Analysis of the drought in Xanten, Germany in 2018 using NDVI, precipitation and temperature data

---

Prof. Dr. Rolf Becker

Clein-Alexander Sarmiento-Castrillon

Group Members:

Alex-Kagwi Mwenja (26497)
Rohit Turankar (26509)
Rajesha Koppa Ramesha(26566)

Hochschule Rhein-Waal
Rhine Waal University of Applied Sciences
Faculty: Communication and Environment
Program: Information Engineering and Computer Science
Module: Geoinformatics, Winter Semester: 2019-2020

Submitted on:
02.03.2020

Table of Contents

Abstract

This paper reports an investigation of the drought that was experienced by Germany in the year 2018. Normalized Difference Vegetation Index (NDVI), precipitation and temperature were studied. These three methods of measurement were chosen to complement each other as it was determined that one measurement alone would represent an incomplete picture. By studying the interactions between the NDVI, the precipitation and the temperature it is possible to explain the reason for the drought. The Xanten region in Germany was chosen for the investigation. By calculating the correlation between the NDVI, precipitation and temperature values for the years 2016, 2017 and 2018 it can be clearly discerned that the correlation between cumulative precipitation and average temperature is -0.95, while the correlation between NDVI and cumulative precipitation is 0.179 and lastly, the correlation between NDVI and the average temperature is -0.560.

1. Introduction

Drought has no specific definition that can be applied across the world, which is why the definition changes depending on the region to be studied. Drought can either be approached in a qualitative way through descriptions or quantitively through statistical analysis. An example of this is the definition by Diaz and McMahon (1982) whereby drought is expressed as the hydrological disparity resulting from a period of arid meteorological conditions and deficiency of rainfall. Tate E.L., Gustard A. (2000) recapitulate this brilliantly by stating:

> *“A comprehensive classification of drought considers its type, intensity, duration and spatial extent in conjunction with its perceived impact, since only through such a broad description can a fully objective picture be obtained.”*

Among the factors considered for drought, there is the Normalized Difference Vegetation Index (NDVI) which is a measure of vegetation vigour which was established by Rouse et al. (1974), this was so that a spectral index could be created to separate green vegetation from its surroundings using satellite imagery. It is expressed through the subtraction of the near-infrared and red bands divided by their sum. This results in a measurement scale ranging from -1 to 1, 0 represents the approximate value of no vegetation with negative values representing non-vegetated surfaces.

In this paper we take an in-depth look at the severe drought experienced by Germany, particularly the region of Xanten, in the year 2018 in an aim to understand the factors that contributed to the drought. Through this, we aim to understand whether the drought was a result of a combination of varied factors working in tandem or one specific factor. We plan to explore factors such as NDVI, precipitation rate and temperature to evaluate several hypotheses:

1. Drought's effect on vegetation can be observed through the Normalized Difference Vegetation Index (NDVI).
2. Drought can be associated with cumulative precipitation and average temperature over the last year.
3. There is a correlation between cumulative precipitation, NDVI, average temperature.

2. Material and Methods

As part of this investigation, we are exploring several regions of interest (ROI) in Xanten which is an area in the state of North Rhein Westphalia in Germany. There are eight ROI each having different land uses as verified using the CORINE land cover project as shown in the figure below (https://land.copernicus.eu/pan-european/corine-land-cover)

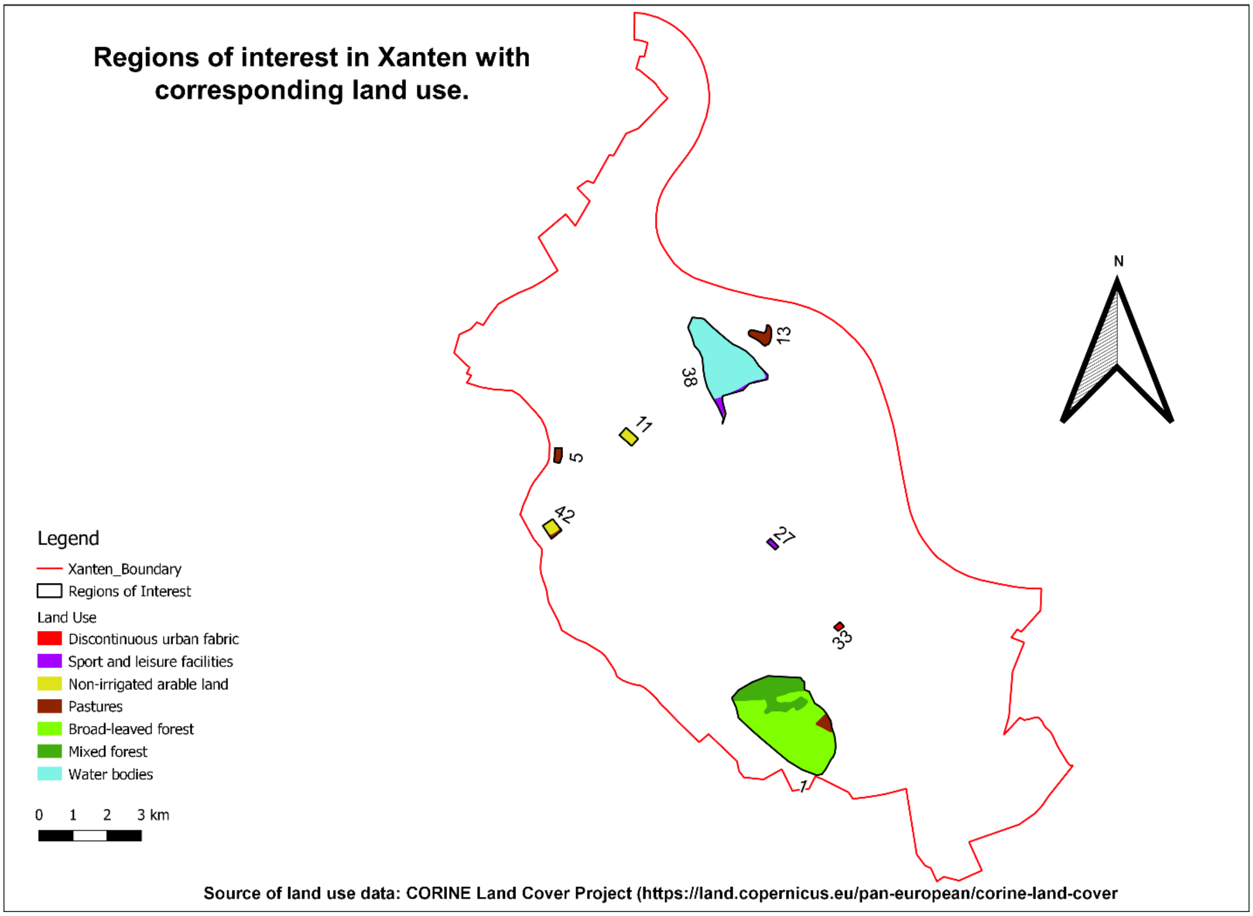

Figure 1: Regions of interest in Xanten with respect to land use

Topographical information was downloaded from the Federal Agency for Cartography and Geodesy (https://gdz.bkg.bund.de/index.php/default/open-data.html) to be used in conjunction with the CORINE land cover data in order to get a clearer picture of the regions of interest for further analysis. The analysis of the CORINE land cover data, the topographical information and Google satellite revealed the following about the regions of interest:

ROI 1 (Broad-Leaved forest and Mixed forest: Die Hees): This region of interest falls under the forest category as it contains trees specifically broad-leaved trees including shrub and bush. The forest name is Die Hees and it is managed by the Rhein-Weser Federal Forest Service which was renatured after having been used for ammunition and detonator warehouses during the Second World War.

ROI 5 (Pasture: Gartenland): This region of interest falls under the pasture classification which has been described by the CORINE Land Cover nomenclature (https://land.copernicus.eu/user-corner/technical-library/corine-land-cover-nomenclature-guidelines) as land that has been used for at least 5 years for fodder production. The region of interest also falls under the Gartenland category which means that the land can only be used as a garden or horticultural use. However, it is noted that the routine agricultural activities performed on the land may sway the growth of natural vegetation.

ROI 11 (Non-Irrigated Arable Land: Ackerland): This region of interest is used as farmland hence the classification of Ackerland. This land is used mainly for agricultural activities such as the growth of non-permanent crops. The plants grown on the land are usually rainfed with periodic irrigation using sprinklers.

ROI 13 (Water Body/Pasture: Reeser Schanz): The region of interest is contained in an area that falls under the pasture category named Reeser Schanz: (http://nsg.naturschutzinformationen.nrw.de/nsg/de/fachinfo/gebiete/gesamt/WES-011) but the region of interest itself is a waterbody fed by a stream from the Rhine river.

ROI 27 (Sport & Leisure: Archaologischer Park): The Archaologischer park in Xanten (https://apx.lvr.de/de/willkommen/willkommen_1.html#) is a leisure park created for exploration by the general public. It is an archaeological open-air museum displaying the Roman city of Colonia Ulpia Triana.

ROI 33 (Discontinued Urban Fabric: Residential Area): This region of interest falls in an urban area containing urban structures as well as transportation networks. The discontinued urban fabric classification means that there are vegetated surfaces present in an intermittent pattern in the residential area.

ROI 38 (Water Body: Xantener Nordsee): This region of interest is a lake in Xanten named Nordsee which is a natural waterbody.

ROI 42 (Non-Irrigated Arable Land: Ackerland): This region is the same as the region of interest 11 in terms of category and land use.

To ensure that the regions of interest were within the Xanten region, boundary data was also downloaded from the Federal Agency for Cartography and Geodesy for the entire Germany region before clipping the Xanten area using the QGIS software.

The temperature and precipitation data have already been measured by the government agencies of Germany thus we would be considering the data in their servers which has been made available to the public. The German Weather Service department has been recording the temperature and precipitation data and it has made it available for analytical purposes on their servers.

Access URL for the server: ftp://opendata.dwd.de

The server consists of yearly, daily, hourly and many other formats of data ranging from the year 1900 to the current date. There is data recorded by several stations across Germany, but the focus is on the regions of interest.

For the analysis, daily temperature and cumulative precipitation data were considered for the years 2016, 2017 and 2018. The reason for considering only these three years is that the drought experienced was heavy in the year 2018, so the comparison will be made with the data of the year 2016 and 2017 in context.

The exact location of the data can be found under below URLs.

For daily temperature data:
ftp://opendata.dwd.de/climate_environment/CDC/observations_germany/climate/daily/kl/historical/

For daily precipitation data:
ftp://opendata.dwd.de/climate_environment/CDC/observations_germany/climate/daily/more_precip/historical/

The calculation of the NDVI involved downloading satellite imagery from the Copernicus Open Access Hub (https://scihub.copernicus.eu/) for the Sentinel-2 satellite. Six sensing dates were chosen as they had low cloud coverage spanning the years 2016, 2017 and 2018 for comparison. QGIS, an application that supports analysis of geospatial data, was used to further process the downloaded satellite imagery. The satellite imagery consisted of several raster images separated by colour bands as shown in table 1 below:

Table 1:Spectral Bands of the Sentinel-2 Satellite.

(Source: https://sentinel.esa.int/web/sentinel/user-guides/sentinel-2-msi/resolutions/radiometric)

Spectral Band	Wavelength [µm]	Resolution [m]
B1 – Coastal Aerosol	0.443	60
B2 – Blue (B)	0.492	10
B3 – Green (G)	0.560	10
B4 – Red (R)	0.665	10
B5 – Red Edge 1 (Re1)	0.704	20
B6 – Red Edge 2 (Re2)	0.741	20
B7 – Red Edge 3 (Re3)	0.783	20
B8 – Near Infrared (NIR)	0.833	10
B8a – Near-Infrared narrow	0.865	20
B9 – Water Vapor	0.945	60
B10 – Shortwave Infrared / Cirrus	1.374	60
B11 – Shortwave Infrared 1 (SWIR1)	1.614	20
B12 – Shortwave Infrared (SWIR2)	2.202	20

The specific bands of interest when it comes to calculation of the NDVI are B4 (Red) and B8 (Near Infrared). The raster calculator tool in QGIS was used to calculate the NDVI using the formula: $NDVI = \frac{NIR(B8)-Red(B4)}{NIR(B8)+Red(B4)}$ The resulting raster was further analysed by applying the pseudocolour scheme. To get specific values, the raster layer statistics tool in QGIS was used to acquire the mean value for the regions of interest.

Tools, methods and technologies used in the analysis are,

1. **QGIS:** is an open-source desktop Geographical Information System (GIS) (https://www.qgis.org/en/site/). It enables the user to view, edit, manipulate, analyse and present the geospatial data. QGIS does this by supporting raster and vector layers with vector data being stored in the form of a line, point or polygon feature. The software also allows for georeferencing of images.
2. **Jupyter Notebook**: It is an open-source web application that allows you to create and share documents that contain live code, equations, visualizations and narrative text. Its

uses include data cleaning and transformation, numerical simulation, statistical modelling, data visualization, machine learning, and much more. (https://jupyter.org/)

3. **Pandas:** is a fast, powerful, flexible and easy to use open-source data analysis and manipulation tool, built on top of the Python programming language. (https://pandas.pydata.org/)
4. **Matplotlib:** is a Python 2-D plotting library which produces publication quality figures in a variety of hardcopy formats and interactive environments across platforms. Matplotlib can be used in Python scripts, the Jupyter notebook, web application servers, and four graphical user interface toolkits. (https://matplotlib.org/)
5. **Pearson's coefficient of Correlation:** This correlation coefficient is a statistical measure of the strength of the relationship between the relative movements of two variables (Sedgwick, 2012). The values range between -1.0 and 1.0. A correlation of -1.0 shows a perfect negative correlation, while a correlation of 1.0 shows a perfect positive correlation. A correlation of 0.0 shows no linear relation.
   The Pearson coefficient is a type of correlation coefficient that represents the relationship between two variables that are measured on the same interval or ratio scale. The Pearson coefficient is a measure of the strength of the association between two continuous variables.

2.1 Temperature data aggregation

The temperature data was downloaded from the FTP server (location given above) using pandas. It was observed that server consisted of historical to present date data from all stations across Germany. Given that the data downloaded was huge, the next step was to drill down into this huge amount of data and extract only what was necessary for the analysis.

The investigation was done for the Xanten area in North Rhein Westphalia state of Germany for the years 2016, 2017 and 2018. Thus, the filters were applied to the downloaded data to get the specific data from Xanten area station. Unfortunately, it was discovered that the Xanten station had no data with respect to the time that we were interested in.

The ideal solution to this was to consider the data from nearby stations of Xanten. To do this, the latitude and longitude of the Xanten (latitude $_{XANTEN,}$ longitude $_{XANTEN}$) area was noted and the stations in the area with latitude and longitude values in the range of latitude $_{XANTEN}$ - 0.5 < **latitude** < latitude $_{XANTEN}$ + 0.5 longitude $_{XANTEN}$ - 0.5 < **longitude** < longitude $_{XANTEN}$ + 0.5 were considered.

With this filter applied, six nearby stations were found to contain the data with respect to our time of interest. In the dataset several varieties of data were available, but the temperature data was found under the column “TMK” with the description as the daily mean temperature in °C. Only this column was extracted. All these operations were performed using the Python programming language along with pandas.

The next question to be addressed involved choosing between one or many stations for continued analysis. Hence, latitude and longitude of these stations were taken and plotted on a map using QGIS to find out the closest stations to Xanten as shown in figure 2 below. This also gave a better visualization which served as a justification for the reason why the specific stations were chosen.

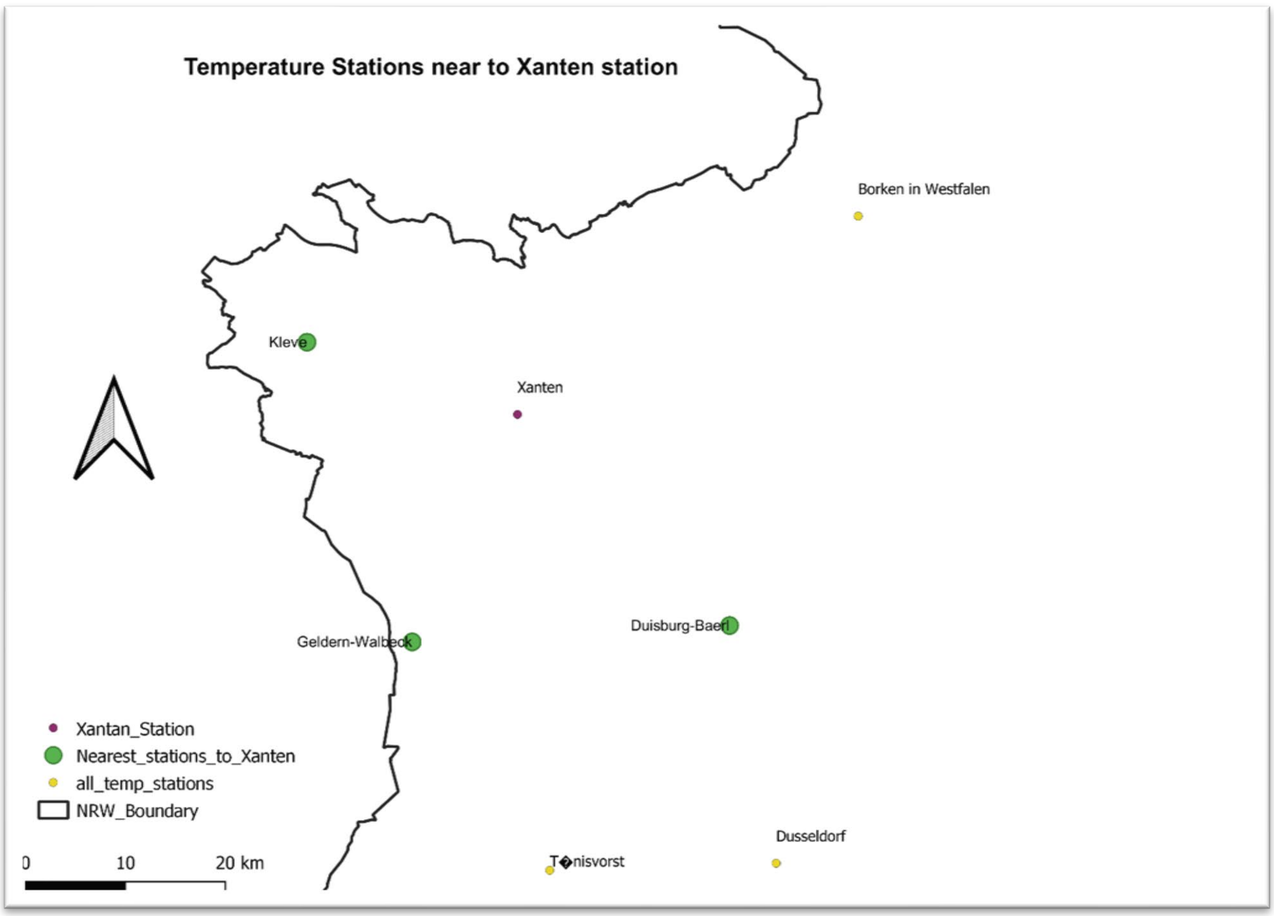

Figure 2: Stations recording temperature of required date range near Xanten

Table 2: Distance matrix for the stations near Xanten

Station Name	Distance from Xanten station (latitude)	Selected
Kleve	22054.49	yes
Geldern-Welbeck	24874.49	yes
Duisburg-Baerl	29719.24	yes
Borken in Westfalen	39008.51	no
Tönisvorst	45429.21	no
Dusseldorf	51526.16	no

From figure 2 and table 2 above, it is evident that stations of Geldern, Kleve and Duisburg are the closest to Xanten area. So, it was assumed that analysing temperature data from these stations would help us study the possible reasons for the drought in the Xanten area.

Temperature data of these stations were analysed further, and it was observed that data had been recorded on an average from 1950 until the present date. However, the investigation was being carried out for just three years. Therefore, data needed to be trimmed to fetch only the necessary dates and their corresponding temperature values.

Having selected three stations with the temperature values recorded, it seemed appropriate to take the average of these three temperature values so that there would be less ambiguity. This data was then exported into a CSV file in order to analyse the correlation with precipitation and NDVI values.

With the average temperature values for the chosen years in hand, graphs were plotted to understand the trend in temperature. Additionally, temperature values were aggregated for each year and then plotted to see an increasing trend over the years.

2.2 Precipitation data aggregation

The next data of interest was the daily precipitation recorded by the stations. Precipitation data had been downloaded from the FTP server using the link mentioned above. Due to the size of the downloaded dataset, extracting the necessary information out of it was important.

It had been assumed that there was a possibility of the data of our time of interest missing in the Xanten station because the temperature data had also been missing. So, it had been decided to follow the previous approach of considering the nearby stations based on latitude and longitude of Xanten. Four stations were discovered containing the necessary data, but it was then observed that one of the four stations were Xanten. As a result, the other stations were discarded straight away while retaining the Xanten station.

From the downloaded data, only the precipitation data was necessary. Daily precipitation was recorded with the column name as “RS” with the description of the column as the daily precipitation height in millimetre. Only the RS column was extracted for further analysis.

Since the data for Xanten was directly available, the job was much easier. The required data from the year 2016 to the end of the year 2018 was extracted from the main dataset. Then it was deemed appropriate to have the cumulative sum of the precipitation rather than taking the daily precipitation values. Pandas was instrumental in the taking of the cumulative sum of the precipitation values.

2.3 NDVI data processing

The NDVI data involved accessing the Copernicus Open Access Hub in order to find the required satellite imagery needed to be processed by QGIS. This satellite imagery needed to correspond to the Xanten area. The satellite imagery was downloaded using the sentinelsat tool (https://www.sentinelsat.readthedocs.io/en/stable/) for the Sentinel-2 satellite. The satellite was chosen because it has the bands required for calculating the NDVI. For comparison, several dates were chosen between 2016 and 2018, these dates were chosen based on cloud cover. The three dates that were chosen are 10-06-2016, 26-05-2017 and 30-06-2016.

Atmospheric correction was deemed unnecessary as the satellite imagery downloaded was in the form of 2A Bottom of Atmosphere (BOA) reflectance images which are derived from 1C Top of Atmosphere (TOA) reflectance images, hence there was no need to correct for the scattering and absorption effects from the atmosphere which would have otherwise interfered with the results. The data was loaded into QGIS for calculation of the NDVI as explained above. ROI 1 was selected based on the analysis done through the CORINE land project. This ROI was chosen as it was listed under the forest category. The rest of the ROI’s consisted of water bodies, pasture and non-irrigated arable land thus were rejected for NDVI calculation. These two ROI’s were then clipped individually against the NDVI rasters for 2016, 2017 and 2018 after which the raster layer statistics tool was used to get the mean value that was used as the NDVI.

2.4 Altitude data breakdown

This analysis was done to determine the altitude variations in the Xanten region including the nearby stations which were selected for the drought analysis. The dataset which contained the temperature and precipitation data for the stations included the altitude value for each station. The change in the altitude values for a specific station changes over the years, however, it’s a process that takes time. As a result, a single altitude value is considered for the study. Below are three approaches which were used to investigate it further: -

1. IDW Interpolation - Inverse distance weighting is an approach to estimate the values of unknown points considering known points in the specific area. It assumes that closer values are more related than further values. It also requires a minimum set of values to properly estimate the unknown points or the value which is calculated will tend to the mean value of known points. This interpolation is done using stations which are near to Xanten and from which a good comparison can be made.

Table3: Selected stations for altitude analysis using IDW interpolation

Latitude	**Longitude**	**Altitude**	**Station Name**
51.8293	6.5365	23	Bocholt-Liedern (Wasserwerk)
51.7217	6.5839	23	Hamminkeln
51.6927	6.5853	25	Wesel
51.697	6.3974	20	Xanten
51.5786	6.5095	22	Alpen (WV NRW)
51.5426	6.315	24	Geldern
51.873	6.8863	47	Borken in Westfalen
51.296	6.7686	37	Dusseldorf
51.4942	6.2463	37	Geldern-Walbeck
51.7612	6.0954	46	Kleve
51.2897	6.4437	37	Tönisvorst
51.5088	6.7018	24	Duisburg-Baerl

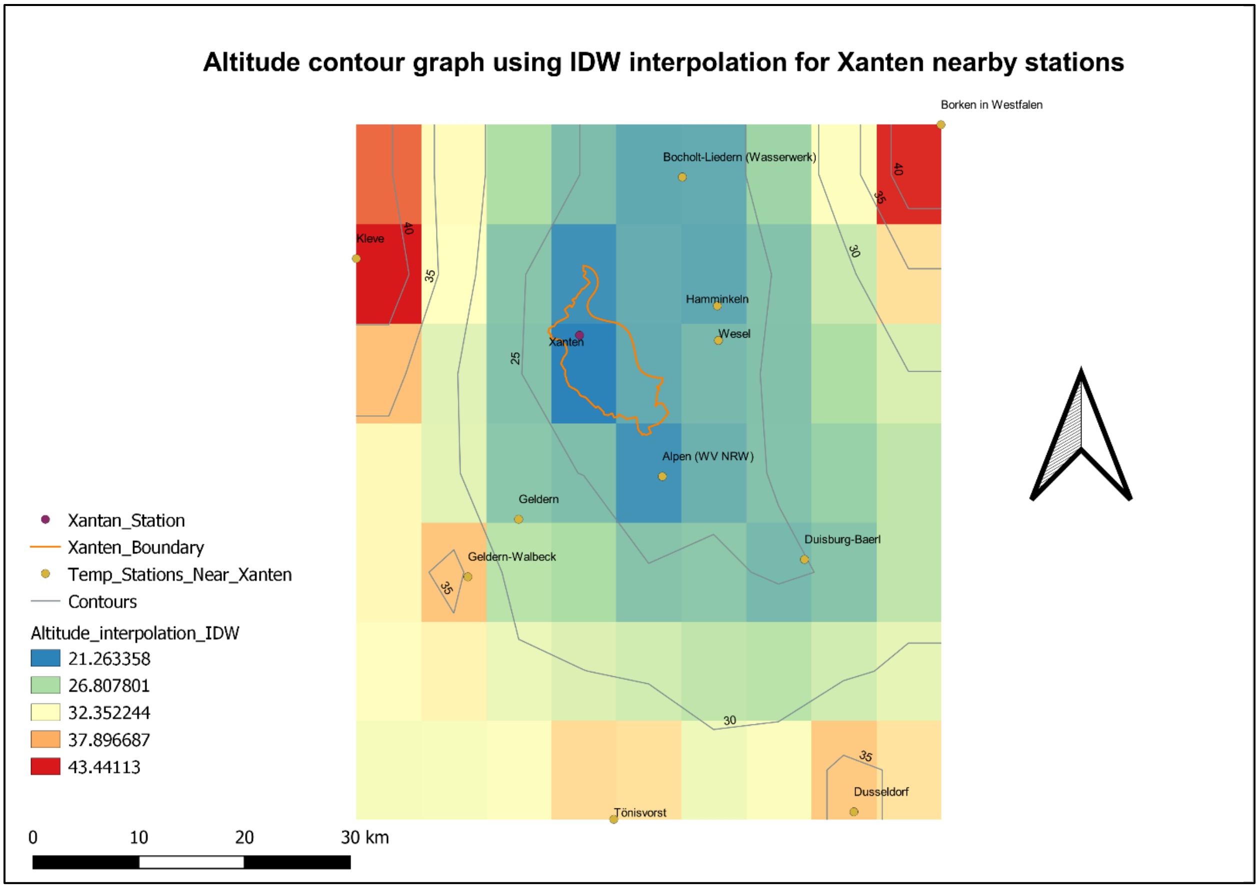

Figure 3:Altitude contour map using IDW interpolation for Xanten nearby stations (2015 to 2018)

From figure 3 above, interpolation of the altitude was done using an inverse distance matrix. It shows the whole Xanten region is not on the same altitude level as we can see 2 different colour grids in the region which is low lying compared to the nearby stations The pixels in figure 3 which are highlighted as Red have the maximum altitude of 43.44113 m while the ones highlighted in blue have the minimum altitude of 21.263358 m. The other values of altitudes are represented by the bands between blue and red. The colour rendering can be changed under the style section of this layer.

2. Digital Terrain Model - A Digital Elevation Model (DEM) is a specialized database, describing a surface relief between known elevation points (Hirt, 2014). It is possible to create a rectangular digital elevation model grid by interpolating known elevation data from sources such as ground surveys and photogrammetric data capture. It is performed to fetch altitude information of the 8 ROI's in our Xanten region. The data from the source was in x-y-z co-ordinate text file format which needs to be changed in GeoTIFF format

(.tif) which is best suited for handling regular geo-referenced grids. This transformation is done using Python language (Pandas and NumPy).
Source - https://www.opengeodata.nrw.de/produkte/geobasis/dgm/dgm1

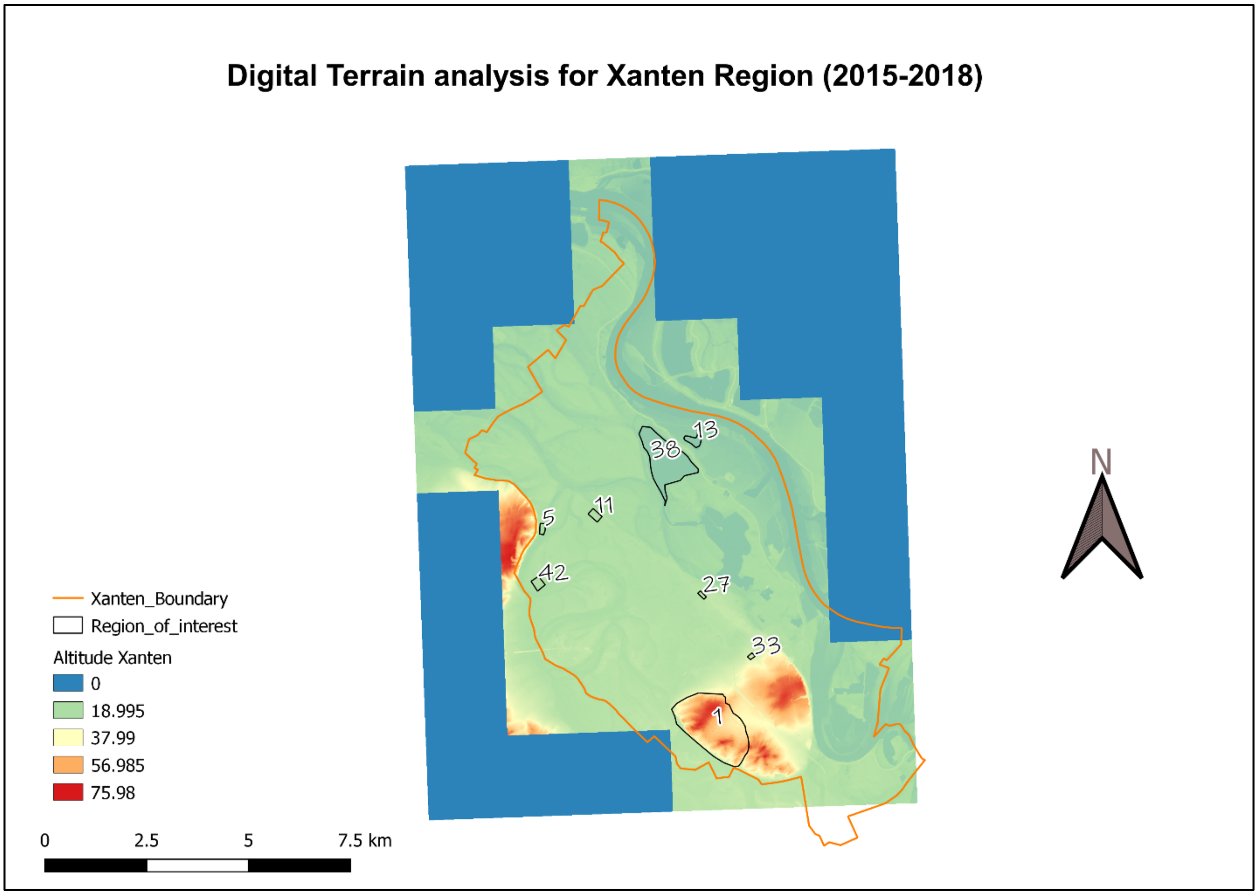

Figure 4:Digital Terrain analysis for Xanten Region (2015-2018) [Source - https://www.opengeodata.nrw.de/produkte/geobasis/dgm/dgm1]

From figure 4 above, altitude analysis of Xanten region is done using Digital Terrain Model (DTM). Each ROI is at different altitude and it is observed that ROI 1 is at the highest altitude. The pixels in figure 4 which are highlighted in Red have a maximum altitude of 75.98 m while the ones highlighted in green have the minimum altitude of 18.995 m. The blue colour band represents no data which is 0 m. The other values representing the altitude are represented by the bands between and green and red. The colour rendering can be changed under the style section of this layer.

3. Results & Discussions

Data required for the analysis was extracted from several sources and aggregated as needed in order to investigate the hypotheses. Furthermore, the graphs were plotted using Python Pandas while Matplotlib was used for observing the trend in the data. NDVI data were calculated using QGIS and the satellite imagery including the respective maps were added. This data was exported into a CSV file for correlation analysis of each data with the rest. Graphs were plotted to provide better visualisation of the data trend for the years 2016, 2017 and 2018.

3.1 Temperature

Temperature data collected for the years 2016, 2017 and 2018 were plotted using line charts. Data was taken for a period of one year using the selected NDVI sensing date as the ending date.

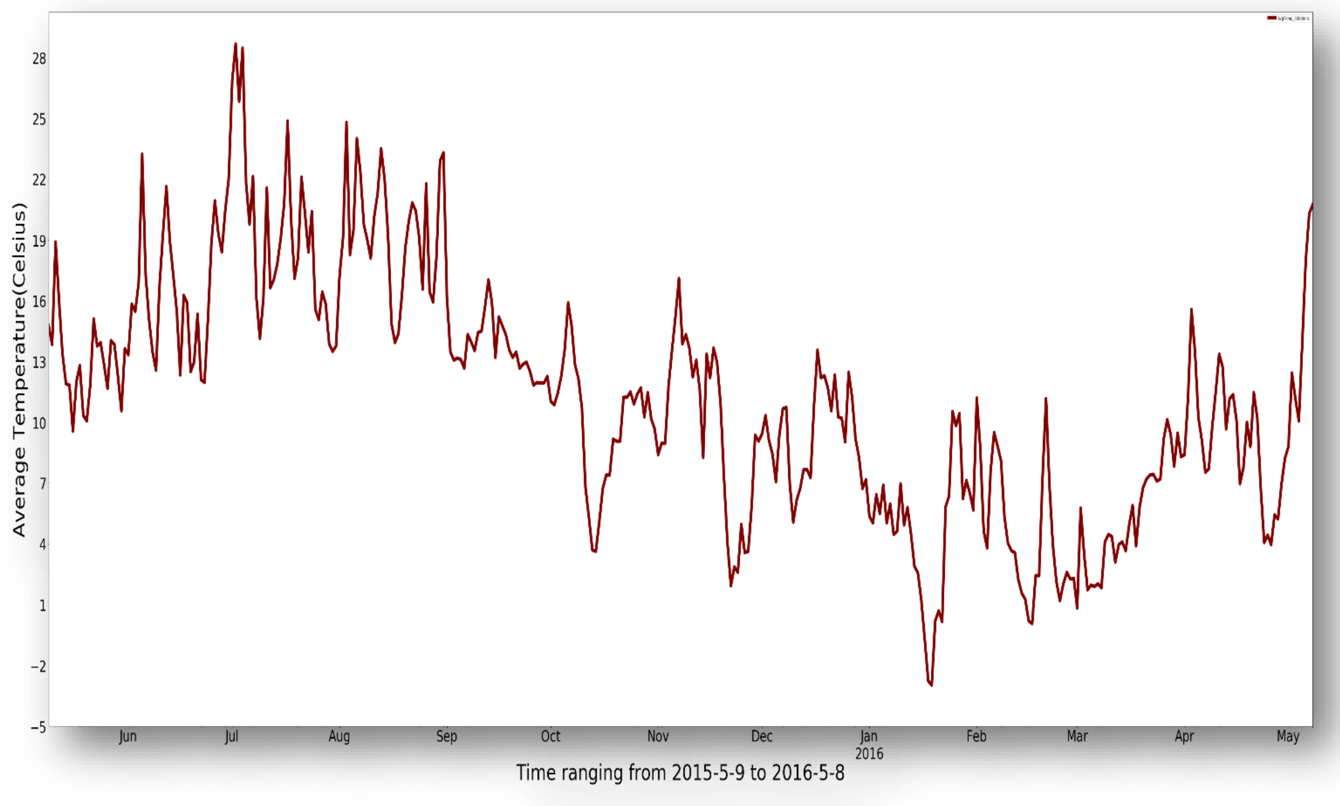

Figure 5:Daily temperature trend for 2015-2016 (9-5-2015 to 8-5-2016)

From figure 5 above, we can see the daily average temperature recorded by three stations around Xanten from 9-5-2015 to 8-5-2016. Temperature varies from -3$^{\circ C}$ to 29$^{\circ C}$. The average temperature was about 11.25 $^{\circ C}$.

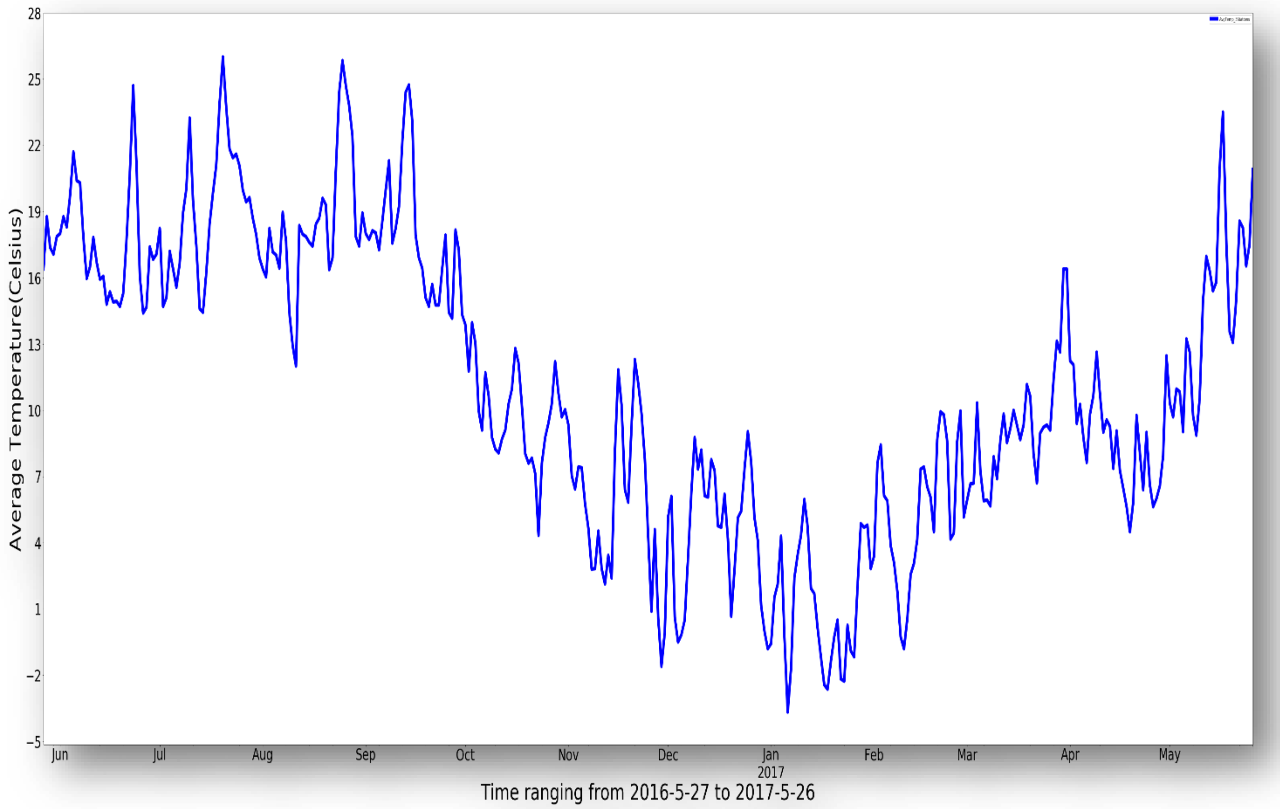

Figure 6:Daily temperature trend for 2016-2017 (27-5-2016 to 26-5-2017)

From figure 6 above, we see the temperature trend recorded from 27-5-2016 to 26-5-2017. The average temperature recorded for this period is around 11.05 $^{\circ}$C. The hottest recorded temperature is roughly 26 $^{\circ}$C. Winter seems to be colder in the 2016 - 2017 period than during 2015 - 2016 as more days have negative temperatures in winter of 2016 - 2017.

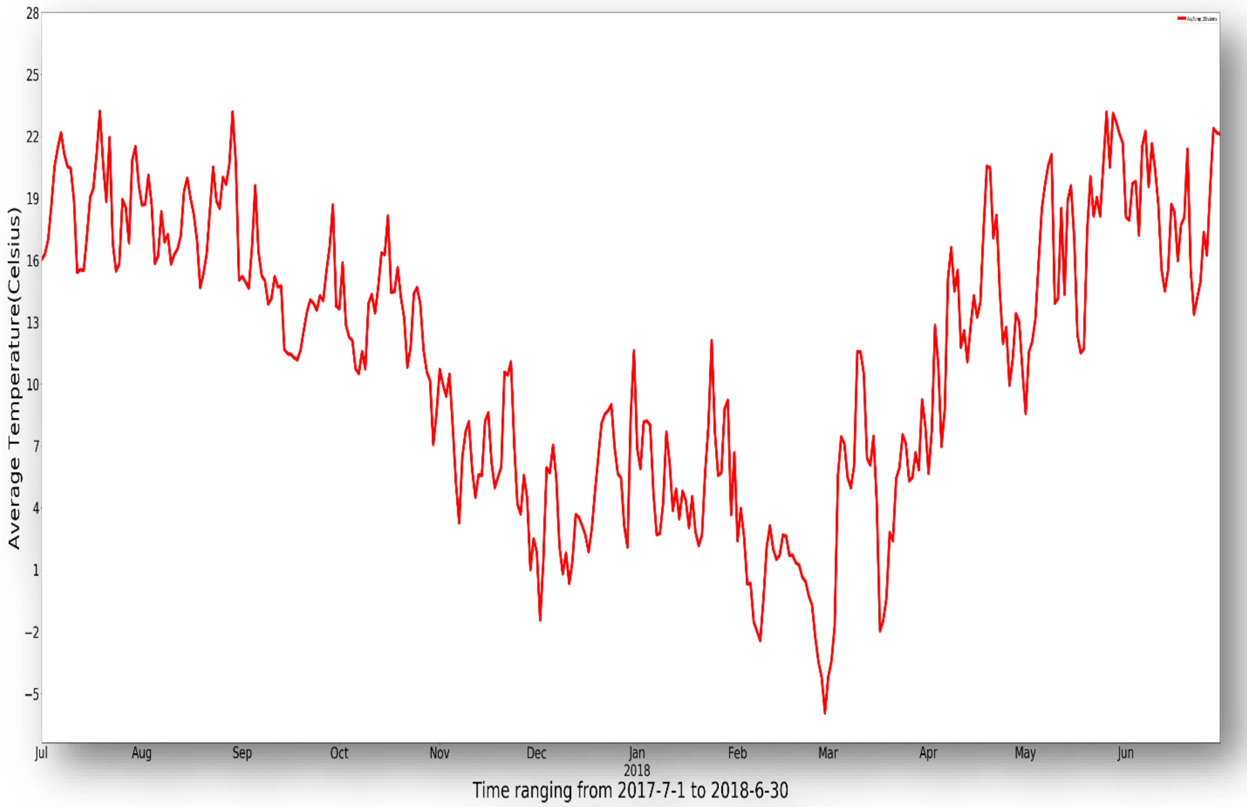

Figure 7:Daily temperature trend for 2017-2018 (1-7-2017 to 30-6-2018)

Figure 7 above shows the temperature trend from 1-7-2017 to 30-6-2018. It can be observed that majority of the days have a recorded temperature exceeding 15 ˚C. Even though the maximum recorded temperature is around 24 ˚C, most of the days during this period are warmer compared to the previous years' accounting for an average temperature of 11.34 ˚C.

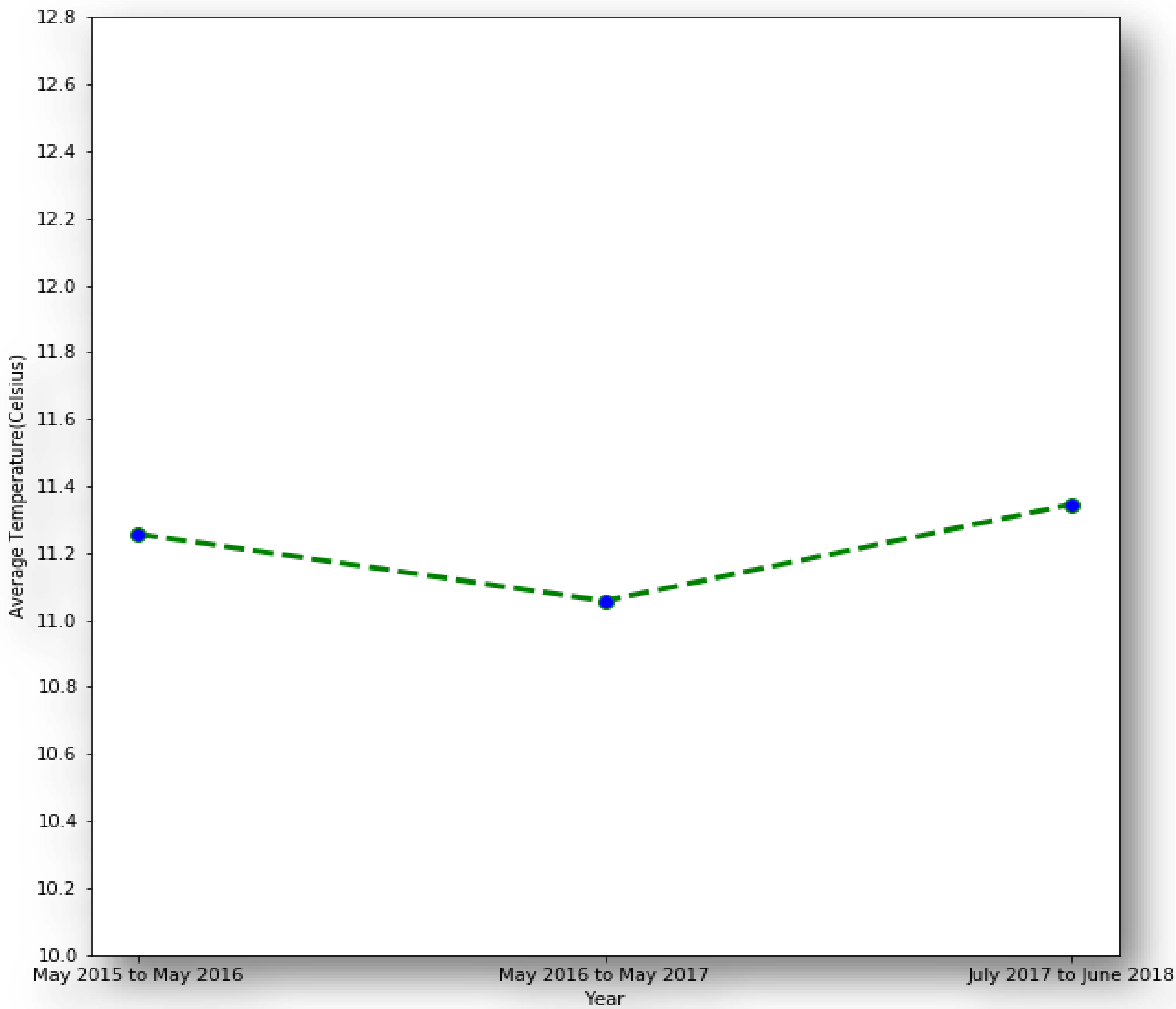

Figure 8:Yearly trend of average temperature from 2016 - 2018

From figure 8 above we see the temperature trend for three years. Average temperature for each year from 2016 to 2018 is plotted. As we observe the average temperature was highest for the period of July 2017 to June 2018.

3.1.1 Interpolation for temperature data

The interpolation activity is performed to estimate the mean temperature value of each year for Xanten station. The input average temperature data for all the 3 nearby stations i.e. Duisburg-

Baerl, Kleve and Geldern-Walbeck is already calculated in the previous analysis. It also helped to determine the difference in interpolated temperature of Xanten station with the mathematically calculated mean value.

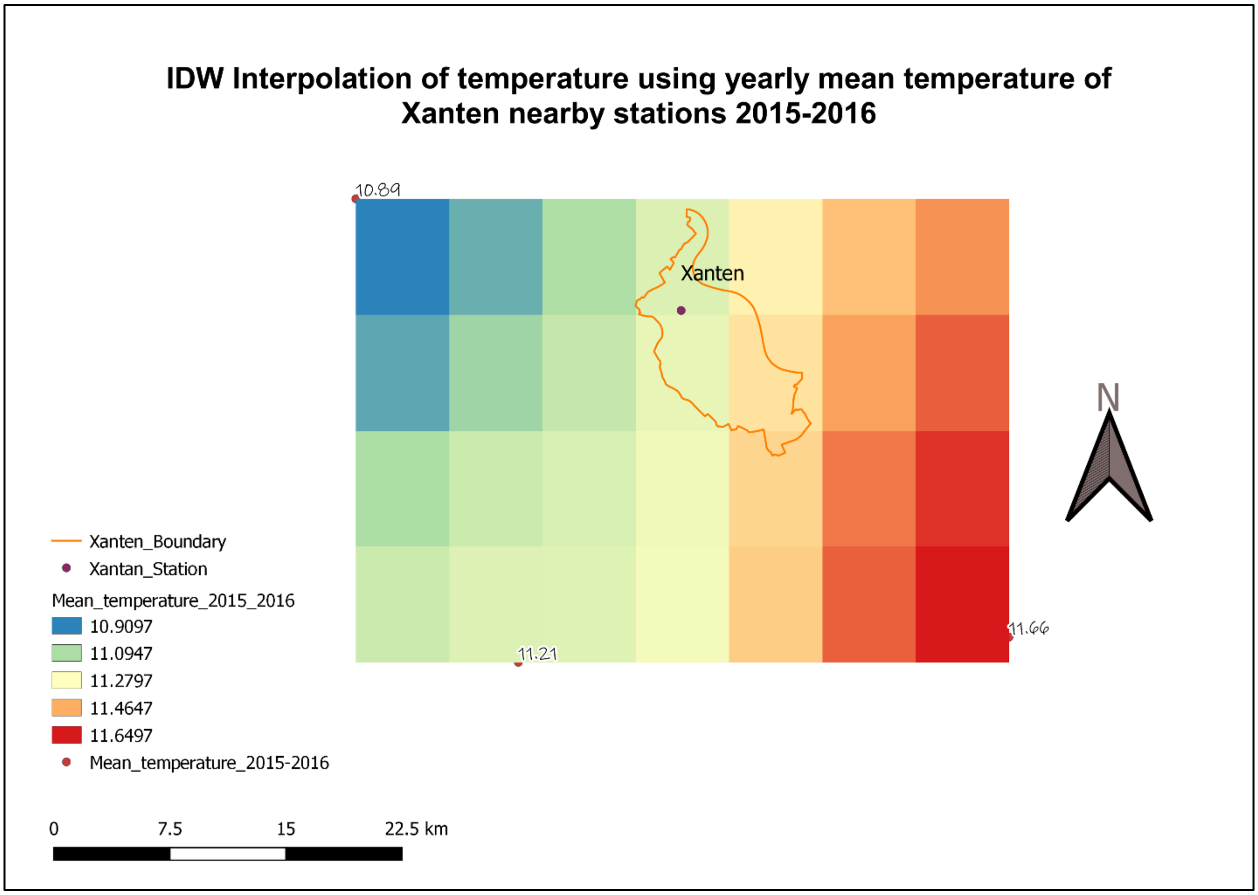

Figure 9: IDW interpolation of temperature using the yearly mean temperature of stations nearby Xanten 2015-2016

From figure 9 above, yearly mean temperature (2015-2016) of Geldern-Walbeck, Kleve and Duisburg-Baerl respectively has been interpolated using IDW interpolation. The pixels in Figure which are highlighted as Red have the maximum mean temperature of 11.6497 °C while the ones highlighted in blue have the minimum mean temperature of 10.9097 °C. The interpolated value of the Xanten station is 11.2027 °C. The other values of mean temperatures are represented by the bands between and blue and red. The colour rendering can be changed under the style section of this layer.

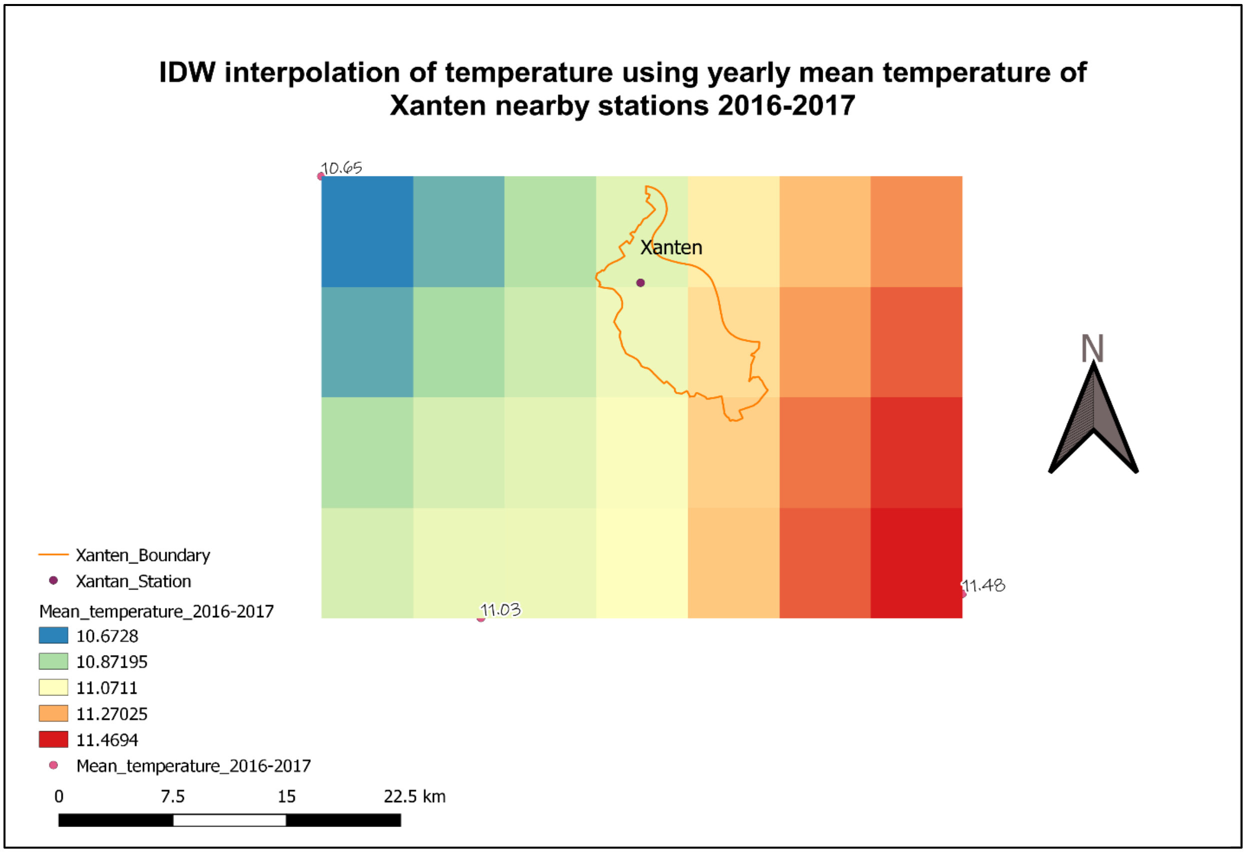

Figure 10:Interpolation of mean temperature using IDW interpolation taking Xanten and nearby stations temperature data (2016 to 2017)

From figure 10 above, yearly mean temperature (2016-2017) of Geldern-Walbeck, Kleve and Duisburg-Baerl respectively has been interpolated using IDW interpolation. The pixels in Figure which are highlighted as Red have the maximum mean temperature of 11.4694 °C while the ones highlighted in blue have the minimum mean temperature of 10.6728 °C. The interpolated value of the Xanten station is 11.00179 °C. The other values of mean temperatures are represented by the bands between and blue and red. The colour rendering can be changed under the style section of this layer.

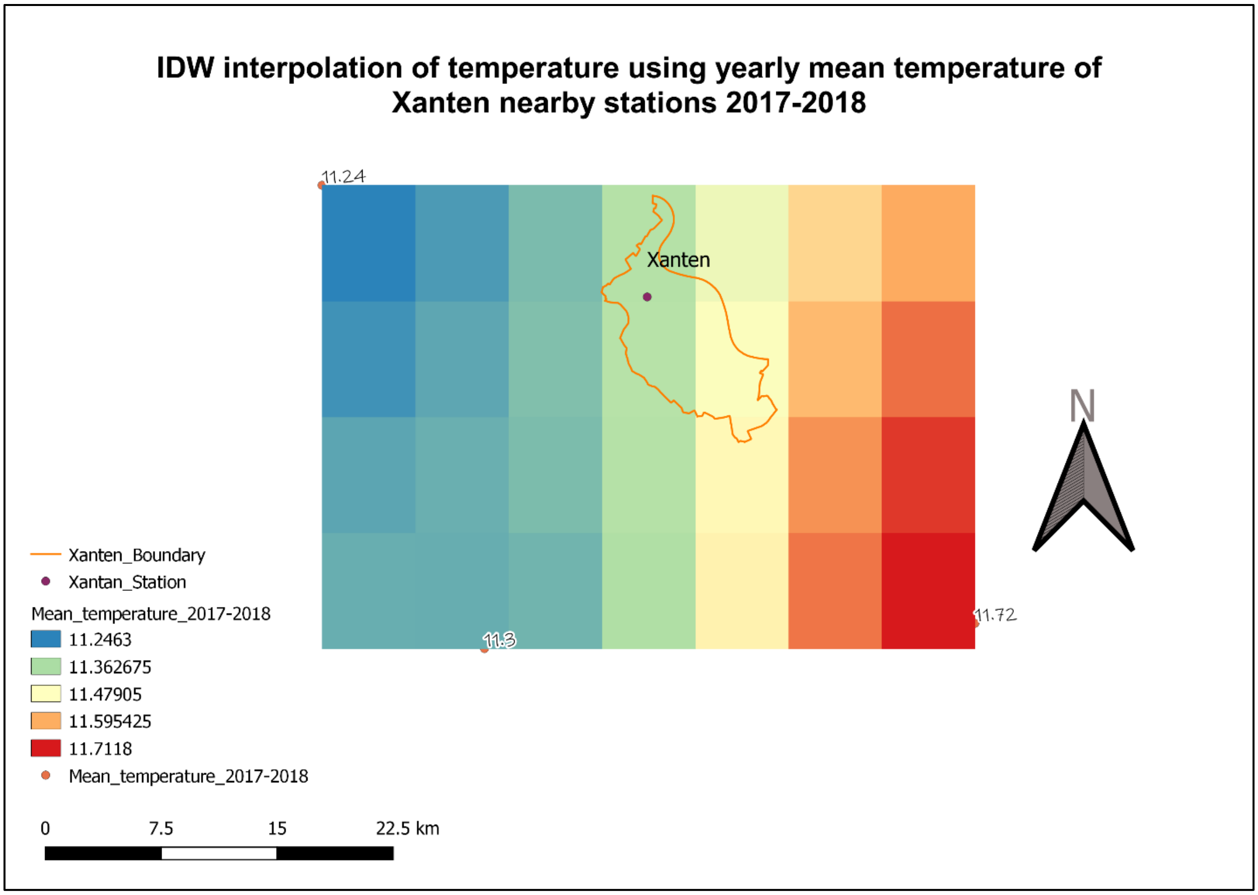

Figure 11:Interpolation of mean temperature using IDW interpolation taking Xanten and nearby stations temperature data (2017 to 2018)

From figure 11 above, yearly mean temperature (2017-2018) of Geldern-Walbeck, Kleve and Duisburg-Baerl respectively has been interpolated using IDW interpolation. The pixels in Figure which are highlighted as Red have the maximum mean temperature of 11.7118 °C while the ones highlighted in blue have the minimum mean temperature of 11.2463 °C. The interpolated value of the Xanten station is 11.37665 °C. The other values of mean temperatures are represented by the bands between and blue and red. The colour rendering can be changed under the style section of this layer.

Table 3:Interpolated Values for Temperature

Time Range	Interpolated mean Temperature °C (Xanten)	Yearly mean Temperature °C (Kleve)	Yearly mean Temperature °C (Geldern-Walbeck)	Yearly mean Temperature °C (Duisburg-Baerl)
2015-2016	11.20276	11.21	10.89	11.66
2016-2017	11.00179	11.03	10.65	11.48
2017-2018	11.37665	11.3	11.24	11.72

The interpolated temperature value for Xanten station is approximately equal to the mathematically calculated temperature value. This is due to the smaller number of inputs i.e. known points given for the interpolation in QGIS. In case of a smaller number of inputs, it is observed that the interpolated and average mean coincides. As we found that the values are almost the same, mean values which were calculated in the previous section were used for the correlation.

The observed graph in figure 8 of the temperature didn't give a clear understanding of whether the temperature was one of the main reasons for the drought. The graph demonstrates that the average temperature value (May 2015 to May 2016) is 11.25 °C and got decreased for the next year to 11.05 °C (May 2016 to May 2017) and again increased for 2018 to 11.34 °C. The rationale for that could be the selection of the stations only considered the distance matrix as important rather than the altitude parameter. The selected stations were on higher altitudes compared to the Xanten region. From figure 3, the altitude is observed that there is a difference between Xanten and Kleve (which is the closest station with the highest influence) is maximum. The magnitude of the influence of altitude on temperature could not be calculated exactly. Therefore, even after the interpolation, the temperature values could highly deviate with respect to the estimated temperature data of Xanten for all the years.

3.2 Precipitation

Precipitation data that was taken from the server was plotted as line graphs for the year 2016, 2017 and 2018. Data was taken for a period of one year prior to the date chosen for the NDVI sensing data for each year. Even though the data is recorded on a daily basis, cumulative precipitation was calculated as it was more suitable for the analysis.

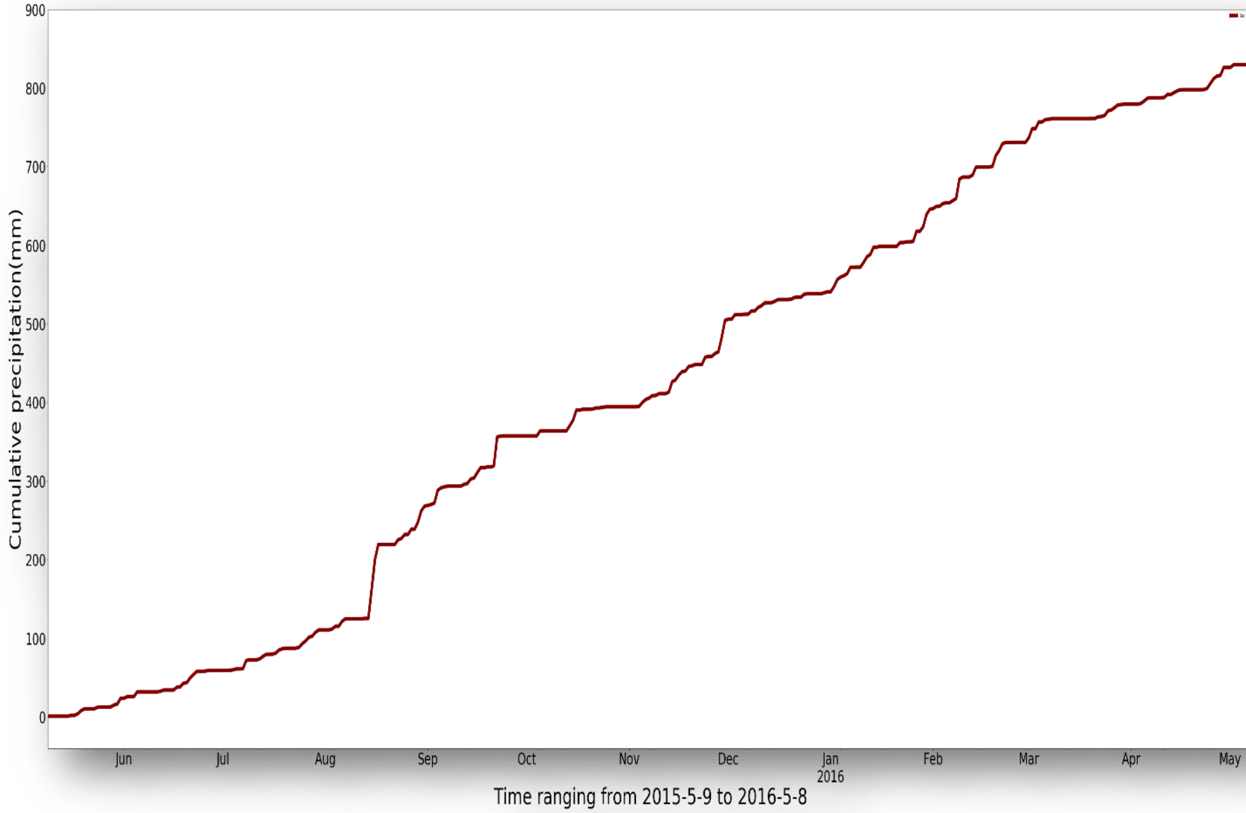

Figure 12: Cumulative precipitation trend for 2015-2016 (9-5-2015 to 8-5-2016)

From figure 12 above, the cumulative precipitation trend from 9-5-2015 to 8-5-2016 can be seen. It can be observed that the precipitation has been good throughout this period accounting for 829 millimetres at the end of the time period.

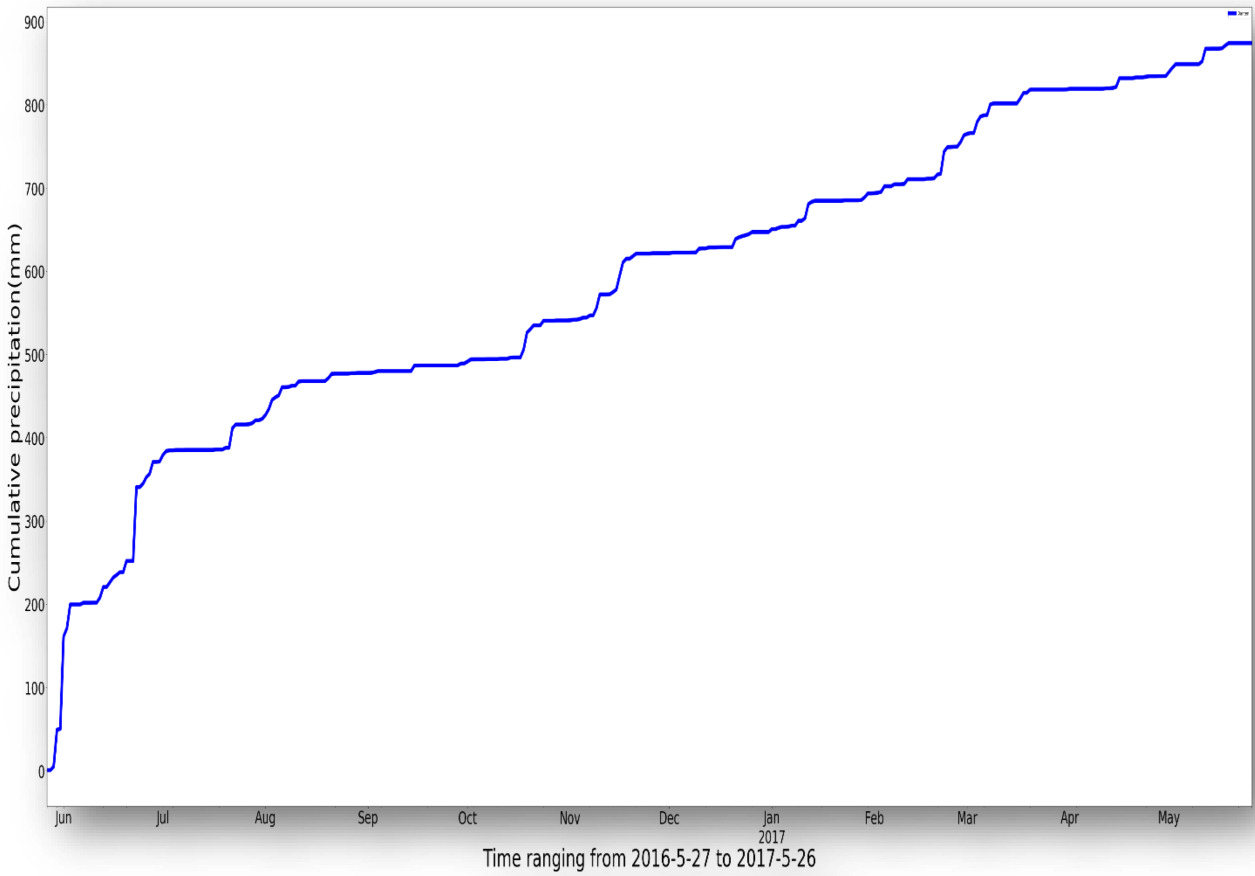

Figure 13:Cumulative precipitation trend for 2016-2017 (27-5-2016 to 26-5-2017)

From figure 13 above, we see the cumulative precipitation trend from 9-5-2016 to 8-5-2017. The cumulative precipitation goes up sporadically in the considered period accounting for 874 millimetres of cumulative precipitation level at the end of this period.

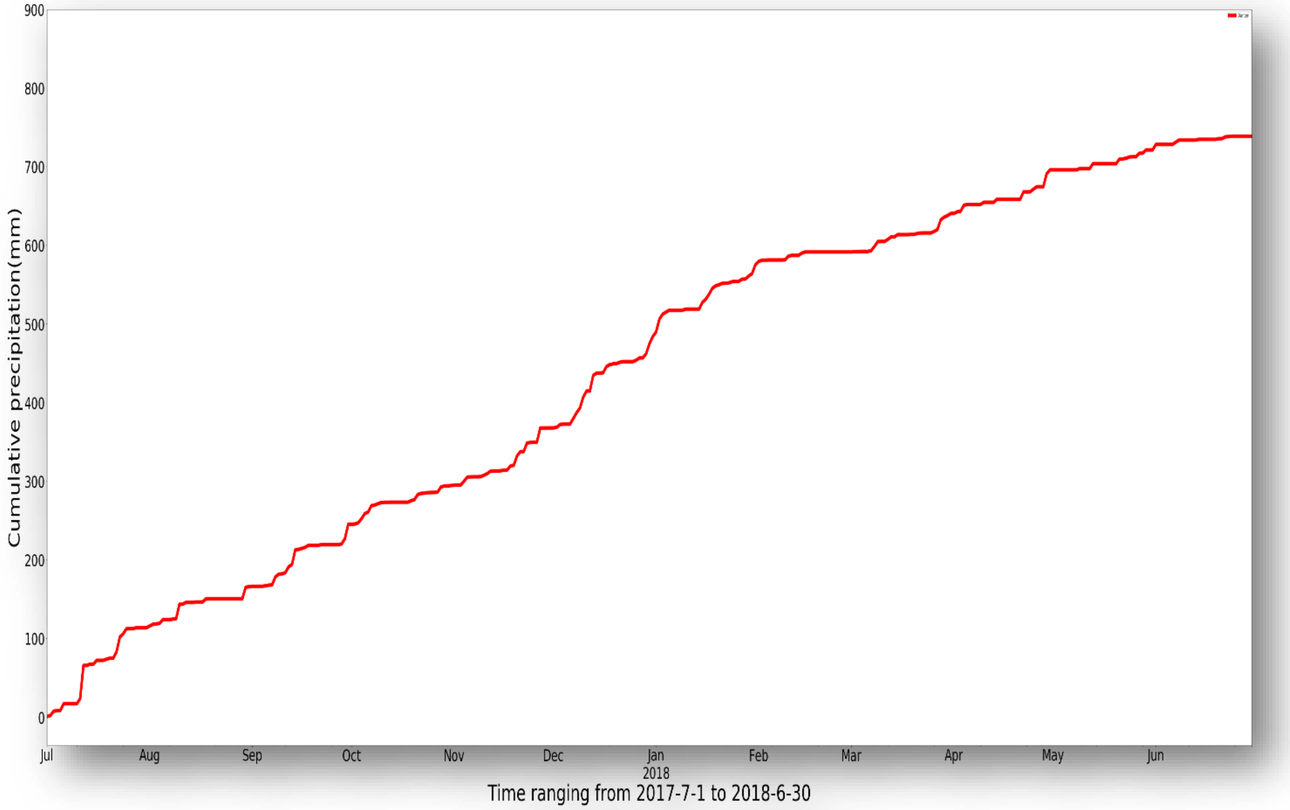

Figure 14:Cumulative precipitation trend for 2017-2018 (1-7-2017 to 30-6-2018)

From figure 14 above, the cumulative precipitation trend can be observed from 1-7-2017 to 30-6-2018. It can be discerned that cumulative precipitation was rising during most of this period but at the end of it there is no significant rise in precipitation during the months of May and June in 2018. As a result, the cumulative precipitation at the end of this period is 738 millimetres, being the least of the three considered time periods.

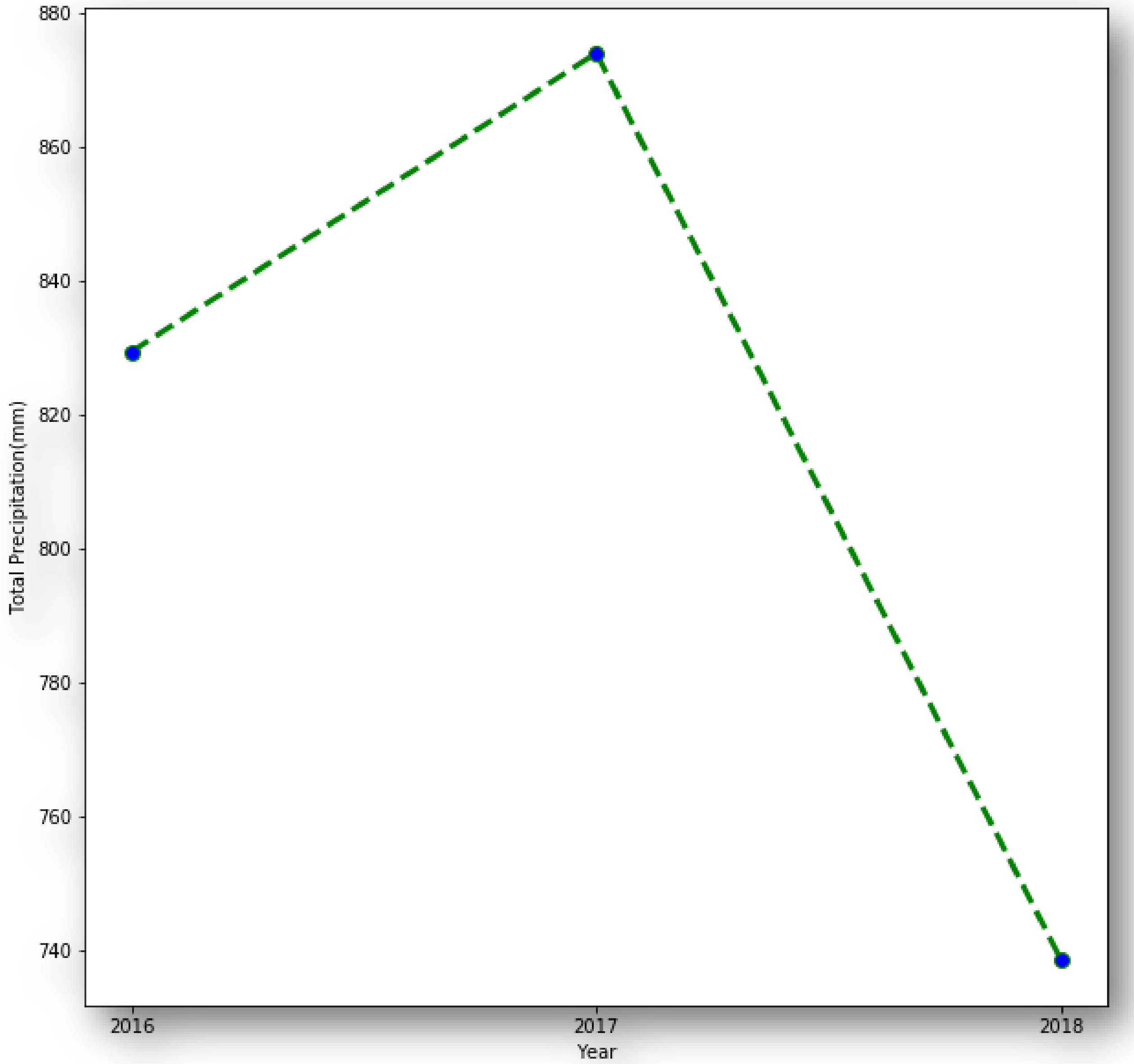

Figure 15:Yearly trend of cumulative precipitation from 2016 - 2018

In figure 15 above, the yearly summer cumulative precipitation is plotted. It is observed that the precipitation undergoes a slight growth from first to second year period but experiences a significant dip in the year 2018.

The graph gives a clear understanding of the drought in 2018. The slope of precipitation value drops significantly from 2017 to 2018. Also, the data is collected for Xanten only which is the plus point for this analysis. It does not require interpolation and there is no effect of altitudes on the precipitation value. But it's also observed from figure 4 that the overall region of Xanten is not on the same altitude. The altitude range is from 18.8 m to 75.98 m. The ROI 1 which is forest according to figure 1 is at the highest altitude. Despite this, it would not deviate the results as the contribution of each ROI to cumulative precipitation will be in the same proportion in the Xanten.

3.3 NDVI data

Normalized Difference Vegetation Index (NDVI) was calculated using QGIS and satellite imagery for the years 2016 to 2018. The range of values for an NDVI is -1 to 1 with negative values (values advancing towards -1) being a positive indication of water. Values near zero (-0.1 to 0.1) largely mean that the area has no vegetation but contains rock, sand, or snow. While low positive values correspond to shrub and grassland (roughly 0.2 to 0.4) high values (values advancing towards 1) suggest temperate and tropical forests (Sentinel-hub.com, n.d.).

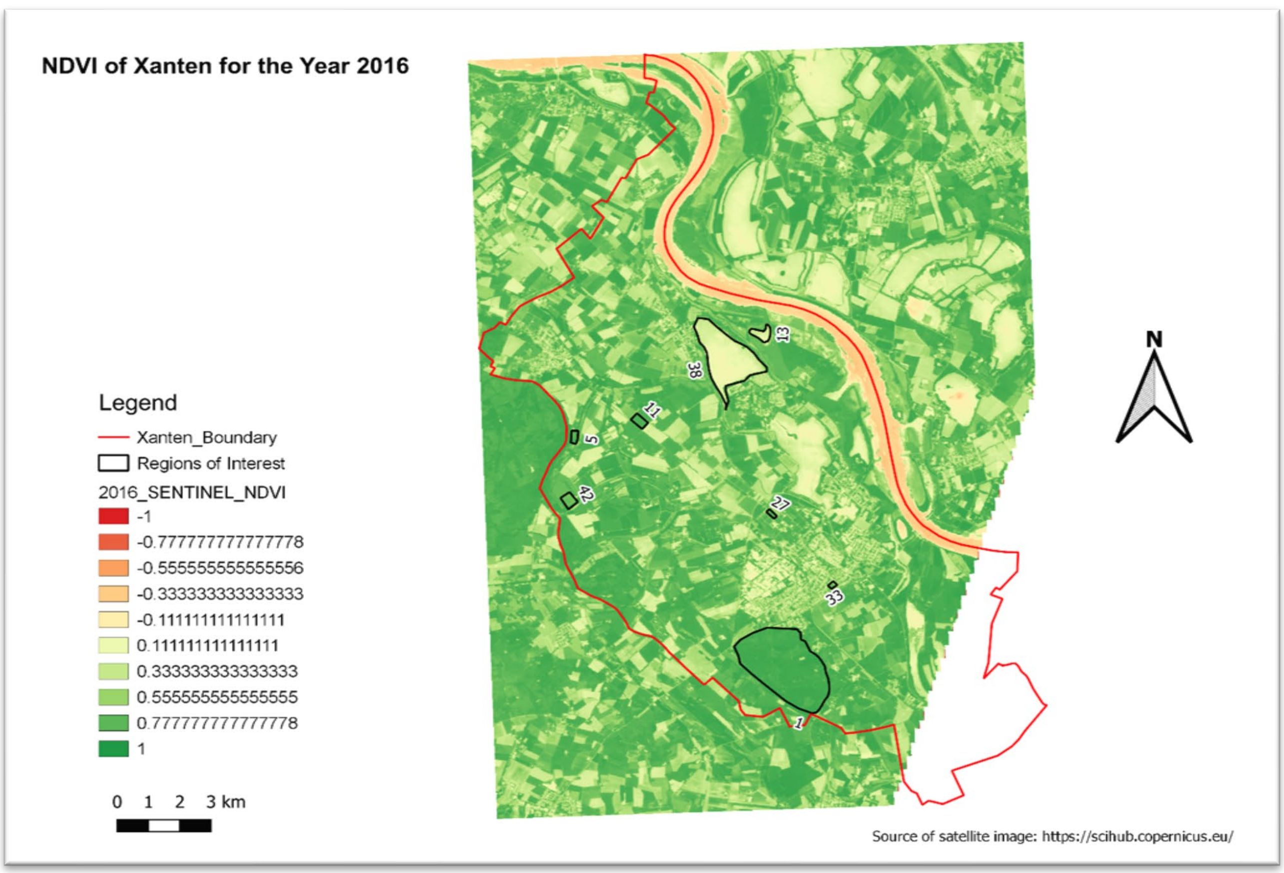

Figure 16:NDVI for the Xanten region with the regions of interest highlighted for the year 2016.

Table 4: NDVI numerical figures for the year 2016 extracted from the map

	ROI 1	ROI 5	ROI 11	ROI 13	ROI 27	ROI 33	ROI 38	ROI 42
08.05.2016	0.845	0.913	0.525	0.140	0.831	0.606	0.081	0.912

From figure 16 and table 4 above it can be observed that for the year 2016 there is a high NDVI associated with the regions of interest labelled 1, 5, 11, 27 and 42. This corresponds with the land cover associated with the regions as was shown in Figure 1 above. It can also be seen that the NDVI for the region 33 is moderate as it is associated with discontinuous urban fabric with the lowest NDVI being for regions 13 and 38 which are water bodies.

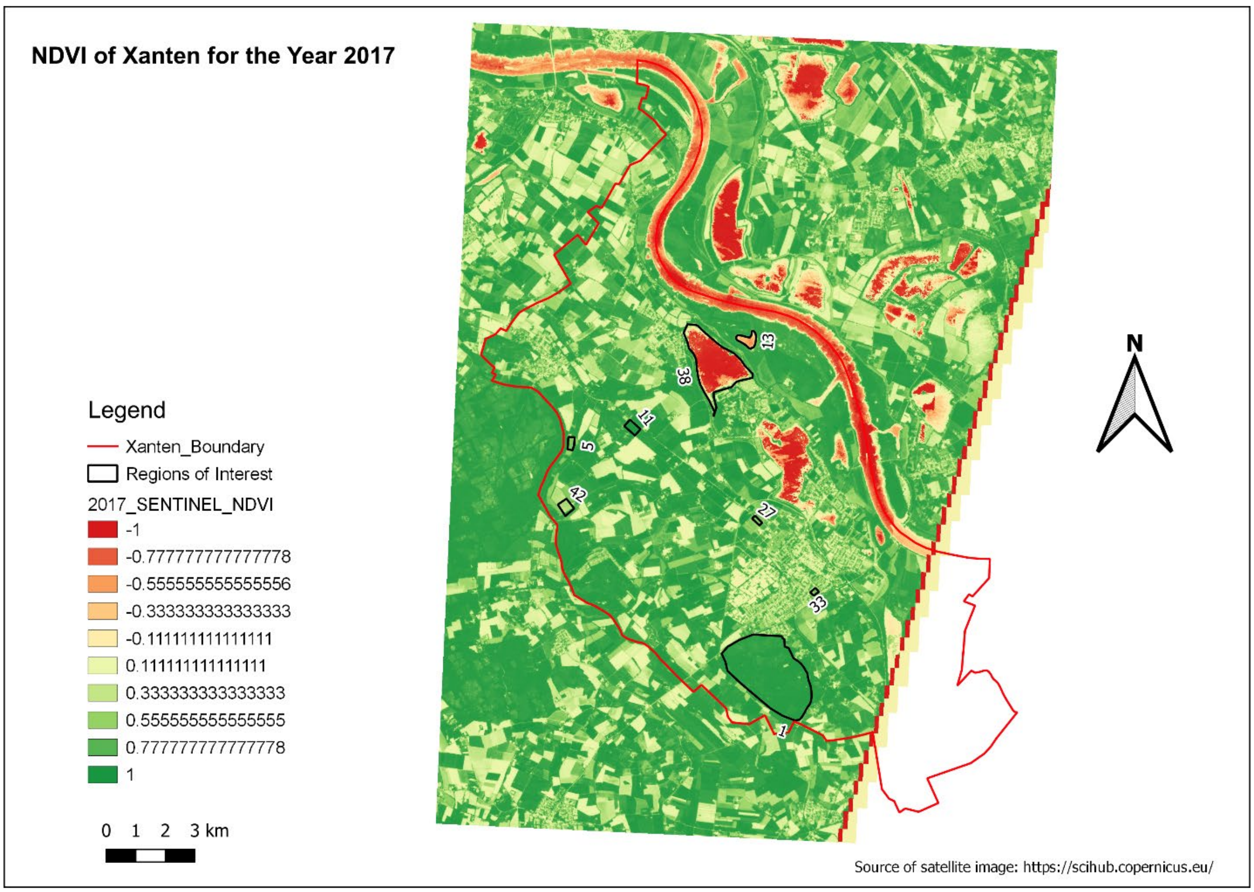

Figure 17:NDVI for the Xanten region with the regions of interest highlighted for the year 2017.

Table 5: NDVI numerical figures for the year 2017 extracted from the map.

	ROI 1	ROI 5	ROI 11	ROI 13	ROI 27	ROI 33	ROI 38	ROI 42
26.05.2017	0.931	0.643	0.973	- 0.343	0.610	0.571	- 0.616	0.519

Figure 17 and table 5 above represent the NDVI for the Xanten region including the regions of interest for the year 2017. Immediately it can be observed that the NDVI value increases for the regions of interest 1 and 11 but decreases for all the other regions.

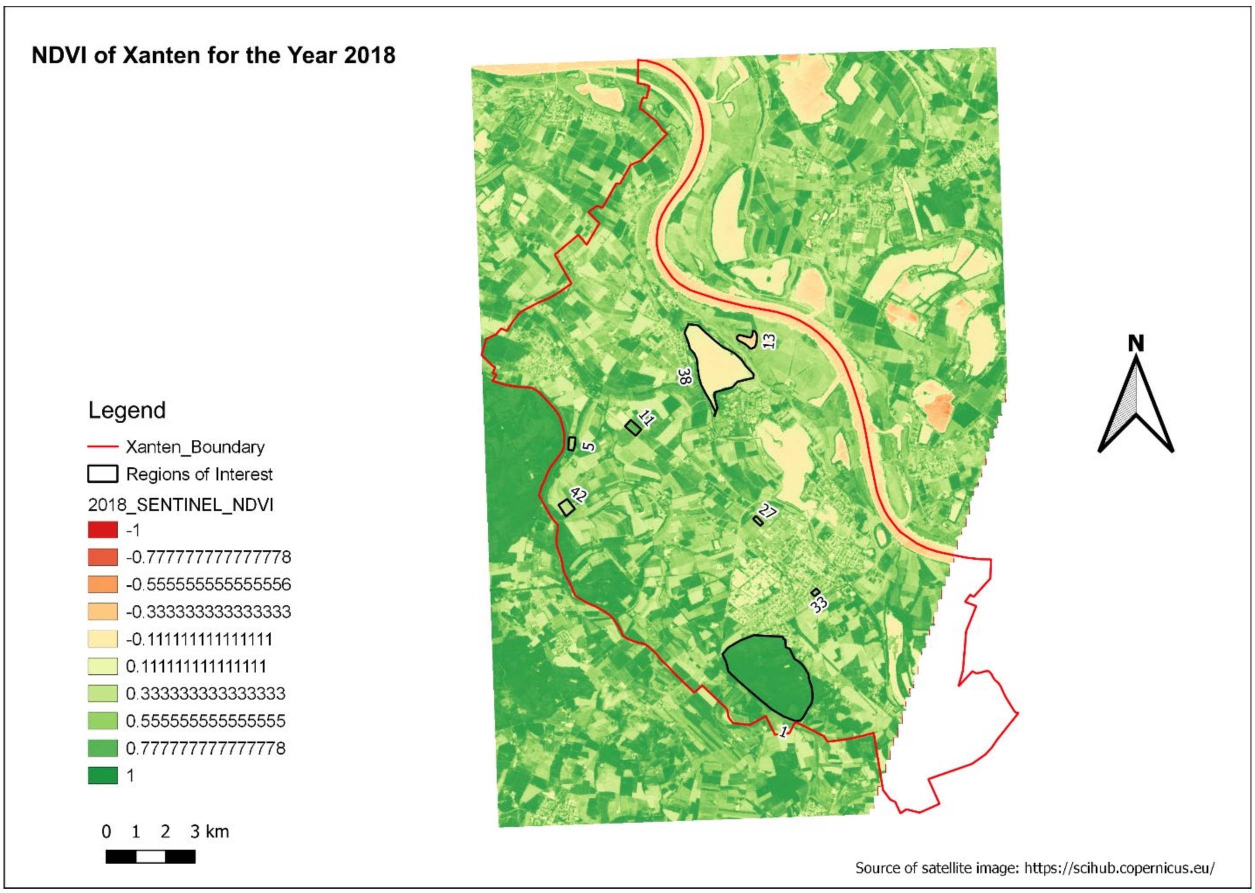

Figure 18:NDVI for the Xanten region with the regions of interest highlighted for the year 2018.

Table 6: NDVI numerical figures for the year 2018 extracted from the map.

	ROI 1	ROI 5	ROI 11	ROI 13	ROI 27	ROI 33	ROI 38	ROI 42
30.06.2018	0.899	0.394	0.737	- 0.169	0.465	0.472	0.021	0.410

Figure 18 contains the NDVI map of Xanten and the regions of interest retrieved from QGIS. The subsequent extracted values are contained in table 6. These values are for the year 2018. It was observed that the NDVI values for the regions of interest associated with land areas experienced a drop in their NDVI values, regions 5 and 11 especially experienced a greater drop compared to the other land area-based regions of interest. However, regions of interest 13 and 38 experienced an increase in their NDVI values despite being water body-based regions of interest.

The NDVI results obtained could not be trusted for the regions of interest except regions of interest 1, 13 and 38. This is because from the topographical analysis it was determined that they had experienced human interference which would affect the NDVI values. Results for regions of interest 13 and 38 were not considered as well as they are waterbodies hence, they have no vegetation. This made using the NDVI data from the region of interest (ROI 1) the logical reason as it is untouched forest. The plummeting of the NDVI value for the region of interest 1 from the year 2017 to the year 2018 suggests that the vegetation contained was under stress which is a good indication of a drought event. However, this data is not enough to confirm or refute the possibility of drought occurrence, but only to shed a light on where the deeper analysis should be conducted. This is because NDVI cannot be used to tell what the cause of the vegetation's stress was. This means that NDVI alone is not enough to prove drought. It must be used in conjunction with other data to draw a conclusion.

3.4 Correlation

The overall motto of this investigation is to verify or disprove the hypotheses that are considered at the beginning of this paper. So, all the data that was extracted through the various operations were merged as in table 6 below to find the correlation between the average temperature, the cumulative precipitation and the NDVI values of the region of interest (ROI 1).

Table 7:Variables and the values considered for correlation

Time	Cum_Precipitation(mm)	Avg_Temperature(°C)	NDVI ROI1 (Forest)
2016	829.4	11.25	0.845702
2017	874.0	11.05	0.930787
2018	738.6	11.34	0.899335

Table 8: Correlation values for temperature, precipitation and NDVI for ROI 1 for 2016, 2017, 2018

Correlation	Cum_Precipitation	Avg_Temperature	NDVI ROI1 (Forest)
Cum_Precipitation	1.0000	- 0.915032	0.178731
Avg_Temperature	- 0.915032	1.0000	- 0.560431
NDVI ROI1 (Forest)	0.178731	- 0.560431	1.0000

The below heat map was generated to visualize the correlation based on table 8 above.

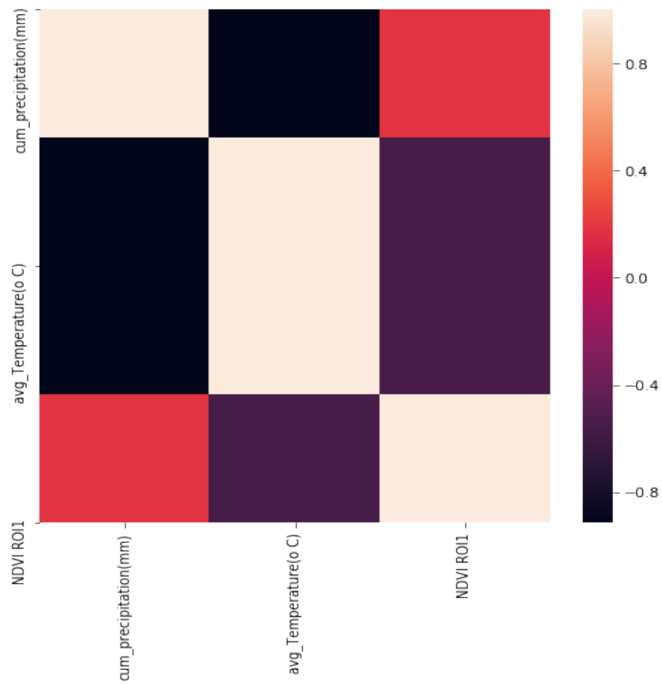

Figure 19:Heatmap for the correlation

From table 8 and figure 19, it is observed that there is a negative strong correlation (-0.915032) existing between cumulative precipitation and average temperature values, this means that when one of them increases then the other decreases. But the change in temperature for progressive years (2016-2018) is so minimal that the correlation value of average temperature with NDVI should not be considered. Moreover, as discussed earlier, the estimated temperature might be highly deviated from the actual temperature value due to the altitude variation which could have resulted in high correlation value between average temperature and precipitation. Additionally, the correlation between average temperature and NDVI is also negative (-0.560431) but the strength of the relationship is moderate and that means with an increase in temperature value, the NDVI value decreases and the change is moderate.

On the other hand, the correlation between the cumulative precipitation and the NDVI value of the ROI 1 is found to be positive (0.178731) but the relationship is weak which implies with increase/decrease in precipitation, the NDVI value increases/decreases too but the change is very minute. Nevertheless, the correlation value is found to be significant as the actual precipitation data is used for cumulative precipitation and correlated with NDVI data for the ROI.

4. Conclusion and Outlook

Several methods for measurement of drought conditions were presented and compared with each other in order to test several hypotheses. NDVI, precipitation and temperature data were acquired for the Xanten region based on several regions of interest. Through comparison of this data, a determination was made as to whether all these factors were responsible for the drought that occurred during the year 2018 or whether only one or two of the factors was responsible. After considering data between the years 2016 and 2018, it was discovered that the NDVI data could not be used by itself as an indicator for drought. This suggested that either more NDVI data sourced from more sensing dates would be required in order to get a more comprehensive picture or the NDVI data would need to be combined with other parameters to acquire better results. The temperature was not considered as the reason for the drought because it only fluctuated slightly between the years when the data was analysed. It was noted that it would have been advantageous if the temperature data would have been sourced from the singular Xanten station instead of using the mean of several stations around Xanten, unfortunately, that data did not exist in the Xanten station. However, when precipitation data were analysed, a large drop was noted between 2017 and 2018 which suggested that the dip in precipitation was the cause for the drought that was experienced.

It was observed that taking a single attribute for estimating climate data with spatial analysis will not be productive as other parameters can enhance or degrade its value. This may lead to inaccurate results.

For future application more NDVI data from multiple sensing dates should be acquired instead of relying on one date to represent the NDVI of a region for an entire year. This data is not enough to get accurate information concerning the explanation behind the plant life status or stress condition.

In order to continue the investigation, soil moisture should also be considered as a factor to be analysed because soil moisture is measured during analysis for drought (Tate and Gustard, 2000). This is especially because agricultural drought is a resultant of the lessening of soil moisture which hinders vegetation growth. This can be used in tandem with the other factors such as NDVI, precipitation and temperature to gain more insight into the investigation.

5. References

1. BKG. (n.d.). BKG - Bundesamt für Kartographie und Geodäsie. [online] Available at: https://gdz.bkg.bund.de/index.php/default/open-data.html [Accessed 27 Feb. 2020].
2. Copernicus Land Monitoring Service. (n.d.). CORINE Land Cover — Copernicus Land Monitoring Service. [online] Available at: https://land.copernicus.eu/pan-european/corine-land-cover [Accessed 28 Dec. 2019].
3. Copernicus Open Access Hub. (n.d.). Open Access Hub. [online] Available at: https://scihub.copernicus.eu/ [Accessed 20 Jan. 2020].
4. Hirt, C. (2014). Digital Terrain Models. Encyclopedia of Geodesy, pp.1-6. DOI: https://doi.org/10.1007/978-3-319-02370-0_31-1
5. Jupyter.org. (n.d.). Project Jupyter. [online] Available at: https://jupyter.org/ [Accessed 6 Jan. 2020].
6. Matplotlib.org. (n.d.). Matplotlib: Python plotting — Matplotlib 3.1.3 documentation. [online] Available at: https://matplotlib.org/ [Accessed 6 Jan. 2020].
7. McMahon, T.A. and Diaz Arenas, A. (1982) Methods of Computation of Low Streamflow, Studies and Reports in Hydrology 36 UNESCO Paris 95.
8. Nsg.naturschutzinformationen.nrw.de. (n.d.). Naturschutzgebiete und Nationalpark Eifel in NRW - Fachinformation - Gebietslisten - Naturschutzgebiete gesamt - NSG Reeser Schanz, Rheinaue zwischen Obermoermter und Vynen, bei Gut Grindt und Haus Luettingen. [online] Available at: http://nsg.naturschutzinformationen.nrw.de/nsg/de/fachinfo/gebiete/gesamt/WES-011) [Accessed 2 Dec. 2019].
9. Pandas.pydata.org. (n.d.). pandas - Python Data Analysis Library. [online] Available at: https://pandas.pydata.org/ [Accessed 6 Jan. 2020].
10. Qgis.org. (n.d.). Welcome to the QGIS project!. [online] Available at: https://www.qgis.org/en/site/ [Accessed 5 Jan. 2020].
11. Rouse, J.W. Jr., R.H. Haas, D.W. Deering, J.A. Schell, and J.C. Harlan, 1974. Monitoring the Vernal Advancement and Retrogradation (Green Wave Effect) of Natural Vegetation.
12. Sedgwick, P. (2012). Pearson's correlation coefficient. BMJ, 345(jul04 1), pp.e4483-e4483.
13. Sentinel-hub.com. (n.d.). NDVI (Normalized Difference Vegetation Index) | Sentinel Hub. [online] Available at: https://www.sentinel-hub.com/eoproducts/ndvi-normalized-difference-vegetation-index [Accessed 27 January. 2020].

14. Tate E.L., Gustard A. (2000) Drought Definition: A Hydrological Perspective. In: Vogt J.V., Somma F. (eds) Drought and Drought Mitigation in Europe. Advances in Natural and Technological Hazards Research, vol 14. Springer, Dordrecht. DOI: https://doi.org/10.1007/978-94-015-9472-1_3
15. Willie, M. (2020). Sentinelsat — Sentinelsat 0.13 documentation. [online] Sentinelsat.readthedocs.io. Available at: https://sentinelsat.readthedocs.io/en/stable/ [Accessed 12 Dec. 2019].

Group member	Contribution
Alex-Kagwi Mwenja (26497)	Introduction, Region of Interest research and documentation, Satellite imagery and NDVI data processing, NDVI results and discussion, Conclusion.
Rohit Turankar (26509)	Altitude data aggregation, Interpolation of temperature data and altitude, Digital terrain model creation and analysis, Documenting Results and Discussion
Rajesha Koppa Ramesha (26566)	Temperature & precipitation data gathering, data aggregation using python, Correlation analysis, Documenting materials and methods, Results and discussion.

DECLARATION OF AUTHENTICITY

We hereby declare that our contribution to the work presented herein is our 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 regarding 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. We will retain a copy of this assignment until after the Board of Examiners has published the results, which we will make available on request.

Alex-Kagwi Mwenja (26497)

Rohit Turankar (26509)

Rajesha Koppa Ramesha(26566)

Kamp Lintfort, 2020