Hochschule Rhein-Waal
Rhine Waal University of Applied Sciences
Faculty of Communication and Environment
Degree Program - Information Engineering and Computer Science (M.Sc.)

Case Study:
Drought analysis and its impact on vegetation in Sauerland region of Germany for year 2017 based on soil moisture, precipitation, and altitude

Submitted to
Prof. Dr. Rolf Becker
Mr. Clein Alexander Sarmiento Castrillon

Group Members
Nikhilesh Munthala (28383)
Sheharaz Abdul Majeed Sheikh (28381)
Thirumal Janakiraman (28364)

Submitted on
28.03.2021

Table of Contents

Abstract

This report investigates the drought and vegetation for the area Sauerland region of Germany for the year 2017. Different parameters such as Precipitation, Soil Moisture Index, and the Normalized Difference Vegetation Index (NDVI) were analyzed for selected regions in Olpe and HSK. Further, the investigation focused on determining the correlation between the chosen parameters and their effects on vegetation and their role in causing drought. Based on the analysis, it is identified that there is no conclusive evidence for stating that precipitation and altitude are correlated to each other. Also, precipitation directly affects the Soil Moisture Index(SMI) of the region. The correlation between SMI and altitude cannot be determined with the available data. Therefore it is not possible to conclude with the given data whether all these factors directly affect drought conditions. With the given precipitation data, estimating the water to drain by the drainage system can be estimated approximately but needs more research.

1. Introduction

Drought is one of the prominent areas of study in natural disasters, even though its impact is mostly understated. It is one of the most important natural disasters due to its impact on flora and fauna, livestock, crops, and many industrial sectors. It is known to have affected the livelihood of 2.2 billion people worldwide between 1950 and 2014 (Matthias Zink. et al., 2016). The behavior of drought differs with the region and is irregular, making it difficult to predict the drought.

Drought is a condition caused due to lack of precipitation or water content and imbalance in the hydrological cycle leading to interruption in activities that requires water (Chandrasekar. et al., 2013). Drought is region-specific and is a natural climate phenomenon. Also, the adverse effect of drought is not observed unless it is cumulative and occurs in consecutive years for the specific region. [ChandrasekarK_Book_Chapter16]: There are three significant types of drought based on the impacted area of applications: agricultural, hydrological, and meteorological drought (Chandrasekar. et al., 2013).

In this study, the cause of drought in Germany's Sauerland region for the year 2017 is investigated. This study aims to analyze the role of altitude, precipitation, and drought condition in the chosen area and how they impacted the vegetation by correlating the Soil Moisture Index and Normalized Difference Vegetation Index.

1.1 Soil Moisture Index

Soil Moisture Index (SMI) is one of the key factors in understanding the hydrological cycle and can determine phenomena like drought and vegetation in an area (Chandrasekar. et al., 2013). It is a crucial climate observation variable and is declared as an "Essential Climate Variable" (ECV) by the Global Climate Observing System (GCOS) (Pellet, C and C. Hauck, 2016). Continuous monitoring of SMI is made mandatory across the globe after 2010 (Pellet, C and C. Hauck, 2016). SMI is defined as the proportion of the difference between the current soil moisture and the permanent wilting point to the field capacity and the permanent wilting point (Chandrasekar. et al., 2013). The role of SMI in causing drought is long-term and cumulative.

> *‘Only when the current soil moisture falls below the 20th percentile in comparison to long-term data, i.e., the value seen in only 20% of the years of a long time series, is it considered to be a drought.’* (Marx, n.d.)

SMI is indicated with values ranging from 0 to 1. While 0 indicates a low soil moisture region or extremely dry (arid) region, 1 indicates a high soil moisture region or extremely wet region (Saha, et al.,2018).

1.2 Precipitation

Precipitation is the primary intermediary in the water cycle that facilitates water transfer from the atmosphere to the Earth. It can occur in many forms, such as snow, hail, and freezing rain. Rain is the most prevalent form of precipitation. Precipitation impacts the soil moisture index and depends linearly on altitude. Higher precipitation will keep high soil moisture, whereas a low soil moisture index will be caused due to low precipitation. (Earth Science Week 2015; Morán-Tejeda et al. 2013)

1.3 Normalized Difference Vegetation Index (NDVI)

Normalized Difference Vegetation Index (NDVI) is the indicator of plant health focusing on how the plant absorbs and reflects the light based on different frequencies (Earth Observing System, n.d.). According to NASA, it states that the NDVI is a strong predictor for identifying the drought that is when water restricts plant growth, the relative NDVI, and density of the vegetation decreases (GIS Geography, 2020).

NDVI is calculated in accordance with the formula (GIS Geography, 2020),

NDVI = (NIR - RED) / (NIR + RED)

Where, NIR - reflection in the near-infrared spectrum.

RED - reflection in the red range of the spectrum.

According to the formula, it quantifies the vegetation by measuring the difference between near-infrared and red light divided by the sum of near-infrared and red light (GIS Geography, 2020).

The NDVI value always ranges from -1 to 1. Where -1 or negative value mostly represents the water bodies that absorb the near-infrared and reflect the red light (Earth Observing System, n.d.). The values which are close to 0 belong to the area with rocks or bare soil with less greenery. The values ranging from 0.2 to 0.3 belong to shrubs or meadows, whereas the values ranging from 0.6 to 0.8 belong to forests which indicates high vegetation index (Earth Observing System, n.d.).

Regions of Healthy vegetation reflects near-infrared and green light whereas regions which belong to water bodies or less vegetation or no vegetation absorbs the near-infrared and reflects the red light (Earth Observing System, n.d.).

1.4 Drainage System in Germany

According to the reports of BMU (Bundesministerium für Umwelt, Naturschutz und nukleare Sicherheit), almost 540,723 km of sewers collect the wastewater from various industries and households across Germany. With 10,000 water treatment facilities, almost 98 percent of the collected wastewater is treated with various biological and mechanical treatments, and the unwanted toxic nutrients are removed. In Germany, the water sewage system working plan is divided into two systems: one for wastewater drainage and another for stormwater drainage. Wastewater from households, industrial buildings are uniform and regular. Hence wastewater drainage is directly connected with wastewater treatment facilities. However, in stormwater drainage, a large water flow can happen due to storms or lousy climate anytime, so it is connected to lakes, streams, and rivers for redirection when necessary. (Bundesministerium für Umwelt 2021)

The German government charges for rainwater according to the area of the property from which rainwater is drained to the municipal system. The charges are normally per cubic meter of the water volume. (Olpe 2021; Berlin* 2021)

1.5 Hypothesis

The various factors influencing drought are used to propose the four hypotheses, and their validity is assessed through qualitative and quantitative analysis. Thus, this study postulates the following hypothesis:

Precipitation in an area is strongly affected by the altitude of that area.

The Soil Moisture Index determining the drought of a region is affected by precipitation.

The Soil Moisture Index determines the drought of a region is affected by the altitude of the region.

NDVI is affected by Soil Moisture Index.

Further, along with the analysis, an investigation is also done to find whether a water drainage system can be built based on existing precipitation data.

2. Study Area

Figure 1 represents the thirteen counties of interests chosen for the investigation. These counties are located within the Sauerland region which is located at south-eastern part of North Rhine Westphalia. It is a hilly region where majority of the land is covered with forests (Wikipedia, 2021).

As the part of investigation, all the thirteen counties of interest are distinguished with the geo package file that includes the administrative boundaries of all counties of North Rhine-Westphalia. The geo package file which is used for this study was downloaded from the git hub link https://github.com/rolfbecker/MIE_2.02_WS2020_II_Exam/tree/main/data/original/counties_municipalities_nrw.

This file is imported as a vector layer in the QGIS software for further investigations.

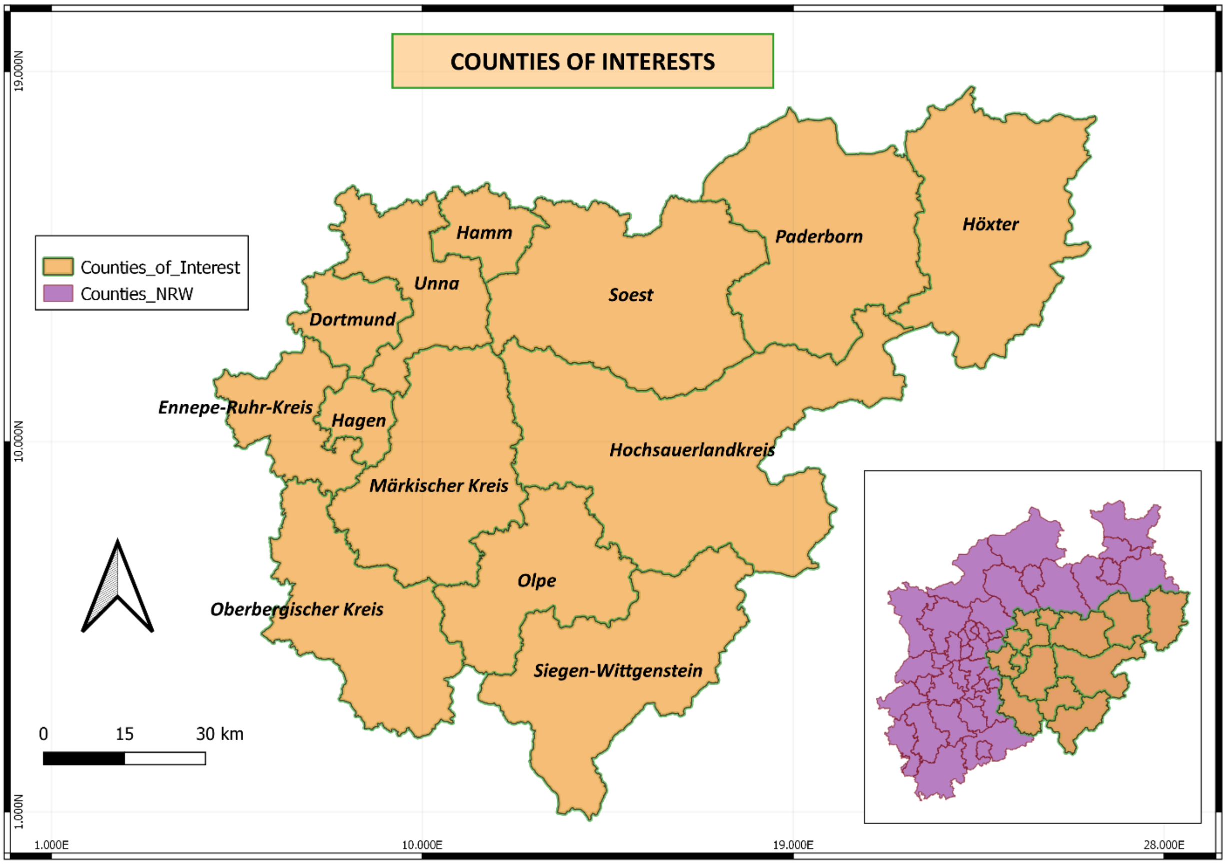


Figure 1: Counties of interest for the investigation

Figure 2 shows the layer defining 13 counties of interest overlayed with the Digital Terrain Model to understand the elevation data of these counties. The least elevation was observed at 44 meters above sea level, represented by blue color, and the maximum elevation was at 830.3 meters above sea level, represented by faded white color. It also shows the markers which represents the 33 precipitation station points which were active during the year 2017 identified within the Sauerland region of Germany. Precipitation station points were identified to fetch the historical precipitation data from the counties of interest, which will be used for further investigation.

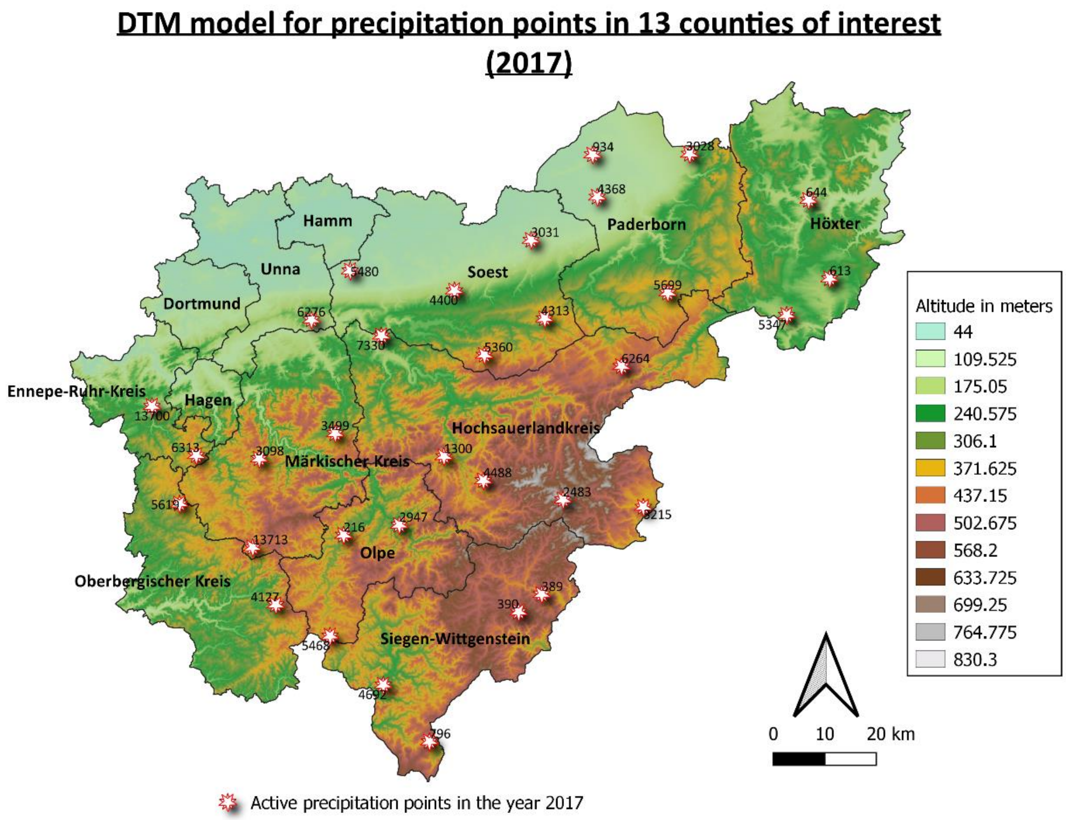


Figure 2: Weather Stations with Elevation Model

3. Methods and Materials

3.1 Regions of Interest (ROI)

The Regions of Interest includes regions within the county of Olpe which is in the south-eastern part of North Rhein Westphalia. This region is geographically rich in forest land as it covers the south-western part of the Sauerland mountains (Sauerland, 2021). These regions are used in the NDVI analysis.

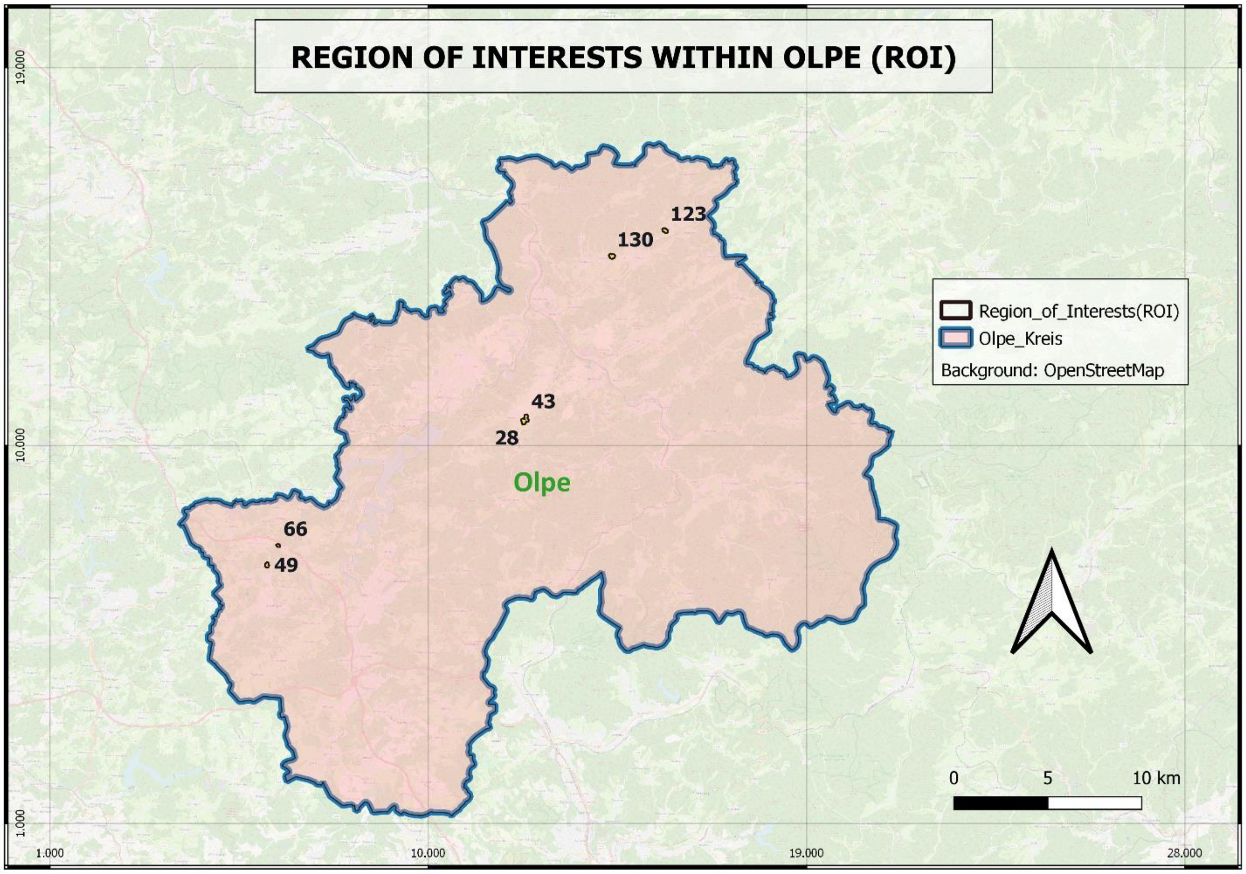


Figure 3:Region of Interests within OLPE

The polygons shown in Figure 3 represents the ROIs. They are categorized into two main types of fields namely Grünland and Ackerland. These regions belong to three different municipalities.

As part of the investigation, the six regions of interest are distinguished by loading geo package file defining regions of interest into QGIS software. The geo package file which is used for this study was downloaded from following git hub link https://github.com/rolfbecker/MIE_2.02_WS2020_II_Exam/tree/main/data/original/agricultural_areas_of_interest .

It displays all the ROIs within the Olpe region and further selected with only polygon IDs which are specific to this study. A new layer with the selected ROIs is created to distinguish the ROIs specific to this study from the original layer. The selected ROIs include six different regions with polygon IDs 28, 43, 49, 66, 123, and 130 as shown in the figure xx.

The layer with selected ROIs is further merged with the boundary of Olpe and Corine Land Cover, as discussed in the next section. This ensures that all the regions of interest are within the boundary of the Olpe. The boundary of Olpe is clipped from the provided geo package file of counties of interest which includes the administrative boundaries of the counties of North Rhein Westphalia. The Corine Land Cover is used and clipped with selected ROIs to determine the type of land cover used by each region of interest.

3.2 CORINE LAND COVER (CLC)

The inventory for Corine Land Cover was proposed and started in the year 1985. It currently consists of five datasets for the five years 1990, 2000, 2006, 2012, and 2018. The Corine Land Cover dataset projects visual information about the land use or land cover for the entire European region. The CLC datasets produced were distinguished as per the satellite images. It covers 44 unique classifications of land use (Copernicus Programme, 2021).

The dataset used was CLC-2018 to distinguish the land usage for the given region of interest. This dataset was downloaded from the Copernicus Land Monitoring Service website, maintained by European Environment Agency. This website allows access to the public for downloading the CLC datasets (Copernicus Programme, 2021).

The downloaded CLC-2018 dataset was processed using the QGIS software, and regions of interest were clipped, i.e., selected polygon IDs, to identify the type of land cover area for each region of interest.

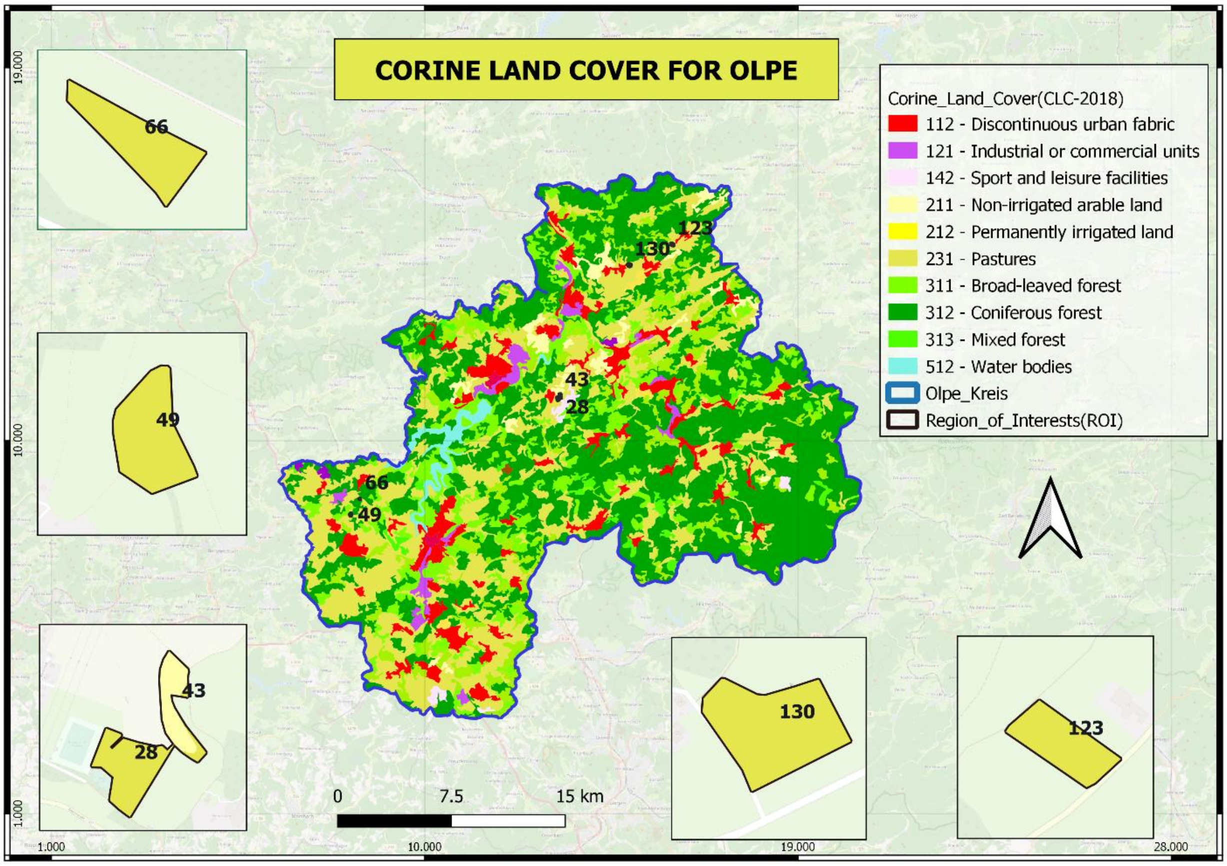


Figure 4: Corine Land Cover for OLPE

Figure 4 shows the Corine Land Cover for the Olpe Kreis. The numbers inside the county of Olpe represent the selected polygon ids that specify the region of interest. The land cover for the entire county of Olpe shows the different classifications of land use, but the land usage for the specified regions of interest that were important for this study belongs to two different classifications of land cover, as discussed below.

- **Pastures:** It is the permanent grassland for agriculture purposes typically used for grazing pastures or harvesting grass meadows to feed domestic animals such as cattle, sheep, etc. and it involves strong human disturbance. From the above map, the pastures completely cover the five regions of interest which include polygon Ids 28, 49, 66, 123, and 130. The polygon ID 43 is very slightly covered with pastures. It denotes the CLC classification Id 231 (Kosztra et al. 2017).
- **Non-Irrigated Arable Land:** It is the cultivated land under crop rotation where non-permanent crops grow with the use of rainwater. It includes the fields with intermittent

sprinkler irrigation and non-permanent irrigation devices to support effective rainfed agriculture. From the above map, the land used for the polygon ID 43 is majorly covered with non-irrigated arable land. It denotes the CLC classification Id 211 (Kosztra et al. 2017)

3.3 Data Collection

3.3.1 SMI Data

The Helmholtz Centre for Environmental Research – UFZ monitors the drought and soil moisture throughout Germany. It provides drought and soil moisture data based on the simulations using the mHM mesoscale Hydrological Model developed at UFZ (Marx, n.d.). The UFZ provides the data for the soil layers in general and the topsoil layer in specific. The uppermost 25cm depth of the soil layer is considered as the topsoil layer, and the soil moisture in the topsoil layer is responsible for the availability of water to the vegetation and indicates the fast change in SMI caused due to recent precipitation (Marx, n.d.). The soil moisture index of the topsoil layer was used in this study.

The original dataset was available for download in NetCDF format on the UFZ Drought Monitor Website. The Geoprocessed GeoTIFF image was downloaded from the following github link https://github.com/rolfbecker/MIE_2.02_WS2020_II_Exam/tree/main/data/original/SMI .

It contained the SMI data of the topsoil layer in GeoTIFF raster format for the 13 counties of interest for the dates 16.05.2017, 16.06.2017, 16.07.2017, and 16.08.2017. This SMI data is the data for the actual date, i.e., 16th of the month, and not the average value of SMI for the respective months.

3.3.2 Precipitation Data

The Deutscher Wetterdienst or DWD is meteorological service provider for Germany which is responsible to provide weather data and services for general use (Deutscher Wetterdienst. 2021). The data is available for access of general public through their ftp portal ftp://opendata.dwd.de/climate_environment/CDC/. The precipitation data is part of the climate data and is structured hierarchically in the portal for easy access. The data is available for both current dates and historical dates and is classified as daily, hourly, monthly, annually and 4 other classifications.

The period of interest for this study that is between April 16th 2017 and August 16th 2017 is classified under historical data. The hourly precipitation data is downloaded into python dataframe from ftp server using ftp library in python. Following processing is performed after the download:

- The data in dataframe is filtered for North Rhine-Westphalia state, since the 13 counties of interest lie in this state.
- The dates are filtered for the intended period of study. Only data between 16th April 2017 and 16th August 2017 is selected and rest of the data is ignored.
- This dataframe is taken as the main dataframe during the study and all the other dataframes are created using this dataframe.

The precipitation data is analysed in different ways and different dataframes are created post creation of main dataframe.

- A dataframe with daily total precipitation value is created for the stations in Olpe and HSK counties using pandas sum function.
- A dataframe for daily cumulative precipitation data is created for stations in Olpe and HSK using pandas cumsum function.

3.3.3 Altitude Data

Altitude is an important parameter for this study and its role in altering precipitation and SMI is analyzed for the counties of interest. There were two sources from where altitude data was obtained for the stations of interest:

- DWD portal metadata

The altitude data for stations was downloaded along with the precipitation data from DWP portal. The data contained the altitude data for all the precipitation stations in Germany. The data was filtered for the 13 counties of interest and the precipitation stations in the underlying area using python. These values were the actual and exact altitude value for the precipitation stations.

- QGIS- Sampling of DTM raster values

The Digital Elevation Model or DTM contains the elevation data for the region. A 50m resolution DTM model was used for this study. The link to download the DTM model required for the study was provided by Prof. Dr. Rolf Becker in the GitHub repository. The link to the GitHub recovery is: https://github.com/rolfbecker/MIE_2.02_WS2020_II_Exam/blob/main/doc/ex1.md.

This link redirects to opendata portal which is an open-source portal for various data. The 50m DTM model was downloaded for North Rhine-Westphalia region in GeoTIFF raster format and uploaded to QGIS as a raster layer.

This model was sampled for raster values using QGIS processing toolbox to obtain the altitude for the 33 stations in the 13 counties of interest. It is important to note that these values are not the exact values of altitude for the stations. The DTM model with a 50m resolution has grid spacing of 50m, which means that all the regions under the same 50m grid would be having same altitude. Thus, this value is not precise value of altitude for the precipitation stations but the altitude value for the 50m grid in which respective stations reside.

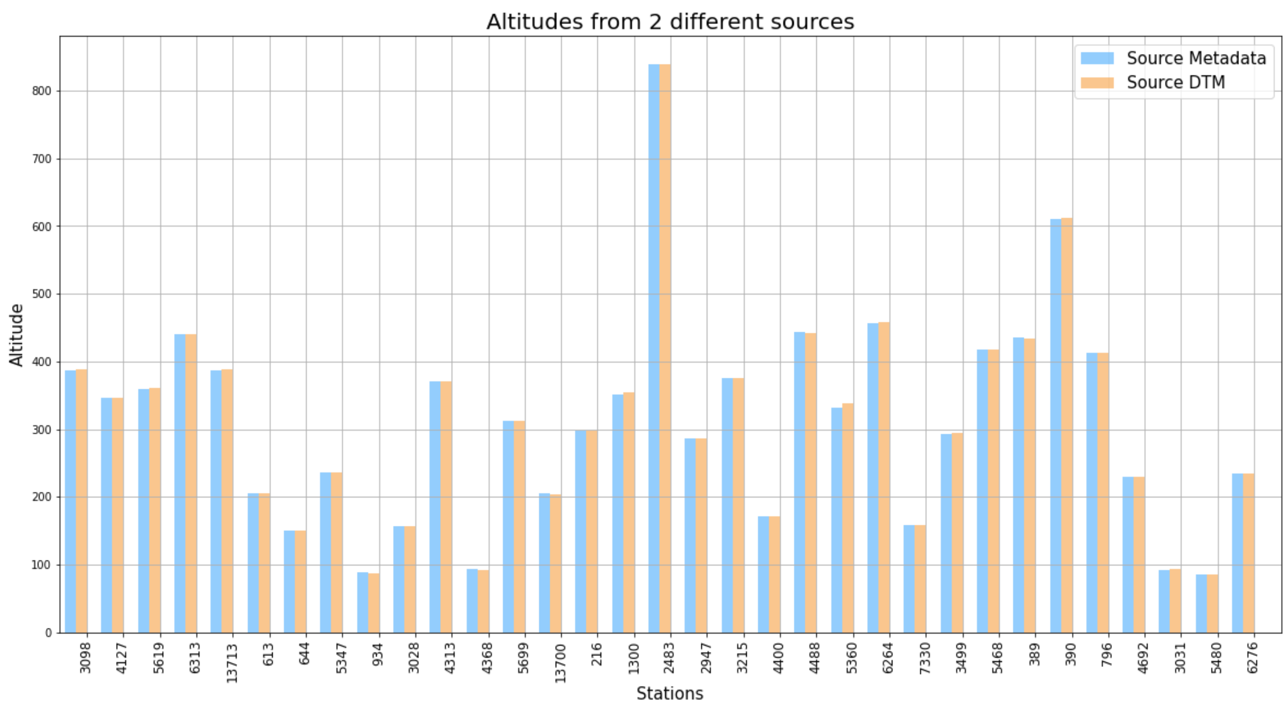


Figure 5: Altitude data comparison of two sources

Comparing the altitudes, it was observed that, there was a small difference in data for altitudes obtained from two different sources for the 33 stations in the 13 counties of interest. Altitude obtained from the metadata source was considered for this study as it was more precise and specific to the precipitation stations.

3.3.4 NDVI Data

For analyzing Normalized Difference Vegetation Index (NDVI), the satellite imagery data used belongs to the Sentinel product of level-2A. The reason for choosing Sentinel-2 is because it provides the data for different bands which include band-8 and band-4 representing near-infrared and red light that are required for calculating the NDVI data. The satellite imagery datasets were downloaded from the Copernicus open access hub website from the following link https://scihub.copernicus.eu/,

It allows the complete access to the data for the public for downloading all the sentinel products. The satellite imagery data was downloaded for the three different sensing dates such as 2017.05.10, 2017.06.19, and 2017.07.19. These datasets include different resolutions of imagery such as 10m, 20m and 60m. These dates have low cloud coverage which does not affect the NDVI calculation. All the three sensing dates belong to the year 2017.

To calculate the NDVI for each sensing date. A 10m resolution imagery dataset is used for this study. It is imported as raster layers into QGIS software. The raster is further clipped using mask layer for band-4 and band-8 with the selected ROIs separately for each sense date and performed raster calculation using NDVI formula with the extracted mask layers of band 8 and band 4, for distinguishing the NDVI. Where band 8 is considered as near infrared and band 4 is considered as red light, as shown in below table 1.

The Table 1 provides the complete list of all imagery bands provided by the Sentinel-2 Level - 2A satellite (Copernicus, n.d.) (Gordana and Avdan, Ugur. 2017).

.

Band	Band	Spectral Band	Central Wavelength (nm)	Spatial Resolution (m)	Objective
B1	VNIR	Coastal Aerosol	443	60	Aerosol Correction
B2		Blue	490	10	Aerosol Correction, Land Measurement Band
B3		Green	560	10	Land Measurement Band
B4		Red	665	10	Land Measurement Band
B5		Red Edge 1	705	20	Land Measurement Band
B6		Red Edge 2	740	20	Land Measurement Band
B7		Red Edge 3	783	20	Land Measurement Band
B8		Near-Infrared	842	10	Water Vapor Correction, Land Measurement Band
B8A		Near-Infrared Narrow	865	20	Water Vapor Correction, Land Measurement Band
B9		Water Vapor	945	60	Water Vapor Correction
B10	SWIR	Shortwave infrared - CIRRUS	1380	60	Cirrus Detection
B11		Shortwave infrared -1	1610	20	Land Measurement Band
B12		Shortwave infrared -2	2190	20	Aerosol Correction, Land Measurement Band

Table 1: List of Sentinel –2 Level 2A Spectral Bands and wavelengths.

Source for above table: https://spaceflight101.com/copernicus/sentinel-2/

Source for Sentinel 2 data: https://scihub.copernicus.eu/

3.4 Interpolation

3.4.1 Nearest Neighbor Interpolation

Let location be l, value to interpolate be v, and the total number of locations be n. Interpolation estimates the unknown value v at a new location l from a given set of known values $\{v_1, v_2, v_3,$

...v_n} at locations {l_1, l_2, l_3, ...l_n}. The Nearest Neighbour Interpolation technique uses the value v_i that is closest to l. Hence, the algorithm tries to determine i such that $|l_i-l|$ is minimized, then the estimate of v is v_i.

The nearest points within the area to l_i forms a polygon which is called Thiessen polygon or Voronoi polygon. This polygon belongs to a region comprising all points nearest to l_i than to any other l. The issue here is that the interpolated areas are discontinuous in this case (Hartmann, K., Krois, J., Waske, B. 2018).

In simple terms, the known value of the nearest location found within a search radius will be assigned to all unknown points within the search radius.

3.4.2 Inverse Distance Weighted Interpolation

The Inverse Distance Weighted (IDW) algorithm is used to interpolate spatial data and find the value in the unknown points based on distance weighting. The IDW formula is as follows:

$$v^* = (w_1v_1+w_2v_2+w_3v_3+\ldots+w_nv_n) / (w_1+w_2+w_3+\ldots+w_n)$$

$$w_i = (1)/(d^{P}_{iv^*})$$

where v^* represents the unknown data value,

n is the number of locations or points in the chosen search radius,

w_i represents the weight (The weight is calculated as the inverse distance of an unknown point to each known point value),

d_{iv^*} is the distance from the known point to the unknown point,

P is the power, a control parameter. There is no predefined value for P, optimal power value can be identified by trial-and-error method (Spatial Interpolation with Inverse Distance Weighting (IDW) Method Explained, 2021).

In simple terms, to determine the value of an unknown point, the weights will be calculated and assigned to each known point based on its distance from the unknown point. Hence, the farther known points have less impact while closer known points will have more impact on the unknown value.

3.5 Tools

- Jupyterlab is a free open-source interactive user-interface which acts as a development environment in multiple programming languages. It is used for the python coding part, visualizations and data used for the analysis performed. (Project Jupyter, 2021)
- Python is a dynamic, cross-platform programming language that enables modularity and code reuse by supporting several modules and packages. It is a high-level object-oriented programming language. (Org, Python,2021)

3.5.1 QGIS

Quantum Geographic Information System, also known as QGIS, is an open-source desktop application that enables to analyze spatial and geographic data. It helps to integrate layers of data into visualizations using maps and 3D models. The software also supports the import of raster and vector layers. It also follows a plugin-based architecture that supports various core and external plugins to make the geo analysis easier. (Welcome to the QGIS project!, 2021)

- **Time Manager -** Time Manager is a plug-in used in QGIS to analyze spatio-temporal data. This plugin allows time control of data in QGIS. Input to the time manager is layer containing data and time series column. The time manager allows data represented on QGIS layer to animate over time and each frame is exported by the time manager as a .png output file.
- **Blender -** Blender is a free open-source software developed by Blender Foundation community. It is used for image rendering, animation, simulation and many other applications. Blender can be used to create an animation or video from image sequences and encode it into standard video formats. [https://www.blender.org/about/]

4. Results and Discussion

After the data was gathered, they were processed from raw data to the data required for different study. The transformed data was used for the visualization and comparison of parameters like precipitation, altitude and NDVI. While discussing this, the intention was focused on drought in the 13 counties of interest indicated by the SMI values.

4.1 SMI Analysis

The change in SMI of top soil layer for 13 counties of interest was observed and analyzed by plotting 4 maps representing SMI data for 16th of each month from May 2017 to August 2017.

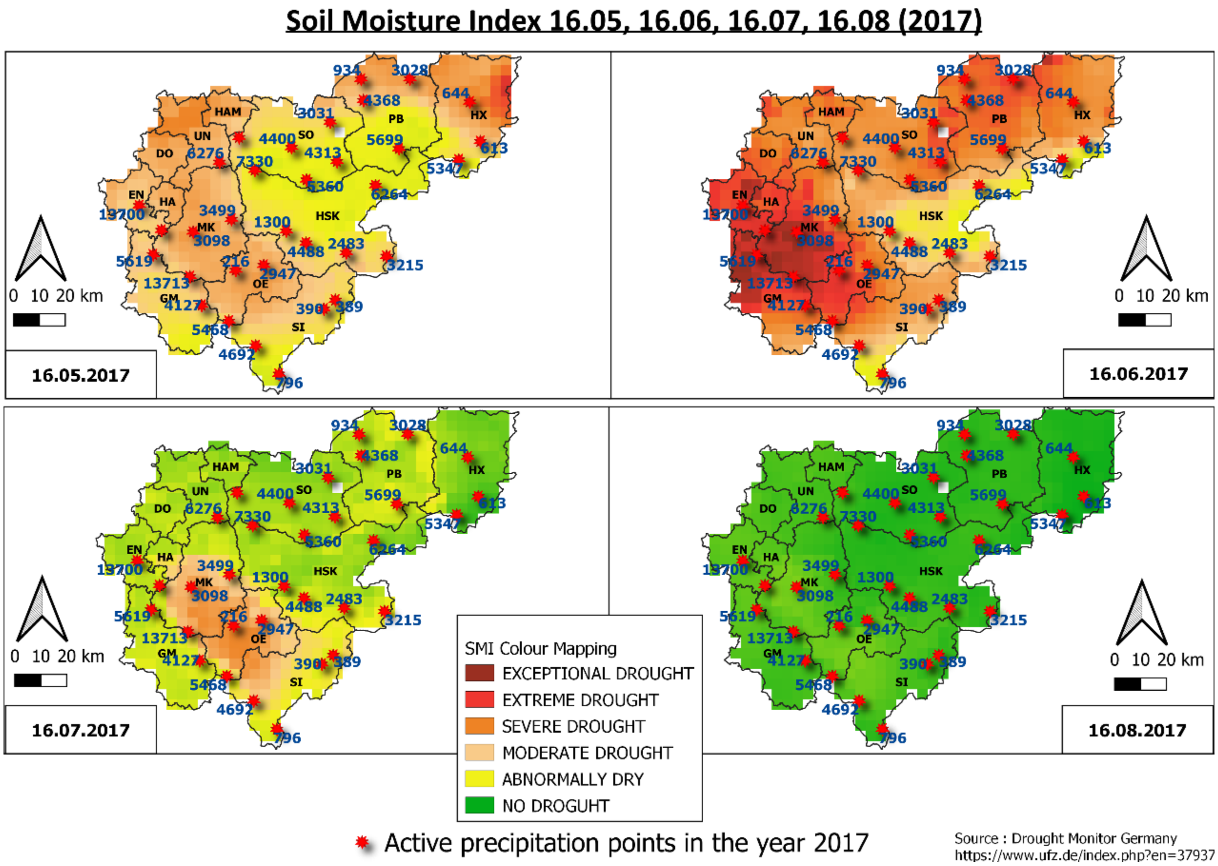


Figure 6:Soil moisture index analysis for year 2017

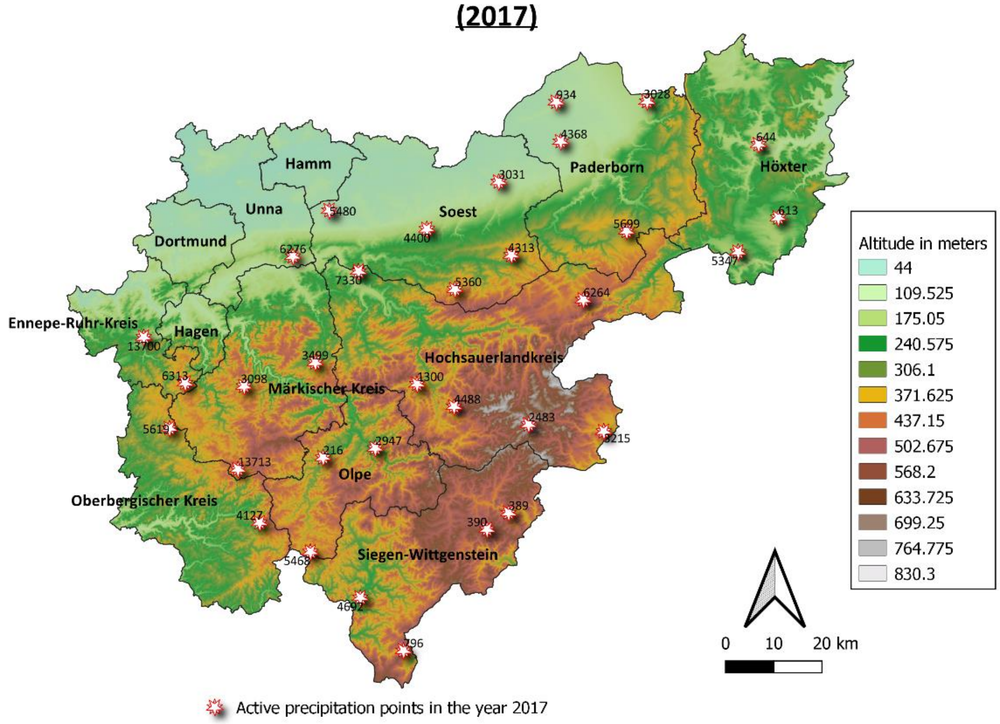
The soil moisture index for the topsoil layer in 13 counties of interest is shown in figure 5. The date of interest was taken as 16th of every month between May 2017 and August 2017 to understand the change in SMI. The soil moisture index data was mapped to 4 maps representing 4 months. It was observed that 16th of June had a minimum soil moisture index indicating drought in major areas belonging to counties of interest. The drought was observed to evade completely on 16th August 2017. The data on 16th of July was observed as transition period from drought to no drought regions between June and August 2017. Most of the counties had already recovered from the drought by 16th of August 2017 except Olpe, Markischer-Kreis and parts of Siegen-Wittgenstein.

It was interesting to identify what caused this fluctuation in SMI during the period. Using the data obtained from DWD portal, precipitation and altitude data were analyzed to see if it had a role in this behavior of change in SMI over 4 months period.

4.2 Altitude Analysis

The precipitation station points in 13 counties of interest were plotted. Before analyzing the effect of precipitation on SMI, the altitude data was analyzed for the 13 counties of interest. The altitude of the stations was realized using hill shade DTM model and color ramp was used to differentiate different range of altitudes. The color band for altitude varied from 44 meters (light blue) to 830.3 meters (light grey).

DTM model for precipitation points in 13 counties of interest (2017)



The counties of Hochsauerlandkreis, Markischer-Kreis, Olpe and Siegen-Wittgenstein were found to be at higher altitudes compared to other counties. HSK was found to have highest altitude region among all the counties. The region surrounding station 2483 which was at the highest altitude among all stations was observed to not have faced drought and only caught abnormally dry SMI in the month of June 2017. Also, the station 5468 in OLPE region recovered from severe drought in the month of July 2017 better than other 2 stations in OLPE. This data suggested that SMI is better at higher elevation. However, when the stations from other regions were analyzed, it was seen that some stations at lower elevation recovered from the drought faster than higher elevation stations. For example, station 934 belonging to Paderborn county, which was one of the lowest altitude stations in all of 13 counties had better SMI in month of July compared to the relatively higher altitude stations of 216 and 2947 belonging to Olpe county. This shows that there is ambiguity in relationship between SMI and altitude.

Further analyzing in-depth, the regions Hoschsauerlandkreis and Olpe form adjacent counties but was observed to have a vast difference of SMI in the month of July 2017. With the ambiguity in altitude and SMI relationship, precipitation data was analyzed to understand the behavioral trend of SMI in the period between 16th of May and 16th of August.

4.3 Precipitation Analysis

A general temporal analysis of precipitation time series data was done for all of NRW and data was fed to the time manager in QGIS to observe the change in precipitation during the time period of 16th of June 2017 to 16th of July 2017. The color ramp in symbology tab of layer properties was used to define the different color ramps for different precipitation value ranges. The color ramp ranges from precipitation value of 0 to 33.6mm. The video created from time manager output frames was investigated for trends in precipitation.

Not much precipitation was observed until 22nd of June 2017 all over NRW. However, a heavy rainfall trend was seen all over NRW on 22nd of June. The rainfall in lower amount continued till 26th of June in some parts of the region. Afterwards, the rainfall was continued only after 28th of June. Heavier and consistent rainfall was observed in some counties like Hoschsauerlandkreis after 1st of July but some counties like Hoschsauerlandkreis received lesser and inconsistent rainfall. A trend of inconsistent precipitation continued till 9th of July after which a consistent and heavy precipitation was observed all over NRW. This partially explains why the 13 counties of interest recovered from drought only after July 2017.

For an in-depth study, the precipitation data was further closely observed by narrowing down the study to two counties and extracting strong evidence regarding the effect of precipitation on SMI. The SMI for the regions Hoschsauerlandkreis and Olpe had a vast and intriguing difference considering they are adjacent set of counties. The precipitation data was extracted and analyzed to understand the reason behind the difference in SMIs for the two counties. The two counties were found to have a total of 9 active precipitation stations in 2017. The recorded data from these precipitation stations was studied to understand if precipitation had a role in difference of SMI during the period of interest.

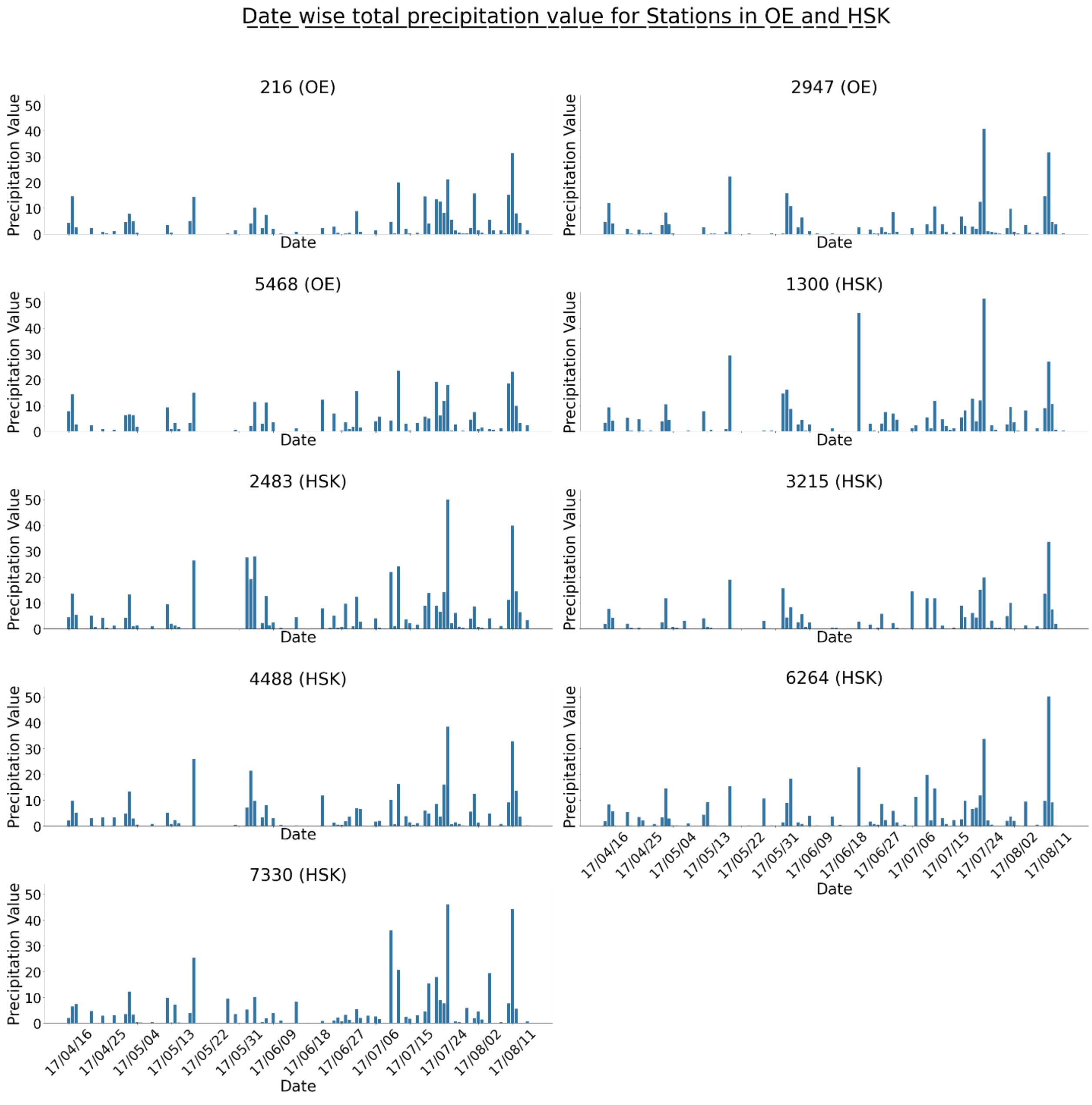


Figure 7: Total precipitation value for stations in OE and HSK

The bar plots in figure 6 show daily precipitation recorded by 9 stations between 16th of April to 16th of August 2017. The highest precipitation in Olpe was observed at station 2947 with value 40.8mm and the highest precipitation in Hoschsauerlandkreis region was observed to be 51.3mm in station 1300. There was negligible precipitation found between 13th of May and 22nd of May 2017 for both Olpe and Hoschsauerlandkreis counties.

To further analyze the shift in precipitation value during the period, cumulative precipitation between the period 16th of April to 16th of August was plotted for 9 stations in OE and HSK.

Cumulative precipitation between the period 16th of April to 16th of August is plotted for 9 stations in OE and HSK to further analyze the shift in precipitation value during the period.

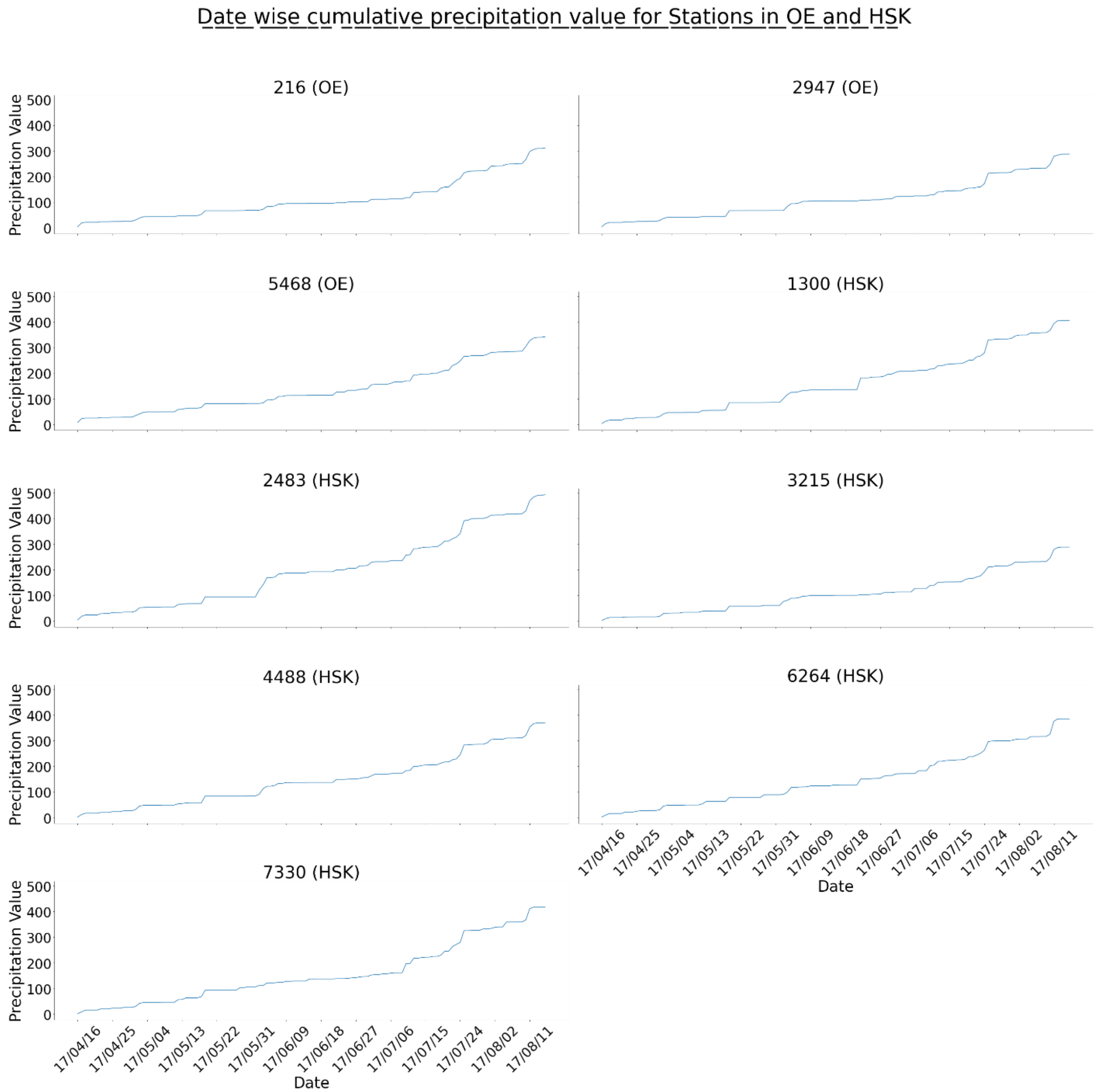


Figure 8: Cumulative precipitation comparison for OE and HSK

It was observed from figure xx that stations 216 and 2947 belonging to OE county had less steady rise in the precipitation value until mid of July 2017. This means that the amount of precipitation was not consistent over the period and only had a small number of precipitations during the period. On the contrary, station 5468 belonging to Olpe region had a steady rise in precipitation value since the beginning of June 2017 until 16th of August 2017. This fairly supports the evidence why stations 5468 belonging to Olpe region had recovered from drought by mid of July 2017, whereas stations 246 and 2947 were still facing severe drought.

A surprising observation is that station 2947 from Olpe had a higher maximum precipitation compared to station 3215 from HSK county. However, it was observed that station 3215 recovered from drought sooner than the station 2947. It could be noted that although station 2947 had a very high rainfall for a short period compared to rainfall in station 3215, there was no rainfall consistency observed for the station 2947. Station 3215 had a more consistent and steadier rise in precipitation during the observed period. This explains that consistency in precipitation also plays a role in SMI for the region.

Further investigating the effect of precipitation on SMI, it was observed from the figure 7 that all the stations in Hoschsauerlandkreis had significantly better rise in cumulative precipitation. Also, stations in Hoschsauerlandkreis had already recovered from drought by the mid of July 2017.

County	Max_prec	Mean	No. of Days
OE	40.8	2.569672	122.0
HSK	51.3	3.222541	122.0

Table 2: OE and HSK stations SMI and Precipitation table

From the table 2, we can see that the average precipitation for Olpe is significantly less compared to average precipitation in Hoschsauerlandkreis for a 122 days period from 16th of April to 16th of August. This further suggests that, precipitation had a role in the difference in SMI index between Olpe and Hoschsauerlandkreis in the month of July 2017.

From the above investigation, it is evident that precipitation had role in SMI difference in Olpe and Hoschsauerlandkreis, although it is naive to conclude that only precipitation plays role in SMI.

4.4 Correlation Altitude and Precipitation

It was found that precipitation effected the SMI but altitude didn’t play a direct role in altering SMI over the period of 4 months for the stations present in 13 counties of interest. Further, it was analyzed if altitude of the stations had any role in total precipitation at stations.

Thus, a plot for relationship between altitude and total precipitation for the period from 16th April 2017 to 16th August 2017 for the 33 stations in Olpe and Hoschsauerlandkreis was created. The relationship was measured using the r squared value and p value for correlation.

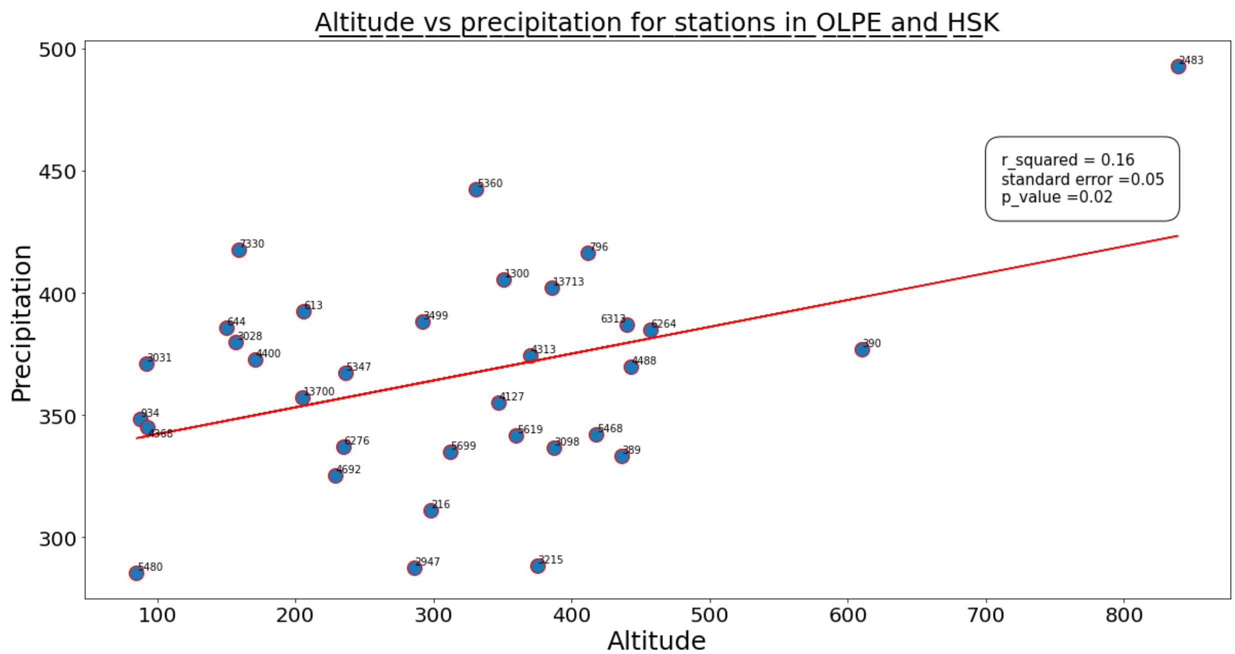


Figure 9: Altitude vs Precipitation for stations in OE and HSK

It was observed that, for the stations in Olpe and Hoschsauerlandkreis, the correlation of altitude and precipitation was low with an r squared value of 0.16 inferring weak positive correlation. The p value was also low with a value of 0.02 strongly suggesting to reject the null hypothesis that altitude is directly correlated with the precipitation.

Further, the total precipitation for the period of interest was plotted for all 13 counties of interest to obtain a strong evidence to disprove the null hypothesis that amount of precipitation depended on the altitude of the region.

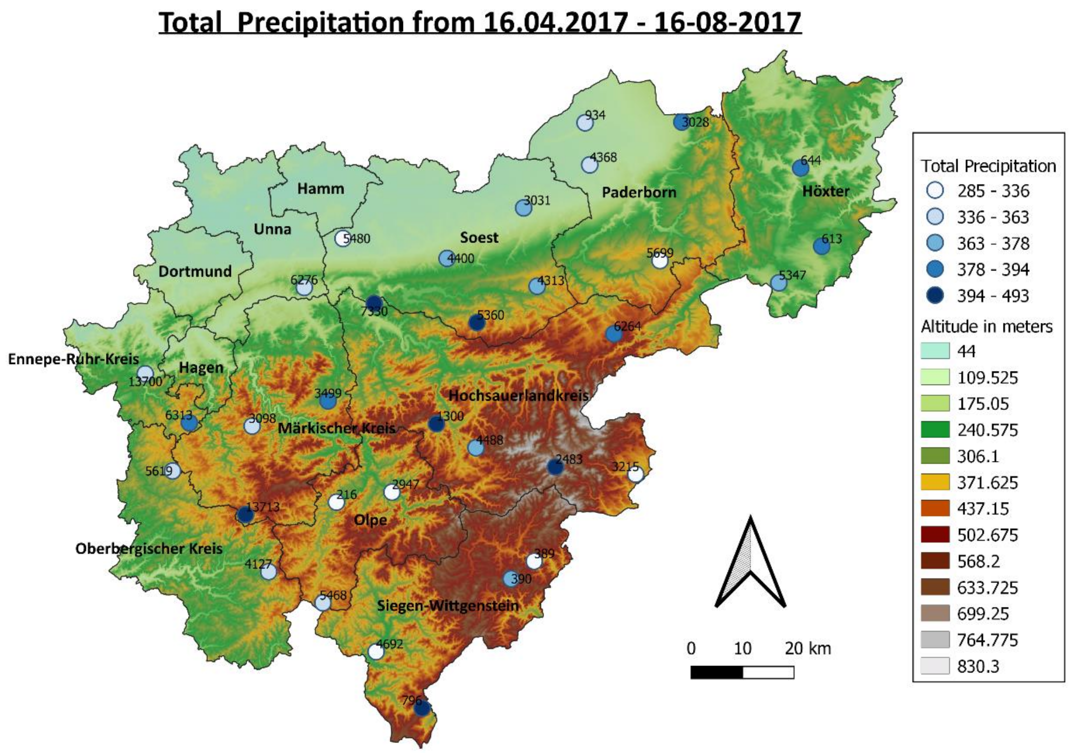


Figure 10: Total precipitation with altitude from 16th April to 16th August of year 2017

It was observed from the figure xx that total precipitation was not complying with the hypothesis. For example, a low altitude station at 159 meters above sea level with station id 7330 had higher total precipitation compared to a higher altitude station at 298 meters above sea level with station id 216. On the other hand, station 2483, which is at highest altitude compared to all the other 33 stations, received the highest precipitation. B

Thus, it could be concluded that for the stations in Olpe and Hoschsauerlandkreis, altitude had minimum role in amount of precipitation during the period 16th of April to 16th of August.

4.5 Discussions

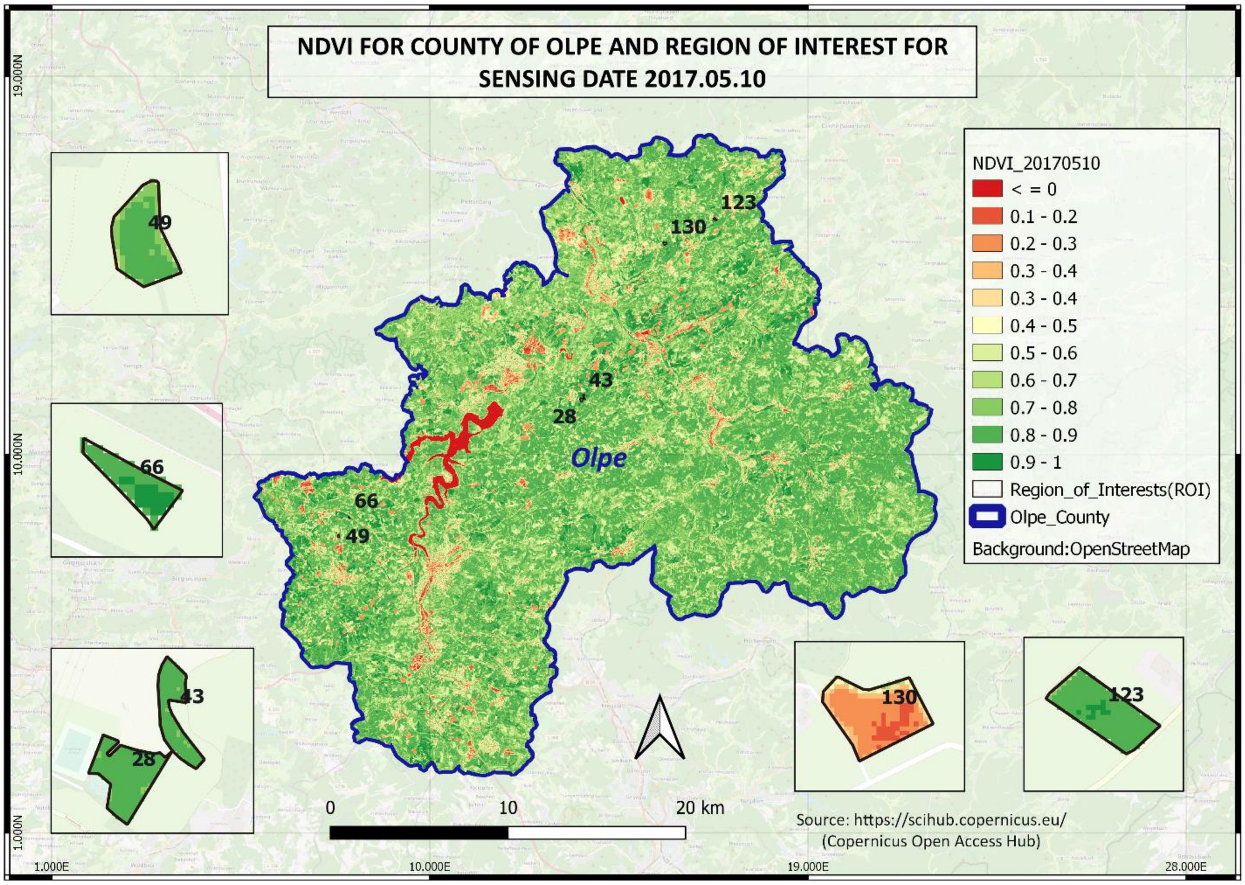
It is observed that counties Hoschsauerlandkreis and Olpe have a strong difference of SMI during the first two months. The precipitation data shows that there was less rainfall between 16th of April and end of June 2017 and counties were facing drought explaining the low SMI. However, the amount of rainfall started to increase at the end of June for HSK and after mid-July for the Olpe county. Thus, the last two months that is July and August 2017 is the period of recovery from drought. This confirms that precipitation has role in SMI and thus effects the drought in area.

Another interesting observation is that county Hoschsauerlandkreis is less effected compared to other 13 counties. The observation from figure xx is that Hoschsauerlandkreis lies at a much higher altitude compared to other 12 counties. This can also be supported by the fact that part of county SE is at higher elevations. It can be observed that for the county SE, the higher elevation places have recovered faster compared to the lower elevation areas. Thus, it can be concluded that altitude plays a role in soil moisture index and was reason for the recovery in the drought and SMI values for Hoschsauerlandkreis.

However, looking at the data from Olpe, even though Olpe is at a higher elevation compared to many counties, it was found that Olpe had less SMI on 16th July 2017 compared to other counties at a lower elevation. Thus, it is practical to infer that SMI for an area are not only dependent on altitude and precipitation, but is also dependent on other factors. Some of the factors that can be noted are temperature, soil quality, soil texture and NDVI. While other factors are out of scope of this study, the following section narrows down the study for relation between NDVI and SMI.

4.6 NDVI Calculation

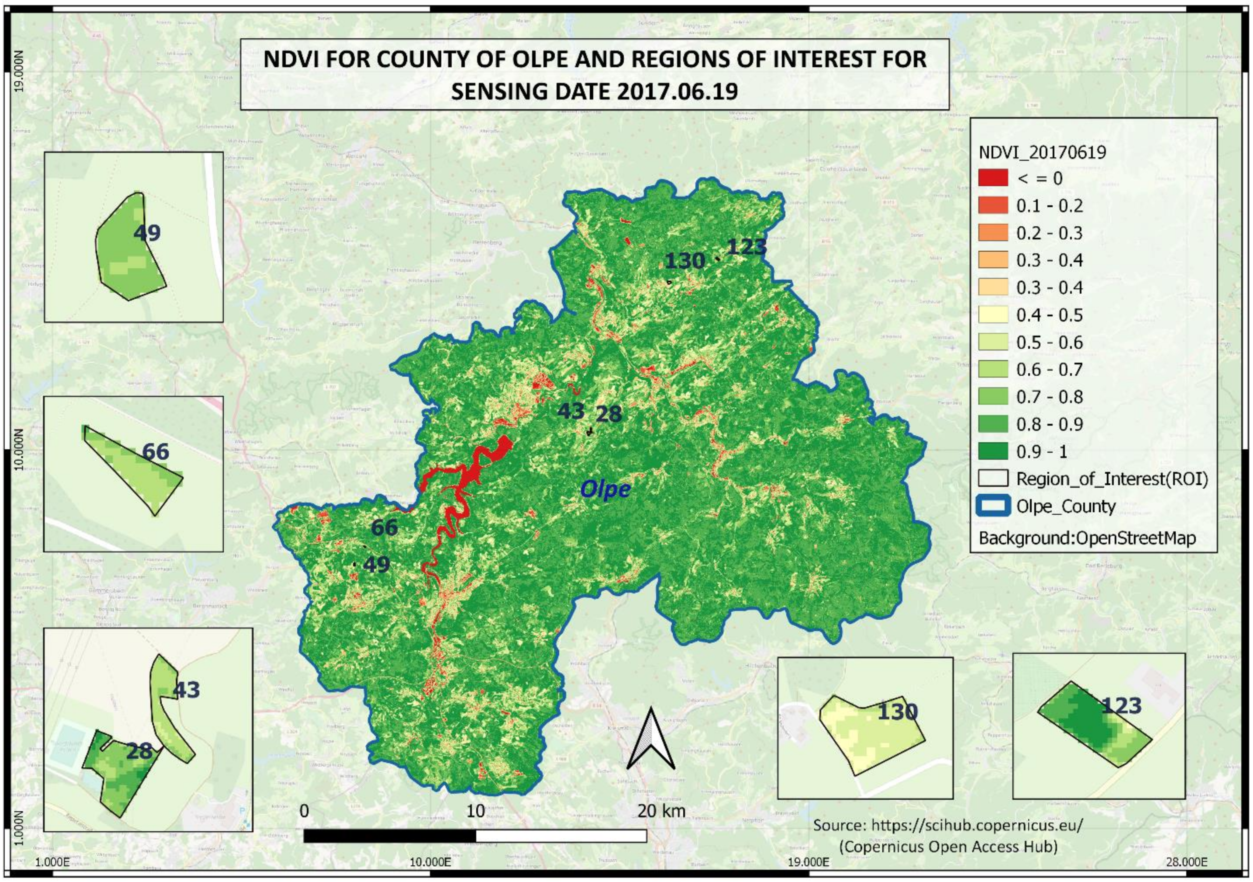
4.6.1 NDVI Analysis for Sensing Date 2017.05.10



As the part of Investigation, from the above figure, the satellite imagery data which is collected is from Sentinel product of level 2A for the date 10.05.2017. The satellite imagery data is downloaded for the requested sensing date from the Copernicus open access hub website. The imagery data is imported as raster layers in QGIS. Further performed raster extraction with clip by mask layer for two different bands that are band-04 and band-08 raster separately with selected regions of interest and performed raster calculation using NDVI formula for determining the NDVI.

For calculating NDVI especially for the plant cover the negative values does not provide much information and the range taken is between 0 to 1, the resulting values display different representation of colors for region of interests. From the above picture, it is observed that the vegetation index for all the polygon Ids except Id 130 is closer to 1 and shows the healthier vegetation result for the month of May in 2017. But the polygon Id 130 shows the lowest value for vegetation among other polygons in this month which is recorded as 0.20 that depicts the drought condition in this region.

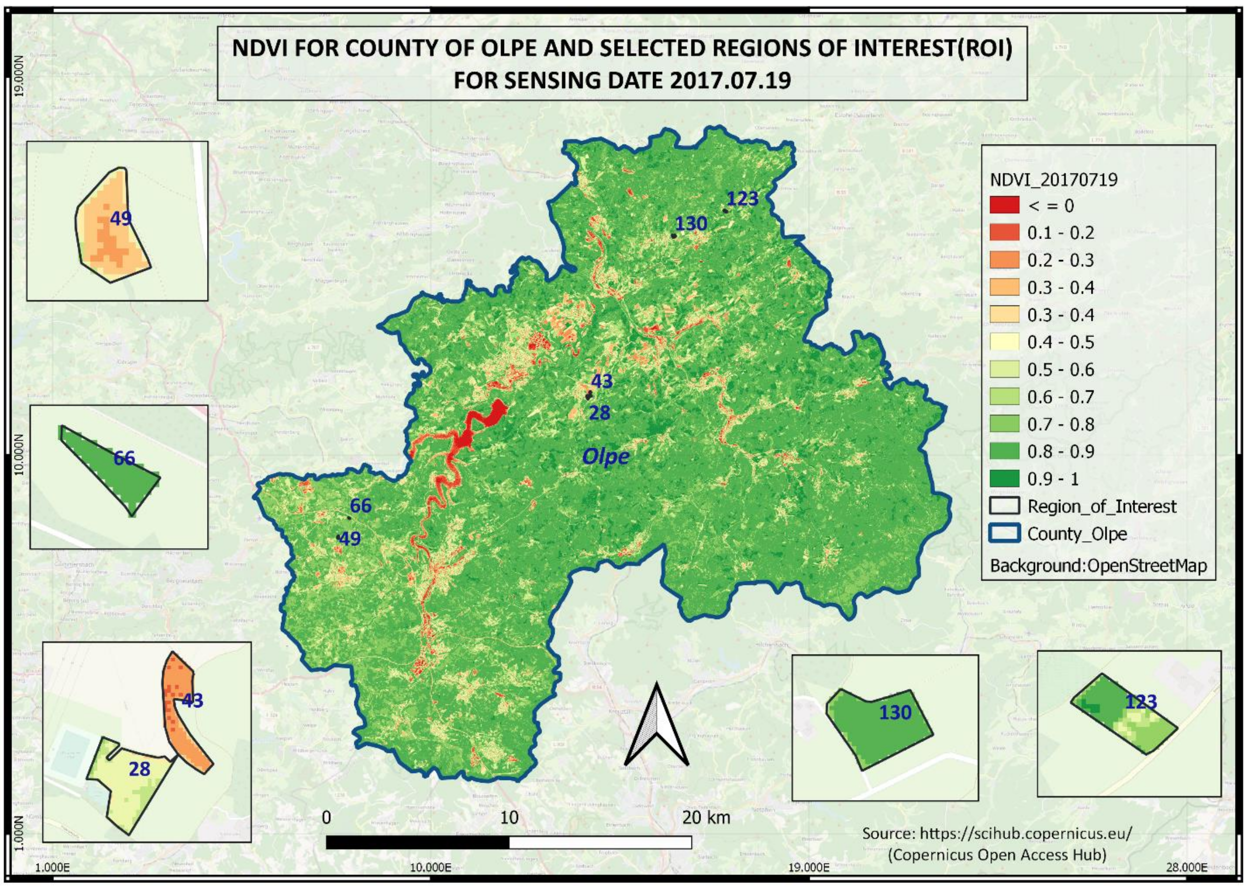
4.6.2 NDVI Analysis for Sensing Date 19.06.2017



As the part of Investigation, from the above figure, the satellite imagery data which is collected is from Sentinel product of level 2A for the date 19.06.2017. The satellite imagery data is downloaded for the requested sensing date from the Copernicus open access hub website, which provides complete open access data of sentinel products. The imagery data is imported as raster layers in QGIS. Further performed raster extraction with clip by mask layer separately for two different bands that are band-04 and band-08 raster with selected regions of interest and performed raster calculation for determining the NDVI.

For calculating NDVI especially for the plant cover the negative values does not provide much information and the range taken is between 0 to 1, the resulting values display different representation of colors for region of interests. From the above picture, it is observed that the vegetation index for the polygon Ids 28, 49 and 123 showed good results. However, there is a slight decrease in vegetation index values for these polygons when compared to previous month. On the other hand, the regions with polygon Ids 43, 66 and 130 are having the moderate vegetation index for this month. But the polygon Ids 43 and 66 share a slight decrease in NDVI, in contrast the polygon Id 130 show improved results of vegetation when compared to previous month.

4.6.3 NDVI Analysis for Sensing Date 2017.07.19



As the part of Investigation, from the above figure, the satellite imagery data which is collected is from Sentinel product of level 2A for the date 19.07.2017. The satellite imagery data is downloaded for the requested sensing date from the Copernicus open access hub website, which provides complete open access data of sentinel products. The imagery data is imported as raster layers in QGIS. Further performed raster extraction with clip by mask layer separately for two different bands that are band-04 and band-08 raster with selected regions of interest and performed raster calculation for determining the NDVI.

For calculating NDVI especially for the plant cover the negative values does not provide much information and the range taken is between 0 to 1, the resulting values display different representation of colors for region of interests. From the above picture, it is observed that the regions with polygon Ids 66, 123 and 130 have good vegetation index in this month out of which the Ids 66 and 130 show improved results of vegetation when compared to last month. On the other side the regions with polygon Ids 28, 43 and 49 have less vegetation index in this month and it is observed that there is continuous decrease in vegetation since past two months (May and June, 2017). The polygon Id 43 share the lowest value of NDVI for this month among other polygons which is recorded as 0.24 and it depicts the drought condition in this region when compared to previous two months.

4.6.4 NDVI Conclusion

Region of Interests	Field Type	Land Cover	NDVI Mean 10-05-2017	NDVI Mean 19-06-2017	NDVI Mean 19-07-2017
28	Grünland	Pastures	0.8426466	0.7817242	0.5664058
43	Ackerland	Non-irrigated arable land	0.8392810	0.6549072	0.2416408
49	Ackerland	Pastures	0.8023854	0.7379732	0.3647198
66	Grünland	pastures	0.8870917	0.6898207	0.8408192
123	Grünland	Pastures	0.8752349	0.8247272	0.7839749
130	Ackerland	Pastures	0.2045788	0.5343183	0.8433808

As the part of investigation, from the above table of NDVI Analysis it is observed that there are significant changes in the NDVI for the region of interest with polygon Ids 28, 43, 49 and 130 when compared among three different sensing dates 10.05.2017, 19.06.2017 and 19.07.2017. The region of interests with polygon Ids 28, 43 and 49 showed healthy vegetation for the first sensing date and there is relatively steady decrease in the vegetation index for the second sensing date. Further it is observed for the third sensing date, there was a second consecutive decrease in the mean NDVI values showing less vegetation when compared to past two months. The mean NDVI values for these polygons show continuous decrease since past two months that depicts the little drought condition especially for the polygons 43 and 49 as of third sensing date.

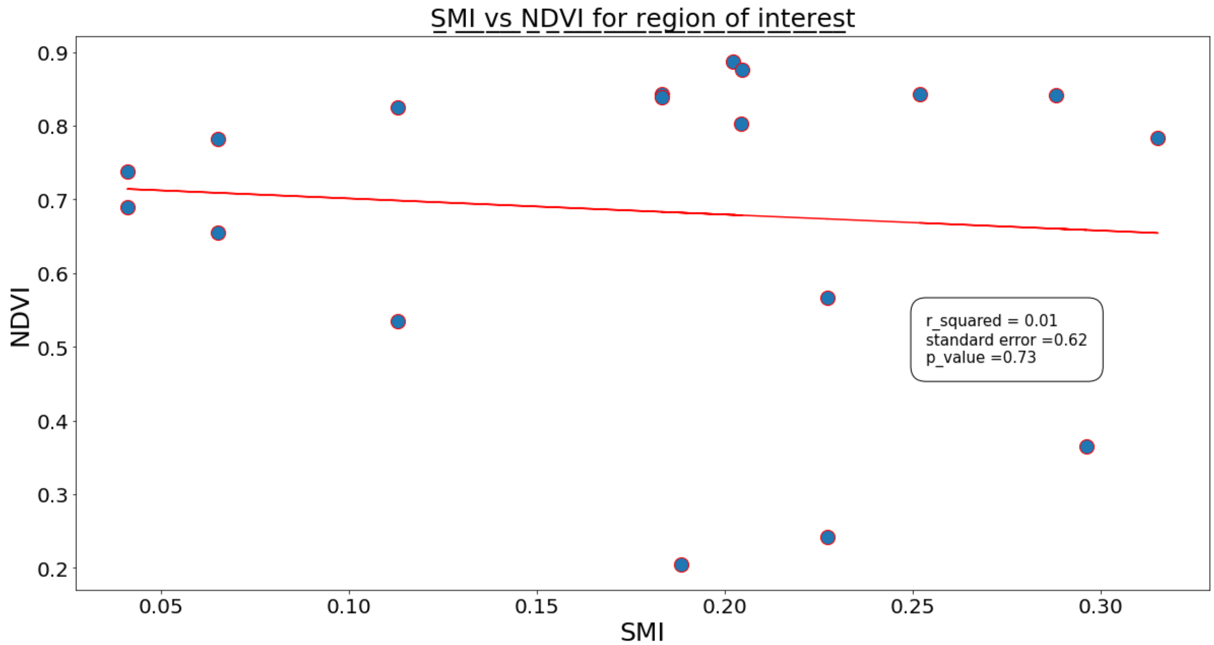
On the contrast, the region of interest with polygon id 130 showed poor results of vegetation index in the first sensing date. For the second sensing date it showed improved results of NDVI with steady increase when compared to previous month. Further it is observed for the third sensing date, there was a second consecutive increase in NDVI that represents good and healthy vegetation when compared to previous month. The mean NDVI values for this polygon show a continuous increase in vegetation index since past two months which depicts the healthy vegetation condition as of third sensing date.

From the above table, there are significant changes observed in NDVI values depending on the field type. For the field type "Ackerland", it showed drastic changes in NDVI values for the three polygon Ids 43, 49 and 130 for three sensing dates. The polygon Ids 43 and 49 showed continuous decrease in NDVI values resulting less vegetation and polygon Id 130 showed continuous increase in NDVI values resulting healthy vegetation for the third sensing date when compared with previous two sensing dates. But in case of field type "Grünland" it is observed slight changes in the NDVI values for polygons 28, 66 and 123 for three sensing dates. However, the polygon Ids 66 and 123 showed good vegetation index throughout the three sensing dates excluding minor changes. Thus, it can be concluded that the NDVI values show significant changes depending on field type over the time period.

4.6.5 Correlation NDVI and SMI

After obtaining NDVI values, SMI data was calculated for 6 ROIs. The NDVI values were calculated for the dates 10.05.2017, 19.06.2017 and 19.07.2017. Thus, the SMI values were calculated for precipitation stations layer by sampling the raster data from SMI raster maps belonging to the dates 16.05.2017, 16.06.2017 and 16.07.2017 respectively. Although, it is not an accurate way to calculate values from different dates, the previous results in this study suggest that SMI value doesn't change within short period of time. Thus, this approach provides with an approximate value of SMI.

After the sampling, the SMI for precipitation stations were obtained. The SMI values for the 6 ROIs were further calculated using IDW interpolation algorithm. These SMI values were plotted against NDVI to find the correlation between the two.



From figure xx, it is observed that the regression line for plot of SMI again NDVI shows negative correlation. Also, a low r squared value suggests that the correlation is weak and doesn't signify a relationship. However, the high p value suggests that there is not enough evidence to disprove the null hypothesis.

Thus, it can be concluded that the available data is not enough for finding the relationship between NDVI and SMI.

5. Water Drainage System for Two Industrial Buildings

This section of the report focuses on whether drought can be reduced manually by improving the drainage system. Two new industrial buildings are planned to be constructed at an undisclosed location in the Olpe Kreis. A drainage system has to be planned for the two new industrial buildings to drain rainwater collected in the rooftop of the two buildings. An analysis is performed whether the drainage system can be accurately planned based on the roof areas of the two buildings and precipitation data from the existing precipitation stations.

The first information and input data available for the analysis are as follows:

- Top view roof plan of the industrial buildings
- Pre-Analysis report mentioning that no precipitation stations available in the selected industrial location.

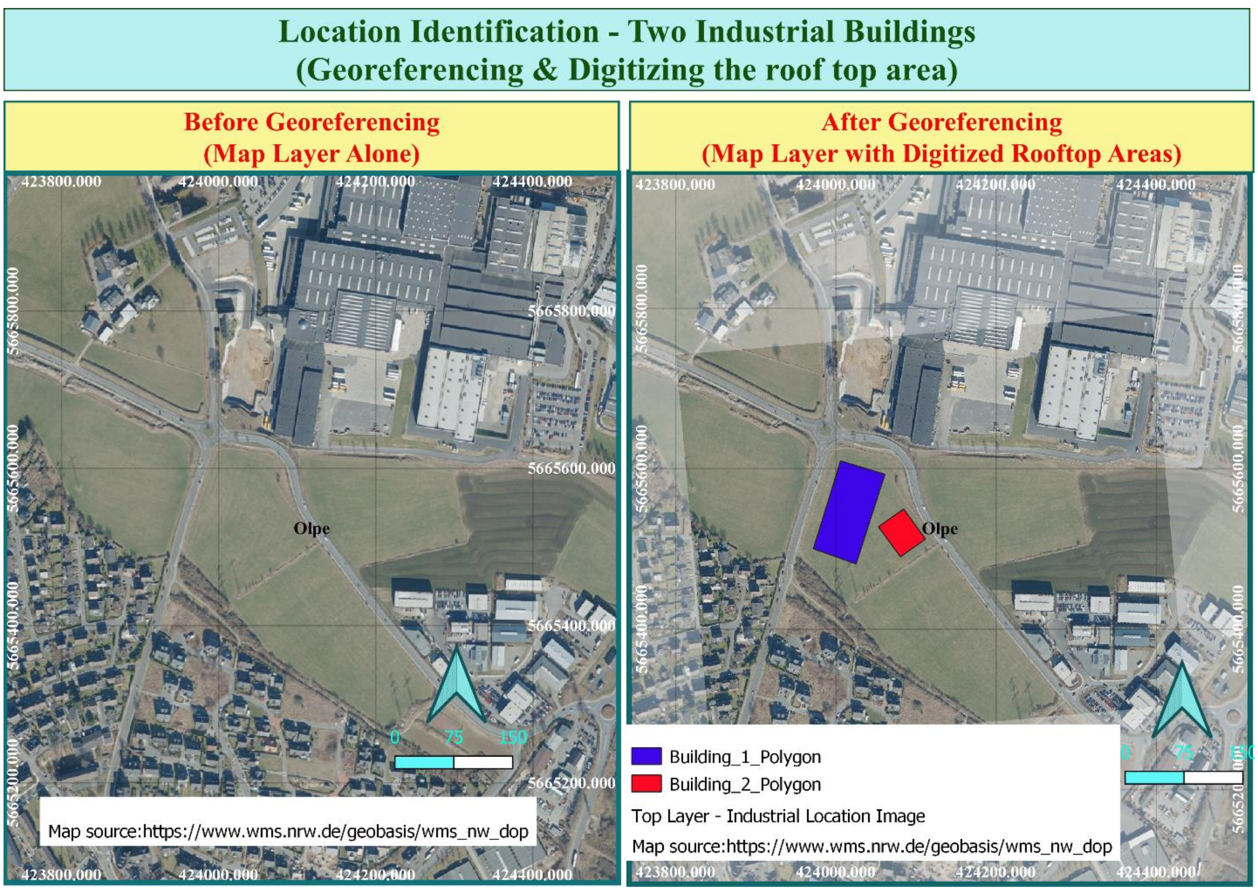
- Precipitation data for the months of May, June, July, and August from all the available stations in the 13 counties for the year 2017 given for previous part of the investigation.

The following information is needed to estimate the total amount of water to drain annually from the industrial buildings.

- Area of the rooftops of the two industrial buildings
- Annual Precipitation in the identified location

5.1 Identifying the rooftop area of the industrial buildings

By applying the Georeferencing algorithm using the "Georeferencer" raster tool in QGIS, the given rooftop image was digitized, and the exact location and dimensions of the rooftops were found.



The two buildings are found to be in "Attendorn," a small town in Olpe Kreis. The buildings were marked as polygons, and the area was calculated. The "After georeferencing" part of figure xx shows that the polygon in blue is building 1, and the polygon in red is building 2. The areas of the two buildings were calculated and mentioned in table xx.

Name	**Colour in Figure xx**	**Area (in m^2)**
Building 1	Blue	6639.284
Building 2	Red	1781.325

5.2 Annual precipitation calculation

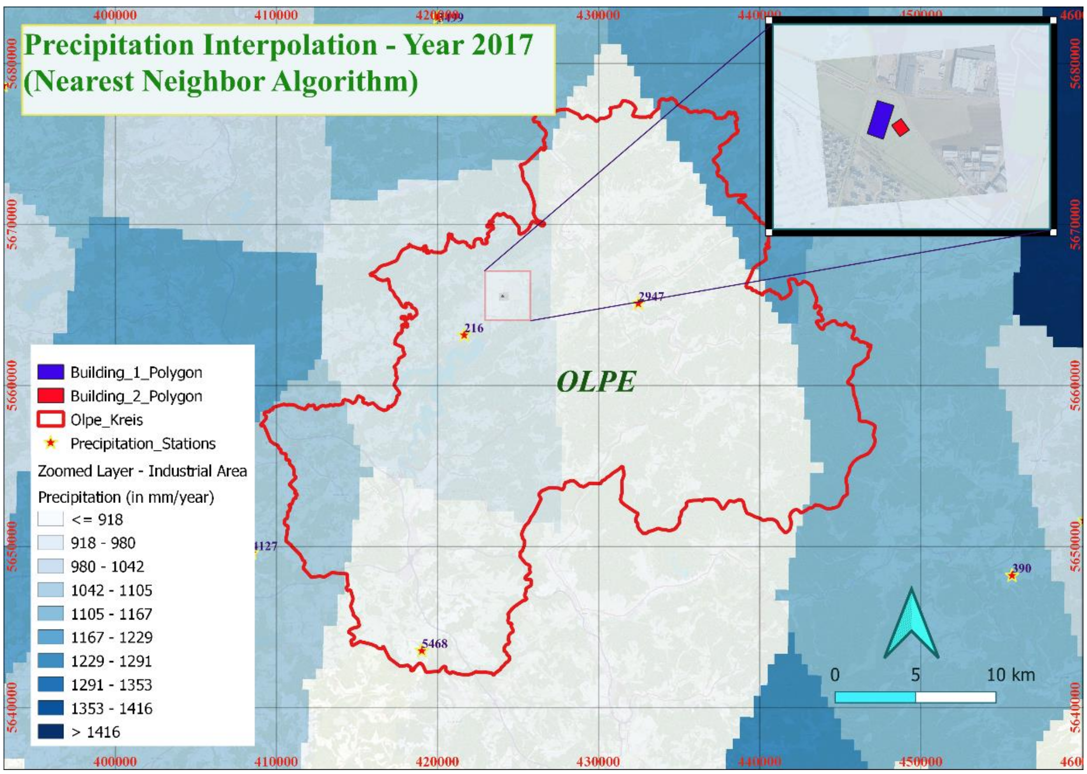
Since there is no precipitation station available at the two industrial buildings of interest, the precipitation value for the building location was estimated using the interpolation algorithms. Inverse Distance Weighted and Nearest Neighbour interpolation were the two algorithms used to determine the precipitation values at the unknown locations from the precipitation values of the nearby stations. These algorithms are applied using the GDAL raster analysis tools from the QGIS software. For this interpolation, the annual precipitation data for all the given 33 stations should be identified.

However, available data is only for the months of May, June, July, August of the year 2017. Hence the total cumulative precipitation is calculated first, and the annual precipitation is considered three times the total cumulative precipitation for all the 33 stations in the 13 counties of interest (assuming that it is varying in the same pattern throughout the year). The annual precipitation data is calculated in mm.

Nearest neighbor Interpolation

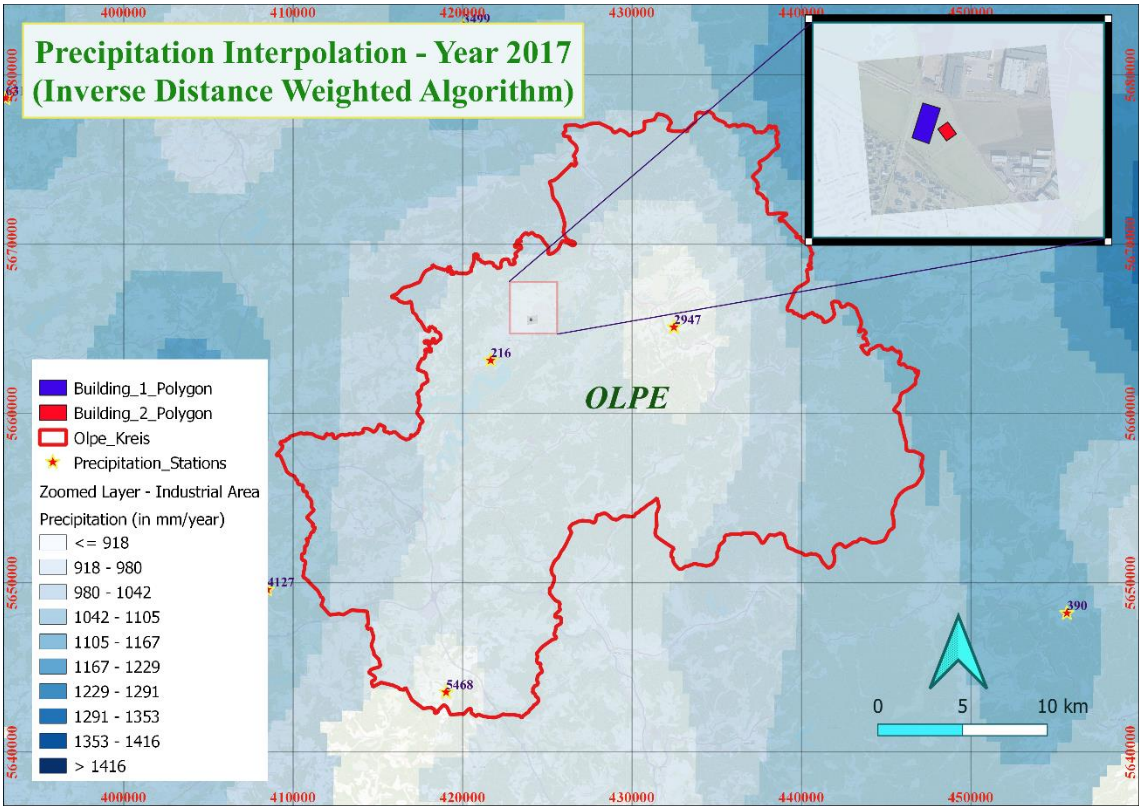
The identified building location is in Olpe, so initially, only three precipitation stations in Olpe Kreis with station ids 216, 2947, and 5468 were chosen for the nearest neighbor interpolation. Using the GRID (Nearest Neighbor) algorithm, the interpolation is performed which creates the 3 Thiessen polygons since 3 precipitation stations were chosen. Then the algorithm allocates the known precipitation value of each station to all the unknown points that lie within the polygon area corresponding to it. However, the two industrial buildings of interest were not lying in any of the Thiessen polygon areas created by the nearest neighbor algorithm. So, the interpolation was repeated using all the 33 stations in the 13 counties of interest.

As we can see in figure xx, the nearest neighbor algorithm is applied for interpolating the precipitation value for all 33 stations in the 13 counties of interest. Now the algorithm creates the 33 Thiessen polygon areas using the known precipitation values from the 33 stations. Then the algorithm allocates the known precipitation value from each station to all the unknown points in the respective polygon area. Figure xx shows that the identified industrial buildings and precipitation station 216 were in the same region. The annual precipitation value of station 216 was allocated to the location of the industrial buildings. So, the annual precipitation value for the buildings of interest after applying Nearest Neighbor algorithm is 933 mm/year.



Inverse Distance Weighted (IDW) Interpolation:

The second algorithm IDW is also applied for determining the unknown precipitation value of the industrial building location of interest. All the 33 stations from the 13 counties of interest were chosen for the IDW interpolation. To apply this IDW algorithm, "Grid (Inverse Distance to a Power)" is used. The algorithm is provided with the control parameter (power or weighting power *P*) value as 2 and distance is calculated for all the 33 stations and corresponding weights were determined. From figure xx, the precipitation value of the industrial location of interest is identified as 962.220 mm/year.



Total water to drain

Based on the results of the two algorithms, the total amount of water to drain from the two buildings were calculated using the following formula:

Total water to drain = (Annual Precipitation * Area) /1000

where Annual Precipitation is in millimeters,

Area is in square meters, and

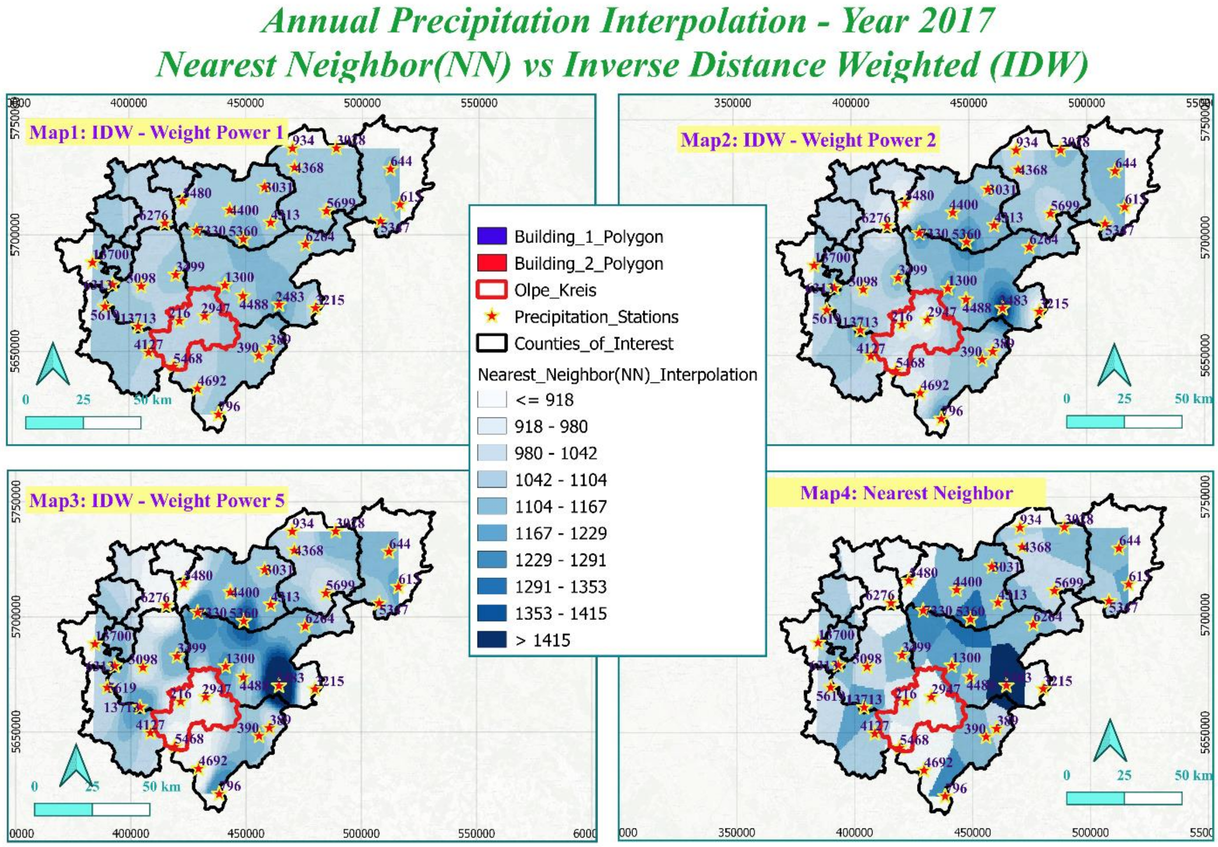
Total water to drain is in cubic meter per year. (1meter = 1000 millimeters).

The total amount of water to be drained annually by each building was calculated and mentioned in table xx. According to this calculation, the average flow rate of the water to be drained is approximately 0.0002 m^3/s from building 1 and 0.00005 m^3/s for building 2. Which means that water can be drained at the rate of 720l/hr for building 1 and 180l/hour for building 2.

	Building 1	**Building 2**
Identified Area (units m^2)	6639.284	1781.325

Interpolated Precipitation (units - mm/year) (Nearest Neighbor Algorithm)	933	933
Interpolated Precipitation (units - mm/year) (Inverse Distance Weighted – Power 2)	962.220	962.220
Total rainwater volume to drain (units m3/year) (Nearest Neighbor Algorithm)	6194.452	1661.976
Total rainwater volume to drain (units m^3/year) (Inverse Distance Weighted – Power 2)	6388.452	1714.027
Total rainwater volume to drain (m^3/s)	0.0002	0.00005

NN and IDW comparison



As mentioned earlier, the interpolated areas are discontinuous and have sharp boundaries in the case of the Nearest Neighbor (NN) algorithm. In NN, only the value of precipitation of the closest station is assigned to all the entire surrounding extent within the Thiessen polygon. Hence the precipitation values obtained at the unknown points in the location of interest could not be accurately identified. However, the IDW interpolation technique uses all the available stations in the given region to estimate the precipitation at the unknown location. The weights are assigned to the stations of known precipitation values according to their distance from the unknown point based on the inverse distance weight formula.

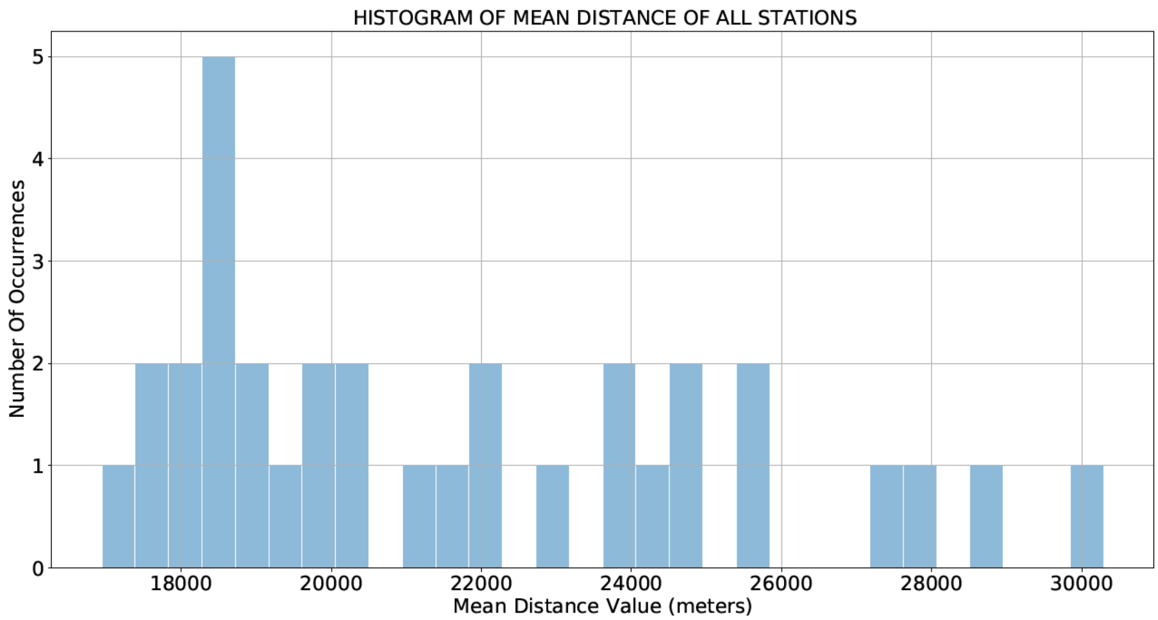
Different weighting power (Power *P*) such as 1, 2, 5 were applied for the IDW algorithm, and their results were analyzed. Map 3 and Map 4 from figure xx show that IDW with weighting power 5 yields almost the same result as the nearest neighbor algorithm. Map 1 from figure xx shows that IDW with weighting power 1, the precipitation values in the unknown locations were not appropriately depicted. Since in IDW, the weighting factor Wi decreases exponentially if the power P increases. Hence, as the known point's distance increases, the precipitation value at that point has a more negligible effect on the estimated value than a nearer known point, which then works similar to the Nearest neighbor technique. This is further analyzed in the following section (Spatial Analysis (Interpolation) — QGIS Documentation documentation, 2021).

Analysis of accuracy of the obtained volume of water to be drained:

	Station 216	**Difference from Actual Value for station 216**	**Station 2947**	**Difference from Actual Value for station 2947**
Actual Annual Precipitation Value	933	-	862.800	-
IDW (Power 1)	1086.347	153.347	1090.182	227.382
IDW (Power 2)	1062.358	129.358	1097.503	235.503
IDW (Power 5)	992.831	59.831	1059.763	196.963

Further analysis was made to cross-validate the accuracy of the IDW algorithm applied. For this purpose, 2 stations (station 216 and station 2947) with known precipitation values were excluded and the interpolation was performed. The values obtained after interpolation are compared with the actual precipitation values in Table xx. This shows that the annual precipitation value obtained by IDW interpolation varies according to the choice of the weighting power used. The deviation from the actual value also varies accordingly. The same analysis can be carried out using all the 33 stations and the Root Mean Square Error can be calculated for the choice of each weighting power value. Using this, the optimum value of the weighting power can be selected corresponding to the least RMSE value obtained. According to (Chen, Feng-Wen; Liu, Chen-Wuing, 2012),the accuracy of the IDW algorithm also depends on the radius of the search ellipse used in the IDW algorithm. The search radius determines the number of stations that would be considered for performing the interpolation. (Chen, Feng-Wen; Liu, Chen-Wuing, 2012), states that the ideal search radius for interpolation of precipitation data is around 10Km-30Km. In our analysis, we have used a 20Km search radius for the IDW interpolation carried out. However, one of the critical disadvantages in the Inverse Distance Weighted Interpolation is that if the sample known points are unevenly distributed, then the interpolation's accuracy will be affected negatively.

Further investigation of the distribution of the precipitation stations in the region was carried out using the distance matrix tool of QGIS software and a histogram. The mean distance for all the 33 stations from the other stations was calculated, sorted in ascending order, and plotted as a histogram with a bin size of 30. Figure xx shows that the spatial distribution frequency of the precipitation stations in our ROI and it does not follow a Normal distribution. There are a larger number of stations within 20Km from the unknown location whereas, there are fewer stations that are farther away from it. There are also some large gaps in between where there are no stations available in that distance range, for instance, between 26Km-28Km range. From this, it can be concluded that the stations are not evenly distributed throughout the region. Hence the data set of these precipitation stations are not enough to give accurate results. Therefore, even though we can interpolate and calculate the total precipitation and estimate the total volume of rainwater drained from the two rooftops, it is not possible to get accurate results using the available data.



Observations:

In our analysis, since the error range is in the order of a few hundred millimetres, it is not very likely to highly impact the total charge. The precipitation value of the industrial buildings is influenced by all the stations in IDW with default weighting power of 2, hence it can be considered for drainage system. So, the total water volume to be drained from the buildings will be approximately 6389 cubic meter per year for building 1 and 1714 cubic meter per year for building 2.

- NN interpolation can be compared to hard decision making or classification while IDW interpolation corresponds more to soft decision making or classification.
- Availability of more evenly distributed data can improve the accuracy of IDW.
- The choice of the interpolation technique depends on the data and how much accuracy is critical for the intended application.

Further, to improve the accuracy of the unknown precipitation values, the optimal power value and the search radius for IDW interpolation should be identified.

6. Conclusion

Different parameters such as SMI, Precipitation, and Altitude were investigated during this study to determine how they influence drought in the region of Sauerland for the year 2017 and its impact on vegetation.

The 13 counties of interest from the Sauerland were identified. SMI analysis was performed for all the 13 counties for 16th of each month from May-Aug 2017. It was seen that the SMI was minimum in June, indicating drought, whereas in August, they turned to no-drought regions. These fluctuations in the SMI deemed further analysis on precipitation and altitude necessary to see if they played a role. The altitude analysis showed that some lower elevation regions recovered from the drought faster than the higher elevation regions. However, this result was contradicting for other regions. Thus it could be finally concluded that altitude had minimum role in variation of SMI from region to region in the 13 counties of interest.

Further, the precipitation analysis performed for OE and HSK showed that Olpe had a lower average precipitation rate compared to HSK during this period. Various correlation analyses were performed to determine the relationship between the different parameters and their role in causing drought. The following conclusions are made from the study:

- There is no conclusive evidence that precipitation and altitude are correlated to each other, and hence this hypothesis is strongly rejected for our region.
- But in case of precipitation and SMI, it is evident that precipitation directly affects the SMI for the given regions and this hypothesis is proven.
- However, the analysis shows SMI is not affected by altitude and this is also proven from the maps.
- The relationship for NDVI and SMI cannot be concluded because there is not enough data to predict the relation.
- The drainage system proposed for the industrial buildings, cannot be accurately planned using the given data but an approximate estimation can be made.

An approximate estimation can be made for the proposed drainage system for the industrial buildings, by performing interpolation using the given data. The results obtained using this technique are plausible but not accurate. More accurate results require more information, such as the tanks' storage capacity and wastewater costs in the Olpe region.

References

Chandrasekar, K. et al. (2013). Agricultural Drought: Assessment & Monitoring. Satellite Remote Sensing & GIS Applications in Water Resources. (Accessed: 22 March 2021)

Zink, M. et al. (2016) *'The German drought monitor'* Environ. Res. Lett. 11 074002. pp1. (Accessed: 22 March 2021)

Saha, A. et al. (2018). '*Assessment and Impact of Soil Moisture Index in Agricultural Drought Estimation Using Remote Sensing and GIS Techniques'.* Proceedings. 7. 5802. 10.3390/ECWS-3-05802. (Accessed: 22 March 2021)

Pellet, C., and C. Hauck. (2016). *'Monitoring soil moisture from middle to high elevation in Switzerland: Set-up and first results from the SOMOMOUNT network'*. Hydrol. Earth Syst. Sci. Discuss.. 10.5194/hess-2016-474. (Accessed: 22 March 2021)

Marx, A. (n.d.) A 'Drought Monitor Germany' available at: https://www.ufz.de/index.php?en=37937 (Accessed: 22 March 2021)

(Marx, n.d.)

Deutscher Wetterdienst. (2021). Available at: https://en.wikipedia.org/wiki/Deutscher_Wetterdienst (Accessed: 22 March 2021)

Earth Observing System (n.d.) *NDVI: Normalized Difference Vegetation Index For Agriculture*. Available at: https://eos.com/ndvi/ (Accessed: 15 March 2021)

GIS Geography (2020) *What is NDVI (Normalized Difference vegetation Index)?* Available at: https://gisgeography.com/ndvi-normalized-difference-vegetation-index/ (Accessed: 14 March 2021)

Sauerland. (2021). Available at: https://en.wikipedia.org/wiki/Sauerland (Accessed: 14 March 2021)

Olpe (district). (2021). Available at: https://en.wikipedia.org/wiki/Olpe_(district) (Accessed: 14 March 2021)

Copernicus Programme (2021) *CORINE Land Cover*. Available at: https://land.copernicus.eu/pan-european/corine-land-cover (Accessed: 21 March 2021)

Copernicus Programme (2021) CLC 2018. Available at: https://land.copernicus.eu/user-corner/technical-library/corine-land-cover-nomenclature-guidelines/html/index-clc-231.html (Accessed: 21 March 2021)

Kosztra, B. et al. (2017) *'Updated CLC illustrated nomenclature guidelines',* European Topic Centre on Urban, land and soil systems; ETC/ELS, European Environment Agency Austria; EAA. Available at: https://land.copernicus.eu/user-corner/technical-library/corine-land-cover-nomenclature-guidelines/html/index-clc-211.html (Accessed: 21 March 2021)

Copernicus (n.d.) Sentinel -2 Spacecraft Overview. Available at: https://spaceflight101.com/copernicus/sentinel-2/ (Accessed: 22 March 2021)

Gordana, J.K., and U. Avdan. (2017). Object-based water body extraction model using Sentinel-2 satellite imagery. European Journal of Remote Sensing. 50. 137-143. 10.1080/22797254.2017.1297540.

Chen, Feng-Wen; Liu, Chen-Wuing (2012): Estimation of the spatial rainfall distribution using inverse distance weighting (IDW) in the middle of Taiwan. In Paddy and Water Environment 10. DOI: 10.1007/s10333-012-0319-1.

Spatial Interpolation with Inverse Distance Weighting (IDW) Method Explained (2021). Available online at https://www.geodose.com/2019/03/spatial-interpolation-inverse-distance-weighting-idw.html, updated on 3/28/2021, checked on 3/28/2021.

Earth Science Week (2015): Soil Moisture. Available online at https://www.earthsciweek.org/classroom-activities/soil-moisture, updated on 6/13/2018-04:00, checked on 3/28/2021.

Morán-Tejeda, Enrique; Moreno, J.; Beniston, Martin (2013): The changing role of temperature, precipitation and elevation on snowpack variability in Switzerland, p. 7280.

Bundesministerium für Umwelt, Naturschutz und nukleare Sicherheit (2021): Waste Water. Available online at https://www.bmu.de/en/topics/water-waste-soil/water-management/wastewater/, updated on 3/28/2021, checked on 3/28/2021.

Spatial Analysis (Interpolation) — QGIS Documentation documentation (2021). Available online at https://docs.qgis.org/3.16/en/docs/gentle_gis_introduction/spatial_analysis_interpolation.html?highlight=interpolation#figure-idw-interpolation, updated on 3/28/2021, checked on 3/28/2021.

Berlin* (2021): Berliner Wasserbetriebe - Overview of Tariffs. Available online at https://www.bwb.de/content/en/html/1720.php, updated on 3/28/2021, checked on 3/28/2021.

Hartmann, K., Krois, J., Waske, B. (2018): E-Learning Project SOGA: Statistics and Geospatial Data Analysis. Department of Earth Sciences, Freie Universitaet Berlin.

Project Jupyter (2021). Available online at https://jupyter.org/, updated on 2/8/2021, checked on 3/28/2021

Welcome to the QGIS project! (2021). Available online at https://qgis.org/en/site/, updated on 3/28/2021, checked on 3/28/2021.

Org, Python (2021): Welcome to Python.org. Available online at https://www.python.org/, updated on 3/28/2021, checked on 3/28/2021.

Annex

Project files

All the python source code and QGIS projects are uploaded to a GitHub repository. The time series analysis video is also found in the same link. The GitHub link is as below.

https://github.com/Sheharaz/GeoInformatics.git

Counties abbreviations

The abbreviations for the counties were taken from the following website.

https://www.kennzeichenking.de/kfz-kennzeichen-liste

Acknowledgment

I, Sheharaz Abdul Majeed Sheikh, Thirumal Janakiraman and Nikhilesh Munthala hereby declare that the work presented herein is my own work completed without the use of any aids other than those listed. Any material from other sources or works done by others has been given due acknowledgement and listed in the reference section. Sentences or parts of sentences quoted literally are marked as quotations; identification of other references with regard to the statement and scope of the work is quoted. The work presented herein has not been published or submitted elsewhere for assessment in the same or a similar form. I will retain a copy of this assignment until after the Board of Examiners has published the results, which I will make available on request.

Sheharaz Abdul Majeed Sheikh (28381)

Thirumal Janakiraman (28364)

Nikhilesh Munthala (28383)