



Master Thesis

submitted within the UNIGIS MSc program
Interfaculty Department of Geoinformatics - Z_GIS
University of Salzburg

Agricultural Change Detection Under Military Conflict Using the Example of Ukraine War

by

B.Sc. Torge Finn August Brunhorn

Student number 106946

A thesis submitted in partial fulfilment of the requirements of
the degree of
Master of Science – MSc

Advisor:

Prof. Michael Leitner

Göttingen, July 28, 2023

*“After all, no one is stupid enough to prefer
war to peace; in peace sons bury their
fathers and in war fathers bury their sons.”*

Herodotus, Greek Geographer, 484 – 425 BC

Acknowledgment

I would like to take a moment to express my gratitude and appreciation to the individuals who have played an integral role in supporting me throughout my master's thesis journey. I thank Professor Michael Leitner for supervising this thesis, whose guidance, expertise, and constructive feedback helped me a lot. The support and encouragement were a constant source of motivation. To my family, friends, and girlfriend, I extend my thanks for their love, encouragement, and unwavering support. Their belief in me, even during moments of self-doubt, was instrumental in helping me persevere through the challenges of graduate school. They provided a source of strength throughout my academic journey. I am grateful to have been blessed with such a supportive network of individuals who have stood by me through the time of writing this thesis. Their contributions have been invaluable, and I could not have achieved this milestone without them.

Göttingen, July 28, 2023

Torge F. A. Brunhorn

Science Pledge

I hereby declare that the thesis is entirely the result of my work. I have cited all sources I have used in my thesis; I have always indicated their origin. This thesis was not previously presented to another examination board and has not been published.

Göttingen, July 28, 2023

Torge F. A. Brunhorn

Content

1. Introduction	1
2. Environment, Land-Use Changes, and Military Conflicts	3
2.1 Literature Review	3
2.2 Historical Overview	7
3. Investigation Sites	9
3.1 Site A “Mezhova”	10
3.2 Site B “Mospyne”	11
3.3 Site C “Rozivka”	12
4. Methods	14
4.1 Climatological Approach	15
4.1.1 Data	15
4.1.2 Processing in R	17
4.2 Remote Sensing Approach	21
4.2.1 Data	21
4.2.2 Object Based Image Analysis	23
4.2.4 Indices	26
4.2.3 Google Earth Engine	29
4.2.5 Procedure for Determining Land Cover Changes	30
5. Results	34
5.1 Climatological Trends	34

5.2 Spatio-Temporal Comparison	42
5.2.1 Validation Results	42
5.2.2 Trends.....	45
5.2.3 Site to Site Comparison.....	49
6. Discussion	53
6.1 Result Interpretation.....	54
6.2 Model Uncertainties	58
7. Conclusion.....	62

Figures

Figure 1: Investigation sites in eastern Ukraine	9
Figure 2: Site A (False color composite).....	11
Figure 3: Site B (False color composite).....	12
Figure 4: Site C (False color composite).....	13
Figure 5: Model of the methods used in this thesis.....	14
Figure 6: Climate stations near the investigation sites	17
Figure 7: Example of the training and validation points in site A	23
Figure 8: Autocorrelation function for the temperature and precipitation	36
Figure 9: Mean temperature trend for all stations	41
Figure 10: Mean precipitation trend for all stations	42
Figure 11: Results of the classification process	45
Figure 12: BSI time series for site A.....	47
Figure 13: NDVI time series for site A	48
Figure 14: NPCRI time series for site A	48
Figure 15: BSI 2019 site A.....	52
Figure 16: NDVI 2019 site A.....	52

Tables

Table 1: Specification of the climate stations	17
Table 2: Correlation matrix of the temperature from the climate stations	35
Table 3: Correlation matrix of the precipitation from the climate stations	35
Table 4: Decomposed time series of temperature and precipitation of chosen stations.....	37
Table 5: Trendline for the last ten years for temperature and precipitation	39
Table 6: Comparison of results for different buffer size and seed spacing.....	43
Table 7: Confusion matrix for the classification results of all investigation sites	44

Abbreviations

OBIA – Object Based Image Analysis

GEE – Google Earth Engine

LULC - Land Use/Land Cover

AOI – Area of Interest

SNIC – Simple Non-Iterative Clustering

PCA – Principal Components Analysis

RF – Random Forest

GLCM – Gray Level Co-Occurrences Matrix

NOAA – National Oceanic and Atmospheric Administration

DWD – Deutscher Wetterdienst (German Meteorological Service)

Abstract

This master thesis explores agricultural change detection during the Ukrainian War, focusing on the impact of conflict on agricultural areas. The study integrates climate trend analysis, land cover classification, and index-based change assessment to determine if the climate is responsible for agricultural changes. By analyzing long-term climate patterns and station data over 30 years, data show the temperature has consistently risen. The data exhibit correlations among stations, especially considering temperature. The climate analysis indicates that the temperature and precipitation in 2022 do not differ significantly from previous years. Object-based image analysis (OBIA) combined with the Google Earth Engine enables precise classification of agricultural fields with over 91% accuracy. Indices such as NDVI, BSI, and NPCRI reveal a significant decline in agricultural vitality, with healthy field areas experiencing a decline of 66% to 94% between 2021 and 2022. The findings also indicate neglected and unfertilized fields. The study demonstrates the potential of integrating climate analysis and remote sensing techniques for investigating agricultural change in conflict regions. It emphasizes the importance of understanding the dynamics of war in the agricultural sector. This research contributes to the existing knowledge by providing evidence of the impacts of conflict on agricultural productivity and the potential of advanced technological tools for change detection. Further exploration in larger areas of Ukraine or application to other conflict zones will be necessary. Obtaining a comprehensive understanding of the environmental and human impacts of armed conflict can aid in minimizing losses and suffering in affected regions.

1. Introduction

War in Europe? Unthinkable. For a long time, this was the mindset of many European nations. Conflicts outside of Europe were observed, and occasionally interventions took place, but the idea of a war in the European neighborhood was inconceivable. This changed in early 2022 when the conflict in Ukraine escalated into a war, marking a watershed moment for Europe. However, the wound inflicted on Europe is nothing compared to the wounds suffered by Ukraine. The psychological and physical wounds on individuals and societies can be long-lasting and may affect generations to come. The toll on human life and well-being is a tragic consequence of war that is impossible to justify or overlook. The suffering caused by direct war impacts is followed by indirect factors such as famine. These effects can be mitigated through targeted organization and foreign aid. However, the implementation of these measures requires reliable information to be effective.

The conflict in Ukraine, which began in 2014, has resulted in enormous damage to infrastructure and disruption of economic activity, including the agricultural sector. The conflict has caused significant displacement of Ukrainian people, with over 2 million people being forced to flee their homes, many of them were employed in agriculture. In addition, the conflict has influenced the supply chain for farmers and disrupted trade, leading to reduced productivity and therefore reduced income for farmers.

Agricultural change detection, or the process of identifying and analyzing changes in agricultural land use, is a vital tool for understanding the impacts of military conflict on local communities and the environment. This is especially true in the case of the ongoing conflict in Ukraine, which has had significant and far-reaching effects on the country's agricultural sector. Agricultural change detection can help to quantify and understand these impacts, as well as identify potential opportunities for intervention and recovery. By analyzing satellite imagery

and other data sources, it is possible to visualize changes in land use patterns, such as the abandonment of farmland. This information can be used to assess the influence of the conflict on agricultural production on which the livelihoods of farmers strongly depend. Potential areas for rehabilitation and recovery can be identified as well.

In addition to its importance for understanding the impacts of the current conflict on the agricultural sector, agricultural change detection also helps to inform policy and decision-making that is related to resource and land use management. By identifying changes in land use patterns over multiple years, it is possible to understand the drivers of these changes and the potential consequences for food security and the environment. This information can be used to influence strategies for sustainable land use and resource management, as well as to identify potential areas for intervention and support.

The aim of this investigation is to determine how agricultural production responds to conflicts by mapping and quantifying changes in Ukraine. The goal is to

- I. examine if changes in agriculture are caused by climatological processes,
- II. classify agricultural fields to separate them from other landcover types,
- III. identify changes in vitality of the vegetation within the classified fields.

This may help to gain a better understanding of war-related changes on land use and food security. To minimize environmental bias, climatological changes must be considered to determine whether changes have originated from war or natural reasons. For a Ukraine-size area, cloud computing is best suited for handling this amount of data. In this thesis “Google Earth Engine” will be used. Landsat-8 and Sentinel-2 Imagery provide continually imagery for the conflict region. The imagery will be trained and classified in an object-based process to achieve the best results for agricultural areas. Additionally, classified data will be used to determine the amount and places of possible changes by calculating multispectral indices.

2. Environment, Land-Use Changes, and Military Conflicts

This literature review will critically analyze the existing literature on the topic of agricultural change detection, with the aim of contributing to the development of a deeper understanding of the subject and identifying avenues for future research.

2.1 Literature Review

For thousands of years, food supply and thus agriculture has been a major concern of mankind. Regarding this background the increasing population and climate change are posing a problem in distribution of resources (Karthikeyan et al., 2020). But not just population and climate change represent a turning point for agriculture but also conflicts. For example, agricultural fields are no longer cultivated because of the combat or can no longer be cultivated due to destruction (Kaplan et al., 2022).

It has long been known that climatic changes and the associated influences on the environment can lead to conflicts and that this will probably intensify in the future (Mach et al., 2019). On the other hand, less research has been devoted to the opposite question of how conflicts affect the environment. Particular attention should be paid to agriculture and its key role as a factor in food security. Comparing publications reveals a similar picture at the global level namely that conflicts significantly affect agriculture.

A recent publication focuses on analyzing the impact of the Nigerian conflict regarding the terrorist group “Boko Haram” on agriculture, with a particular emphasis on the direct effects on output, input, infrastructure, and human capital. This means that multiple aspects were investigated. The findings show that the intensity of Boko Haram attacks has a significant negative effect on total agricultural output and productivity, while land use remains unaffected. Specific staple crops such as sorghum, cassava, soya, and yam are also negatively impacted by

the conflict. Additionally, the conflict leads to a reduction in hours of hired labor for both men and women (Adelaja & George, 2019). Other publications from Myanmar (Aung, 2021) and Syria (Hazaymeh et al., 2022; Mohamed et al., 2020) support the hypothesis that conflicts lead to a decline in agriculture.

Also, the Ukraine is facing environmental problems due to the war. There are several environmental consequences, e.g., on soil, water, or the ecosystem but also on agriculture (Rawtani et al., 2022). The intense deforestation has resulted in a drastic impact on different ecosystem services and biodiversity, which may have potential implications for wildlife. Soil degradation and landscape morphology are likely to be negatively affected by bombing, trench, and tunnel excavations. This is of particular concern in the Ukraine, because of the fertile Chernozem soil, as it impacts food production (Pereira et al., 2022).

A recent study aimed to understand the impact of the on-going conflict in the Ukraine on the country's agriculture, particularly in the Donetsk and Luhansk regions. The rural regions have been significantly affected by the conflict since 2014, resulting in a decline in industry and an increased reliance on agriculture. Unfortunately, reliance is accompanied by a decline in agricultural production. Using satellite data, the researchers mapped cropland areas in 2013 and 2018 and found that there were cropland losses in the regions, with more substantial losses occurring in areas not under the control of the Ukrainian government and within a buffer zone along the conflict border line. The losses in this areas are between 22% - 46% (Skakun et al., 2019). On a global scale a decline in wheat export can be measured (Mottaleb et al., 2022).

But there were research innovations not only in the thematic field of agriculture and conflicts, but also on the technical level, especially in remote sensing. Satellite-based remote sensing, i.e., the provision of data of the earth's surface through satellite images, is a key technology for detecting changes in agriculture that helps examining large areas (Huang et al., 2018). In

particular, the evaluation of optical, multispectral sensor data is often used to quantify agricultural yields (Weiss et al., 2020). During the past decades, there has been a significant increase in the use of remote sensing tools for various purposes in agriculture (Sishodia et al., 2020). Also, there is an increasing popularity in the field of remote sensing publications as the impact factor of different journals indicate e.g. in *Remote Sensing* (Remote Sensing, 2023). The availability of high-resolution satellite images, with advanced capabilities in terms of spectral, spatial, and temporal resolution, has played a crucial role in promoting the use of remote sensing for many purposes. These include nutrient application, crop monitoring, and disease management, among others (Sishodia et al., 2020).

The identification of spatial changes from satellite images taken at different times, caused by natural or man-made events, is referred to as change detection. This process is critical in the field of remote sensing, as it allows the monitoring of ecological changes and land cover alterations. Over time, numerous methods have been developed for analyzing remote sensing data, with newer approaches constantly being developed. The timely and accurate detection of changes in Earth's surface features provides a foundation for evaluating the interactions and relationships between human and natural phenomena, resulting in better resource management. Generally, change detection employs multi-temporal datasets to quantitatively examine the temporal impacts of the phenomenon (Asokan & Anitha, 2019). The methodology of classifying remote sensing data and the identification of changes strongly dependent on deep learning processes which has become a focus in research over the last years. Especially, fully supervised learning, fully unsupervised learning, and techniques based on transfer learning are the main deep learning methods used today (Khelifi & Mignotte, 2020). In the future there might be a significant change in the approach from detecting changes by comparing two points in time, to continuous monitoring of changes over time. This shift brings numerous advantages, such as the ability to monitor almost instantly in real-time. Foreseen developments involve the

use of multiple sensors for monitoring, greater emphasis on temporal precision of findings, extension of applications to larger geographical regions, and operational utilization of time series analysis (Woodcock et al., 2020).

In recent years, the availability of remote sensing data as open data has increased significantly. Data are becoming more extensive, more detailed, and available over longer periods of time. The larger the amount of data, the more computing power is required. For this reason, cloud-based methods are required. Google has provided an instrument for this since 2010, referred to as the Google Earth Engine (GEE). This cloud computing platform is also an opportunity for change detection in agricultural research (Amani et al., 2020). While most Land Use/Land Cover (LULC) focus on approved machine learning techniques such as random forest (RF), other methods appear that are also usable in the GEE. The Support Vector Machine (SVM) was identified as a best practice method for LULC classification in Iraq with an accuracy of 90% compared to ground control points (Feizizadeh et al., 2023). This is also confirmed by a study from Italy that shows advantages of SVM compared to RF. This research also highlights that a pixel-based approach is slightly better suited than an object-based approach using Landsat-8 imagery for the specific Area of Interest (AOI) mentioned in the paper. This might be due to the spatial resolution of Landsat-8 compared to the other systems used in this study and the very complex, small-size landcover mosaic of the AOI. Nevertheless, object-based image analysis is also a current trend in remote sensing and change detection. The authors highlight that the clustering method of Gray-Level Co-Occurrence Matrix (GLCM) produces the best results in object-based methods (Tassi & Vizzari, 2020). While using higher spatial resolution the results of the object-based classification become clearly better for LULC classification (Tassi et al., 2021).

A current conflict in Europe happens in the Ukraine. There has been internal political tension and military conflicts since 2013. As already mentioned, parts of the Ukrainian territory have

been surveyed for agricultural changes (Skakun et al., 2019). However, a comprehensive study of the agricultural changes regarding the invasion of the Russian army in February 2022 is not yet available. A further investigation into this topic will provide insights concerning the current food security. Regarding the topicality of this field of study, this thesis research is supposed to assess the effects of military conflicts on agricultural landcover/land use.

2.2 Historical Overview

Since the time this thesis is being written, the Ukraine has been looking back on a conflict-ridden past. Understanding this past is initially important to be able to comprehend the spatial aspects of the conflict. Unfortunately, the topicality of this conflict leads to a lack of scientific literature concerning the course of events in the Ukraine.

In November 2013, the Ukraine's government, led by pro-Russian President Viktor Yanukovich, made the decision to not sign a previously planned association agreement with the European Union. Yanukovich's decision to back out of the long-planned agreement leads to protests that erupted in Kyiv, known as the 'Euromaidan' demonstrations. These protests turned violent in early 2014, prompting European foreign ministers to mediate a compromise involving a unity government and early elections. However, on February 22, 2014, a power-sharing agreement collapsed, and President Yanukovich fled to Russia. Subsequently, a new government was installed by the Ukrainian parliament. In late February 2014, unidentified military personnel, later confirmed to be Russian forces, surrounded Crimea's airports. Pro-Russian forces then took over the Crimean autonomous assembly. In March 2014, the assembly declared independence, followed by a referendum leading to a union with Russia. Since then, Russia has maintained control over Crimea and has supported pro-Russian separatist forces that took over parts of the Donetsk and Luhansk regions of eastern Ukraine (the Donbas) in 2014. Despite the negotiation of the Minsk Agreements in 2014/2015, which called for a ceasefire,

the withdrawal of all foreign armed groups, and constitutional reforms, fighting between Russian-supported separatists and Ukrainian government forces has continued in the Donbas for the last eight years (Walker, 2023).

On February 24, 2022, Russia initiated a military operation in Ukraine by crossing the borders from Belarus in the north, Russia in the east, and Crimea in the south. President Wladimir Putin called it a "special military operation" aimed to protect the people of the Donbas by demilitarizing and denazifying Ukraine. He stated that Russia had no intention of occupying Ukrainian territory or using force to impose changes. However, over the past year, Russian forces have launched a large-scale assault on the Ukraine. In early October 2022, Russia signed annexation treaties, declaring Donetsk, Luhansk, Kherson, and Zaporizhzhia as part of the Russian Federation, even though these regions are not entirely under Russian control. In response, Ukraine, with the help of Western military assistance, has launched a counteroffensive and has made some territorial gains. Both sides are currently preparing for new offensives, with Ukraine vowing to recapture all its sovereign territory, including Crimea. The Kremlin, on the other hand, has declared that the annexed regions will remain part of Russia indefinitely. The conflict in the Ukraine has had significant regional and international consequences. It has resulted in economic sanctions being imposed on Russia by the United States and the European Union, and it has led to a deterioration of relations between Russia and the West (Walker, 2023).

The invasion has caused widespread unrest, leading to significant economic consequences including disruptions to supply chains both regionally and globally. Despite their relatively small economies, Russia and Ukraine are major exporters of essential commodities such as agricultural products. The 2022 invasion has already had significant negative effects on the economy, particularly in commodity markets, resulting in rising prices for commodities like oil, gas, and wheat (Kotoulas & Pusztai, 2022).

3. Investigation Sites



Figure 1: Investigation sites in eastern Ukraine

In this short chapter a brief overview of the study areas will be provided. The regional differences regarding the impact of the conflict are part of this investigation. Therefore, a brief geographic contextualization is necessary. There are three areas of interest that differ from each other by their localization. All three study areas are in the eastern region of Ukraine, mainly in

the Donetsk oblast (oblast = district). One study area is in the so-called separatist territories, another one is in the areas that have been occupied by the Russian army since its invasion. The last study area is located in a previously unoccupied part of Ukraine. Care was taken to ensure that the three study areas are located close to each other to prevent regional differences from causing distortions. These differences could be, for example, in land use, land surface structure, or mesoclimate between the individual areas of interest. All three study areas have a side length of 50 km, which result in an area of 2,500 km². Thus, each study area is approximately the size of the state of Luxembourg. All three study areas are visualized in Figure 1. This map also shows the areas under Ukrainian and Russian control. The individual areas of interest are briefly described below.

3.1 Site A “Mezhova”

This area, visualized in Figure 2, is located approximately 80 km northwest of Donetsk. The area is characterized by agricultural activities, with vast farmlands and rural settlements. The terrain is relatively flat, with low-lying hills and fertile soils. The region is also known for its mineral resources, such as coal, which has contributed to the development of the local economy. The two largest settlements are Mezhova and Udachne. The investigation site is mainly located in the Donetsk oblast and to a small extent in the Dnipropetrovsk oblast.

This area has always been within the sphere of influence of the Ukrainian government and has not yet been occupied by foreign military forces. Nevertheless, this area is very close to the frontline of the conflict and might be influenced by combat operations.

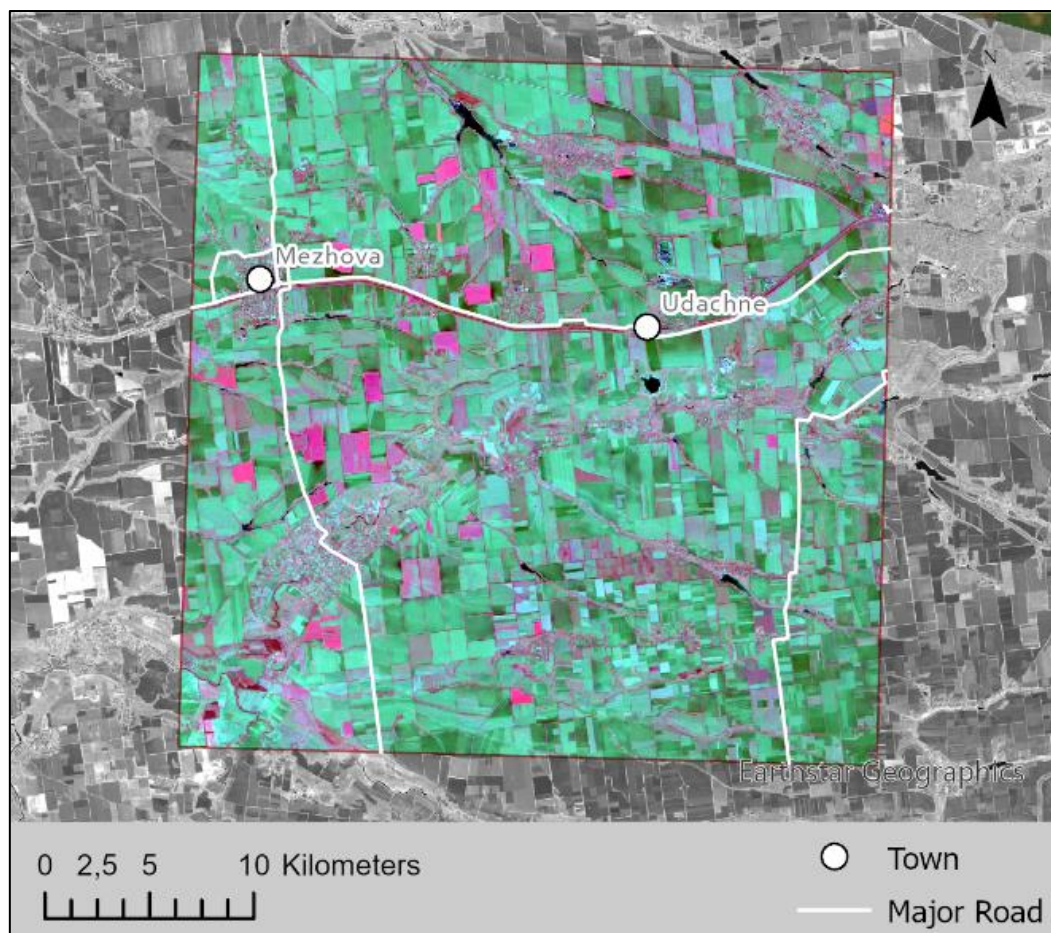


Figure 2: Site A (False color composite)

3.2 Site B “Mospyne”

Site B is located only about 20 km southeast of the city of Donetsk and is also strongly characterized by agriculture, although there are significantly more settlements due to its proximity to the city. The AOI is located exclusively in the Donetsk oblast and borders the Russian Federation in close proximity.

This area was taken over by pro-Russian separatists in 2014 and has since been under their influence, and therefore also under the influence of Russia. Figure 3 shows a map of the Mospyne investigation site.

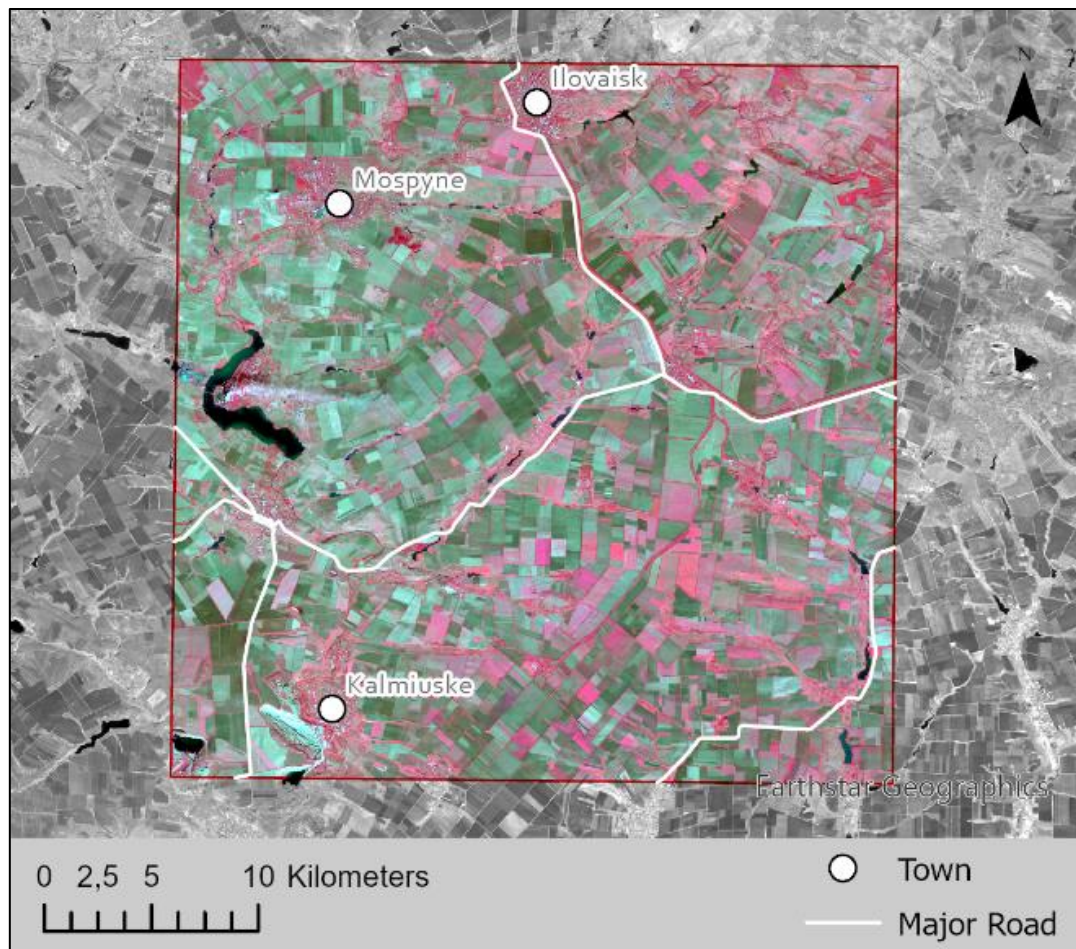


Figure 3: Site B (False color composite)

3.3 Site C “Rozivka”

This study area is in the Donetsk oblast and to a small extent in the Zaporizhzhia oblast. It is located approximately 90 km southwest of Donetsk, near the city of Mariupol, and thus not far from the Sea of Azov. In this agriculturally characterized area, there is only one significantly large city named Rozivka.

The area around Site C was under the control of the Ukraine for a long time until it was finally occupied by Russian troops in early 2022. Figure 4 represents the area around Rozivka.

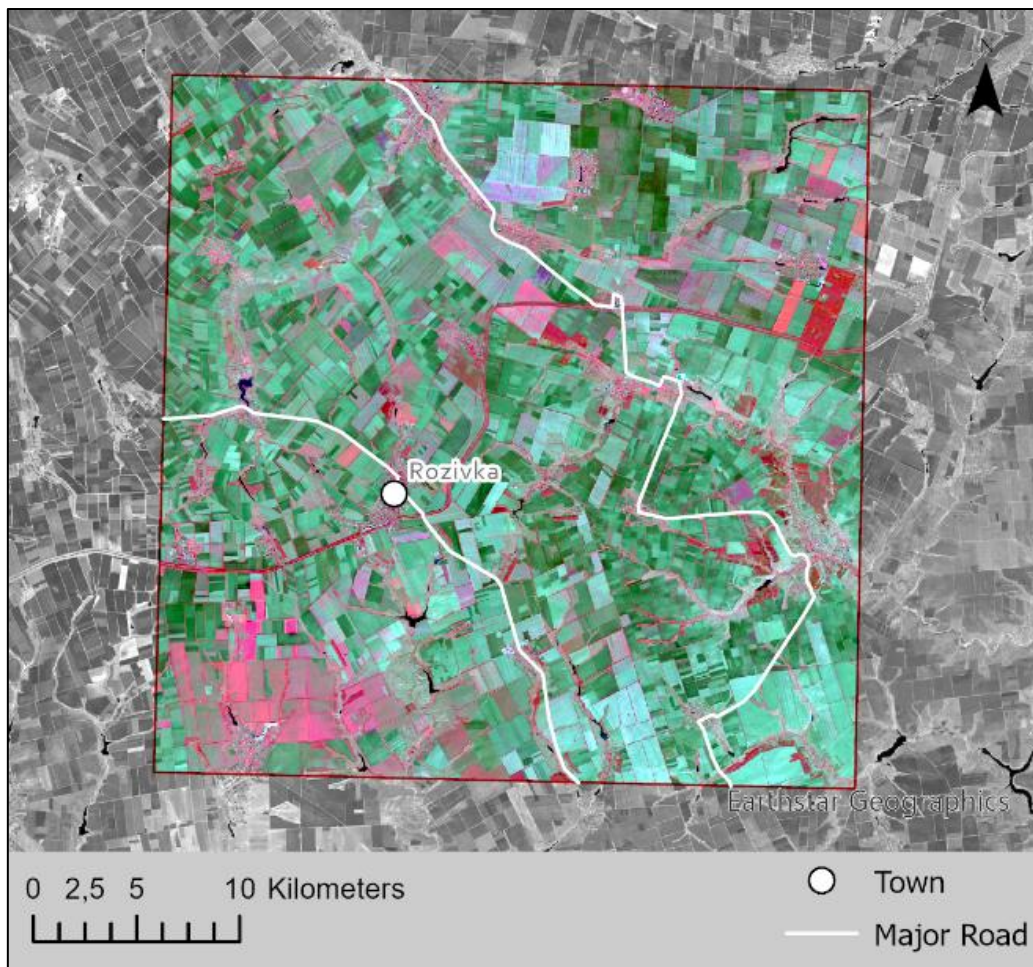


Figure 4: Site C (False color composite)

4. Methods

This chapter provides a closer examination of the methodological processes employed in this study. It encompasses both the theoretical principles and the practical implementation using the Google Earth Engine (GEE). This chapter is divided into two main methodological sections: the investigation of climate developments and the quantification of land use changes. The following graphic (Figure 5) provides an overview of the methodology underlying this study. The investigation of land use changes comprises data acquisition and preparation, identification through classification of agricultural areas, and the detection of changes within these areas. Each of these points will be further elaborated in the subsequent sections of this chapter.

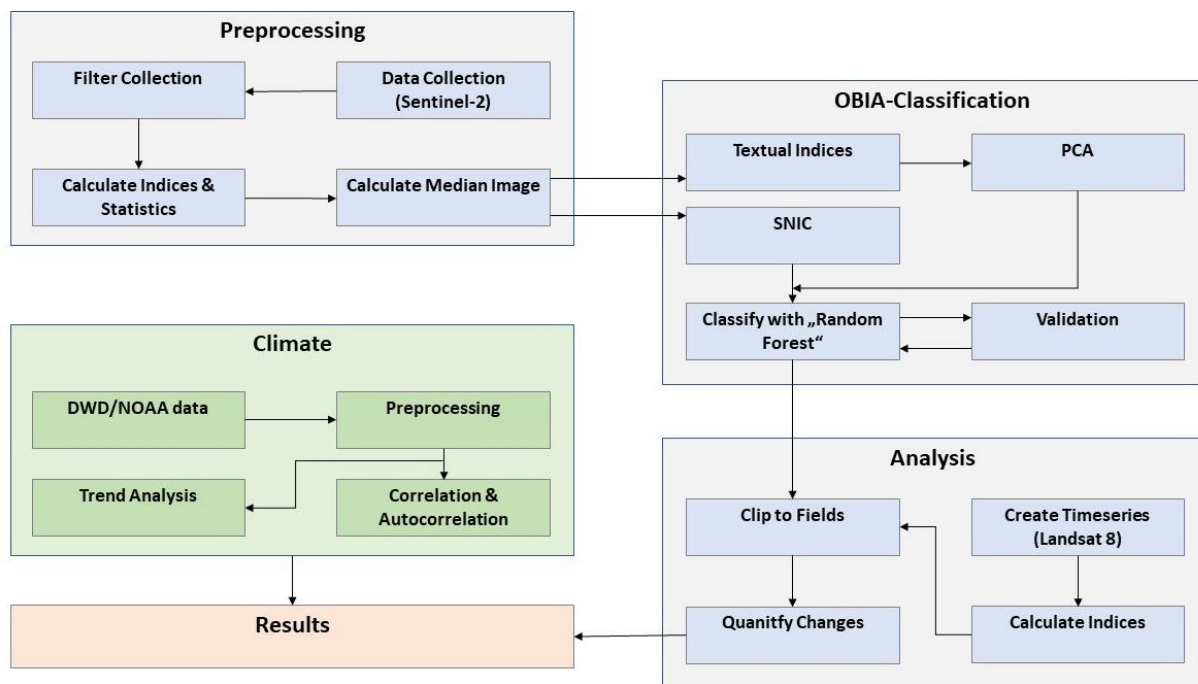


Figure 5: Model of the methods used in this thesis

4.1 Climatological Approach

In this chapter the development of climatological features is under investigation. Environmental bias must be precluded by assessing climate related changes. It is important to determine whether a change in land use is caused by conflict or altered climate and weather conditions. The two main climatological factors affecting crop yields are temperature and precipitation, which provide optimal growth conditions for plants. Therefore, this work will focus on precipitation and temperature. The climatic analysis will be based on time-series analysis methods in R with data from different climate stations. The purpose of this analysis is to gain an overview to set changes into relation to each other.

4.1.1 Data

The search for suitable data proved to be difficult. As the study areas are located in sparsely populated regions, there are few measurement stations available. Nevertheless, some stations in the vicinity of the study areas were identified. The fact that official climate stations worldwide have an identification number, the so-called World Meteorological Organization (WMO) Station Number, helped in this process. Using the number and a map, the stations can be well delimited. The stations are in the east of the Dnipro river, on the Ukrainian territory and on the Russian territory next to the border.

Another requirement for the data is that they cover a period of 30 years. The 30-year period is important because it is only after this time frame that a climate period can be considered representative of a trend. Shorter periods can be influenced by temporally limited weather phenomena. In total, nine climate stations were selected, some of which better meet the criteria than others. Only four stations provide continuous data for the entire 30-year period (Dnipropetrovsk, Kharkiv, Taganrog and Certkovo). Many stations have only been in operation for a shorter period of time or have no data available.

The data can be obtained from the freely accessible databases of the German Meteorological Service (The German translation is “Deutscher Wetterdienst” or “DWD”) for temperature (DWD Climate Data Center, 2017a, 2023a) and precipitation (DWD Climate Data Center, 2017b, 2023b) and the National Oceanic and Atmospheric Administration (NOAA) which is the US-American counterpart to the DWD (NOAA, 2023). These can be downloaded from there as Comma-separated values (CSV) files. Unfortunately, the availability of data is often very low. There are many gaps in the datasets of the DWD and particularly in those of NOAA. The gaps usually consist of a few months. Only data that had undergone prior control by the respective authority were used. Raw data were not used.

The data were initially processed and converted into a uniform format. For example, DWD has a column-wise notation of monthly values, while NOAA provides them row-wise. The datasets were transferred to EXCEL, checked, and compared with each other. Gaps in the DWD data were filled in with NOAA data where possible. The decimal separators were adapted to the European notation and then checked. In addition, an initial overview was obtained. The data were organized into three columns representing the date, the average temperature, and the cumulative precipitation. The date is on a monthly basis, so every year contains 12 rows. Subsequently, these data were further processed statistically in R. In figure 6 the locations of the climate stations are highlighted. Table 1 summarizes the specifications of all stations including the WMO Number and the date range.

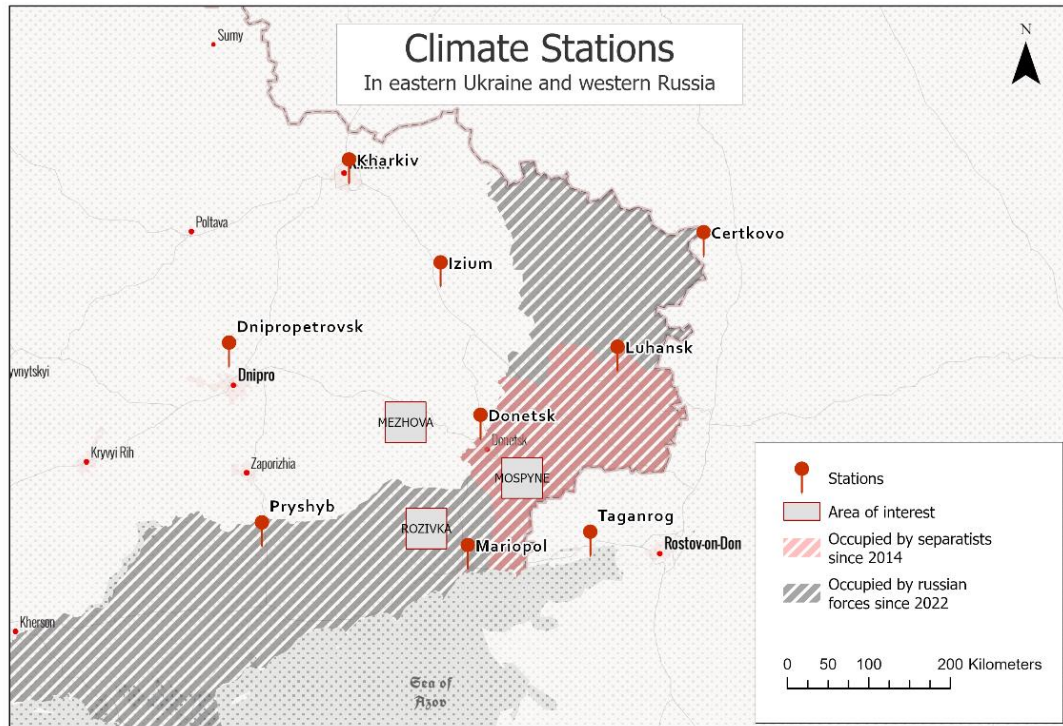


Figure 6: Climate stations near the investigation sites

Table 1: Specification of the climate stations

List of Stations				
Station	WMO number	Date range	Latitude	Longitude
Dnipropetrovsk	34504	01/1991 – 12/2022	48,60	34,97
Kharkiv	34300	01/1991 – 01/2021	49,92	36,29
Taganrog	34720	01/1991 – 12/2022	47,20	38,95
Certkovo	34432	01/1991 – 06/2019	49,40	40,20
Donetsk	34519	01/1991 – 06/2014	48,07	37,74
Luhansk	34523	01/1991 – 04/2014	48,57	39,25
Izium	34415	01/2004 – 02/2022	49,18	37,30
Mariopol	34712	01/2004 – 02/2022	47,10	37,60
Pryshyb	34607	04/2004 – 01/2022	47,27	35,33

4.1.2 Processing in R

In this section, the methodology concerning the analysis of the climatological data is under investigation. The processing will be performed with R software. R is a popular programming

language for data processing due to its tools and its flexibility. Data processing in R involves cleaning, manipulating, and organizing data in a way that makes it usable for analysis and visualization. R has a large and active community of users, so there is a wealth of resources available online for those looking to learn more about data processing (R Core Team, 2022).

One of the first steps in data processing is importing data into R. This can be done from a variety of sources, including CSV files or Excel sheets. R has several built-in functions for reading in data from these sources, such as `read.csv` and `read.xlsx`. After importing the data, it must be checked for data accuracy to be able to use it in any subsequent analysis. The exact procedure for data evaluation in R can be found in the attached script in Appendix 1. First, all data were imported into R, packages were imported, and the working directory was set. A package is a software extension that contains functionalities that are not in the basic R product.

After the data were checked for inconsistencies, especially for gaps and incorrect values, data can be manipulated and analyzed in R in different ways (Appendix 1). R has a wide range of functions and packages for manipulating data, including correlation analysis, and working with time series. First, the question arises to what extent the trend of an individual station can be extrapolated to the entire study area. It is therefore necessary to check the extent to which the data from one station matches the data from the other stations. This is important to ensure that there is a trend for the entire study area, despite the distances between individual measurement stations. For this purpose, a simple Pearson correlation can be used. Correlation describes the relationship between two variables. The correlation coefficient value provides information on the strength of the relationship. The Pearson coefficient refers to interval- and ratio-scaled variables and represents linear relationships.

A correlation matrix represents the correlation of several variables in a table format. This is also used in R. All stations are correlated with each other in a matrix, so that one correlation

coefficient is the output between every pair of stations. Here, the missing data present a problem because a correlation needs an equal number of data pairs. Otherwise, a correlation with "NA" values will also be "NA" as a result. This can be overcome by setting `use = "complete.obs"`. This eliminates all "NA" through pairwise deletion (R Core Team, 2022).

Differences between temperature and precipitation can also result from the extent to which these variables are subject to randomness. For example, if precipitation amounts vary greatly from one month to the next, the correlation may be affected, as the amount of precipitation can vary greatly not only between stations, but also for a single station. A simple way to check this is to use temporal autocorrelation, which can be accessed in R using the `acf()` function. Autocorrelation is well-known in spatial terms but can also be applied to time series. Temporal autocorrelation correlates (as the name suggests) a time series with itself - shifted by a certain time period. This shows the extent to which the values of the time series are dependent on each other. If the temporal autocorrelation lies within the confidence interval, it can be concluded that the value of one month is to some extent dependent on the previous months. If the temporal autocorrelation is outside of the confidence interval, it is evident that the value of a particular month has nothing to do with the previous month (R Core Team, 2022).

To analyze the data of each station, a time series analysis can be performed. R provides a powerful tool for this with the "decompose" function. A time series always consists of several factors: On the one hand, it is determined by seasonality. In the case of temperature, for example, this is the change of seasons. On the other hand, a time series is influenced by randomness, i.e., unpredictable changes. Finally, a time series has the trend component, which is the general development over several years. The decompose function breaks down a time series into these individual factors. This method can be used to quickly access the trend that is required for this work. First, the trend component is determined by the algorithm in R through a moving average. Then, seasonality is determined by calculating the mean for each time unit

and then centering the values. The randomness component is determined by subtracting the trend and seasonality from the original time series (R Core Team, 2022).

A problem arises due to gaps in the time series. A trend analysis requires complete time series, hence missing values need to be replaced. Various statistical methods are available for interpolating time series. For example, simple methods that linearly interpolate the value before and after the gap can be used. However, there are also more complex methods that use polynomial functions (such as Spline) or search for patterns in the time series to apply to these gaps. In general, the method of interpolation depends on the data and the size of the gaps. For example, it is not useful to choose linear interpolation for strongly seasonal data with a gap of > 6 months, as an entire seasonal cycle would be "cut off" in this way. The function `na.StructTS()` uses a Kalman filter to fill the data gap (Zeileis & Grothendieck, 2005).

Data can be stored in R as `ts` or `xts` elements. Both are data structures specifically designed for time series. This has advantages in processing, as each value is assigned a specific date (Ryan & Ulrich, 2023). Since both the 30 years overall trend and the 10 years study period are of particular interest, the trend is trimmed to the period after 2013 for clarity. This way, relevant fluctuations in temperature and precipitation can be more easily identified.

To obtain a clear representation of the trend, the measurement data from the nine stations were summed per month and divided by the available number of monthly measurements. This yields the average of all stations for each month. This averaged time series is not suitable for determining the exact temperatures or precipitation, but only for estimating the trend. This is because there are sometimes large differences in values between stations. However, as described at the beginning, a positive correlation allows this method to be used to estimate the trend. A positive side effect of averaging the values is that all gaps are closed, and a

homogeneous time series is created. This time series can be converted into an xts object for visualization and representation.

4.2 Remote Sensing Approach

The following section explains the methodology regarding the remote sensing part of this thesis. It describes both the data and methods used, as well as the approach taken.

4.2.1 Data

The data foundation of this study relies on freely accessible satellite images obtained from two different systems: Landsat 8 and Sentinel-2.

Landsat is an earth observation system developed by the National Aeronautics and Space Administration (NASA) and the United States Geological Survey (USGS). It has been collecting images for many decades. Currently, Landsat 8 is in orbit, providing continuous imagery of the Earth's surface with a resolution of 15 meters in panchromatic range and 30 meters in the visible and infrared range. Landsat 8 has been operational since 2013. Various levels of data processing are available (Roy et al., 2014). In this thesis, the LANDSAT/LC08/C02/T1_L2 collection was used, which includes images since March 18, 2013. The collection is geometrically and atmospherically corrected (*Landsat 8 Level 2, Collection 2, Tier 1*, 2023).

The Earth observation program of the European Space Agency (ESA) is called Copernicus and has been continuously providing images since 2015 through the Sentinel-2 satellites. Unlike Landsat 8, Sentinel-2 offers better spatial resolution of 15 meters in the bands relevant to this study (Phiri et al., 2020). The COPENICUS/S2_SR collection was used, which includes data from March 28, 2017, and is also geometrically and atmospherically corrected (*Sentinel-2 MSI*, 2023).

The classification of agricultural fields was performed using Sentinel-2 data, as they possess higher resolution and are considered more suitable for the methodology (Tassi & Vizzari, 2020). To establish a robust classification basis, multiple images are required. Particularly, images with high cloud cover cannot be utilized. To create a classification image, a composite was generated from several images by calculating the median value. The median value is resilient to outliers in the data (such as clouds), thus generating a cloud-free and temporally averaged image that serves as the basis for field classification.

For classification purposes, training and validation data are essential. A total of 500 randomly distributed points were created per study area using the ArcGIS Pro software. The entire set of 1,500 points was manually inspected, compared with satellite images, and divided into two groups: field and non-field. The resulting dataset was further divided into 300 training points and 200 validation points per study area. To enhance accuracy, a buffer area 7 to 8 meters was generated around the points. This buffer area calculates an average value from the surrounding pixels, thereby improving the classification by reducing pixel value deviations. An example of the point distribution in site A can be found in figure 7.

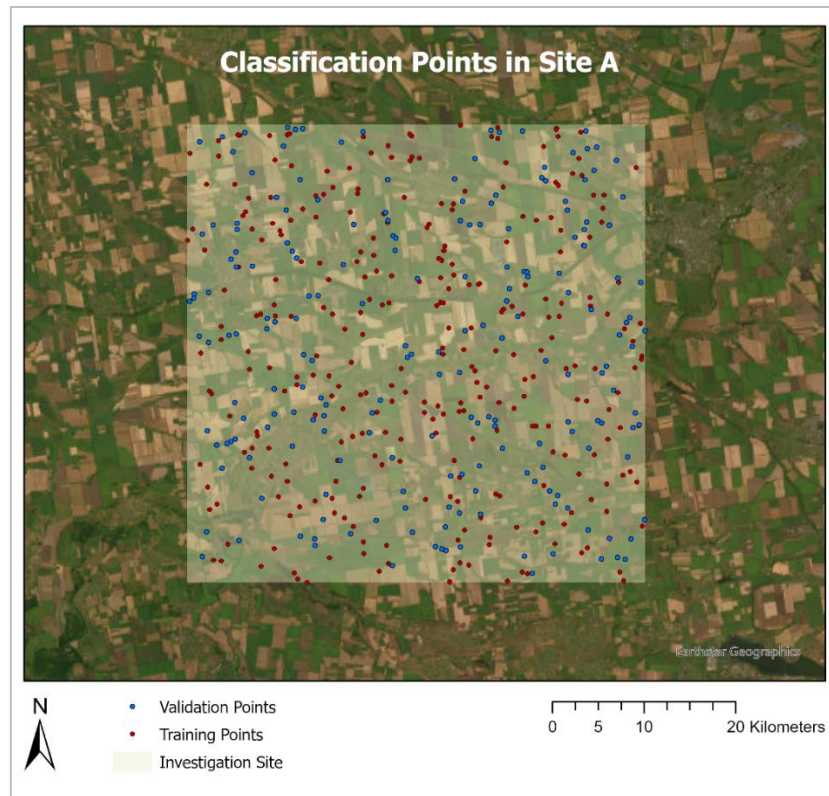


Figure 7: Example of the training and validation points in site A

The data were directly integrated using the Google Earth Engine, which provides access to a vast collection of remote sensing products. Additionally, the study area polygons were loaded into the GEE as a shapefile. The remote sensing data can be easily clipped to these study areas.

To assess changes, the Landsat data were initially utilized, as they have been available since 2013, which coincides with the beginning of the conflict. Sentinel data became available at a later stage and were subsequently included in the study for comparative purposes.

4.2.2 Object Based Image Analysis

Object-based image analysis (OBIA) is a method for analyzing remote sensing imagery. It involves partitioning the image into distinct objects or segments, based on the characteristics of the pixels in the image. So, this is an approach that groups pixels into meaningful objects based on their properties. Unlike traditional pixel-based approaches, OBIA considers image objects as the fundamental units of analysis. These objects can range from individual trees to entire

land cover classes, and they are characterized by their shape, size, texture, context, and spectral information (Blaschke, 2010).

OBIA can be seen as the counterpart to the pixel-based approach. Compared to traditional pixel-based approaches OBIA is becoming increasingly popular, because it allows a detailed and accurate analysis of the features and patterns in the image. OBIA is superior in classifying imagery with a high spatial resolution. By partitioning the image into discrete objects, OBIA can capture and analyze the characteristics of individual features in the image, rather than just the overall patterns and trends. This can be especially useful for analyzing images with a high degree of variability, such as those with a mix of different land covers or urban and rural areas. A lot of details will make it difficult to separate homogeneous features. For example, an agricultural field may contain some areas without any vegetation due to rainfall or wind. These small areas will be classified as features that are not a field in the pixel-based approach despite it is nevertheless part of a field (Hossain & Chen, 2019).

The first step in order to perform an OBIA is pre-processing the image. It means preparing the image for analysis by correcting any distortions or errors, such as atmospheric conditions or sensor noise. There are two main phases in the workflow of OBIA. Firstly, segmentation, which involves dividing an image into groups. Secondly, feature extraction and classification. Especially the segmentation is crucial for the quality of the results obtained.

The Simple Non-Iterative Clustering (SNIC) algorithm is used for segmentation of the imagery. Its advantage as a non-iterative algorithm offers high performance and makes it particularly suitable for large image files, like in remote sensing applications. Initially, the algorithm evenly distributes points, called "seed points," across the image. These seed points serve as the basis for individual clusters. Each seed point is moved to the adjacent pixel with the highest gradient, resulting in an ascending gradient. This process occurs within a specified range. Subsequently,

the distance from the center seed to the pixels is calculated, and all pixels within the corresponding radius are assigned to the same cluster. This is done simultaneously for all points and repeated until convergence is reached. In this way, the entire image is divided into individual segments (Achanta & Susstrunk, 2017).

In preparation for the classification, a good data foundation is necessary. For this purpose, an image is created to simplify the classification process. The required step to create such an image is the Gray Level Co-occurrence Matrix (GLCM), which relies on a previously generated grayscale image. This algorithm calculates relevant information from the grayscale levels of an image, referring to the precomputed segments (Sebastian et al., 2012). The algorithm computes seven relevant metrics (Angular Second Moment, Contrast, Correlation, Entropy, Variance, Inverse Difference Moment, and Sum Average) from the grayscale image. These metrics are then combined into a single image using Principal Component Analysis (PCA) (Tassi & Vizzari, 2020).

This PCA attempts to generate an image from multiple remote sensing bands using statistical methods. This image then contains the most important features of all bands. Thus, through this process, multidimensional data are reduced to a single dimension. The covariance is computed from the corresponding pixels of the bands and stored in a matrix. Subsequently, the eigenvector-eigenvalue component is calculated. Eigenvectors and eigenvalues provide a way to identify important directions (eigenvectors) within a geodata set and measure the variability (eigenvalue) along those directions (Eklundh & Singh, 1993).

Then, the eigenvectors are sorted based on the magnitude of their eigenvalues and reduced to the required number of bands. For instance, if there are 10 spectral bands, the newly computed band that accounts for e.g., 95% of the total variance is selected. This ensures that the most important components are retained without significantly affecting the variability of the original

bands. Consequently, an image is obtained that incorporates the most relevant information of the original data at the cluster level.

The next step involves the classification of the individual segments. This is typically done through a combination of algorithms and manual classification. In this thesis this is accomplished with a supervised classification. The training areas, which were manually classified in a previous step, are now used by the algorithm to compare them with the individual segments. The classification algorithm used in this thesis is based on decision trees, which is a method of machine learning.

Random Forest (RF) is a method used to implement these decision trees. A large number of decision trees are created, which can be compared to tree diagrams (Belgiu & Drăguț, 2016). For example, a tree is built for each individual cluster, checking if the cluster values match those of the training data. For instance, the tree might ask: Do the information from this cluster correspond to the "Field" class? If the answer is "No," a new branch is formed, and it asks whether it belongs to the "Non-field" class. In this way, the corresponding land cover class is assigned to each cluster. At the end, a classified image is obtained, which can be further processed.

4.2.4 Indices

Indices in remote sensing are a very useful tool to combine different wavelengths and their specific information into a single and meaningful factor. They can be calculated by using a formula that combines the values of the different spectral ranges. In the following, three of these indices will be described shortly.

The normalized difference vegetation index (NDVI) is a remote sensing tool used to assess the health and productivity of vegetation. It uses reflectance values of the red and near-infrared (NIR) wavelength.

It is effective to understand the status of a vegetation cover as well as to quantify certain attributes. It is one of the oldest spectral indices going back to the year 1969. Nowadays NDVI is the most popular Index. It is very suitable to identify landcover types such as forests, buildings, and agricultural fields. And it can also be used to optimize irrigation and fertilization practices in agriculture. Water has a low reflectance in NIR, and leaves have a low reflectance in red light because of photosynthesis. NDVI relies on the principle that green vegetation absorbs more red light and reflects more NIR wavelengths, while non-vegetation surfaces reflect more red wavelengths and less NIR. So, a quotient from both highlights healthy and unhealthy vegetation and can be used to differentiate between vegetation and non-vegetation (Huang et al., 2021). NDVI values range from -1 to 1, with higher values indicating greater amounts of green vegetation. Values below 0 are waterbodies. Values between 0 and 0,2 are buildings, bare soil, and rocks. Values between 0,2 and 0,5 indicate grassland, sparse or unhealthy vegetation and a value above 0,5 represents a vegetation that is dense and healthy. The NDVI can be calculated using this formula:

$$NDVI = \frac{(NIR - RED)}{(NIR + RED)}$$

The Bare Soil Index (BSI) is a spectral index commonly used to detect bare soil and exposed land surfaces in satellite imagery or other remote sensing products. It relies on the principle that bare soil exhibits distinct spectral properties such as the near-infrared (NIR) and shortwave infrared (SWIR) regions of the electromagnetic spectrum. A higher BSI value indicates a higher likelihood of bare soil or exposed land surfaces, while a lower BSI value suggests the presence of vegetation or other land cover types (Nguyen et al., 2021).

The Bare Soil Index helps to distinguish between non-agricultural areas and agricultural areas. Therefore, the BSI can help classify agricultural fields and increase the model's quality. The BSI is calculated by taking the difference between the reflectance values of the Red, blue, SWIR, and NIR bands and dividing it by their sum (Tassi & Vizzari, 2020). The BSI can be separated into three groups. The range from -1 to 0 consists of dense vegetation. 0 to 0,3 represents sparse vegetation and values above 0,3 can be considered bare soil. The formula for the BSI is as follows:

$$BSI = \frac{(RED + SWIR) - (NIR + BLUE)}{(RED + SWIR) + (NIR + BLUE)}$$

The Normalized Pigment Chlorophyll Ratio Index (NPCRI) is an index to estimate the chlorophyll content in vegetation. In some publications the name Normalized Pigment Chlorophyll Index (NPCI) is used. In contrast to the NDVI, the NPCRI focuses on the chlorophyll content. This has the advantage of not only examining chlorophyll itself but also investigating the prerequisites for chlorophyll formation. Chlorophyll is synthesized within the chloroplasts of the plant. In addition to sunlight and water, it also requires nutrients for its formation. The availability of nutrients directly influences the production of chlorophyll. Greater availability of nutrients results in higher chlorophyll production. Therefore, the NPCRI can provide an assessment of the nutrient supply to plants.

The NPCI values range from -1 to 1. Higher positive values indicate higher chlorophyll content and healthier vegetation. The formular for the NPCRI contains just the visible parts of the light in the red and blue spectrum of light (Hatfield & Prueger, 2010).

$$NPCRI = \frac{(RED - BLUE)}{(RED + BLUE)}$$

4.2.3 Google Earth Engine

The Google Earth Engine (GEE) is a cloud platform developed for geodata by Google that was launched in 2010. It can be used to examine a wide range of applications. It is a powerful and versatile platform for analyzing and visualizing geodata. Via Application Programming Interface (API) the user can access the Google Server and its computing power to work with vast amounts of geospatial data. The API can be used by JavaScript and Python commands and therefore covers two of the most common programming languages. Everyone can use the GEE for scientific purposes because it is available free of charge. The open-source character and the enormous computing power are very useful and highlight GEE's uniqueness in the field of geospatial cloud computing (Zhao et al., 2021).

Gomes et al. 2020 conclude that another cloud system, Open Data Cube (ODC) has advantages but is not easy to handle and not so commonly used in terms of community and publications (Gomes et al., 2020).

Mutanga and Kumar 2019 discuss the use of the Google Earth Engine platform for various applications. The authors highlight that GEE has a user-friendly interface for data exploration and algorithm development. It allows users to access their own data while utilizing Google's cloud resources for processing. This helps to analyze big amounts of data in regions that are not accessible. It also allows to conduct advanced spatial analysis without having massive computing power or specific software on its own PC. Therefore, GEE also provides access to stakeholders that have not yet access to data or to a specific software. The Earth Engine provides access to a huge archive of datasets, vector data, including satellite imagery, social and demographic data, or climate data layers. There is a range of topics connected with the GEE, including agricultural applications, vegetation monitoring, land cover mapping, and disaster management. Within these areas the GEE can be used for estimating biodiversity variables,

mapping vegetation degradation and mitigation efforts, monitoring rangelands, analyzing land cover dynamics, assessing agricultural productivity, and addressing disaster response. One of the key applications of the GEE is environmental monitoring and analysis. The platform can be used to track changes in land cover and land use, and assess the impacts of environmental factors, such as climate change or pollution. There are many advanced algorithms and machine learning techniques for tasks such as land cover classification, crop yield estimation, flood prevention and response, drought monitoring, and mapping various environmental features like wetlands and mining areas (Mutanga & Kumar, 2019).

Also, the Google Earth Engine platform offers numerous applications in the field of object-based classification and time series analysis. One such application is the utilization of GEE in the agricultural domain for analytical purposes, including the classification of agricultural fields based on crop types and the investigation of long-term trends in crop rotation and land use patterns. By leveraging GEE's capabilities, researchers have been able to study and understand the dynamics of agricultural systems over extended periods of time (Luo et al., 2021).

4.2.5 Procedure for Determining Land Cover Changes

This subsection elucidates the methodology of spatiotemporal analysis employed within this thesis. As the main tool for the spatial analysis GEE will be used. Via API, data can be analyzed using the Google server infrastructure. Visualization and other steps will be executed in ArcGIS Pro. The scripts can be found in appendix 2, where a link to the Google Earth Engine workspace is provided.

The classification of agricultural fields is based on Sentinel-2 data. These data were initially preprocessed. The study areas in the Shapefile format were loaded into the Google Earth Engine, and a function was created to mask the clouds in the images, i.e., to assign a value of 0 to the pixels where clouds are present. Subsequently, an Image Collection was created, which

includes all images for the study area since 2017 with a cloud cover below 10%. Finally, the cloud masking function was applied to this collection. Statistical measures necessary for classification were then calculated, and an additional band containing these measures was added to the dataset. Ultimately, an image was created from all the images in the collection by taking the median of the collection. The median is less sensitive to outliers compared to the mean, ensuring that the resulting image represents an average where outliers, such as the masked areas where clouds were present, are less heavily weighted. This median image is saved as an asset for use in the next step.

The next step is the classification, which requires the previously generated image, as well as the Area of Interest and the training points or validation points. Firstly, several variables were defined. A buffer with an 8-meter radius was applied to the training points to improve the results. Then, the variables for the SNIC (Simple Linear Iterative Clustering) segmentation were determined. The grayscale image for the GLCM (Gray-Level Co-occurrence Matrix) was prepared, and the PCA (Principal Component Analysis) was performed as described above. Finally, the classification was carried out using the random forest algorithm with all created variables. The classified data were saved as an asset.

The validation of the classification is performed using the validation points generated beforehand. The quality of the classification can be determined using a confusion matrix. In this method, the points are used as "ground truth" and compared with the classification model. When the ground truth matches the classification, it is referred to as true positives or false negatives, indicating that the model aligns with reality. False positives and true negatives indicate errors in the classification. By comparing true positives and false negatives with true negatives and false positives, the overall accuracy of the model, known as the overall accuracy, can be obtained. A value of 1 indicates that 100% were correctly classified, while a value of 0.5

indicates that 50% were correctly classified. The closer the value is to 1, the better the classification performance.

In the next step, the indices are calculated as a time series. To achieve this, a collection of images is created like the classification step, where clouds are masked. Subsequently, the three indices are computed, clipped to the field extents, and added as bands to each individual image in the collection. Finally, a time series is generated from the image collection using the time series function in the Google Earth Engine. The resulting time series for each study area and for each index are stored locally.

This process is performed for both Sentinel data and Landsat data. Additionally, the results of the Sentinel time series are exported as a map by creating a composite image from all images within the most productive month in terms of vegetation. This is done to have a comparable monthly image instead of a snapshot of vegetation at a specific moment.

The comparison of results can be done in various ways, with particular interest in the temporal and spatial components. To assess temporal changes, a time series is useful, where values are recorded at each time point and presented as a line graph. Each time point represents the average value of observations in the study area for a specific time. The time series enables a quick understanding of fluctuations in the annual cycle. It also allows comparison of the current development with previous years. These time series reveal the most productive time points, indicating when the agricultural fields are most vital. It is advisable to compare these time points to each other, focusing on the annual optimum. Images can be generated for these time points, allowing for spatial visualization. Additionally, the distribution can be visualized using a histogram. For generating histograms, the R software was used (Appendix 6). It has advantages compared to the ArcGIS visualization of distributions especially the configuration of the lag size and the color scheme. Therefore, a short script was made in R to visualize the histograms.

The distribution reveals areas where a particular index has a higher frequency of values. Changes in the distribution from one year to another indicate compression. Multi-peaked distributions can also provide insights into the development. A comparison of the discrete groups generated for the BSI and NDVI also provides insights into the developments. For instance, the percentage can be determined by which the highly vital areas have decreased.

5. Results

In this chapter, the results are examined to gain a deeper insight into the climatological and remote sensing outcomes in this thesis. A detailed overview is necessary to understand the processes and to put them in perspective.

5.1 Climatological Trends

First, the climatological patterns will be examined. Table two and three show the correlations between the stations. Table three pertains to precipitation, while table two pertains to temperature. Since the values in the matrix are mirrored, correlations are always represented twice (e.g. Certkovo - Donetsk, Donetsk - Certkovo). Since both values for each comparison have the same significance, it suffices to focus on the color-coded area. The diagonal represents the correlation of each station with itself, which is equal to 1. The cells with strong correlations (≥ 0.5) are highlighted in green, while those with medium and weak correlations are highlighted in red (< 0.5). The statistical significance of the correlation coefficients is < 0.05 (Appendix 19). It is assumed that there is a statistically significant correlation.

Initially, the temperature results are of interest. It can be observed that an extremely high positive correlation prevails between the stations, exceeding 0.99. This is an almost perfect correlation. The results suggest that there is a very strong relationship between all stations regarding temperature. In the case of precipitation, the results are more complex. All values exhibit a positive correlation, but it can be observed that the correlations in this matrix are weaker. Out of 36 correlation pairs, 10 are below the threshold of 0.5. Although the correlation is significantly weaker, it is clear that the stations exhibit a relationship.

Table 2: Correlation matrix of the temperature from the climate stations

<i>Correlation-Matrix Temperature</i>									
	<i>Dnibr.</i>	<i>Kharkiv</i>	<i>Taganrog</i>	<i>Certkovo</i>	<i>Donetsk</i>	<i>Luhansk</i>	<i>Izium</i>	<i>Mariopol</i>	<i>Pryshyb</i>
<i>Dnibr.</i>	1,0000	0,9983	0,9970	0,9969	0,9988	0,9977	0,9981	0,9965	0,9987
<i>Kharkiv</i>	0,9983	1,0000	0,9942	0,9977	0,9971	0,9973	0,9986	0,9932	0,9953
<i>Taganrog</i>	0,9970	0,9942	1,0000	0,9962	0,9982	0,9976	0,9953	0,9991	0,9983
<i>Certkovo</i>	0,9969	0,9977	0,9962	1,0000	0,9981	0,9989	0,9983	0,9946	0,9959
<i>Donetsk</i>	0,9988	0,9971	0,9982	0,9981	1,0000	0,9990	0,9979	0,9974	0,9987
<i>Luhansk</i>	0,9977	0,9973	0,9976	0,9989	0,9990	1,0000	0,9989	0,9960	0,9975
<i>Izium</i>	0,9981	0,9986	0,9953	0,9983	0,9979	0,9989	1,0000	0,9938	0,9966
<i>Mariopol</i>	0,9965	0,9932	0,9991	0,9946	0,9974	0,9960	0,9938	1,0000	0,9983
<i>Pryshyb</i>	0,9987	0,9953	0,9983	0,9959	0,9987	0,9975	0,9966	0,9983	1,0000

Table 3: Correlation matrix of the precipitation from the climate stations

<i>Correlation-Matrix Precipitation</i>									
	<i>Dnibr.</i>	<i>Kharkiv</i>	<i>Taganrog</i>	<i>Certkovo</i>	<i>Donetsk</i>	<i>Luhansk</i>	<i>Izium</i>	<i>Mariopol</i>	<i>Pryshyb</i>
<i>Dnibr.</i>	1,0000	0,6413	0,5722	0,3860	0,7167	0,5621	0,6920	0,5774	0,7347
<i>Kharkiv</i>	0,6413	1,0000	0,4117	0,4935	0,5909	0,5264	0,7413	0,3529	0,4686
<i>Taganrog</i>	0,5722	0,4117	1,0000	0,4692	0,5416	0,6255	0,4167	0,5791	0,5345
<i>Certkovo</i>	0,3860	0,4935	0,4692	1,0000	0,5346	0,6268	0,4442	0,6012	0,2742
<i>Donetsk</i>	0,7167	0,5909	0,5416	0,5346	1,0000	0,6551	0,6884	0,6916	0,6898
<i>Luhansk</i>	0,5621	0,5264	0,6255	0,6268	0,6551	1,0000	0,6060	0,5322	0,5230
<i>Izium</i>	0,6920	0,7413	0,4167	0,4442	0,6884	0,6060	1,0000	0,3914	0,5394
<i>Mariopol</i>	0,5774	0,3529	0,5791	0,6012	0,6916	0,5322	0,3914	1,0000	0,5308
<i>Pryshyb</i>	0,7347	0,4686	0,5345	0,2742	0,6898	0,5230	0,5394	0,5308	1,0000

To investigate the extent to which temperature and precipitation undergo monthly variations, the temporal autocorrelation functions can be examined. In the graph, the covariance is shown on the Y-axis, while the lag is displayed on the X-axis. The lag is the time unit by which the time series is shifted, i.e., the higher the lag, the greater the temporal distance. The dashed line represents the significance level. Values below this level indicate that a value is not dependent on the previous month's value.

For temperature, a cyclic autocorrelation predominates. This is evident from the temperature trend over seasons. It can be observed that there is a highly positive temporal autocorrelation, followed by a high negative temporal autocorrelation. It can be assumed that the temperature

in the study area closely matches the temperature of the previous month and the temperature of the previous season.

In contrast, precipitation exhibits a very low temporal autocorrelation, which is below the significance level. Consequently, the amount of rainfall in a month cannot be necessarily attributed to the rainfall in the previous month. This could be an explanation for the weaker correlations between the stations with respect to precipitation. Temporal autocorrelation functions were determined for the averaged values as well as for the Dnipropetrovsk station. In both cases, the described cyclic and non-cyclic pattern is clearly visible (Figure 8).

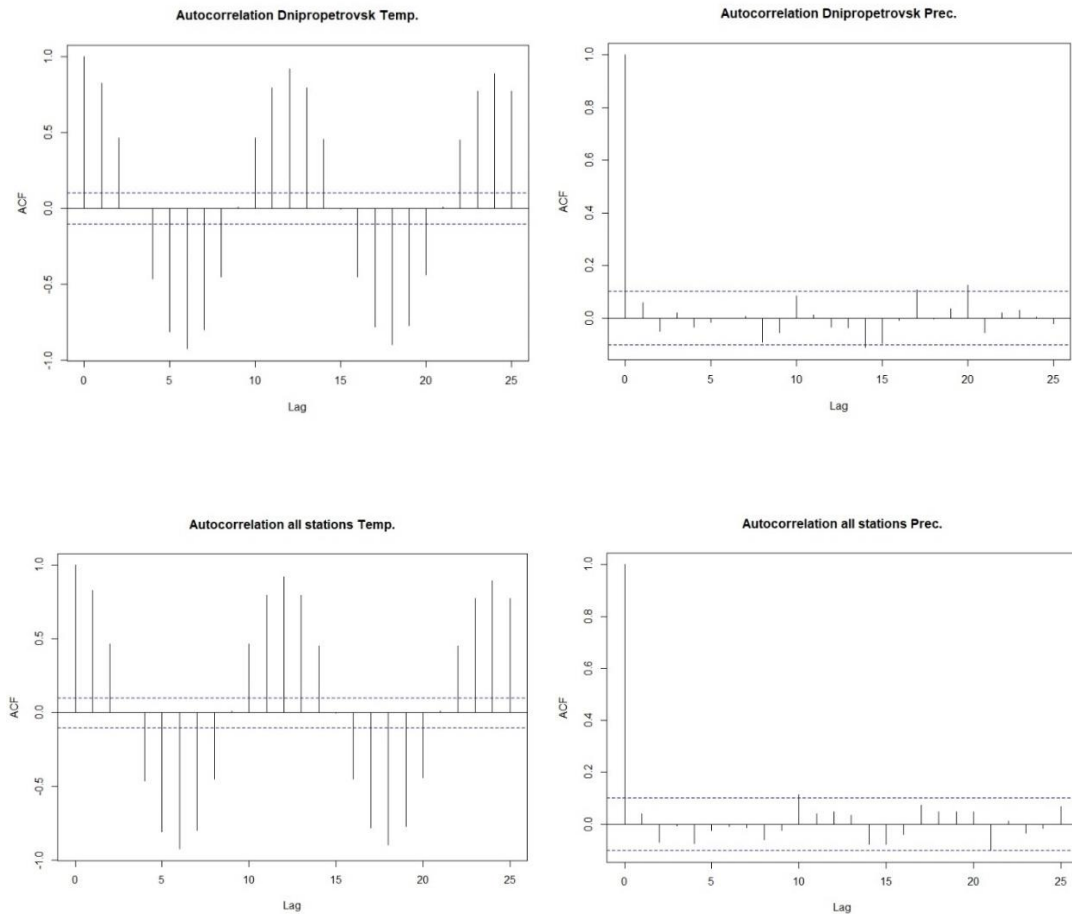
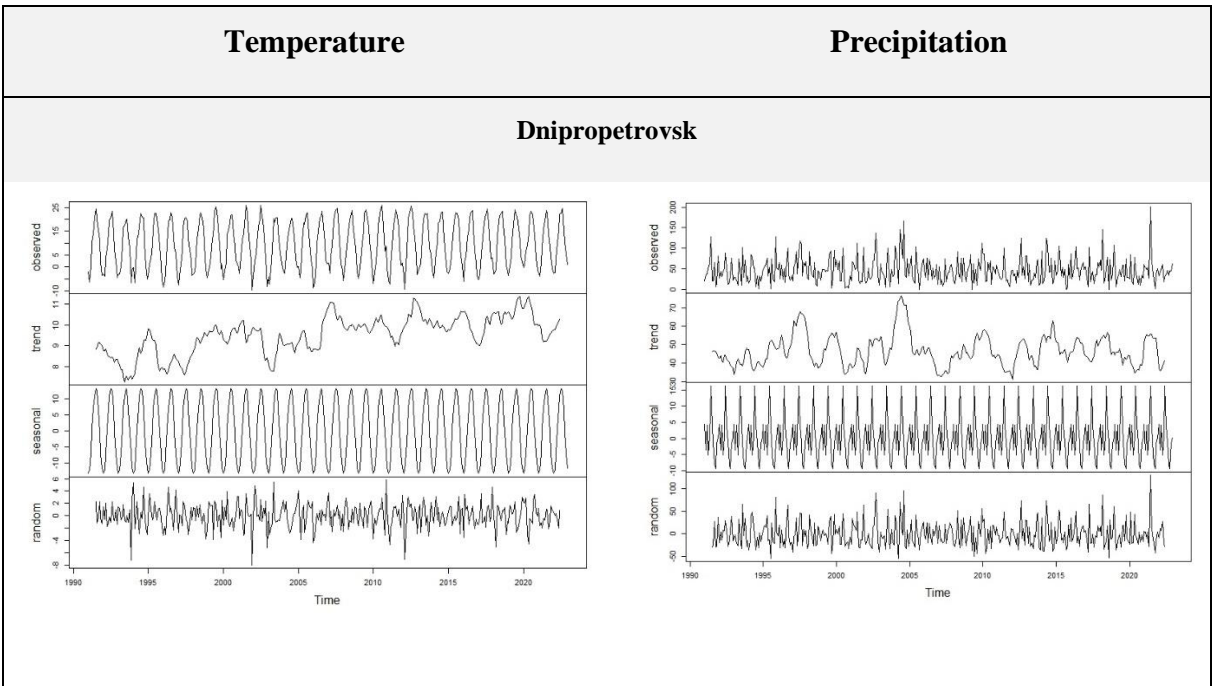


Figure 8: Autocorrelation function for the temperature and precipitation

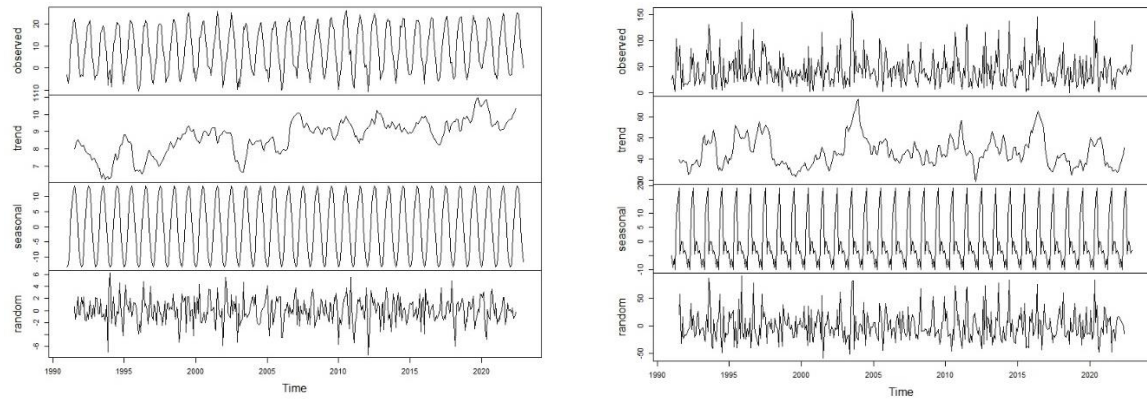
The time series analysis focused on the most meaningful stations that have no significant data gaps since 1990. The "decomposition" representation shows the time on the X-axis and the individual time series components with their respective units on the Y-axis. The figure is divided into the original time series, followed by the trend, the seasonal component, and the random component. The decomposition is exemplified here using complete time series. The stations Dnipropetrovsk, Kharkiv, Taganrog, and Certkovo were analyzed in terms of temperature and precipitation using this approach because they cover the time range between 1990 and 2022.

For all four temperature graphs, a slight increase in temperature over the last 30 years is evident. Therefore, it can be assumed that there was an increase in the average monthly temperature in recent decades. On the other hand, the precipitation trend time series remained relatively stable during the same period. It is characterized by fluctuations in the amount of precipitation, but no long-term trend is evident. Therefore, it can be concluded that the precipitation has not changed significantly in the last 30 years in terms of trend. All graphs are visualized in Table 4.

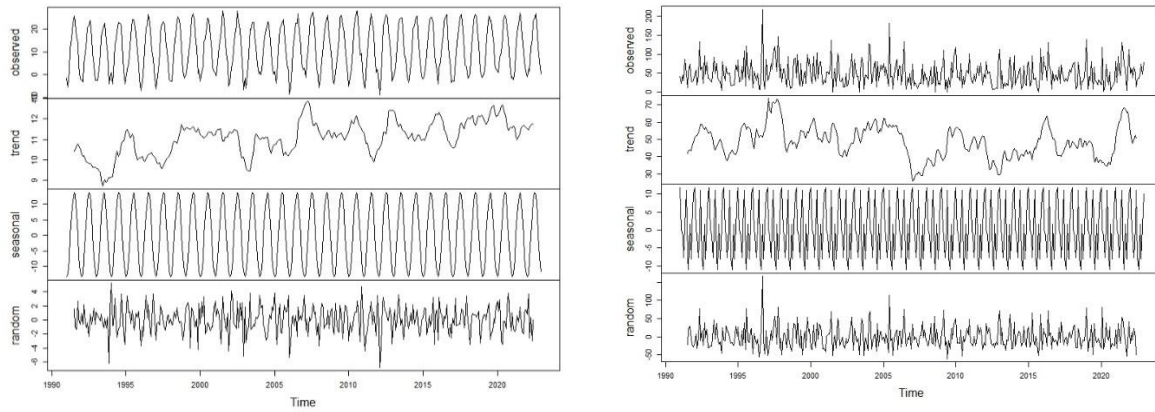
Table 4: Decomposed time series of temperature and precipitation of chosen stations



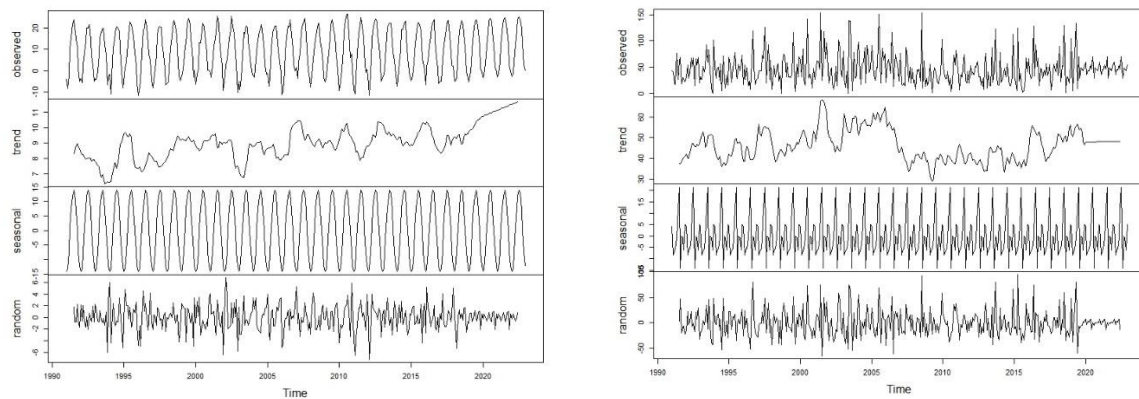
Kharkiv



Taganrog

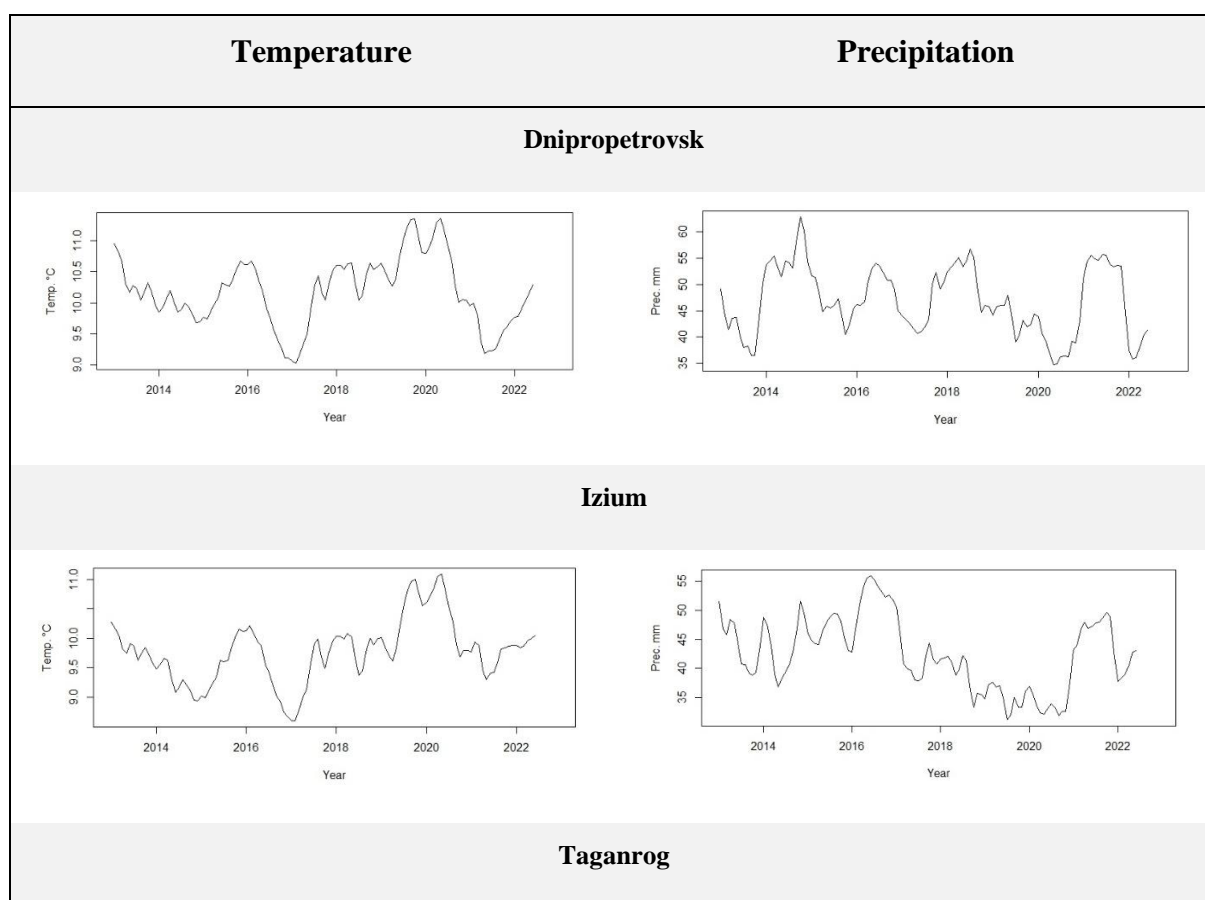


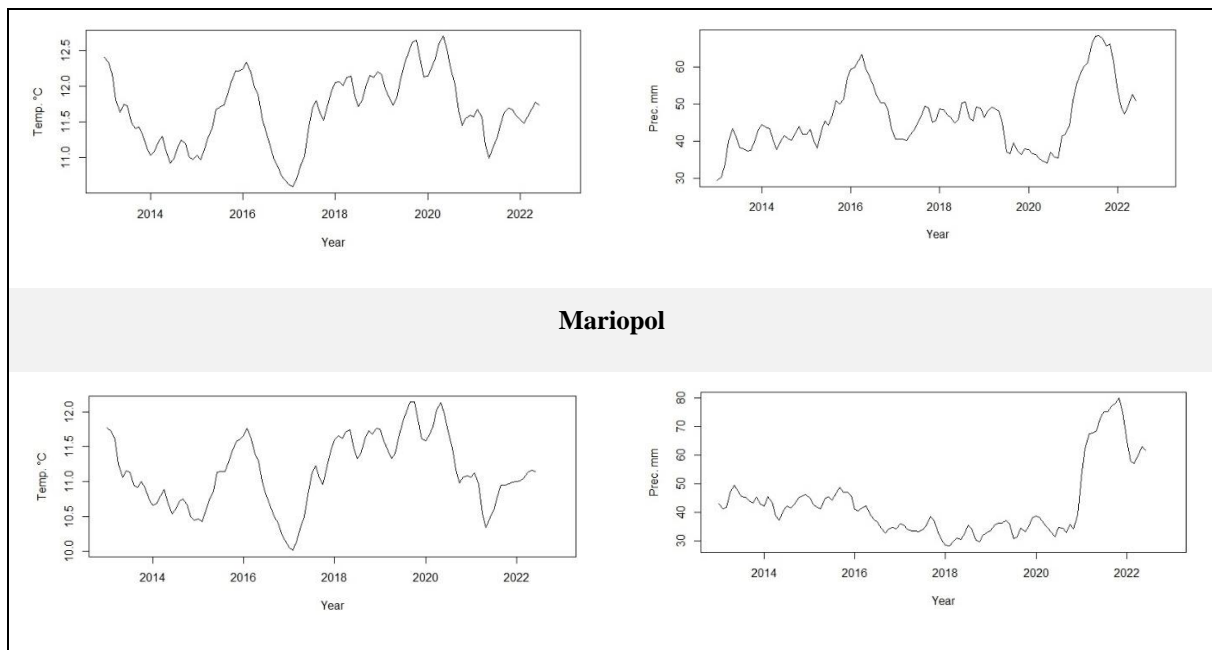
Certkovo



Of particular interest is the period since 2013, which serves as the primary investigation period in this study (Table 5). The extremes in temperature and precipitation are particularly important in identifying periods that provide unfavorable conditions for agriculture. This section also focuses on the stations with the best data density. Similar patterns can be identified in both temperature and precipitation at the four stations of Dnipropetrovsk, Iziium, Taganrog, and Mariopol. From 2013 to 2015, a decrease of about one degree Celsius in temperature can be observed. This value rises back to its original level between 2015 and 2016. In 2017, there was a rapid drop in temperature, representing the biggest cut in this ten-year period. The highest temperatures are recorded in 2019 and 2020.

Table 5: Trendline for the last ten years for temperature and precipitation





For precipitation, a stronger variation can be observed between years. Some years have low precipitation, while others have high precipitation. Particularly, 2021 and 2016 can be considered as very rainy, whereas 2020, 2017, and 2013 are particularly dry. These trends can be observed at individual stations (Table 5) as well as in the averaged values (Figure 9 and 10).

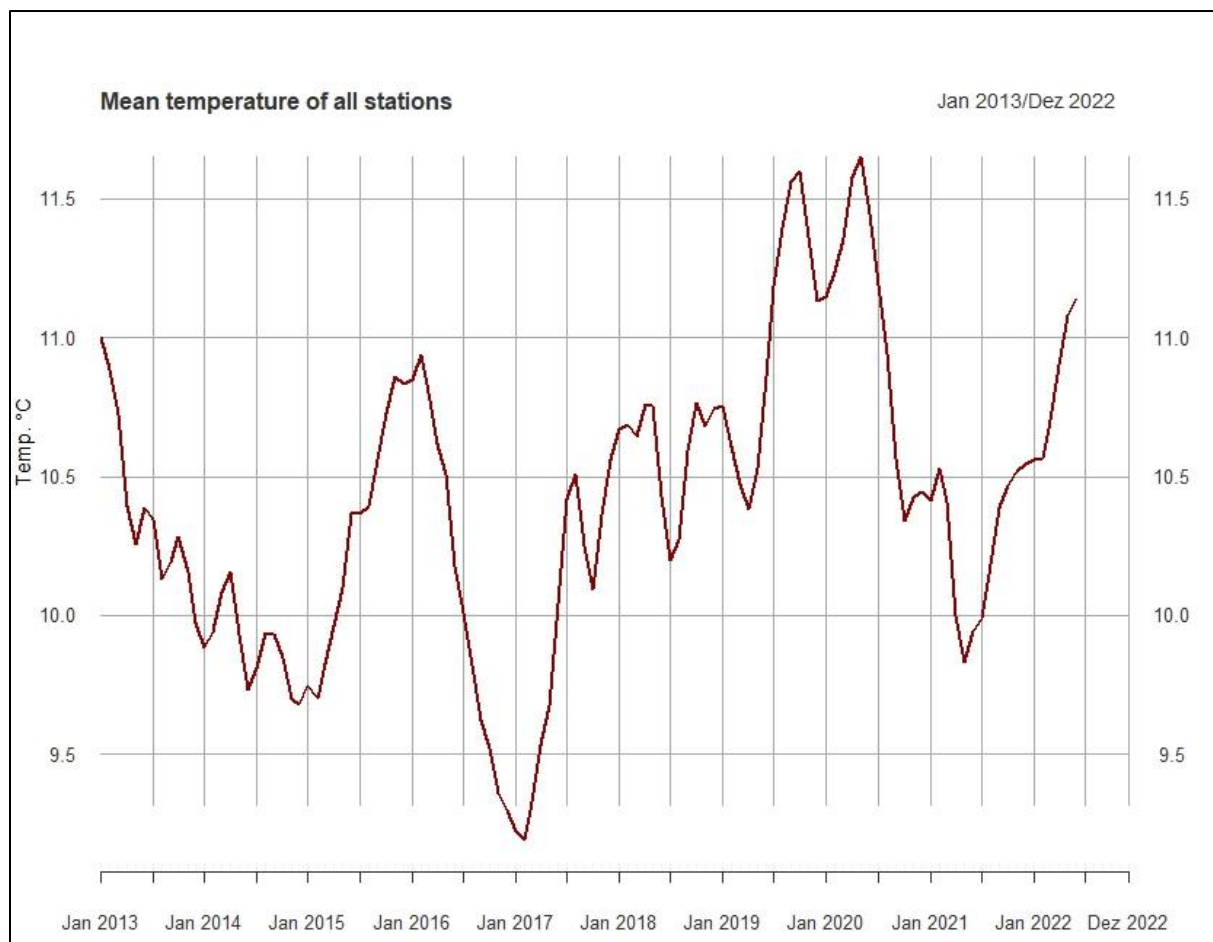


Figure 9: Mean temperature trend for all stations

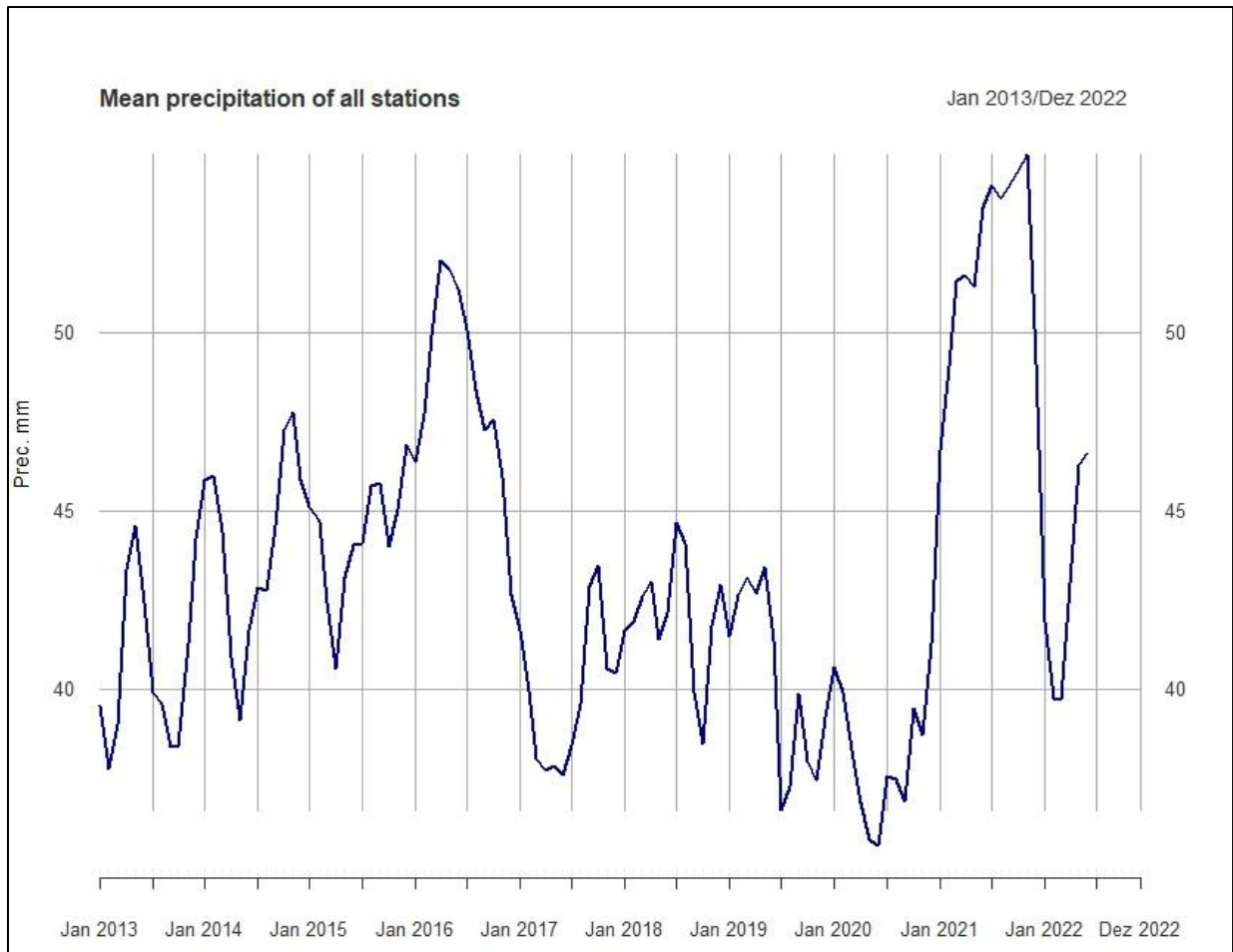


Figure 10: Mean precipitation trend for all stations

5.2 Spatio-Temporal Comparison

The following chapter will focus on spatial aspects. The results of the investigation conducted in the Google Earth Engine will be presented and described. The emphasis lies on the classification results and the trend development of the indices.

5.2.1 Validation Results

Following the principle of "No classification without validation," a validation of the results is performed after each individual classification run. As described above, the segmentation process is particularly relevant to the model quality. The script allowed for the manipulation of

numerous variables, which had varying degrees of influence on the outcome. These variables include the buffer size around the training points, the compactness factor, the connectivity factor, neighborhood size, and seed spacing.

Buffer size and seed spacing have an impact on the outcome. The remaining factors had only a marginal influence on the results, or, as in the case of the connectivity settings, worsened the outcome. In the case of study area A, Table 6 exemplifies the overall accuracy under various combinations of seed spacing and buffer size. It can be observed that a buffer size setting between 8 and 9 meters is most suitable, as well as a seed spacing setting ranging from 5 to 15 meters. Different combinations of input parameters were tested for all three study areas, resulting in good to very good outcomes.

Table 6: Comparison of results for different buffer size and seed spacing

Buffer Size (m)	2	5	7	8	9	10	15
Seed Spacing							
5	0,895	0,92	0,93	0,93	0,94	0,935	0,92
10	0,92	0,93	0,925	0,935	0,92	0,93	0,92
15	0,905	0,925	0,915	0,94	0,935	0,925	0,915
20	0,92	0,93	0,92	0,92	0,91	0,92	0,905
25	0,885	0,89	0,915	0,925	0,915	0,92	0,92

Table 7 (confusion matrices) provides the validation results. Study area A generated the best result with an accuracy of 94 %. Only 12 out of 200 validation points correspond to incorrectly classified areas. Study area B has an overall accuracy of 91%, and area C achieves an accuracy of 93%. Thus, the results in this study are significantly more accurate than those of Tassi et al. 2021, which served as a reference.

Table 7: Confusion matrix for the classification results of all investigation sites

Site A Confusion Matrix			
	<i>Actual Positive</i>	<i>Actual Negative</i>	<i>Overall Accuracy</i>
<i>Predicted Positive</i>	134	5	
<i>Predicted Negative</i>	7	45	
			0,94

Site B Confusion Matrix			
	<i>Actual Positive</i>	<i>Actual Negative</i>	<i>Overall Accuracy</i>
<i>Predicted Positive</i>	120	5	
<i>Predicted Negative</i>	14	61	
			0,905

Site C Confusion Matrix			
	<i>Actual Positive</i>	<i>Actual Negative</i>	<i>Overall Accuracy</i>
<i>Predicted Positive</i>	147	5	
<i>Predicted Negative</i>	9	39	
			0,93

The final classification results can be observed in Figure 11. The images clearly depict the segmented individual fields. The gaps where settlements, roads, and railway tracks are located are also evident. The model shows room for improvement in the areas of river floodplains and meadows, which were mostly identified as "non-field" but still exhibit remnants mistakenly classified as an agricultural field. Additionally, in the settlement areas, small areas classified as fields can be found, which are likely gardens. In summary, it can be concluded that the classification result is highly satisfactory and can serve as a solid basis for further work.

The area of each of the three study areas is 2,500 km². In study area A, the area of fields within the study area amounts to 1853.72 km² (74%), in area B it is 1773.01 km² (71%), and in area C it is 1963.88 km² (79%). It can be concluded that the field areas within the study areas are similarly sized, with only an 8% difference.

Site A



Site B



Site C



Figure 11: Results of the classification process

5.2.2 Trends

The indices were calculated for the classified field areas and presented as time series. The Landsat time series covers the period since 2013, while the Sentinel time series covers the period since 2019.

The time series were created for each study area, representing the three indices. The x-axis represents the chronological dates, and the y-axis represents the index values. A gray vertical

line within the chart indicates the zero value. In this diagram, the blue line represents the development of measurement values for each index. The values at each time point were interpolated to generate the line.

Initially, it is noticeable in the Landsat results that no trend development is discernible for any of the indices (Appendix 7-9). However, seasonal fluctuations are clearly evident. A closer examination reveals that many of the results are implausible. For example, the BSI does not exceed the threshold of 0.1, which would indicate that the area is fully vegetated. This is illogical since many areas are bare during winter. Similarly, an average NDVI value reaching only up to 0.3 is unrealistic, as healthy agricultural field areas exhibit significantly higher values during the growing season. Moreover, the NDVI occasionally assumes negative values, which cannot be attributed even to a snow cover, for which the NDVI is approximately 0.1 - 0, but rather suggests widespread flooding of the fields, which did not occur. The NPCRI data range is 0.13, with values fluctuating between 0.09 and -0.04. This extremely narrow range is also suspicious since the NPCRI should be significantly higher (>0.3) during the growth phase, and in the absence of water bodies within the fields. It should not exhibit values below 0. Therefore, it can be concluded that the outcomes from the Landsat data after being critically examined, are unsuitable for meaningful interpretation. Consequently, the analysis was repeated using Sentinel-2 data.

The results of the Sentinel investigation yield entirely different outcomes (Appendix 10-12). The chart representation is similar to that of the Landsat data, with the exception of a shorter period displayed. The results are more coherent. The BSI varies between ± 0.3 . The highest values occur in spring and autumn, before planting and after harvest, respectively. The lowest BSI values are typically found between May and September. There is an annual fluctuation of values, temporarily reaching the negative range. However, since these time points fall within winter, it is not indicative of a high land cover rate. Instead, it is presumed to be an effect caused

by snow cover, which aligns with findings from the climatic investigation. The fluctuations of the BSI are consistent across all three study areas. It is noteworthy that the BSI in 2022 is lower than in previous years, indicating higher land coverage during that year. The data for 2019 and 2020 exhibit similar BSI values in all study areas. However, in 2021, it can be observed that the low BSI values begin later. This can be explained by incorporating climatic data. 2021 was a very wet year with lower average temperatures, likely influencing the vegetation period. The example of the BSI time series for site A can be seen in figure 12.

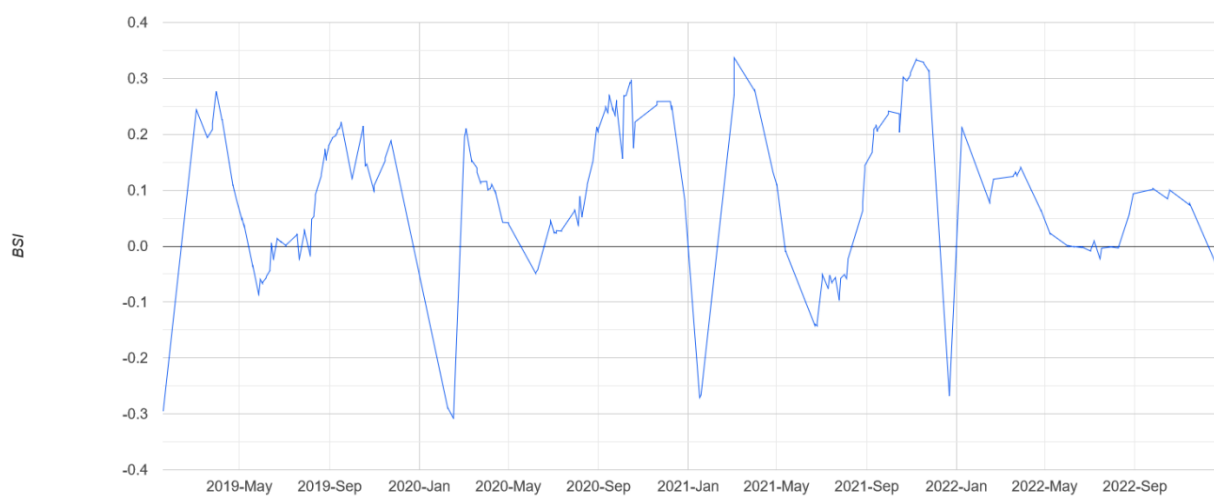


Figure 12: BSI time series for site A

Similar trends can also be observed in the NDVI as described above. 2019 and 2020 show similarities, while 2021 demonstrates higher NDVI values, and 2022 shows significantly lower values (Figure 13). The NDVI peak occurs in June and July, with values reaching up to 0.7. Two declines are observed during the year, one between September and October, and another around the turn of the year. The decrease in NDVI in September is likely attributed to the harvest. The low NDVI values in December, January, and February can be easily explained by unfavorable winter growth conditions, representing the natural phenology.

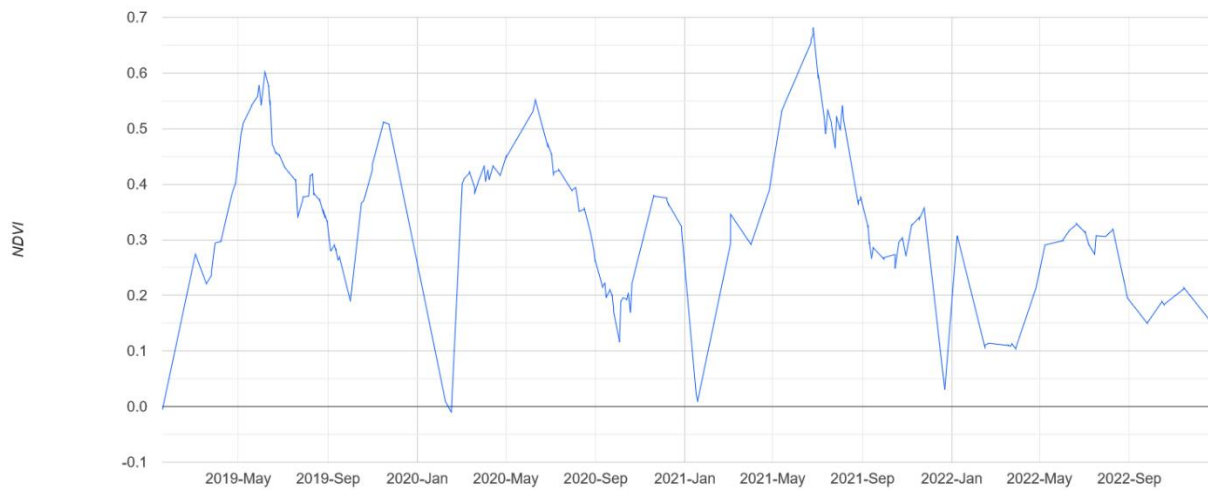


Figure 13: NDVI time series for site A

The NPCRI in figure 14 exhibits more pronounced fluctuations, but the trends from the previous time series can still be identified. The maximum NPCRI value is found in November, with a slight increase also observed in spring. The minimum value occurs around the turn of the year. Of particular interest is that in 2022, where the NPCRI values are extremely low and show minimal fluctuations. The chlorophyll content significantly decreased in that year. In February and March 2021, the NPCRI is unusually high, possibly due to the dry and mild winter of 2020.

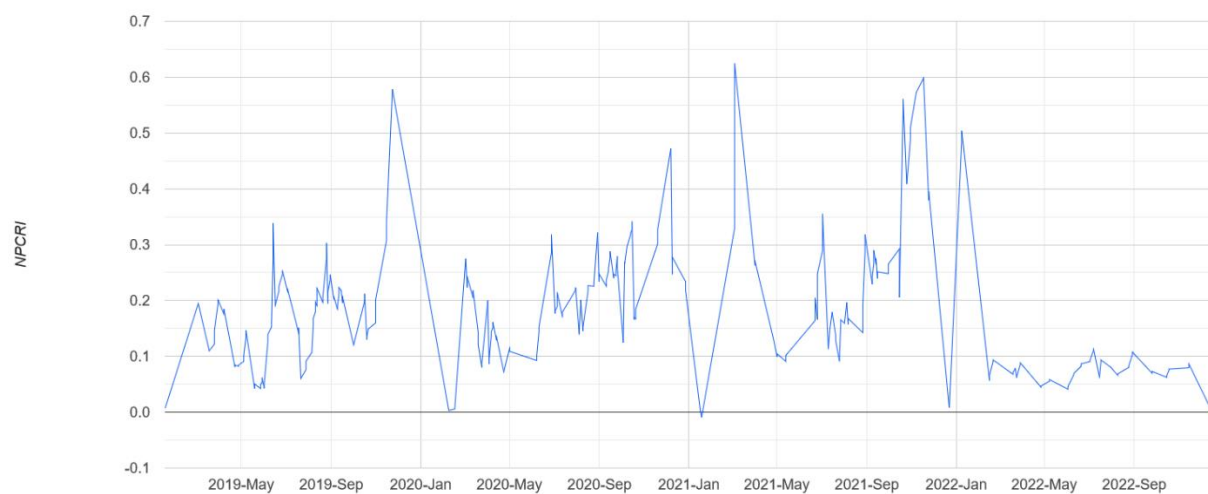


Figure 14: NPCRI time series for site A

The reason for the high index values in January 2022 is not easily discernible. Normally, NDVI and BSI exhibit opposite trends since more bare soil would logically result in less vegetation. Therefore, it is challenging to understand why both values are unusually high in January. This discrepancy could potentially be attributed to an error in the input data.

5.2.3 Site to Site Comparison

The results of the time series were also visualized in a map (Appendix 13-15). For this purpose, the month with the highest NDVI values was selected, and all images from that month were combined to create a comparable image. In this way, the BSI and NDVI were represented. The visualizations show the predefined classes of both indices, each delineated by distinct colours. Of particular interest is whether spatial patterns can be observed in the study areas. In the spatial distribution of the classes, no obvious patterns are observed in the sense that, for example, there would be more healthy vegetation in the eastern part of a study area. The results do not exhibit a clear pattern, and the classes are homogeneous across the study area. Figure 15 and 16 shows the NDVI and BSI in a map for study area A in the year 2019.

However, there are notable anomalies regarding smaller features that reflect the findings from the time series analysis. The NDVI shows minimal visual changes in its manifestation from the images of 2019 to 2021. In contrast, the NDVI is significantly lower in 2022 compared to previous years. In the case of the BSI, the images from 2019 to 2021 always contain areas with a BSI of ≥ 0.3 , indicating areas not covered by vegetation. However, in 2022, no empty areas are present, which is initially surprising considering the lower NDVI. These developments are consistent across all areas.

An examination of the distributions of the images provides additional information. Firstly, it confirms the observations from the maps. Secondly, it reveals the index ranges with a high frequency of values. The data are not evenly distributed but exhibits a multimodal manifestation. The NDVI exhibits two peaks, one around 0.2 and another around 0.7. This is particularly evident in study areas A and C. The NDVI for 2021 deviates in the sense that it is more pronounced in the range between 0.6 and 0.8. Study area B, on the other hand, exhibits an overall compressed distribution. The BSI shows a three-peaked distribution. The first peak is in the range of -0.4 to -0.2 and can be attributed to vegetated areas. The second peak is between 0.1 and 0.2, representing areas with moderate vegetation cover. Finally, there is a small peak at 0.3, corresponding to non-vegetated areas. The NPCRI is bimodal with peaks at 0.1 and 0.3. The distribution could also be used to establish meaningful class divisions, although this was not pursued in this study with respect to the NPCRI.

In general, all conclusions drawn from the examination of the maps can also be observed in the distributions (Appendix 3-5). Particularly interesting is the fact that study area B exhibits distributions that are less informative compared to the other two study areas. It is more difficult to identify individual class boundaries, and the distribution is noticeably flatter. In the peripheral areas of the distributions, few pixels deviate from what would be expected in agricultural areas in terms of index values. This can be attributed to minor inaccuracies in the classification results. However, these boundary areas are minimal, which further supports the high quality of the classification results.

The classified values can be used to obtain the number of pixels per class. In this way, changes can be quantified, and the percentage of the study area belonging to each class can be measured (Appendix 16-18). For the NDVI, the values in the negative range (misclassified pixels) are below 0.1% for all study areas. The results of study area A confirm that the year 2021 was particularly productive. Fifty percent of the agricultural fields were covered with highly

vigorous vegetation in that year. However, in the following year, this percentage dropped to only 16%. Similar trends can be observed in the other study areas. Areas with non-vital vegetation developed inversely. In all study areas, these areas accounted for less than 15.5% of the total area. In 2021, it was even below 6.27%, whereas in 2022, more than 41% of the area fell into the category of less vigorous vegetation. The decline from 2021 to 2022 in the NDVI class with active vegetation is interesting. In study area A, the decline is -66%, in B it is -94%, and in C it is -86%.

Regarding the BSI, it can be observed that between 2019 and 2021, a maximum of 3.6% of the area was not covered by vegetation. Study area A has the highest number of non-vegetated areas, followed by area C. Area B has the fewest non-vegetated areas. Except for 2022, there were non-vegetated areas in all years and all study areas. However, in 2022, there are no non-vegetated areas in any of the study areas. All fields are covered with vegetation.

In summary, in all areas, there is a significant decline in the NDVI peak values in 2022, while the percentage of non-vegetated fields has dropped to 0%. From 2021 to 2022, the particularly healthy vegetation has decreased by more than two-thirds. The chlorophyll content has also decreased significantly in all areas in 2022.

Site A Bare Soil Index (BSI) 2019

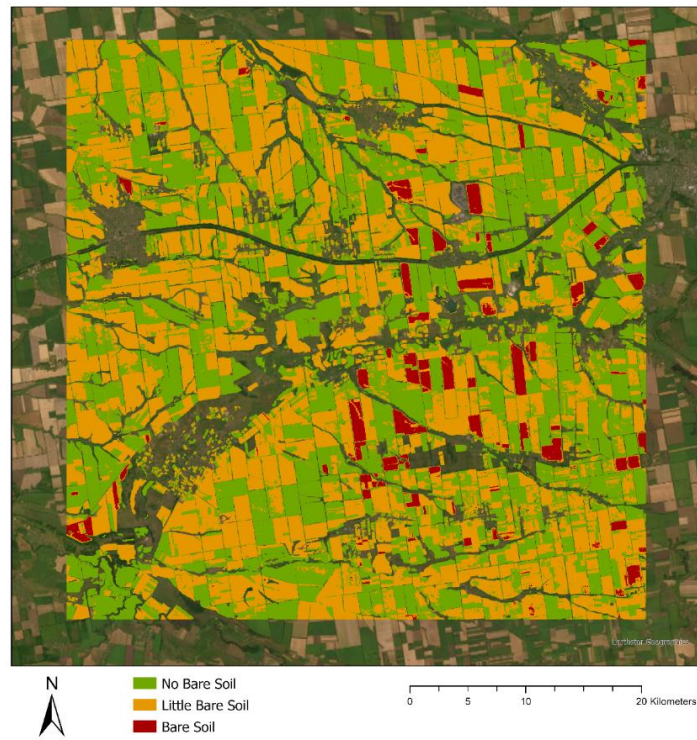


Figure 15: BSI 2019 site A

Site A Normalized Difference Vegetation Index (NDVI) 2019

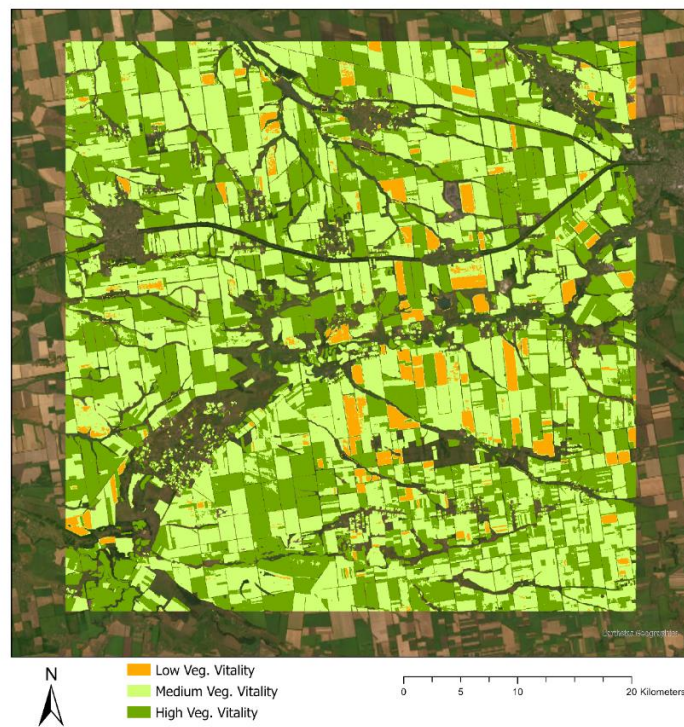


Figure 16: NDVI 2019 site A

6. Discussion

The aim of this thesis is to determine how agriculture in Ukraine has been affected by the impacts of war. To achieve this, three sub-objectives have been formulated. Firstly, a climatic assessment will be conducted to estimate the effects of climate factors on potential changes in agriculture. Secondly, an object-based classification will be performed to differentiate between fields and non-fields. These results will then be used to observe changes using multispectral indices. It is also of interest to examine how the influences of the war differ in various study areas. Various methods and techniques will be employed, including a time series analysis in R and the Google Earth Engine. The overall goal is to generate an assessment of the agricultural situation that helps interpreting war impacts.

The expectations regarding the results were not clearly defined in advance. However, the assumption that agriculture is strongly negatively affected by war seems logical but needs to be scientifically proven. Theoretically, it could also be the case that troop provisioning leads to increased cultivation. Alternatively, the problem could lie in logistics, such as transportation, storage, and export of crops, rather than cultivation. However, due to media coverage, the presumption that there must be a decline in agricultural production was plausible. The fact that agricultural fields are often battlefields and farmers are drafted into the armed forces, unable to tend to their fields, supports this assumption. The literature review on this topic also revealed that a decline in agriculture is to be expected.

The results of my research generally confirmed the previous expectations. However, the assumptions have been expanded with further insights, providing a significantly more detailed assessment of the agricultural situation.

6.1 Result Interpretation

The analysis of climate data constituted the first step. It should be noted that the stations exhibit a nearly perfect correlation with each other in terms of temperature. The correlation for precipitation is also high, although not as high as for the temperature. Assumably, climatic factors represent a continual spatial phenomenon. This means that the developments at one station can also be traced at the other stations. The lower correlations in precipitation can be attributed to the temporal discontinuity of the data, which are subject to greater randomness. This randomness ultimately affects the correlations between individual stations.

The time series were decomposed into their components to observe any trends. There has been no trend in precipitation over the past 30 years, whereas a clear increase is observed in temperature. Whether this can be attributed to climate change cannot be determined and requires further investigation. However, this is a plausible hypothesis. The most important findings lie in the representation of weather conditions during the study period. The fluctuations in temperature and precipitation are natural within the short study period. However, the magnitude of these fluctuations helps interpret remote sensing data results. In summary, the climate investigation proved to be meaningful and helped to assess the results. Furthermore, it generated results that invite further investigation as they go beyond the scope of this study.

The aim of the remote sensing investigation was twofold: firstly, to correctly classify the agricultural field areas, and secondly, to monitor the changes within the identified field areas. The classification of the field areas proceeded surprisingly well. With an accuracy of $> 91\%$ compared to the validation points, the fields were correctly classified. The existing script by Tassi et al. 2020 was easily adapted by removing irrelevant parts and adding necessary commands. After making some parameter adjustments to the segmentation, excellent results were achieved. This method misclassified a minimal amount of areas. The margin of error is

unsurprising for supervised classifications. These could be manually edited post-processing to improve the classification. The results for Site B are less satisfactory than those of the other areas. This could be attributed to two factors. Firstly, the input data on which the classification relies encompasses the entire period during which there was the military conflict in Site B. In other words, the classification data already comes from a timeframe when the conflict was ongoing within the study period. This conflict potentially influenced agriculture, resulting in the fields being neglected and more difficult to distinguish from the surrounding land cover types. Secondly, the result could be explained by other spatial conditions specific to the study area. Soil properties or the terrain can cause spatial patterns of land cover to differ from those in the other study areas. An investigation of the pedological and geomorphological conditions could shed light on this hypothesis.

The time series were examined to identify trends and changes in the field areas. The motivation behind this analysis was to detect and understand patterns to grasp the developments in the Ukrainian War. The investigation of the time series revealed that Landsat 8 data was not usable as it produced poor and illogical results. In contrast, Sentinel-2 data proved to be suitable for analysis. It was found that the NDVI was significantly lower in 2022 compared to previous years. The decline was particularly pronounced in the highly vital vegetation. The BSI indicated that there were no bare areas in the study areas in 2022. However, the NPCRI indicated a substantial decrease in chlorophyll content in plants that year.

The abrupt decline in NDVI is a clear indication of the impacts of the Ukrainian war on agriculture, especially considering that weather conditions did not change significantly compared to previous years. Healthy plants with good nutrient supply and sufficient water exhibit high NDVI values. Therefore, the decline can be attributed to the non-cultivation or inadequate cultivation of the fields. The NPCRI provides evidence by measuring the chlorophyll content, which is essential for chlorophyll formation. The NPCRI experienced a

sharp drop, likely due to the fields not being fertilized. When nitrate levels are reduced, less chlorophyll is formed. The lack of fertilization could be attributed to various factors, including direct combat actions, a shortage of people to tend to the fields or a lack of fertilizer. It is likely a combination of these three factors. During times of war, logistics prioritize essential goods for the war, and other goods, such as fertilizers, have lower priority. Furthermore, a large portion of the population has been recruited into the armed forces, involved in other crucial positions, or has left the country. Additionally, the management of the agricultural fields is hindered for personal safety reasons when there is a risk of landmines or nearby combat. Interestingly, the BSI indicates that all agricultural areas are vegetated. At first glance, this may seem illogical given the decrease in NDVI. However, the question is: What vegetation forms are present in the fields? It can be assumed that the fields are only partially cultivated with agricultural crops. A significant portion may be covered with grass, shrubs, bushes, and weeds. This explanation accounts for both the NDVI and the BSI. In summary, it can be assumed that the fields have largely been left uncultivated. They have not been sown, fertilized, or plowed. As a result, the nutrient supply has declined, leading to decreased vegetation vitality. The abandoned areas have been taken over by pioneer vegetation.

The comparison to the previous year is significant, as 2021 was an excellent year for agriculture. This is because the conflict had not yet escalated that year, and optimal growth conditions prevailed. The mild winter of 2020/2021 likely contributed to this as well. As a result, the vegetation phase started earlier, and fertilizers could probably be applied earlier since the soil was no longer frozen. These interpretations are not possible without the inclusion of climate data. They demonstrate the importance of a climatological assessment for evaluating agricultural areas.

Here too, it can be observed that poorer results are evident in Site B. This can also be attributed to the factors of war influences and structural spatial differences described earlier. However, no

significant differences can be observed between Site A and C. In conclusion all three study areas were equally affected by the war in 2022. In Site B, however, conditions have been worse for a longer period. It can be assumed that eastern Ukraine was equally affected by the outbreak of the war in 2022, regardless of whether the areas are under permanent Russian or Ukrainian control. The areas held by the Ukrainians fared the best in comparison. The separatist-held areas, on the other hand, are the most severely affected as the conflict has been simmering for years.

Finally, the question arises as to why the results from Landsat 8 are so poor. There is no definitive answer, but some explanatory approaches can be identified. Firstly, Landsat 8 and Sentinel-2 differ in the spectral ranges covered by each band. This difference can result in different spectral information, such as between the NIR band in Landsat 8 and Sentinel-2. There are also differences between the two systems in terms of spatial and temporal resolution. Another possibility is an error in the data source. However, the likelihood that the issue lies with the script in the GEE is low, as it was used with minimal modifications for both systems. Only the data source and the names of the spectral bands were adjusted. Another explanation is that no scaling factor was applied to the Landsat data. This factor is not necessary for the Sentinel-2 Collection. The scaling factor is used to adapt the measured, radiometric values of the remote sensing systems to a specific scale. This is done to convert the original data into an interpretable format. In the case of indices, however, this is hardly relevant since an index is a relative value. Even if the values of the individual pixels are changed by the scaling factor, the ratio is still the same. However, it could influence the results.

In summary, the war has had a significant impact on the agricultural sector in the Ukraine. War strongly influences agriculture, regardless of the occupation status of an area.

6.2 Model Uncertainties

Despite careful scientific work, there are always uncertainties in the design and implementation of scientific studies. Errors can also occur that are only noticed at the end. This chapter aims to explain these limitations concerning this study. Furthermore, it will highlight the additional steps necessary to further expand this work.

Since most online climate databases also use data from NOAA or DWD, the results overlap with those in this study. Therefore, a direct comparison of the data is difficult to implement without the appropriate independent data sources. However, since the data used has already been verified by the DWD, it can be assumed that they correspond to reality and comparability can be ensured.

The question must also be addressed as to whether more accurate models can be used to represent historical climate changes. Although the goal of the study was only to provide a rough assessment of the developments to have an additional layer of interpretation, rough overview studies can also be subject to errors that may not be immediately apparent. A detailed examination of the literature regarding climate analysis would be an important step in this regard.

The data density must also be critically questioned. Due to the limited availability of data in the region, only nine stations were examined for this study. These stations represent the weather in their neighborhood. However, weather is a spatially continuous phenomenon and can vary between stations. In this regard, it must be noted that some stations in Ukraine are located near the Black Sea and are therefore influenced by maritime effects. In contrast, some stations are located far inland and are more influenced by continental factors. The data density of each individual station is also questionable. There were often large gaps in the data. Many stations did not exist before 2003, some have not reported data since the start of the war, and others

have significant gaps during the collapse of the Iron Curtain. Although these gaps can be statistically interpolated or the corresponding areas eliminated, they always influence the results. Furthermore, the question arises to what extent a trend picks up on the real values. A trend works with average values, which means that outliers are not always considered. For example, if a month was extremely dry and the vegetation suffered greatly during this period, this outlier month will be poorly represented if the following months were cooler. It can be concluded that climate data and their associated analyses must always be carefully considered. It is also evident that many factors play a more or less important role in the analysis.

The selection of study areas is randomly based on administrative and geographical considerations. The study areas are all located near the front line within the previously described areas of influence of Russian or Ukrainian forces. This selection was done to ensure that the spatial differences between the areas are not too large, allowing for meaningful comparisons. For example, if a study area in western Ukraine were chosen as a comparison, there could be completely different agricultural spatial conditions. However, due to the selection, all study areas are directly affected by the war and may not necessarily be representative for the entire Ukraine. It is also not meaningful to try to identify patterns in the selected areas as they were chosen randomly, and therefore, it is not possible to contextualize any patterns within a larger framework. A reference area, which has remained unaffected by the influences of war, would also be meaningful for this study. Consequently, the three study areas could have been analyzed not only in relation to each other but also in comparison to a reference area. This aspect should be pursued in a subsequent step.

In the processing within the GEE, there is also the question of how much masking of clouds leads to distortions in the images and how much compositing a data collection distorts the spatial patterns. Since a landscape undergoes continuous, natural changes, compositing over a

very long period would result in a blurred and inaccurate image. Therefore, the time period should be chosen wisely to be both representative and sufficiently accurate.

This work is also dependent on the accuracy of the input data. Remote sensing sensors degrade in space and need to be calibrated regularly. These calibrations are the responsibility of the organizations operating the satellites and are difficult to trace, opaque, and cannot be influenced by oneself. If there are errors in the calibration or the data set, there is a high probability that they will not be directly noticed. This potentially leads to incorrect results and erroneous conclusions.

Many detailed questions in this work have not yet been discussed in the scientific community, and there are few papers on them. The integration of climate and remote sensing data is also an area where much research is still possible.

Furthermore, disregarded in this study are other possibilities and techniques of classification. In the field of remote sensing, there are numerous classification methods, each with their own advantages and disadvantages. An unsupervised classification or a pixel-based classification approach could have been performed. The number of validation points and training points can also be discussed. Generally, a higher number of points leads to more accurate results. However, the manual assignment of points to land cover is a time-consuming and monotonous process. As the results reveal, the accuracy is excellent and more than sufficient for this study. A problem in assessing the training and validation points arose from the circumstance that it is not always possible to determine the land cover type only based on the remote sensing images. In this thesis, it was relatively easy as there were only two classes (field and non-field). However, three areas were identified that are difficult to assess. Firstly, gardens that exhibit the spectral characteristics of an agricultural field but are not strictly fields. Secondly, meadows where it is unclear whether they are green fields or pasture areas. Finally, the border areas between two

fields (field boundaries or field edges) are difficult to determine. The resolution of the remote sensing data often does not allow for a clear distinction between whether a classification point belongs to the field or the edge. Consequently, the points that could not be clearly assessed had to be retained. Using various satellite images, an estimation was made as to which group they could belong to. Areas presumed to be meadows were classified as non-field. Gardens were also classified as non-field.

Lastly, the question arises as to the restrictions of this study. Only three study areas were used in this research, representing just a small portion of the absolute area of Ukraine. It would be interesting to conduct further investigations encompassing the entire territory of Ukraine to provide a comprehensive assessment. However, such a study would likely exceed the scope of this work. Another approach for future research is to integrate climate data into a model with remote sensing data. In the literature, indices are often used to monitor changes, but the inclusion of climate data is not frequently found. Statistical regression methods could be employed to examine the extent to which index values depend on climate data. The comparability with other regions needs to be investigated to test the transferability to other conflict areas. Classifying different crops could also be of interest. These could be analyzed in conjunction with radar or lidar data to determine, for example, vegetation heights. With this information, it would be possible to quantify potential agricultural product yields. There is, therefore, a great deal of research potential following this study.

This study further demonstrates the impact of armed conflict on agriculture and opens up additional perspectives for subsequent research. Further investigation of the remaining questions could shed light on the effects of war and contribute to technological advancements in the automated detection of agricultural changes.

7. Conclusion

Humans and the environment are in constant interaction with each other. War and destruction strain these fragile connections. This also applies to agriculture. This master thesis explored the topic of agricultural change detection under military conflict, focusing on the example of the Ukrainian War. Through a comprehensive analysis that involved climate trends, land cover classification, and index-based change detection insights were gained regarding the impact of the conflict on the agricultural landscape. The beginning of this thesis set the stage by highlighting the significance of studying how war changes agriculture in the specific context of the Ukrainian War. The literature review provided a comprehensive overview of existing scientific literature and theories related to the topic of change detection, conflicts, and the interplay between climate, war, and agriculture. This chapter examines the impact of war and highlights the influence of violence on the people and the environment in Ukraine. To provide a historical context, an overview of the war in Ukraine was presented, with a focus on its key events. This historical background provided the necessary foundation for understanding the subsequent analysis by explaining the course of action.

The methods employed in this study show the integration of advanced technological tools regarding the analysis of climate and remote sensing data. The results of the analysis shed light on the complex relationship between military conflict, climate, and agriculture. They revealed significant changes in agricultural patterns. The findings presented valuable evidence of the impacts of conflict on the agricultural sector. Climate trends analysis in R allowed the examination of climate patterns over long periods, enabling the examination of their influence on the agricultural sector. The time series analysis in R for nine stations over a period of 30 years has revealed interesting results. The trend was calculated and analyzed to determine how weather influences agricultural development. During the study period, the year 2021 stood out as it provided the most optimal growth conditions for agriculture. From a long-term perspective,

it is evident that the temperature has been consistently rising over the past 30 years. Therefore, it can be inferred that the temperature in the study area has increased. The data from the stations exhibit correlation, particularly concerning temperature. In contrast to temperature, precipitation is a less temporally continuous phenomenon. This results in weaker correlations. Nonetheless, it is assumed that the stations are interconnected, and the findings of the nine investigation sites can be applied to the entire study area.

In remote sensing there are a lot of different techniques concerning the identification of agricultural areas and quantifying changes. Object based image analysis, combined with the fascinating capabilities of the Google Earth Engine, enables researchers to analyze big areas without visiting them. To achieve good results different methods were used. The classification was performed using an Object-Based Image Analysis approach, which successfully distinguished fields from non-fields with an accuracy of over 91%. For these fields, indices were computed, including NDVI, BSI, and NPCRI. An examination of the results indicates that the war has a significant impact on agricultural productivity. The NDVI has shown a significant decline, indicating a decrease in the vitality of crops. From 2021 to 2022, the decline in particularly healthy field areas ranged from 66% to 94%. The fields were likely left largely unattended and no longer fertilized, as evidenced by the results of BSI and NPCRI. These processes can be observed consistently across all study areas. Furthermore, these data also reveal that the year 2021 was particularly favorable for crop cultivation.

This research contributes to the existing body of knowledge by showcasing the potential of combining climate analysis and remote sensing techniques to investigate agricultural change in conflict regions. This thesis has demonstrated the effectiveness of the chosen methodologies and their potential for using them in conflict-ridden areas for determining changes. This master thesis has successfully explored and analyzed the agricultural change detection under military

conflict using the example of the Ukrainian war. The findings highlight the importance of understanding war dynamics in the agricultural sector.

Further investigation of larger areas in Ukraine or the application of the methodology to other conflict zones would be potential next steps worth exploring. Obtaining a current situation assessment and, consequently, a comprehensive understanding of the impacts of armed conflict can aid in limiting losses, both in terms of the environment and the suffering of people in the affected regions.

References

- Achanta, R., & Susstrunk, S. (2017). Superpixels and Polygons Using Simple Non-iterative Clustering. *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 4895–4904. <https://doi.org/10.1109/CVPR.2017.520>
- Adelaja, A., & George, J. (2019). Effects of conflict on agriculture: Evidence from the Boko Haram insurgency. *World Development*, 117, 184–195. <https://doi.org/10.1016/j.worlddev.2019.01.010>
- Asokan, A., & Anitha, J. (2019). Change detection techniques for remote sensing applications: A survey. *Earth Science Informatics*, 12(2), 143–160. <https://doi.org/10.1007/s12145-019-00380-5>
- Aung, T. S. (2021). Satellite analysis of the environmental impacts of armed-conflict in Rakhine, Myanmar. *Science of The Total Environment*, 781, 146758. <https://doi.org/10.1016/j.scitotenv.2021.146758>
- Belgiu, M., & Drăguț, L. (2016). Random forest in remote sensing: A review of applications and future directions. *ISPRS Journal of Photogrammetry and Remote Sensing*, 114, 24–31. <https://doi.org/10.1016/j.isprsjprs.2016.01.011>
- Blaschke, T. (2010). Object based image analysis for remote sensing. *ISPRS Journal of Photogrammetry and Remote Sensing*, 65(1), 2–16. <https://doi.org/10.1016/j.isprsjprs.2009.06.004>
- DWD Climate Data Center. (2017a). *Historical dataset: Monthly mean air temperature for stations worldwide, version qcV002*. https://opendata.dwd.de/climate_environment/CDC/

- DWD Climate Data Center. (2017b). *Historical dataset: Monthly precipitation total for stations worldwide, version qcV002*. https://opendata.dwd.de/climate_environment/CDC/
- DWD Climate Data Center. (2023a). *Recent dataset: Monthly mean air temperature for stations worldwide*. https://opendata.dwd.de/climate_environment/CDC/
- DWD Climate Data Center. (2023b). *Recent dataset: Monthly precipitation total for stations worldwide, version recent*. https://opendata.dwd.de/climate_environment/CDC/
- Eklundh, L., & Singh, A. (1993). A comparative analysis of standardised and unstandardised Principal Components Analysis in remote sensing. *International Journal of Remote Sensing*, 14(7), 1359–1370. <https://doi.org/10.1080/01431169308953962>
- Feizizadeh, B., Omarzadeh, D., Kazemi Garajeh, M., Lakes, T., & Blaschke, T. (2023). Machine learning data-driven approaches for land use/cover mapping and trend analysis using Google Earth Engine. *Journal of Environmental Planning and Management*, 66(3), 665–697. <https://doi.org/10.1080/09640568.2021.2001317>
- Gomes, V. C. F., Queiroz, G. R., & Ferreira, K. R. (2020). An Overview of Platforms for Big Earth Observation Data Management and Analysis. *Remote Sensing*, 12(8), Article 8. <https://doi.org/10.3390/rs12081253>
- Hatfield, J. L., & Prueger, J. H. (2010). Value of Using Different Vegetative Indices to Quantify Agricultural Crop Characteristics at Different Growth Stages under Varying Management Practices. *Remote Sensing*, 2(2), Article 2. <https://doi.org/10.3390/rs2020562>
- Hazaymeh, K., Sahwan, W., Al Shogoor, S., & Schütt, B. (2022). A Remote Sensing-Based Analysis of the Impact of Syrian Crisis on Agricultural Land Abandonment in Yarmouk River Basin. *Sensors*, 22(10), Article 10. <https://doi.org/10.3390/s22103931>

- Hossain, M. D., & Chen, D. (2019). Segmentation for Object-Based Image Analysis (OBIA): A review of algorithms and challenges from remote sensing perspective. *ISPRS Journal of Photogrammetry and Remote Sensing*, 150, 115–134. <https://doi.org/10.1016/j.isprsjprs.2019.02.009>
- Huang, S., Tang, L., Hupy, J. P., Wang, Y., & Shao, G. (2021). A commentary review on the use of normalized difference vegetation index (NDVI) in the era of popular remote sensing. *Journal of Forestry Research*, 32(1), 1–6. <https://doi.org/10.1007/s11676-020-01155-1>
- Khelifi, L., & Mignotte, M. (2020). Deep Learning for Change Detection in Remote Sensing Images: Comprehensive Review and Meta-Analysis. *IEEE Access*, 8, 126385–126400. <https://doi.org/10.1109/ACCESS.2020.3008036>
- Kotoulas, I. E., & Pusztai, W. (2022). *GEOPOLITICS OF THE WAR IN UKRAINE. Report No. 4.*
- Landsat 8 Level 2, Collection 2, Tier 1.* (2023). Google for Developers. https://developers.google.com/earth-engine/datasets/catalog/LANDSAT_LC08_C02_T1_L2
- Luo, C., Qi, B., Liu, H., Guo, D., Lu, L., Fu, Q., & Shao, Y. (2021). Using Time Series Sentinel-1 Images for Object-Oriented Crop Classification in Google Earth Engine. *Remote Sensing*, 13(4), Article 4. <https://doi.org/10.3390/rs13040561>
- Mach, K. J., Kraan, C. M., Adger, W. N., Buhaug, H., Burke, M., Fearon, J. D., Field, C. B., Hendrix, C. S., Maystadt, J.-F., O’Loughlin, J., Roessler, P., Scheffran, J., Schultz, K. A., & von Uexkull, N. (2019). Climate as a risk factor for armed conflict. *Nature*, 571(7764), Article 7764. <https://doi.org/10.1038/s41586-019-1300-6>

- Mohamed, M. A., Anders, J., & Schneider, C. (2020). Monitoring of Changes in Land Use/Land Cover in Syria from 2010 to 2018 Using Multitemporal Landsat Imagery and GIS. *Land*, 9(7), Article 7. <https://doi.org/10.3390/land9070226>
- Mottaleb, K. A., Kruseman, G., & Snapp, S. (2022). Potential impacts of Ukraine-Russia armed conflict on global wheat food security: A quantitative exploration. *Global Food Security*, 35, 100659. <https://doi.org/10.1016/j.gfs.2022.100659>
- Mutanga, O., & Kumar, L. (2019). Google Earth Engine Applications. *Remote Sensing*, 11(5), Article 5. <https://doi.org/10.3390/rs11050591>
- Nguyen, C. T., Chidthaisong, A., Kieu Diem, P., & Huo, L.-Z. (2021). A Modified Bare Soil Index to Identify Bare Land Features during Agricultural Fallow-Period in Southeast Asia Using Landsat 8. *Land*, 10(3), Article 3. <https://doi.org/10.3390/land10030231>
- NOAA. (2023). *Climate Monitoring. Global Summaries of the Month.* <https://www.ncei.noaa.gov/cdo-web/datatools/findstation>
- Pereira, P., Bašić, F., Bogunovic, I., & Barcelo, D. (2022). Russian-Ukrainian war impacts the total environment. *Science of The Total Environment*, 837, 155865. <https://doi.org/10.1016/j.scitotenv.2022.155865>
- Phiri, D., Simwanda, M., Salekin, S., Nyirenda, V. R., Murayama, Y., & Ranagalage, M. (2020). Sentinel-2 Data for Land Cover/Use Mapping: A Review. *Remote Sensing*, 12(14), Article 14. <https://doi.org/10.3390/rs12142291>
- R Core Team. (2022). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>
- Rawtani, D., Gupta, G., Khatri, N., Rao, P. K., & Hussain, C. M. (2022). Environmental damages due to war in Ukraine: A perspective. *Science of The Total Environment*, 850, 157932. <https://doi.org/10.1016/j.scitotenv.2022.157932>

Remote Sensing. (2023). <https://www.mdpi.com/journal/remotesensing/stats>

Roy, D. P., Wulder, M. A., Loveland, T. R., C.E., W., Allen, R. G., Anderson, M. C., Helder, D., Irons, J. R., Johnson, D. M., Kennedy, R., Scambos, T. A., Schaaf, C. B., Schott, J. R., Sheng, Y., Vermote, E. F., Belward, A. S., Bindschadler, R., Cohen, W. B., Gao, F., ... Zhu, Z. (2014). Landsat-8: Science and product vision for terrestrial global change research. *Remote Sensing of Environment*, 145, 154–172. <https://doi.org/10.1016/j.rse.2014.02.001>

Ryan, J., & Ulrich, J. (2023). *_xts: EXtensible Time Series_. R package version 0.12.2*. <https://CRAN.R-project.org/package=xts>

Sebastian, B., Unnikrishnan, A., & Balakrishnan, K. (2012). *Gray Level Co-Occurrence Matrices: Generalisation and Some New Features* (arXiv:1205.4831). arXiv. <https://doi.org/10.48550/arXiv.1205.4831>

Sentinel-2 MSI: MultiSpectral Instrument, Level-2A | Earth Engine Data Catalog. (2023). Google for Developers. https://developers.google.com/earth-engine/datasets/catalog/COPERNICUS_S2_SR

Sishodia, R. P., Ray, R. L., & Singh, S. K. (2020). Applications of Remote Sensing in Precision Agriculture: A Review. *Remote Sensing*, 12(19), Article 19. <https://doi.org/10.3390/rs12193136>

Skakun, S., Justice, C. O., Kussul, N., Shelestov, A., & Lavreniuk, M. (2019). Satellite Data Reveal Cropland Losses in South-Eastern Ukraine Under Military Conflict. *Frontiers in Earth Science*, 7, 305. <https://doi.org/10.3389/feart.2019.00305>

Tassi, A., Gigante, D., Modica, G., Di Martino, L., & Vizzari, M. (2021). Pixel- vs. Object-Based Landsat 8 Data Classification in Google Earth Engine Using Random Forest: The

- Case Study of Maiella National Park. *Remote Sensing*, 13(12), Article 12.
<https://doi.org/10.3390/rs13122299>
- Tassi, A., & Vizzari, M. (2020). Object-Oriented LULC Classification in Google Earth Engine Combining SNIC, GLCM, and Machine Learning Algorithms. *Remote Sensing*, 12(22), Article 22. <https://doi.org/10.3390/rs12223776>
- Walker, N. (2023). *Conflict in Ukraine: A timeline (2014 – present)*.
<https://commonslibrary.parliament.uk/research-briefings/cbp-9476/>
- Woodcock, C. E., Loveland, T. R., Herold, M., & Bauer, M. E. (2020). Transitioning from change detection to monitoring with remote sensing: A paradigm shift. *Remote Sensing of Environment*, 238, 111558. <https://doi.org/10.1016/j.rse.2019.111558>
- Zeileis, A., & Grothendieck, G. (2005). zoo: S3 Infrastructure for Regular and Irregular Time Series. *Journal of Statistical Software*, 14(6). <https://doi.org/10.18637/jss.v014.i06>
- Zhao, Q., Yu, L., Li, X., Peng, D., Zhang, Y., & Gong, P. (2021). Progress and Trends in the Application of Google Earth and Google Earth Engine. *Remote Sensing*, 13(18), Article 18. <https://doi.org/10.3390/rs13183778>

Appendix

A1: R Code for the climate analysis

```
library(zoo)
library(xts)
library(imputeTS)
library(readxl)
library(openxlsx)

setwd("C:/Users/PC/OneDrive - stud.sbg.ac.at/Desktop/Uni/Master-Thesis/Data/climate")

dnipropetrovsk <- read_excel("finaldata.xlsx", na = "NA", sheet = 2)
kharkiv <- read_excel("finaldata.xlsx", na = "NA", sheet = 3)
taganrog <- read_excel("finaldata.xlsx", na = "NA", sheet = 4)
certkovo <- read_excel("finaldata.xlsx", na = "NA", sheet = 5)
donetsk <- read_excel("finaldata.xlsx", na = "NA", sheet = 6)
luhansk <- read_excel("finaldata.xlsx", na = "NA", sheet = 7)
izium <- read_excel("finaldata.xlsx", na = "NA", sheet = 8)
mariopol <- read_excel("finaldata.xlsx", na = "NA", sheet = 9)
pryshyb <- read_excel("finaldata.xlsx", na = "NA", sheet = 10)
all.mean <- read_excel("finaldata.xlsx", na = "NA", sheet = 11)

#####

### Pearson correlation matrix of the stations for precipitation and temperature
cor.df.prec <- data.frame(dnipropetrovsk$prec, kharkiv$prec,
                          taganrog$prec, certkovo$prec, donetsk$prec,
                          luhansk$prec, izium$prec, mariopol$prec, pryshyb$prec)
cor.tab.prec <- data.frame(cor(cor.df.prec, use = "complete.obs"))
write.xlsx(cor.tab.prec, file="cor_prec.xlsx", overwrite = TRUE, asTable = TRUE)

### Function to create a p-value table
cor.test.p <- function(x){
  FUN <- function(x, y) cor.test(x, y)[["p.value"]]
  z <- outer(
    colnames(x),
```

```

colnames(x),
  Vectorize(function(i,j) FUN(x[,i], x[,j]))
)
dimnames(z) <- list(colnames(x), colnames(x))
z
}

### calculating and exporting p-value table
test1 <- cor.test.p(cor.df.prec)
test1 <- data.frame(test1)
write.xlsx(test1, file="cor_prec_test.xlsx", overwrite = TRUE, asTable = TRUE)

cor.df.temp <- data.frame(dnipropetrovsk$temp, kharkiv$temp,
  taganrog$temp, certkovo$temp, donetsk$temp,
  luhansk$temp, izium$temp, mariopol$temp, pryshyb$temp)
cor.tab.temp <- data.frame(cor(cor.df.temp, use = "complete.obs"))
write.xlsx(cor.tab.temp, file="cor_temp.xlsx", overwrite = TRUE, asTable = TRUE)

### calculating and exporting p-value table
test2 <- cor.test.p(cor.df.temp)
test2 <- data.frame(test2)
write.xlsx(test2, file="cor_temp_test.xlsx", overwrite = TRUE, asTable = TRUE)
#####

### Precipitation: As function to automize analysis
decomp.ts.prec <- function(file) {
  # Create a ts-object and interpolate na-values. A trend decomposit will be displayed
  ts.data <- ts(file$prec, frequency=12, start=c(1991,1))
  ts.data <- decompose(na.StructTS(ts.data))
  plot.ts <- plot(ts.data)

  # Trend for the examination timerange
  ts.part <- window(ts.data$trend, 2013)
  plot.ts.part <- plot(ts.part, lwd = 1.5, xlab = "Year", ylab = "Prec. mm")
}

### Temperature: As function to automize analysis
decomp.ts.temp <- function(file) {

```

```

ts.data <- ts(file$temp, frequency=12, start=c(1991,1))

ts.data <- decompose(na.StructTS(ts.data))

plot.ts <- plot(ts.data)

ts.part <- window(ts.data$trend, 2013)
plot(ts.part, lwd = 1.5, xlab = "Year", ylab = "Temp. °C")
}

#####
### Insert station name in function
decomp.ts.prec(pryshyb)
decomp.ts.temp(pryshyb)

#####
### Autocorrelation for Dnipropetrovsk and all stations
dnip.df <- data.frame(dnippetrovsk$date[1:372], dnippetrovsk$prec[1:372], dnippetrovsk$temp[1:372])
acf.dnip.p <- acf(dnip.df$dnippetrovsk.prec.1.372., main="Autocorrelation Dnipropetrovsk Prec.")
acf.dnip.t <- acf(dnip.df$dnippetrovsk.temp.1.372., main="Autocorrelation Dnipropetrovsk Temp.")

all.df <- data.frame(all.mean$date[1:372], all.mean$p_mean[1:372], all.mean$t_mean[1:372])
acf.all.p <- acf(all.df$all.mean.p_mean.1.372., main="Autocorrelation all stations Prec.")
acf.all.t <- acf(all.df$all.mean.t_mean.1.372., main="Autocorrelation all stations Temp.")

#####
### Overall Trend
ts.temp <- ts(all.mean$t_mean, frequency=12, start=c(1991,1))
ts.temp <- decompose(ts.temp)
ts.temp <- window(ts.temp$trend, 2013)
ts.temp <- as.xts(ts.temp)

ts.prec <- ts(all.mean$p_mean, frequency=12, start=c(1991,1))
ts.prec <- decompose(ts.prec)
ts.prec <- window(ts.prec$trend, 2013)
ts.prec <- as.xts(ts.prec)

plot(ts.temp, main = "Mean temperature of all stations", xlab = "Year", ylab = "Temp. °C", col = "darkred")

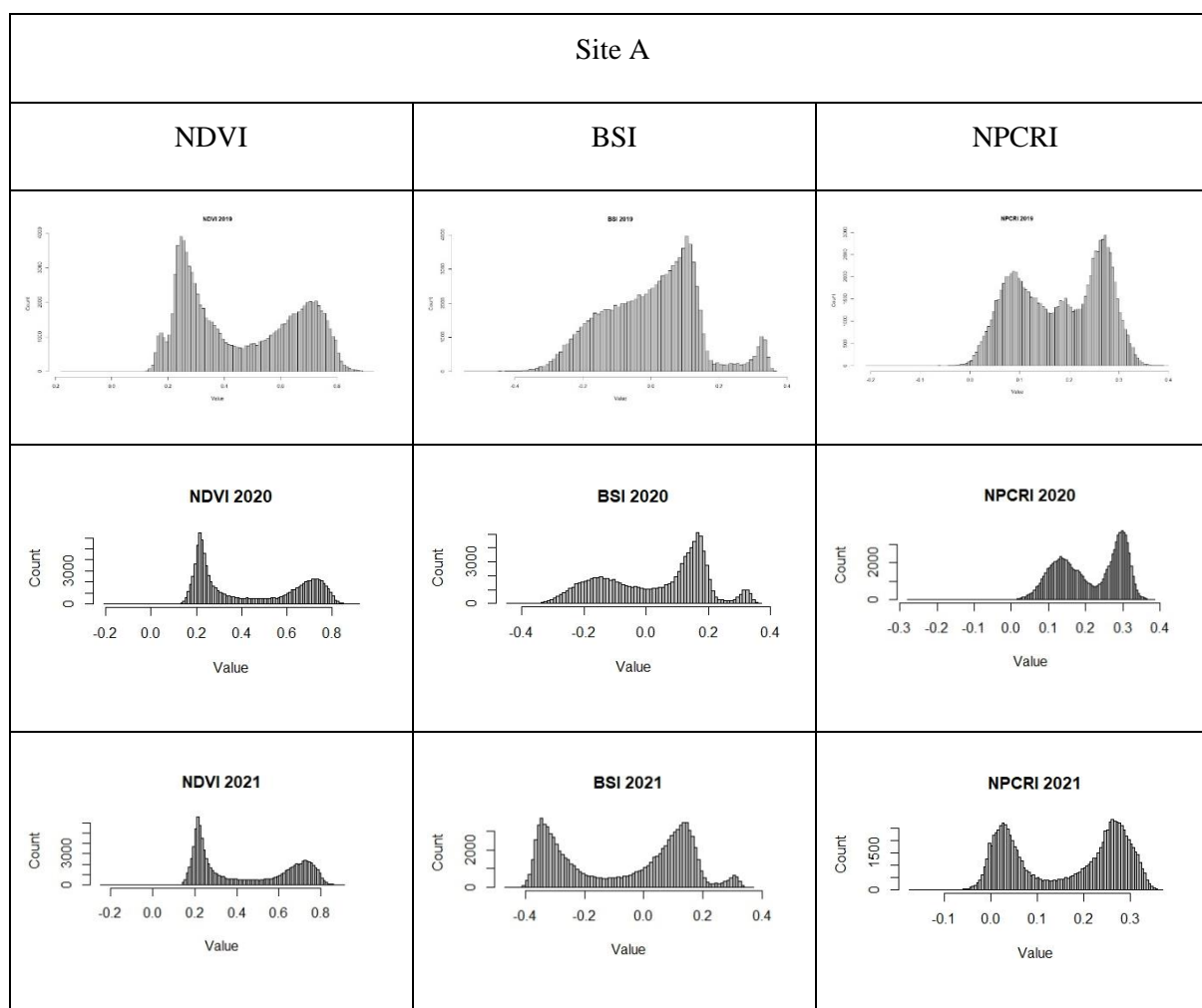
```

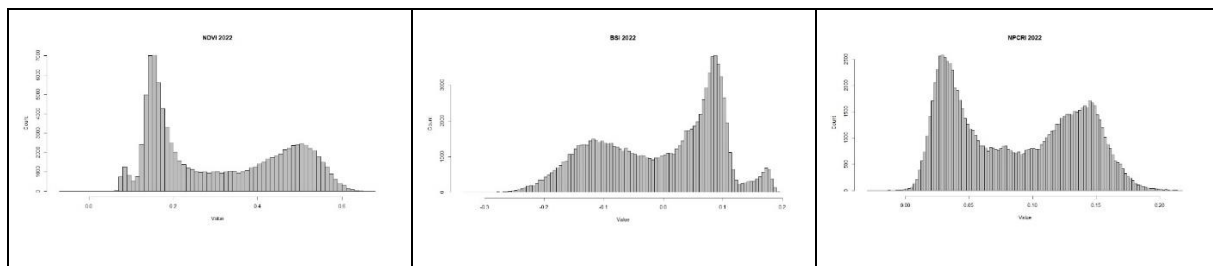
```
plot(ts.prec, main = "Mean precipitation of all stations", xlab = "Year", ylab = "Prec. mm", col = "darkblue")
```

A2: Link to the Google Earth Engine Code

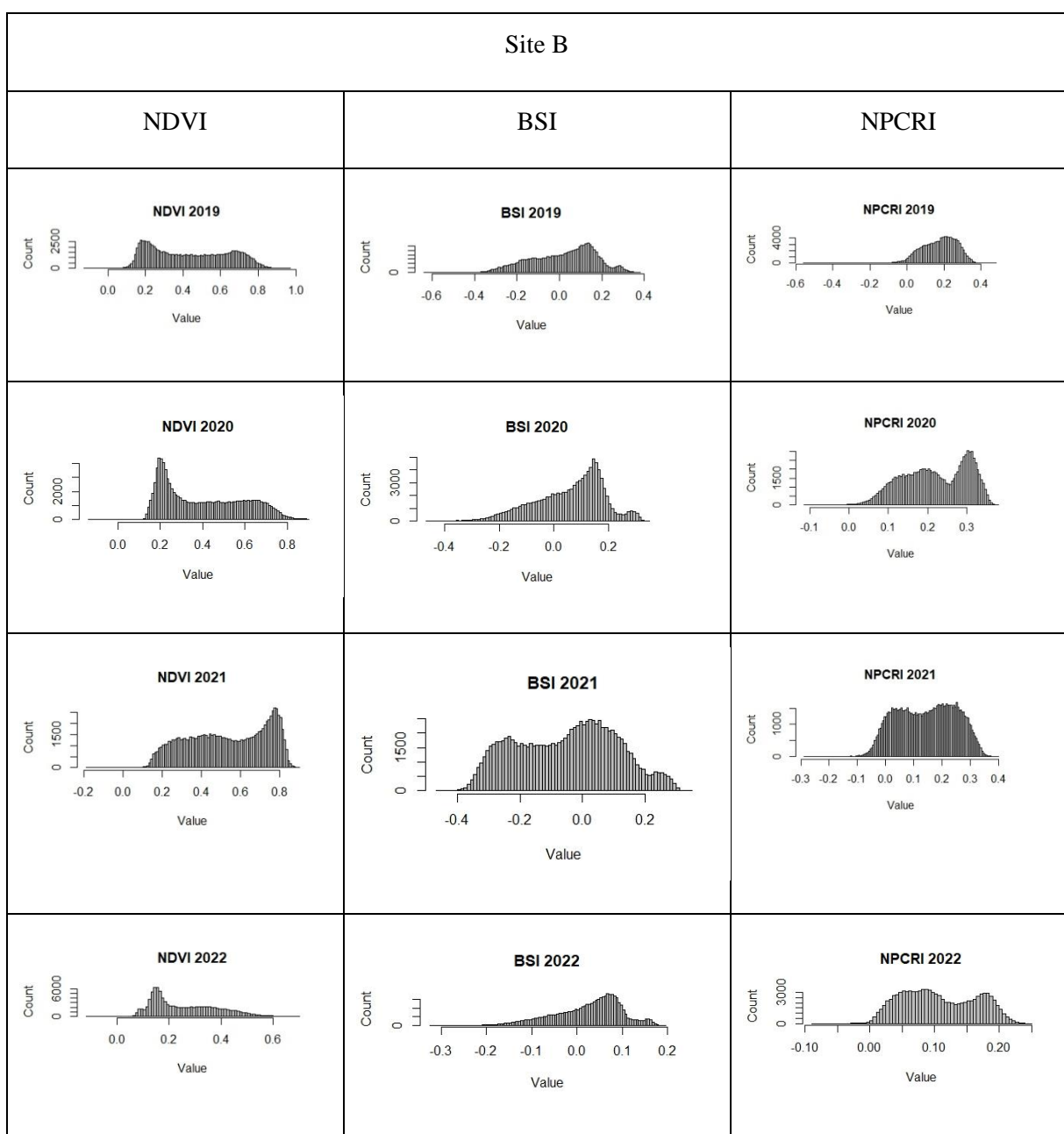
https://code.earthengine.google.com/?accept_repo=users/t_brunhorn/dataset

A3: Distributions of indices in site A

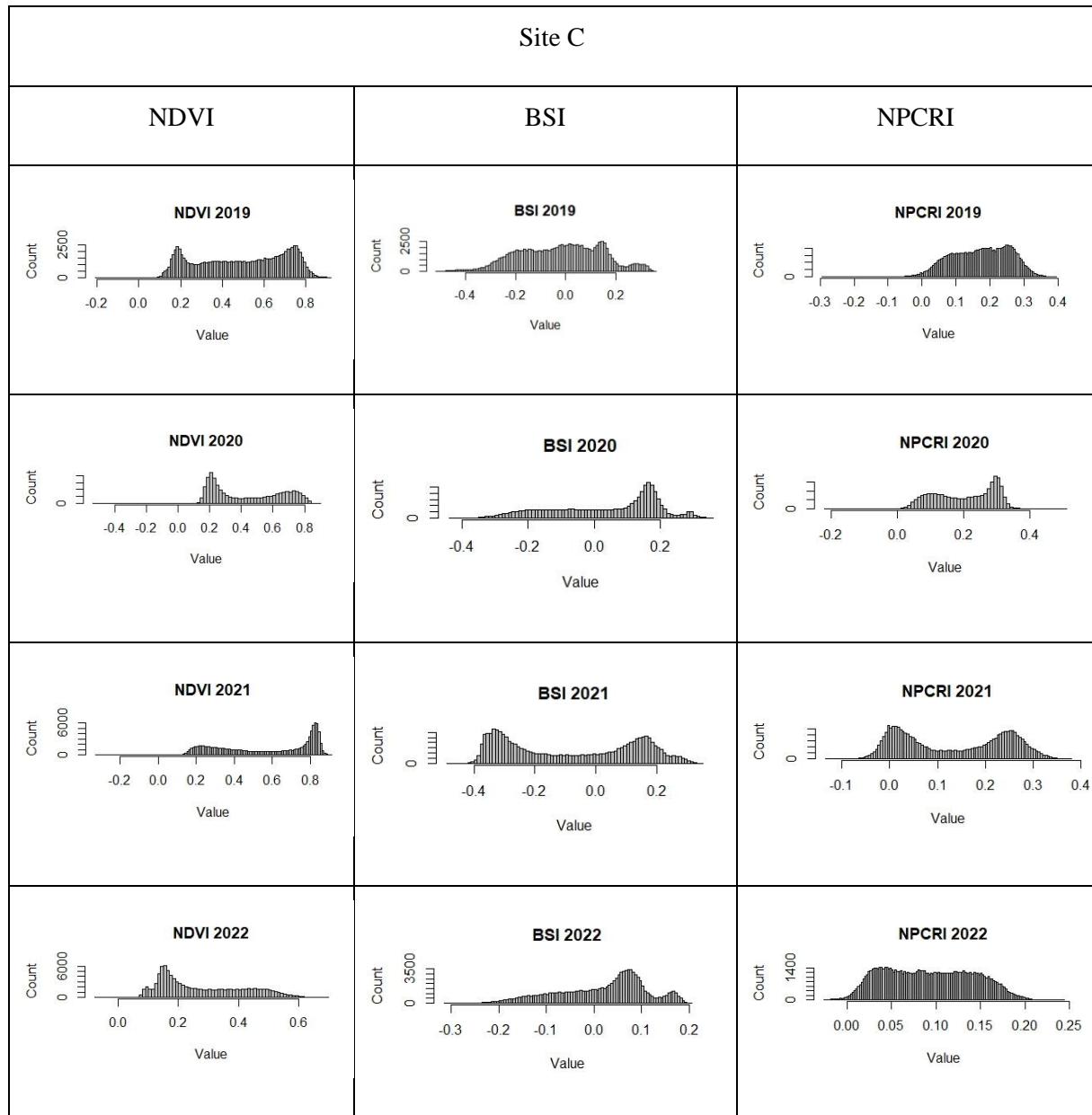




A4: Distributions of indices in site B



A5: Distributions of indices in site C



A6: R Code for creating the distributions

```
library(raster)
```

```
setwd("C:/Users/PC/OneDrive - Universität Salzburg/Desktop/Uni/Master-  
Thesis/Data/indexes/")
```

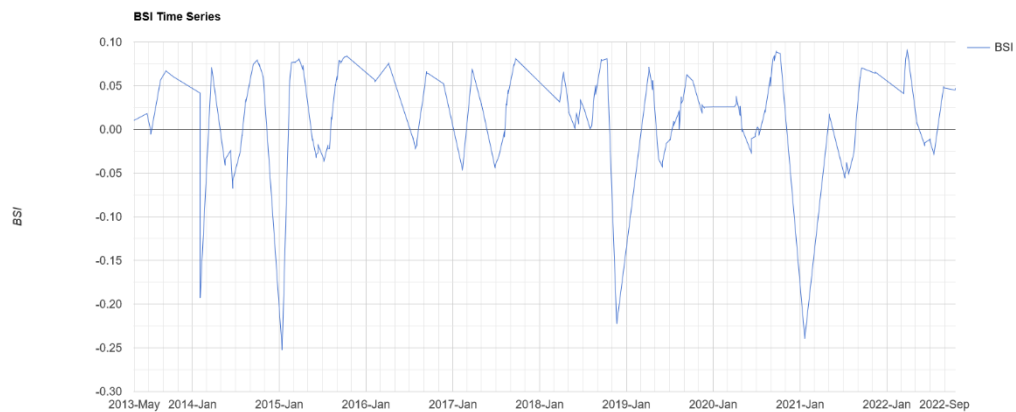
```

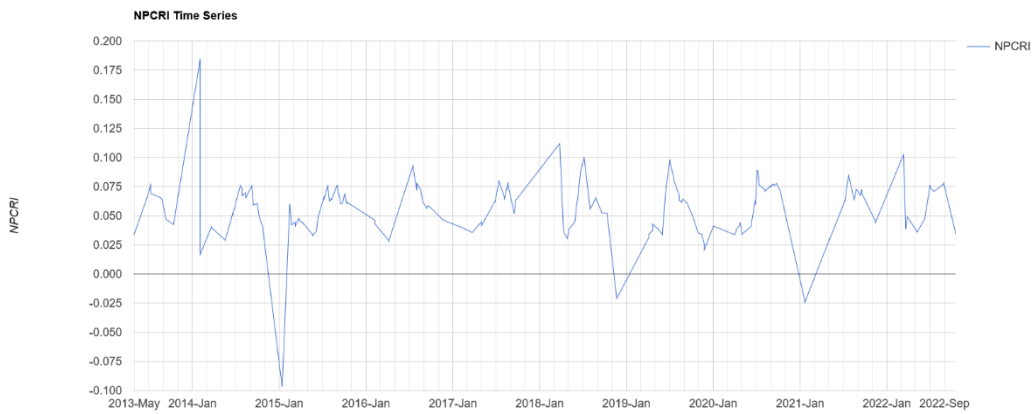
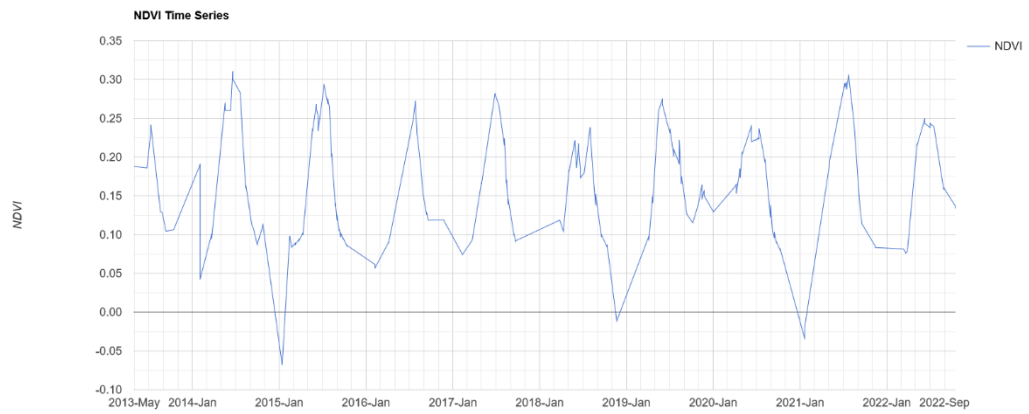
c_ndvi_2020 <- raster(x = "site_c/c_NDVI_2020.tif")
c_ndvi_2021 <- raster(x = "site_c/c_NDVI_2021.tif")
c_bsi_2020 <- raster(x = "site_c/c_BSI_2020.tif")
c_bsi_2021 <- raster(x = "site_c/c_BSI_2021.tif")
c_npcri_2020 <- raster(x = "site_c/c_NPCRI_2020.tif")
c_npcri_2021 <- raster(x = "site_c/c_NPCRI_2021.tif")

hist(c_ndvi_2020,
     main = "NDVI 2020",
     xlab = "Value", ylab = "Count",
     col = "grey",
     breaks = 100)

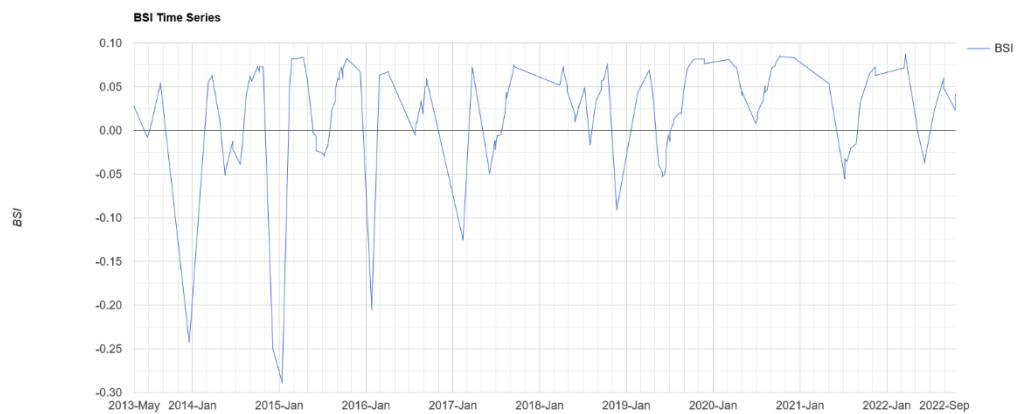
```

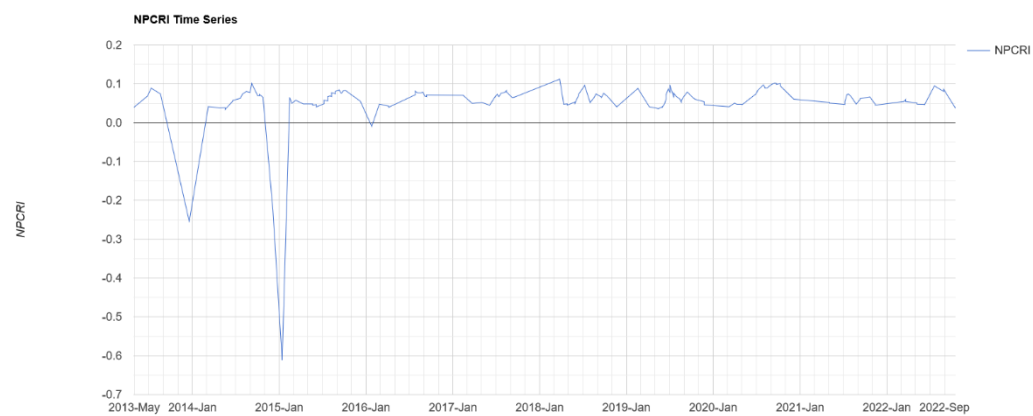
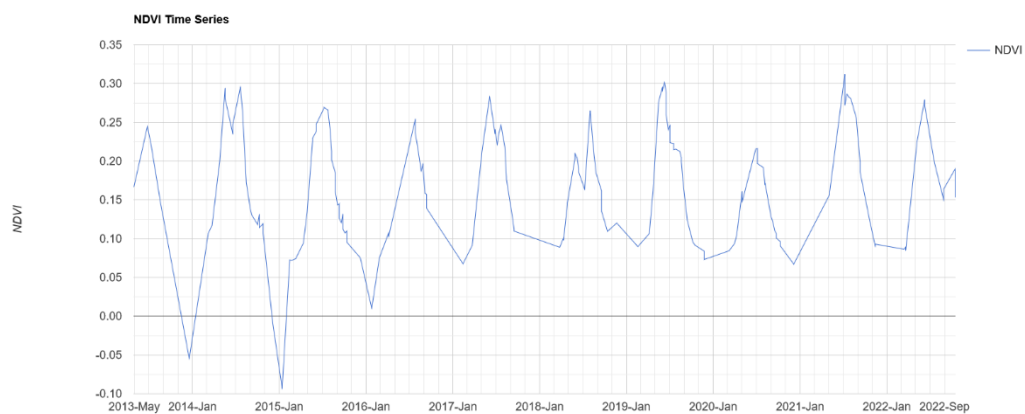
A7: Resulting time series site A Landsat 8



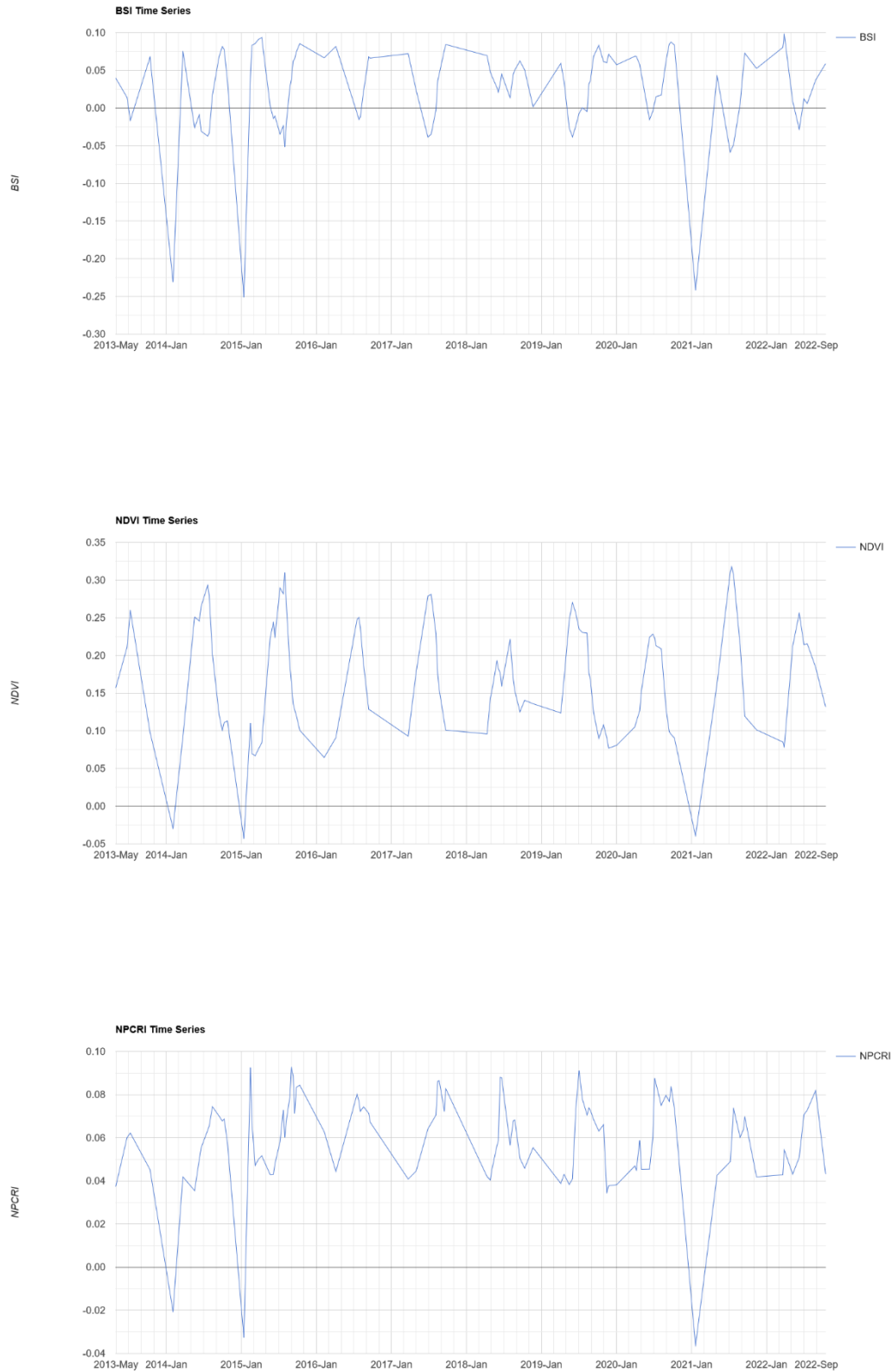


A8: Resulting time series site B Landsat 8





A9: Resulting time series site C Landsat 8



A10: Resulting time series site A Sentinel-2



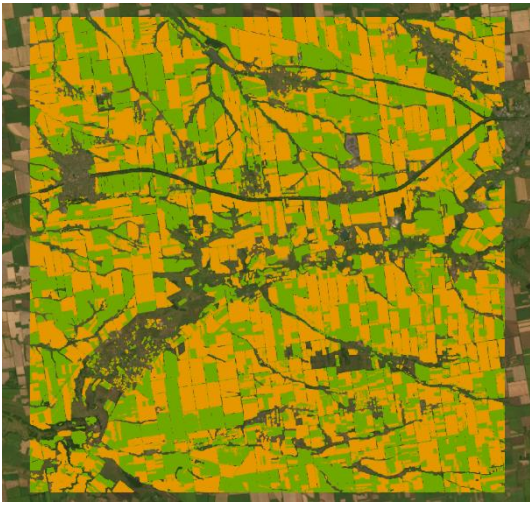
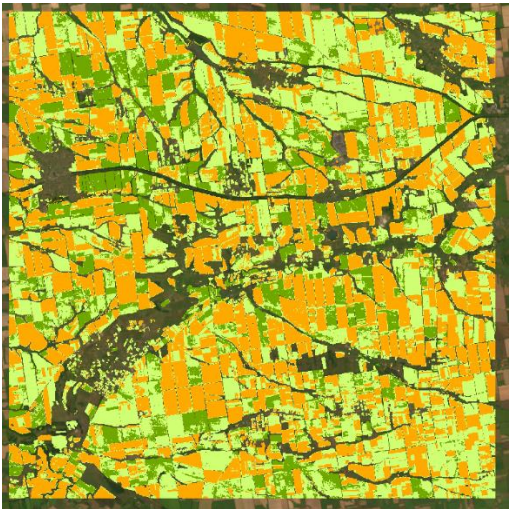
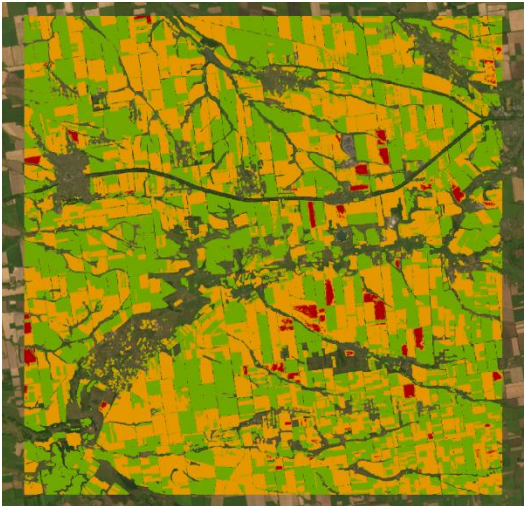
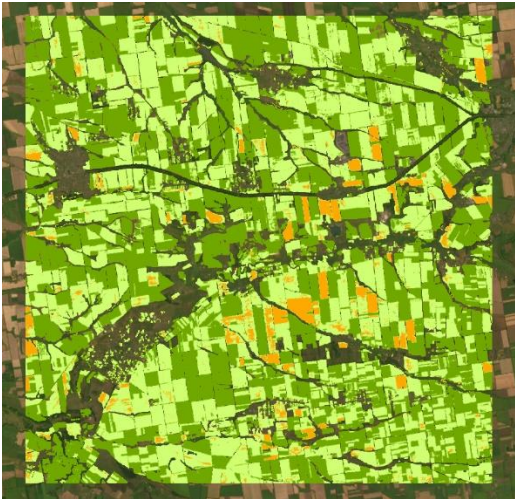
A11: Resulting time series site B Sentinel-2

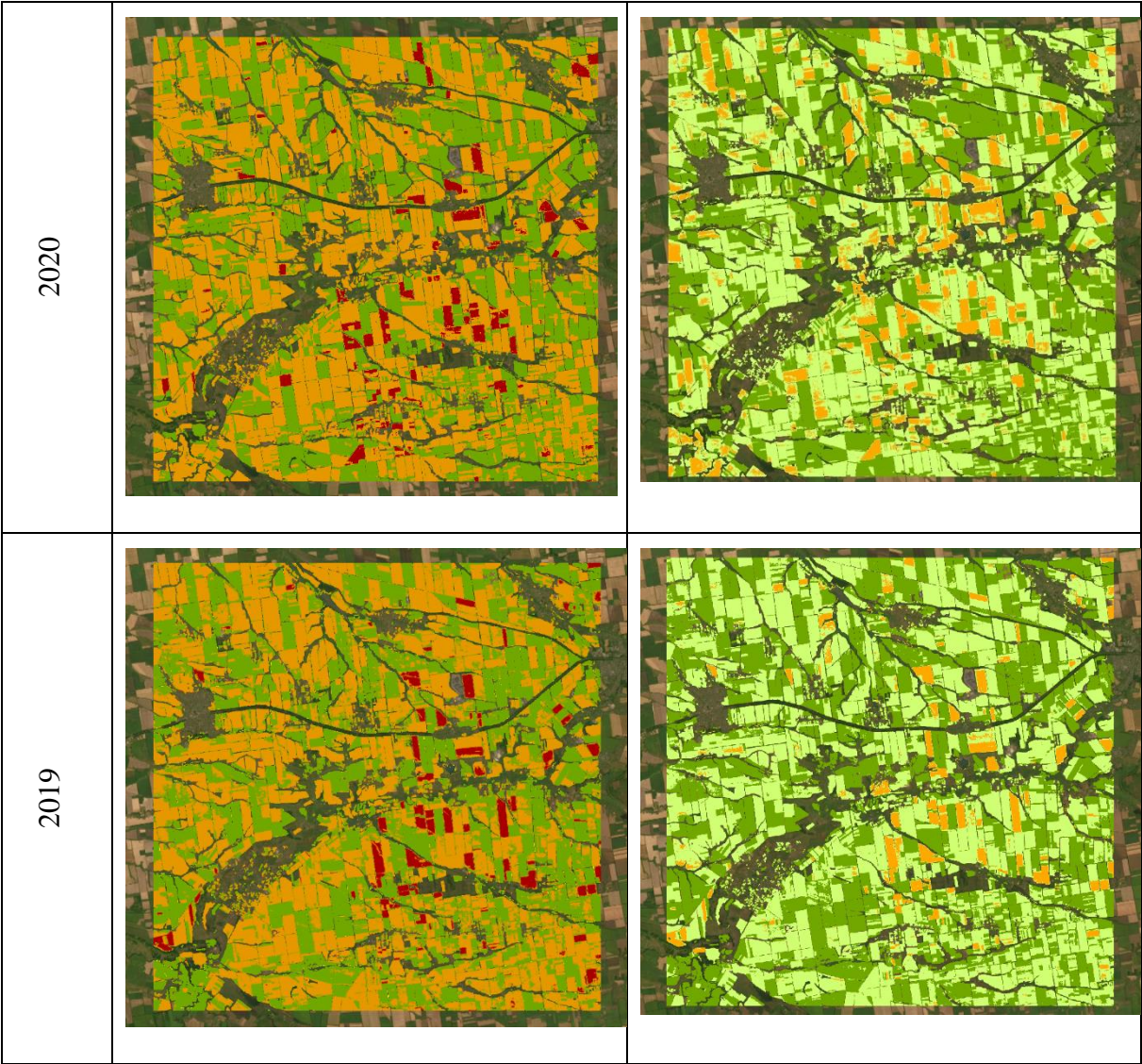


A12: Resulting time series site C Sentinel-2






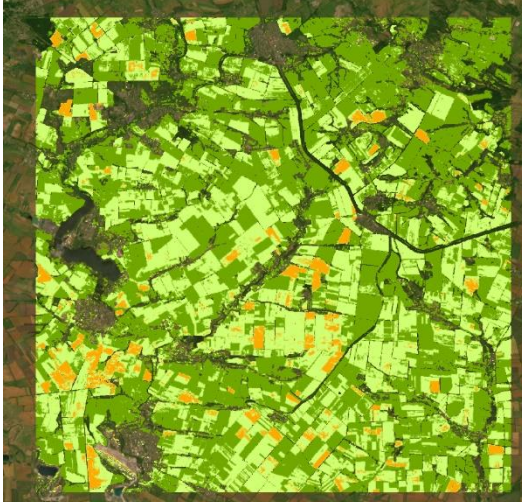
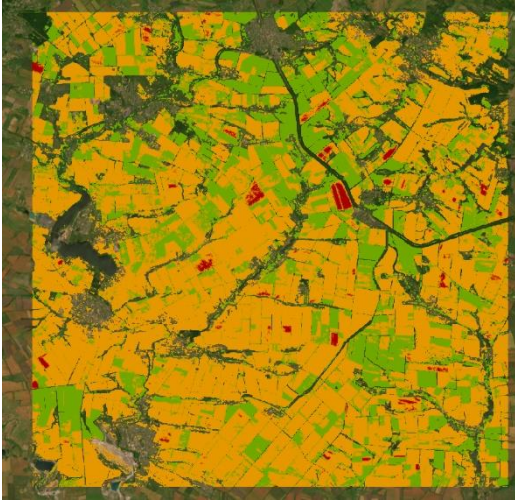

A13: Visualization of the BSI and NDVI Site A

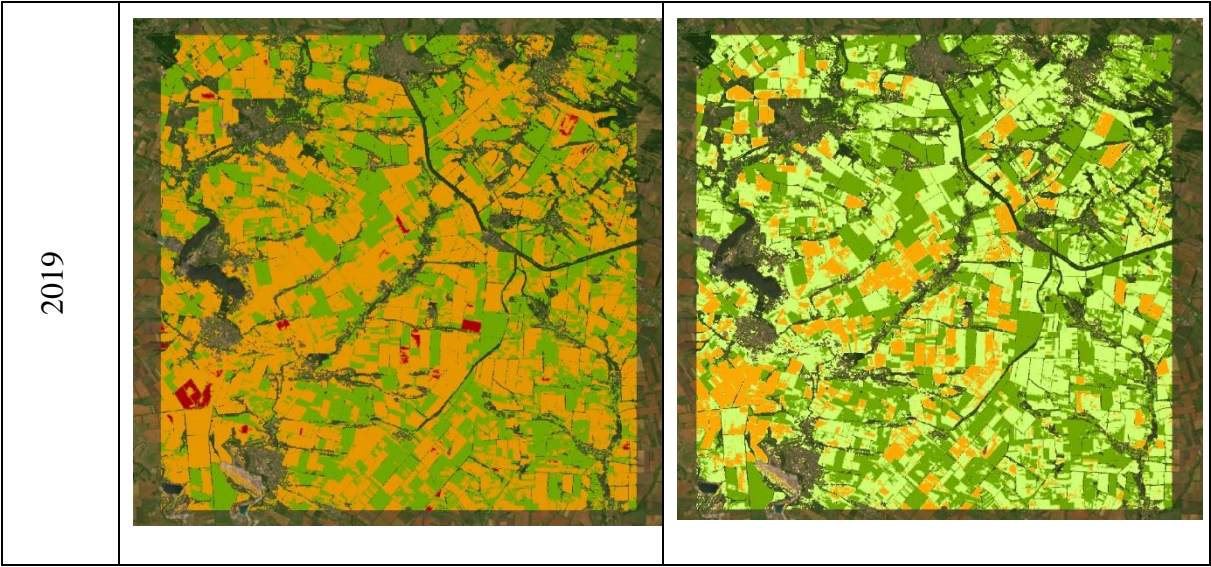
Site A		
	BSI	NDVI
2022		
2021		



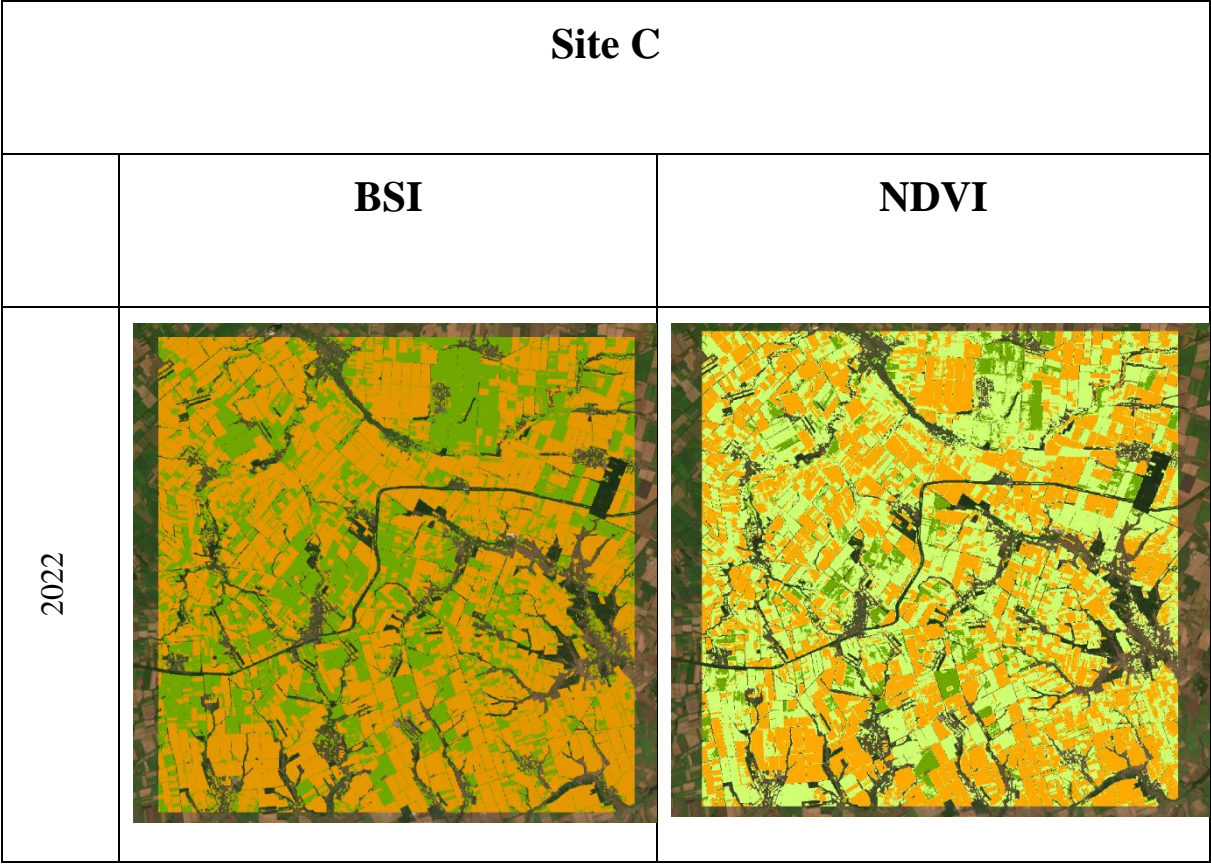
A14: Visualization of the BSI and NDVI Site B

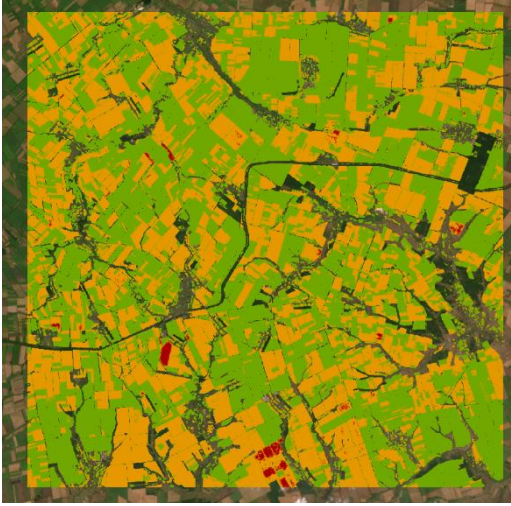

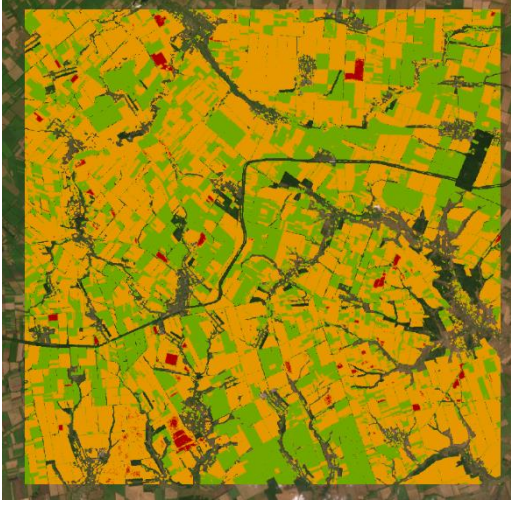

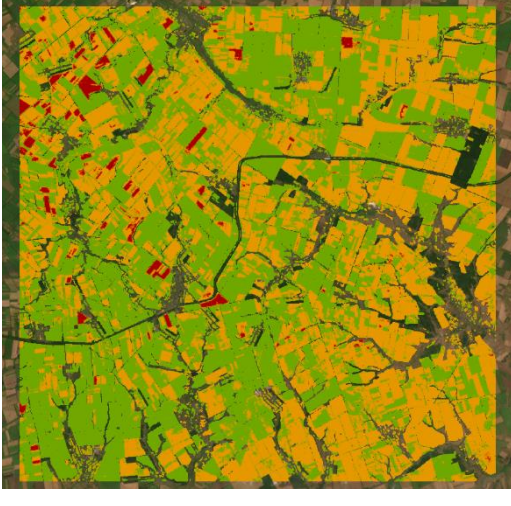
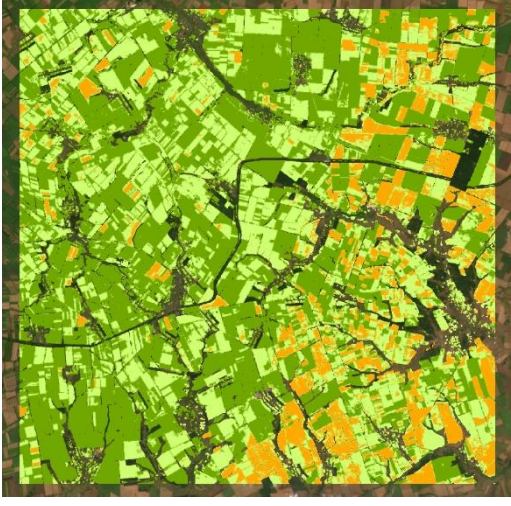
Site B		
	BSI	NDVI

2022		
2021		
2020		



A15: Visualization of the BSI and NDVI Site C



2021		
2020		
2019		

A16: Percentage of index classes inside the field areas (Site A)

NDVI								
2019		2020		2021		2022		
Class	Count	Percentage	Count	Percentage	Count	Percentage	Count	Percentage
Water	974	0,07	862	0,06	931	0,06	955	0,07
Low	70880	4,88	147511	10,15	67279	4,63	598923	41,22
Medium	744104	51,21	685883	47,20	663164	45,64	615641	42,37
High	637146	43,85	618848	42,59	721730	49,67	237585	16,35
Sum	1453104	100	1453104	100	1453104	100	1453104	100
BSI								
2019		2020		2021		2022		
Class	Count	Percentage	Count	Percentage	Count	Percentage	Count	Percentage
Vegetation	679409	46,76	597828	41,14	737749	50,77	663632	45,67
Sparse Veg.	721317	49,64	802626	55,24	691023	47,56	789472	54,33
Bare Soil	52254	3,60	52565	3,62	24323	1,67	0	0,00
Sum	1452980	100	1453019	100	1453095	100	1453104	100

A17: Percentage of index classes inside the field areas (Site B)

NDVI								
2019		2020		2021		2022		
Class	Count	Percentage	Count	Percentage	Count	Percentage	Count	Percentage
Water	37	0,00	50	0,00	50	0,00	46	0,00
Low	217196	15,02	221716	15,33	68826	4,76	642632	44,43
Medium	654394	45,25	741352	51,25	576906	39,89	756885	52,33
High	574431	39,72	483296	33,41	800624	55,35	46852	3,24
Sum	1446058	100	1446414	100	1446406	100	1446415	100
BSI								
2019		2020		2021		2022		
Class	Count	Percentage	Count	Percentage	Count	Percentage	Count	Percentage
Vegetation	599359	41,45	433851	29,99	839064	58,01	470784	32,55
Sparse Veg.	832235	57,56	996520	68,90	604978	41,83	975631	67,45
Bare Soil	14214	0,98	16044	1,11	2371	0,16	0	0,00
Sum	1445808	100	1446415	100	1446413	100	1446415	100

A18: Percentage of index classes inside the field areas (Site C)

NDVI								
2019		2020		2021		2022		
Class	Count	Percentage	Count	Percentage	Count	Percentage	Count	Percentage
Water	1531	0,10	884	0,06	1170	0,07	1266	0,08
Low	191642	12,22	228873	14,60	98238	6,27	668812	42,66
Medium	569654	36,33	689778	44,00	547615	34,93	773219	49,32
High	805055	51,35	648294	41,35	920776	58,73	124610	7,95
Sum	1567882	100	1567829	100	1567799	100	1567907	100

BSI								
2019		2020		2021		2022		
Class	Count	Percentage	Count	Percentage	Count	Percentage	Count	Percentage
Vegetation	825727	52,66	579858	36,98	889402	56,73	575232	36,69
Sparse Veg.	712360	45,43	970078	61,87	670177	42,74	992675	63,31
Bare Soil	29820	1,90	17970	1,15	8327	0,53	0	0,00
Sum	1567907	100	1567906	100	1567906	100	1567907	100

A19: P-values of the correlation analysis

p-Value-Matrix Precipitation									
	Dnibr.	Kharkiv	Taganrog	Certkovo	Donetsk	Luhansk	Izium	Mariopol	Pryshyb
Dnibr.	0	3,83576E-30	2,85294E-30	3,86087E-17	1,0869E-28	3,63391E-14	2,07638E-30	3,17411E-18	1,99432E-34
Kharkiv	3,8358E-30	0	1,98851E-15	2,83791E-28	2,53345E-27	5,25759E-18	1,88464E-30	1,11668E-06	1,20781E-11
Taganrog	2,85294E-30	1,98851E-15	0	2,90589E-29	7,84368E-34	7,40587E-29	8,87363E-15	8,44854E-25	2,10666E-14
Certkovo	3,86087E-17	2,83791E-28	2,90589E-29	0	1,61514E-29	1,60712E-31	4,85224E-12	7,61954E-15	0,000237792
Donetsk	1,0869E-28	2,53345E-27	7,84368E-34	1,61514E-29	0	5,59641E-39	2,0788E-21	5,33785E-17	1,90874E-18
Luhansk	3,63391E-14	5,25759E-18	7,40587E-29	1,60712E-31	5,59641E-39	0	4,42426E-13	2,12494E-09	3,54565E-10
Izium	2,07638E-30	1,88464E-30	8,87363E-15	4,85224E-12	2,0788E-21	4,42426E-13	0	8,04028E-14	1,36622E-14
Mariopol	3,17411E-18	1,11668E-06	8,44854E-25	7,61954E-15	5,33785E-17	2,12494E-09	8,04028E-14	0	6,44691E-16
Pryshyb	1,99432E-34	1,20781E-11	2,10666E-14	0,000237792	1,90874E-18	3,54565E-10	1,36622E-14	6,44691E-16	0

p-Value-Matrix Temperature									
	Dnibr.	Kharkiv	Taganrog	Certkovo	Donetsk	Luhansk	Izium	Mariopol	Pryshyb
Dnibr.	0	0	0	0	0	0	1,4999E-257	1,0758E-211	6,8653E-216
Kharkiv	0	0	0	0	0	0	4,7969E-270	8,141E-205	3,66E-195
Taganrog	0	0	0	0	0	0	3,7936E-240	4,9715E-223	2,6242E-211
Certkovo	0	0	0	0	0	0	6,2134E-226	1,3079E-176	1,4449E-162
Donetsk	0	0	0	0	0	0	2,3211E-155	3,2091E-127	1,5577E-109
Luhansk	0	0	0	0	0	0	2,2086E-162	1,4976E-126	1,2021E-108
Izium	1,4999E-257	4,7969E-270	3,7936E-240	6,2134E-226	2,3211E-155	2,2086E-162	0	2,5866E-235	7,6129E-229
Mariopol	1,0758E-211	8,141E-205	4,9715E-223	1,3079E-176	3,2091E-127	1,4976E-126	2,5866E-235	0	2,6127E-221
Pryshyb	6,8653E-216	3,66E-195	2,6242E-211	1,4449E-162	1,5577E-109	1,2021E-108	7,6129E-229	2,6127E-221	0