in figure 29.

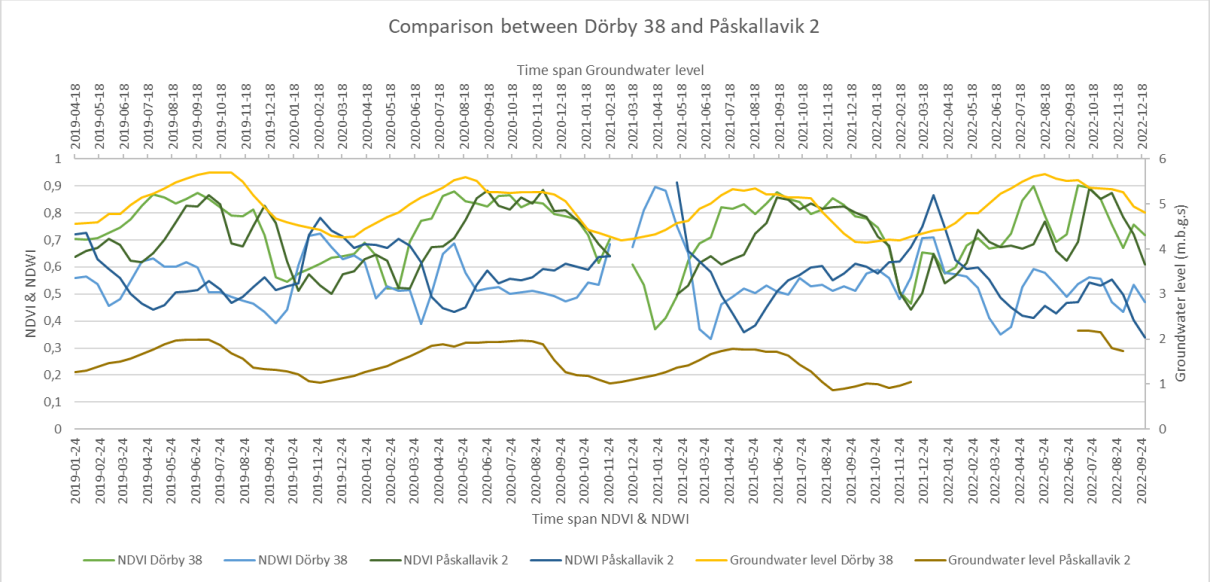


*Figure 29, Comparison of Dörby 38 and Påskallavik 2 time series, displaying similar index and groundwater patterns while possessing different geospatial features*

Further it is difficult to establish any evident spatial distribution pattern that would indicate any regional differences found amongst the 16 stations showing rs >= 0.6 between NDVI and groundwater levels as well as the five stations rs <= -0.6 between NDWI and groundwater levels (see figure 11 and 12). If also counting measurement stations that show rs-values >= 0.5 and rs-values <= -0.5, or even rs-values above or below +-0.4, large parts of the extent for the 57 measurements stations used within Kalmar County would be covered. Instead, the geospatial variation makes it mark on a local scale.

### 7.2.3 Missing precipitation data

While geospatial features analyzed in this study has implications on NDVI and NDWI correlating with groundwater levels, a more obvious common trigger is likely precipitation based on effective precipitation being the main driver for groundwater recharge (Eveborn et al. 2017). Adding precipitation data as an additional variable to this type of study could present more informative results which explores precipitation as a potential trigger for both vegetation health and moisture as well as groundwater level variation. Zhu, et al. (2015) found a strong correlation between precipitation and NDVI during periods of low NDVI and groundwater correlation.

## 7.3 Methodology and input parameters

### 7.3.1 Single value parameter, gaps and outliers

The maximum value parameter was chosen aiming at ensuring reliable remote sensing data. When extracting Sentinel 2 level-1C data over a larger time span there is a risk of images being unusable due to high cloud cover. While the issue of cloud cover was mitigated through interpolation, the high cloud presence in a humid climate such as that of the county of Kalmar can affect the results of the interpolation negatively, making it a limited option. Hence, using a single value parameter such as the maximum value aimed at ensuring one realistic index value, even in the risk of other pixels being faulty. An alternative approach would further have been to use averaged index values. This in order to properly represent the full extent of each study area. However, this requires much more carful data collection, assuring that each image has no cloud cover and displays proper values. This is very time consuming and limits data availability in humid climates such as Sweden's. In order to be time-sufficient, the data collection process was automated through JavaScript used in GEE.

Building on this, gaps in the index time series, as seen at Påskallavik 2 (figure 30, in appendix) between November 2020 and January 2021 for example, indicate that no data adequate data could be collected during this time. This, likely due to extensive cloud cover. Similar gaps occur at multiple measurement stations, indicating widespread cloud cover over Kalmar County. Gaps in groundwater time series are likely due to inactive loggers.

Another source of potential error in the time series data is the presence of certain peaks in the index values. After the gap of the Påskallavik 2 time series, the reappearing index values start just above 0.9. This is likely a faulty index value, creating outliers which have been observed in several of the samples time series. Outliers within the collected data potentially is the result faults within the collection of data form Google Earth Engine, a risk present when working with large amounts of digitalized data and automated scripts. It could also be a question of remaining atmospheric effects, present within Top-of-atmosphere images.

While the ratioed images can mitigate the atmospheric effects (Lillesand, et al. 2015), optimal images with Bottom-of-atmosphere perspective would have been used. Such are found within the Sentinel 2 level-2A image collection. However, this collection does go as far back in time as Sentinel 2 level-1C, starting off around 2018 instead of 2015 (ESA, n.d, a).

Another alternative approach that removes problems of cloud cover is using radar data. Radar waves are not affected by clouds and haze and can therefore collect data not depending on proper weather conditions. However, using radar-data in this context requires more complex preprocessing. Utilizing radar data can be due to this aspect requiring more time and resources (Meijerink, et al. 2007).

### 7.3.2 Biweekly basis

Looking at the biweekly basis parameter, this was chosen with the aim of ensuring frequent measurements for further analysis. Results produced from more measurements rather than less, conceivably present a more realistic result. The two-week basis was deemed appropriate based on the relatively short time spans occurring within several time serie. While a lower frequency of measurement, e.g., monthly basis, are easier to manage and potentially reduces the need for data modification, it affects the results the shorter time spans found in this study.

### 7.3.3 Selection of focus samples

Selection of focus samples was based on their statistical outcomes as well as time spans cover. However, the geospatial features should have been more attention within this selection making sure that features occurred more similarly over the focus samples, possibly providing a better result. Also, the ten focus samples were chosen since ten was perceived as a manageable amount. However, choosing an number of focus samples that, when divided by three, result in an integer would potentially make comparisons over the three soil type and topographic classes more even.

### 7.3.4 Addressing chosen variables

The amount and characteristics of the related geospatial features analyzed within this study, while affecting groundwater conditions in their respective way, are arguably too complex to be properly evaluated within a study of this extent. Conceivably they would be more accurately analyzed if studied more extensively and exclusively in relation to vegetation index and groundwater correlation. As an example, geological features such as soil type and topographic location possess attributes best described though physics, which have not been addressed in this study. Instead, these features have been analyzed on a very general basis. Also, the impact of specific vegetation types would be interesting to explore further to establish patterns found for specific vegetation type extents. Hence, this study could have benefitted from more evident

delimitation and being divided into two or more separate studies. One that focus on the effect of geological feature and one the focus more on the effect of different vegetation types.

Further, while NDVI has established itself within groundwater studies, as found in previous research, NDWI is not as extensively used. Gao (1996) described NDWI as more of a supplement to NDVI in analyzing vegetation conditions and it was added to the study aiming at seeing how it compared to NDVI in relation to groundwater relations. At the end of this study, however, the number of variables used arguably have been too many to properly consider. Results regarding both indices in relation to groundwater could potentially be more informative if studied separately.

### 7.3.5 Regression and seasonal variety

Separating the variables used and examining them within different seasonal contexts could provide findings of interest. For example, only using NDVI values during spring when they increase, and groundwater levels decrease within a linear regression analysis could conceivably present a higher $R^2$-value than looking over the full extent of the time series. However, this was realized at a too late stage of the study.

Using linear regression for seasonally varying time series data is not optimal, while very accessible. Alternatives were attempted, for example a bagged tree regression, however the results of this were difficult to apply to the study. With more time and experience a better regression analysis could have been performed.

### 7.3.6 More use of imagery

Lastly, it is beneficial that when using automated scripts within Google Earth Engine, involving many images, one should look more closely at individual images. This will help in the understanding of connections between vegetation index values and specific vegetation types. Further, using this knowledge to get a deeper understanding of variations within the time series as well as correlation to groundwater.

# 8. Conclusions

Sixteen of the original 57 measurement stations within the county of Kalmar showed $r_s$-values >= 0,6 between NDVI and groundwater levels. 27 of 57 showed $r_s$-values >= 0,5 between NDVI and groundwater levels. Five of the original 57 measurement stations showed $r_s$-values <= -0,6 between NDWI and groundwater levels. 22 of 57 showed $r_s$-values <= -0,5 between NDWI. Time series interpretation within this study has helped provide better indication of correlation than only statistical correlation thresholds.

Delay in groundwater recharge displayed by time-lag of vegetation indices was affected by local variations of index values and not only geological characteristics, showing peak correlation after varying monthly durations. Three-month time-lag showed strongest correlation over most sub-samples between the vegetation indices and groundwater levels.

Separate variable analysis showed:

- Median groundwater level, as a measure for aquifer depth below ground surface, showed $R^2 = 0.6352$ in relation to NDWI correlation found at each focus area.
- Highest averaged correlation was found within gavel or coarser soil types, then sand and lastly moraine, aligning which the high to low permeability found at each soil type class.
- While able connect most index values to broader vegetation type classes, the effect of these on the correlation to groundwater levels remain unsure.
- Results related to topographic location showed low confidence.

Combining variables, while related to each other, makes patterns less certain, suggesting that future studies should focus on each features effect on NDVI and NDWI correlations separately. Preferably, this is done with remote sensing data with preprocessed radiometric correction, such as Sentinel 2 level-2A. Alternatively using Radar to work around problems caused by cloud cover. Also, with more careful data collection, seasonal focus and use of visualization.

Additional precipitation data to NDVI and NDWI correlations, as well as related geospatial features could conceivably provide more accurate relations. That being said, the amount of variables used in this study have been argued to reduce the quality of results.

# 9. References


- Adams, K., Reager, J., Rosen, P., Wiese, D., Farr, T., Rao, S., . . . Rodell, M. (2022). Remote Sensing of Groundwater: Current Capabilities and Future Directions. Water Resources Research, 58(10), N/a.
- Bergström, S. (2001). Sveriges Hydrologi - grundläggande hydrologiska förhållanden. Svenska Hydrologiska Rådet. SMHI.
- Bhanja, S., Malakar, P., Mukherjee, A., Rodell, M., Mitra, P., & Sarkar, S. (2019). Using satellite-based vegetation cover as an indicator of groundwater storage in natural vegetation areas. Geophysical Research Letters, 46(14), 8082-8092.
- Campbell, J., Wynne, R., Thomas, V., & ProQuest. (2022). Introduction to Remote Sensing. (6th ed.).
- Chen, D., Jackson, T., Li, F., Cosh, M., Walthall, C., & Anderson, M. (2003). Estimation of vegetation water content for corn and soybeans with a normalized difference water index (NDWI) using Landsat Thematic Mapper data. IGARSS 2003. 2003 IEEE International Geoscience and Remote Sensing Symposium. Proceedings (IEEE Cat. No.03CH37477), 4, 2853-2856 vol.4.
- Condon, L., & Maxwell, R. (2015). Evaluating the relationship between topography and groundwater using outputs from a continental-scale integrated hydrology model. Water Resources Research, 51(8), 6602-6621.
- Entekhabi, D., Njoku, E., O'Neill, P., Kellogg, K., Crow, W., Edelstein, W., . . . Van Zyl, J. (2010). Soil Moisture Active Passive (SMAP) Mission. Proceedings of the IEEE, 98(5), 704-716.
- EOS Data Analytics. (n.d, a). NDMI (Normalized Difference Moisture Index). Available: https://eos.com/make-an-analysis/ndmi/ (Retrieved: 2023-02-29)
- ESA. (n.d, a). SENTINEL-2 MISSION GUIDE. Available: https://sentinel.esa.int/web/sentinel/missions/sentinel-2 (Retrieved: 2023-02-10)
- ESA. (n.d, b). Cloud Masks. Available: https://sentinel.esa.int/web/sentinel/technical-guides/sentinel-2-msi/level-1c/cloud-masks (Retrieved: 2023-03-02)
- Eveborn, D., Vikberg, E., Thunholm, B., Hjerne, C., Gustafsson, M. (2017). Rapportering av regeringsuppdrag: kunskapsunderlag om grundvattenbildning Grundvattenbildning och grundvattentillgång i Sverige. SGU.
- Gandhi, U. (2021). Temporal Gap-Filling with Linear Interpolation in GEE. Available: https://spatialthoughts.com/2021/11/08/temporal-interpolation-gee/

- Gao, B. (1996). NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space. Remote Sensing of Environment, 58(3), 257-266.
- Gomez, B., & Jones, J. (2010). Research methods in geography a critical introduction (Critical Introductions to Geography).
- Google Earth Engine. (2021). Earth Engine Code Editor. Available: https://developers.google.com/earth-engine/guides/playground (Retrieved: 2023-03-01)
- Grip, H., Lehman, I., & Rodhe, A. (1994). Vattnets Väg Från Regn till Bäck. 3., Rev. Uppl. ed.
- Haraldsson, O. (2015). Topografins inverkan på markvattenhalten i ett jordbruksdominerat avrinningsområde - The topographic impact on soil water content in an agricultural catchment. Institutionen för mark och miljö. SLU.
- Harrie, L. (2020). Geografisk informationsbehandling : Teori, metoder och tillämpningar (Sjunde upplagan ed.).
- Holden, J. (2017). An introduction to physical geography and the environment (4th rev. ed.).
- Johansson, P. (1998). En jämförande studie av grunt grundvatten i skog och hygge. Department of Physical Geography, Gothenburg University.
- Jones, H., & Vaughan, R. (2010). Remote sensing of vegetation : Principles, techniques, and applications
- Karlsson, C., Sohlenius, G., & Peterson Becher, G. (2021). Handledning för jordartsgeologiska kartor och databaser över Sverige. SGU.
- Kuhn, M., & Johnson, K. (2013). Applied predictive modeling. New York, NY: Springer.
- Körner, S., & Wahlgren, L. (2015). Statistisk dataanalys (5. uppl. ed.).
- Lantmäteriet. (n.d). Laserdata Nedladdning, skog. Available: https://www.lantmateriet.se/sv/geodata/vara-produkter/produktlista/laserdata-nedladdning-skog/ (Retrived: 2023-03-28)
- Lillesand, T., Kiefer, R., & Chipman, J. (2015). Remote sensing and image interpretation (7.th ed.). Hoboken, N.J.: Wiley.
- Lindström, M., Lundqvist, J., Calner, T., Sivhed, M., Lundqvist, Thomas, Calner, Mikael, & Sivhed, Ulf. (2011). Sveriges geologi från urtid till nutid (3. [rev.] uppl. ed.)
- Lång, L., Norström, N., Maxe, L.; & Lindeberg, C. (2022). Fördjupad utvärdering av Grundvatten av god kvalitet 2023. SGU-rapport 2022:13. SGU.

- Mason, J. (2016). Physical geography: The global environment (Fifth ed.).
- McFeeters, S. K. (1996). The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features, International Journal of Remote Sensing, 17:7, 1425-1432.
- Meijerink, A.M.J., Bannert, D., Batelaan, O., Lubczynski, M.W., & Pointet, T. (2007). Remote sensing application to groundwater. UNESCO.
- Naturvårdsverket. (2018). Nationella marktäckedata. Available:https://metadatakatalogen.naturvardsverket.se/metadatakatalogen/GetMetaDataById?id=8853721d-a466-4c01-afcc-9eae57b17b39 (Retrieved: 2023-03-28)
- Olsson, H., Reese, H., & Nyström, M. 2018. FOUNDATIONS OF REMOTE SENSING. Skogshushållningsserien, Remote Sensing of Forests. Ljungberskompendium. SLU.
- Press, W. H., Teukolsky, S. A., Vetterling, W.T., & Flannery, B. P. (2007). Numerical Recipes - The art of Scientific Computing. 3rd edition. Cambridge.
- Rouse, J. W., Jr., Haas, R. H., Deering, D. W., & Schell, J. A. 1973. Monitoring the vernal advancement and retrogradation (green wave effect) of natural vegetation. RSC-1978-2, E74-10113, NASA-CR-136103. Goddard Space Flight Center Greenbelt, Maryland 20771.
- Reese, H., Nordkvist, K., & Olsson, H. (2018). Methods for analysis of remote sensing data. Skogshushållningsserien, Remote Sensing of Forest. Ljungbergskompendium. SLU.
- Scalgo. (n.d). Analysis - Depression-Free Flow. Available: https://scalgo.com/en-US/scalgo-live-documentation/analysis/depression-free-flow (Retrieved: 2023-03-28)
- Schober, P., Boer, C., & Schwarte, L. (2018). Correlation Coefficients: Appropriate Use and Interpretation. Anesthesia and Analgesia, 126(5), 1763-1768.
- Shumway, R., & Stoffer, D. (2006). Time series analysis and its applications : With R examples (2.nd ed., Springer texts in statistics). New York: Springer.
- SGU. (1994). Grundvattnet i Sverige. SGU serie Ah nr 17.
- SGU. (2022). Grundvattenmagasin. Available: https://www.geodata.se/geodataportalen/srv/swe/catalog.search#/metadata/703e7a68-714d-477e-a508-a482f291a69b (Retrieved: 2023-05-18)
- SGU. (2023, a). Jorddjupskartan. Available: https://apps.sgu.se/kartvisare/kartvisare-jorddjup.html (Retrieved: 2023-02-29)

- SGU. (2023, b). Beskrivning av grundvattenförekomster. Available: https://www.sgu.se/anvandarstod-for-geologiska-fragor/vattenforvaltning-av-grundvatten/sgus-foreskrifter-om-kartlaggning-och-analys-sgu-fs-2013-1/inledande-kartlaggning/beskrivning-av-grundvattenforekomster/ (Retrieved: 2023-02-29)
- SMHI. (2022). Vattenbalans. Available: https://www.smhi.se/kunskapsbanken/hydrologi/vattnets-kretslopp/vattenbalans-1.124695 (Retrieved: 2023-04-21)
- Song, G., Jin-ting H., Bo-han N., Jia-wei W., & Lei, Z. (2021). Effects of Groundwater Level on Vegetation in the Arid Area of Western China. China Geology 4.3: 527-35. Web.
- Vattenmyndigheten Södra Östersjön. (2016). Del 1 Introduktion: Vattenförvaltning och dess verktyg i Sverige. Förvaltningsplan 2016–2021, Södra Östersjöns vattendistrikt. Länsstyrelsen Kalmar län.
- Vattenmyndigheten Södra Östersjön. (2022). Delförvaltningsplan med åtgärder mot vattenbrist och torka 2021—2027 Södra Östersjöns vattendistrikt. Vattenmyndigheterna i Sveriges fem vattendistrikt.
- Zhu, L., Huili, G., Zhenxue, D., Tingbao, X., & Xiaosi, S. (2015). Integrated Assessment of the Impact of Precipitation and Groundwater on Vegetation Growth in Arid and Semiarid Areas. Environmental Earth Sciences 74.6: 5009-021. Web.

# 10. Appendix

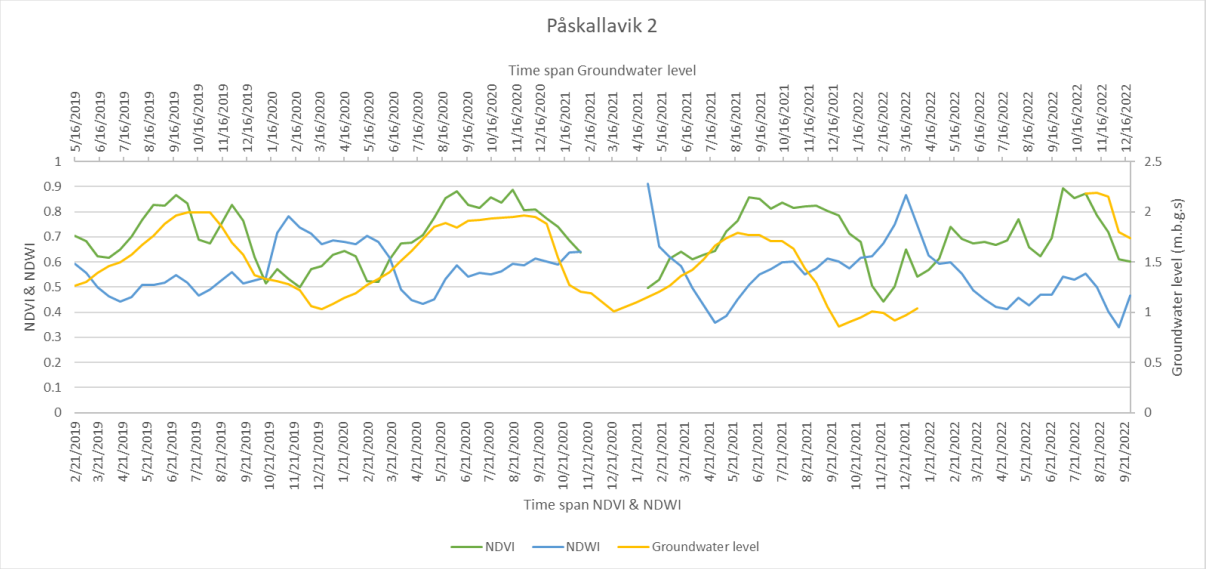


*Figure 30, Påskallvik 2 time serie, displaying NDVI and NDWI values related to the lower time span and groundwater level measurements related to the upper time span.*

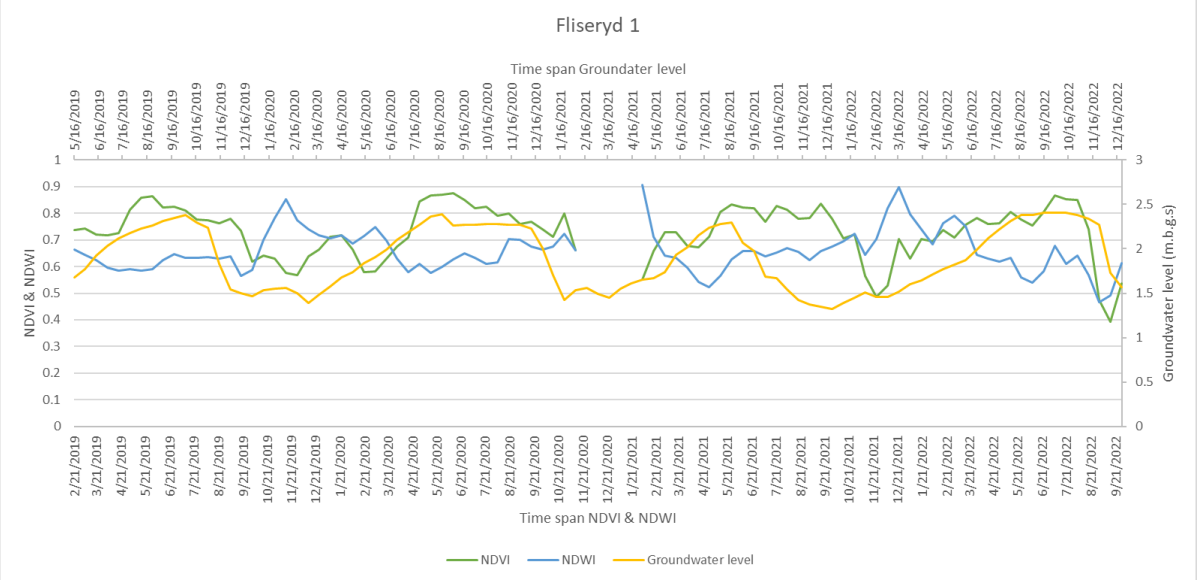


*Figure 31, Fliseryd 1 time serie, displaying NDVI and NDWI values related to the lower time span and groundwater level measurements related to the upper time span.*

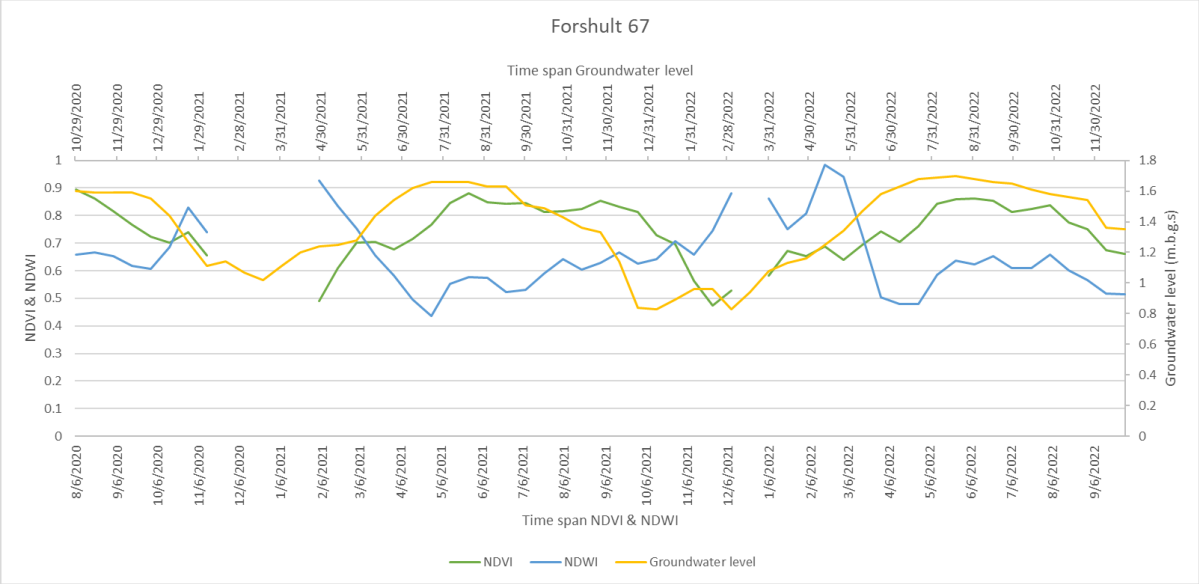


*Figure 32, Forshult 67 time serie, displaying NDVI and NDWI values related to the lower time span and groundwater level measurements related to the upper time span.*

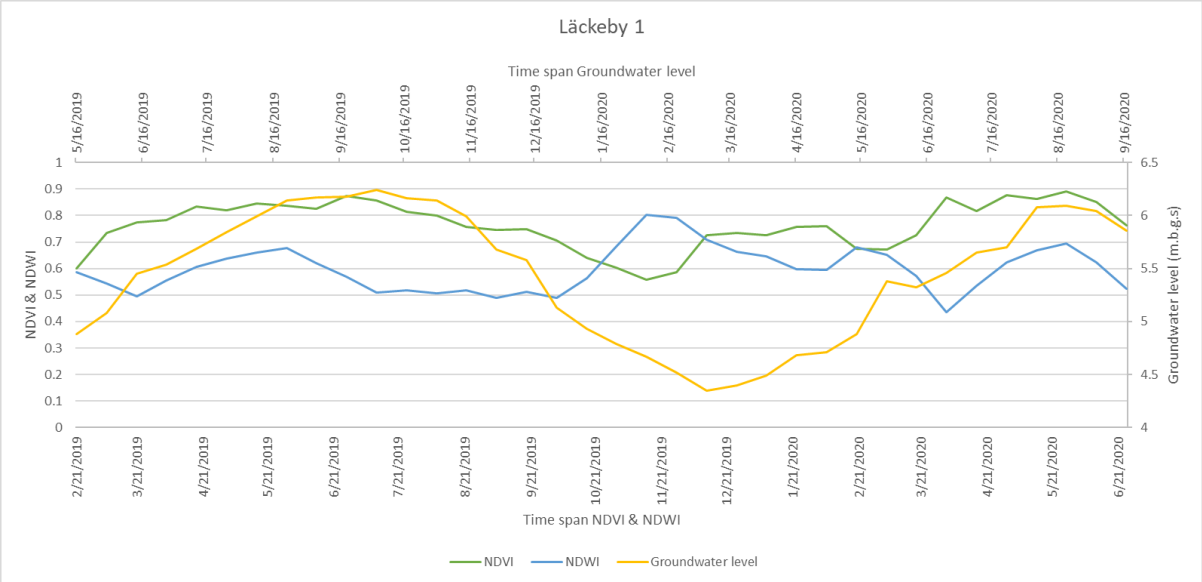


*Figure 33, Läckeby 1 time serie, displaying NDVI and NDWI values related to the lower time span and groundwater level measurements related to the upper time span.*

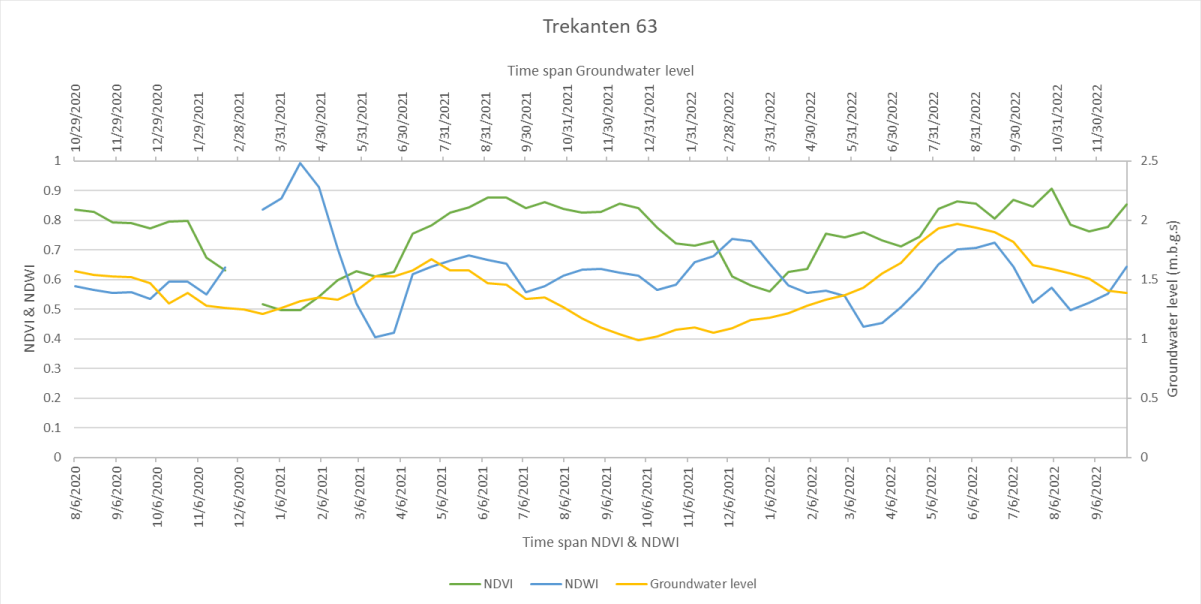


*Figure 34, Trekanten 63 time serie, displaying NDVI and NDWI values related to the lower time span and groundwater level measurements related to the upper time span.*

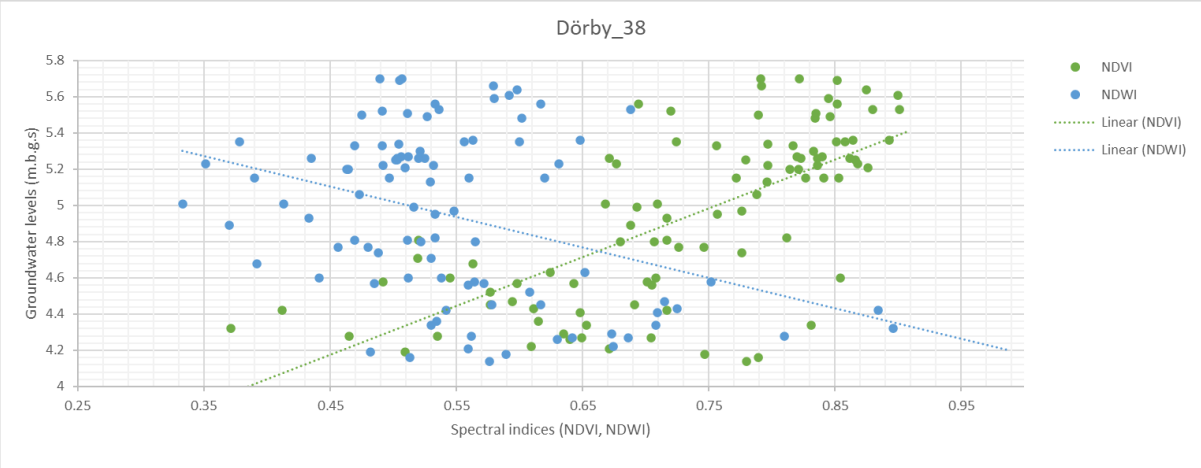


*Figure 35, Dörby 38 regression analysis.*

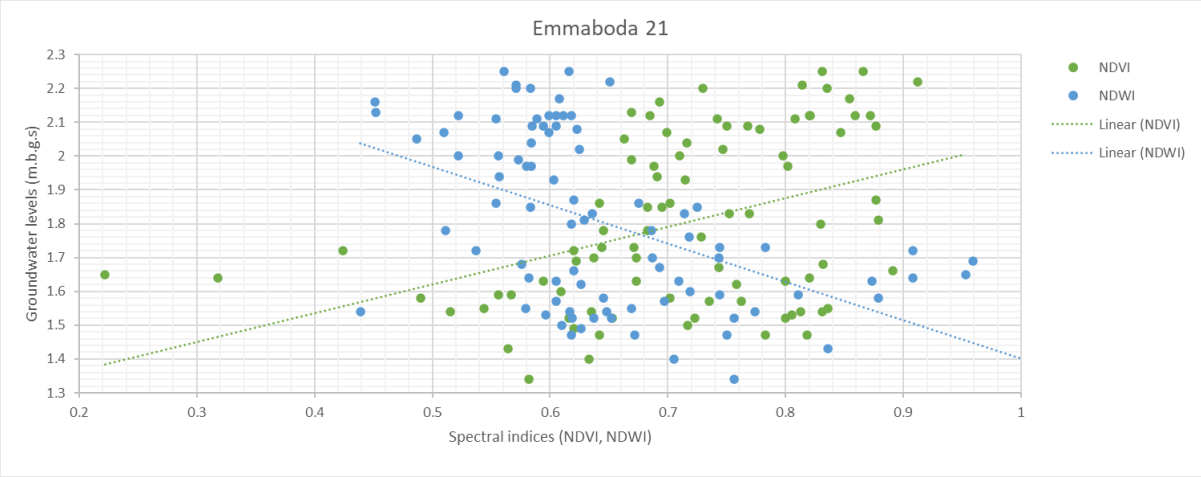


*Figure 36, Emmaboda 21 regression analysis.*

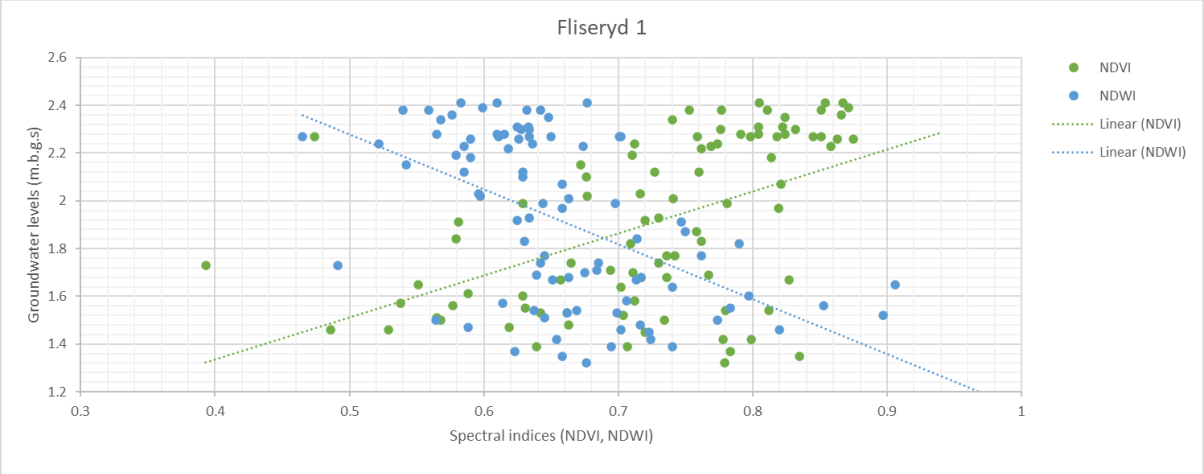


*Figure 37, Fliseryd 1 regression analysis.*

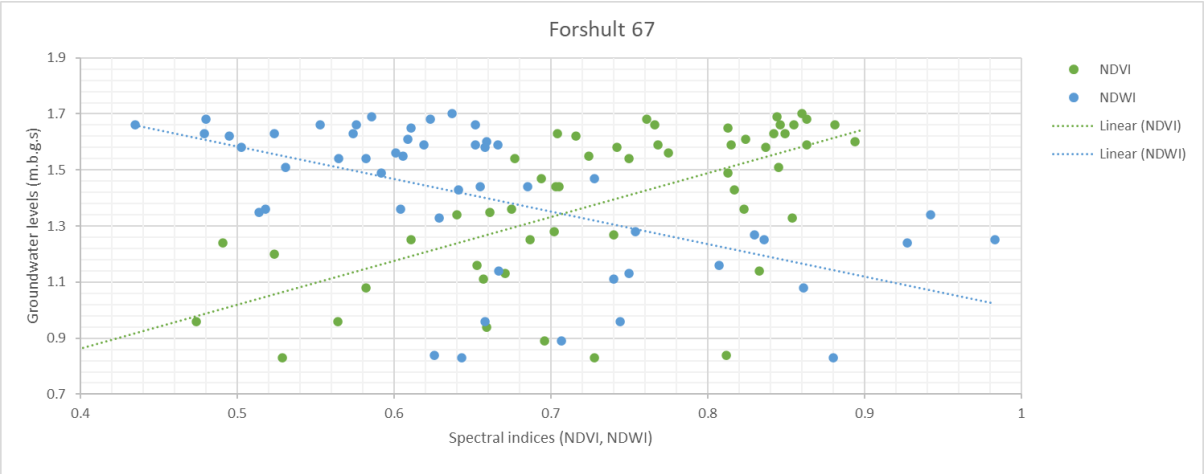


*Figure 38, Forshult 67 regression analysis.*

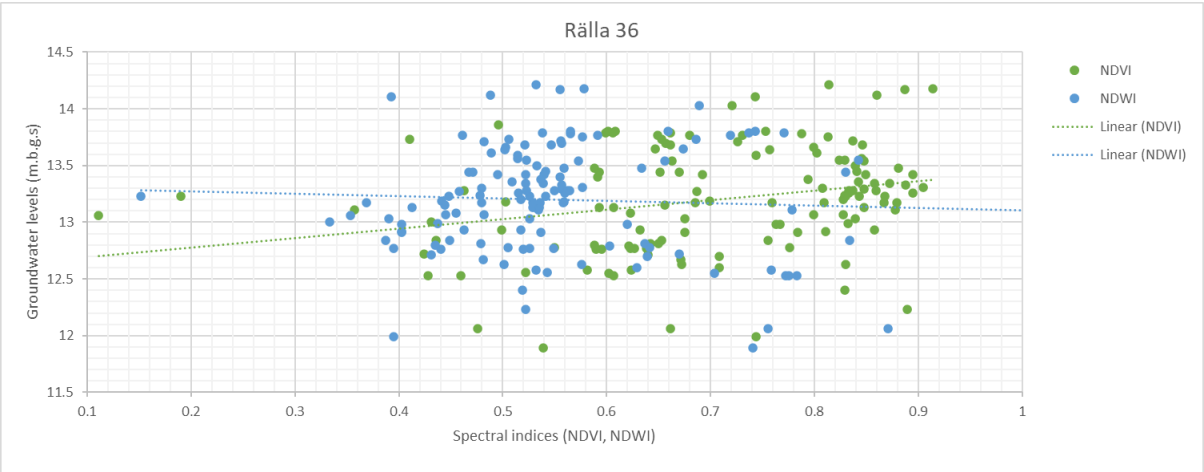


*Figure 39, Rälla 36 regression analysis.*

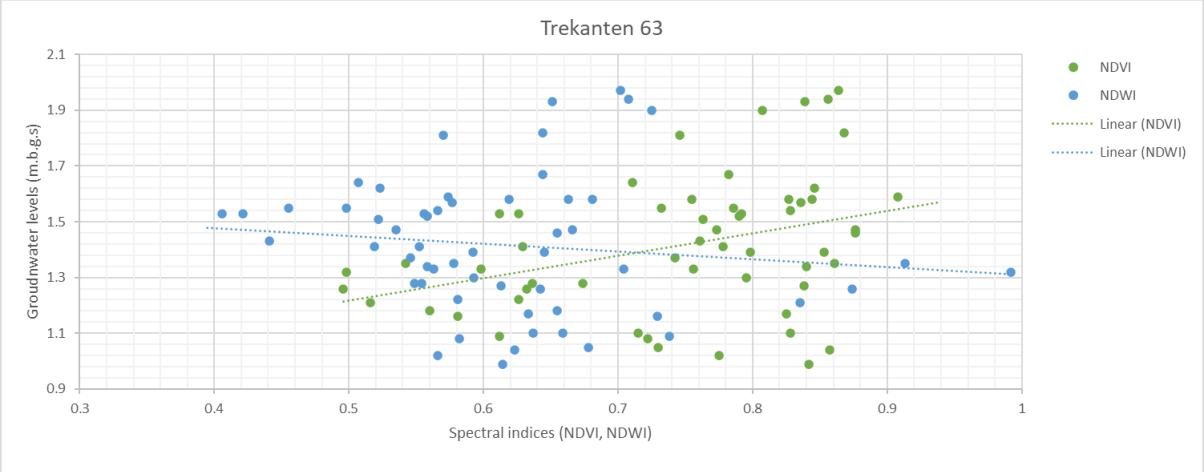


*Figure 40, Trekanten 63 regression analysis.*

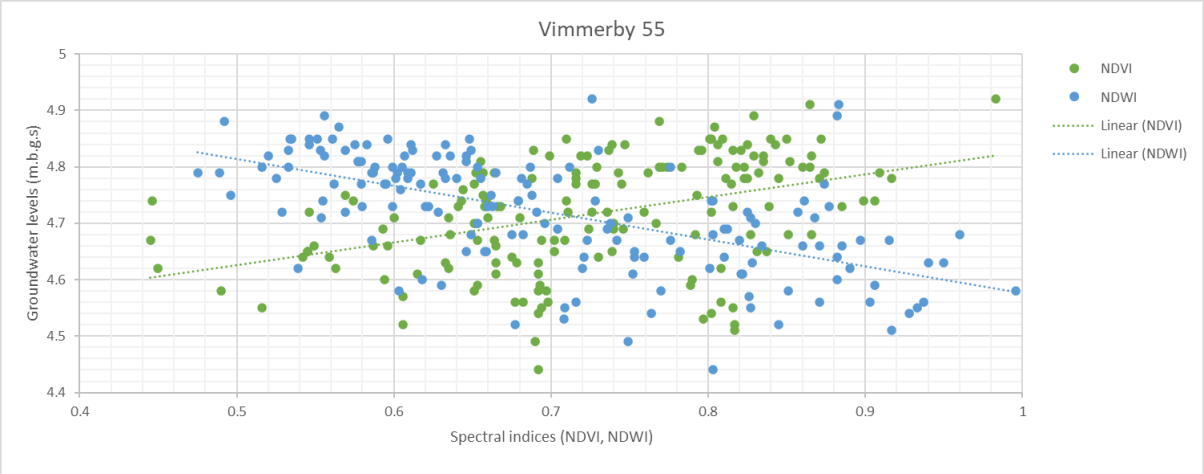


*Figure 41, Vimmerby 55 regression analysis.*

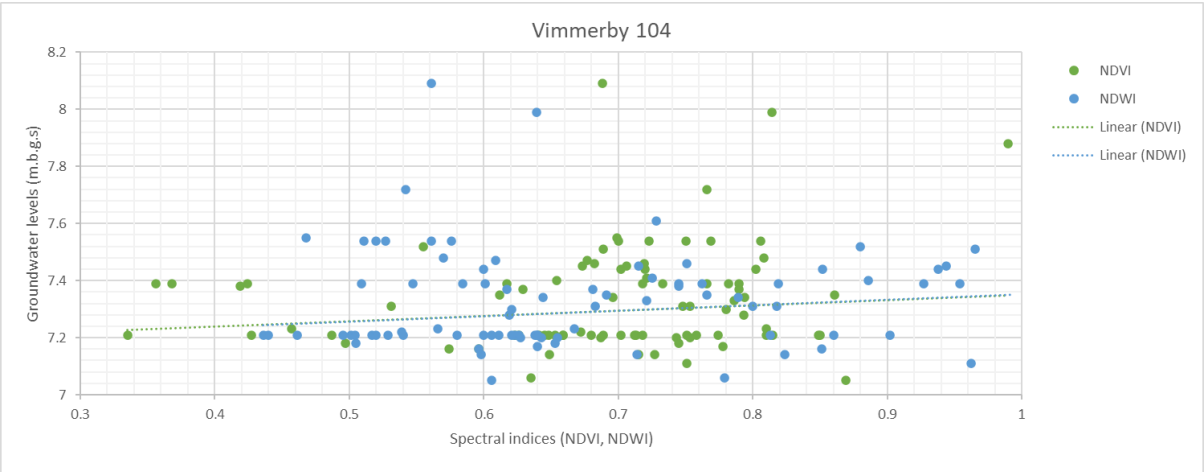


*Figure 42, Vimmerby 104 regression analysis.*