Satellite-Based Analysis of Vegetation Trends in Europe

V. B. Verhoeven

Supervised by
Dr. I. C. Dedoussi

TU Delft

Satellite-Based Analysis of Vegetation Trends in Europe

by

V. B. Verhoeven

to obtain the degree of Master of Science
at the Delft University of Technology,
to be defended publicly on Friday January 24, 2020 at 09:30 AM.

Student number:	4391594	
Project duration:	March 11, 2019 – January 24, 2020	
Thesis committee:	Dr. Ir. I. C. Dedoussi,	TU Delft, supervisor
	Dr. Ir. M. Snellen,	TU Delft
	Ir. J. A. Melkert,	TU Delft

An electronic version of this thesis is available at http://repository.tudelft.nl/.

TU Delft

Preface

This document describes the work performed for the *Satellite-Based Time Series Analysis of Vegetation Trends in Europe* Master thesis. The aim of this thesis project is to investigate the temporal changes of vegetation in Europe using statistical analysis and machine learning. As such, the intended audience consists of those who want to learn more about the application of these fields to a large and long-term remotely gathered data set, rather than people interested in learning more about the biological aspects. A basic knowledge of statistics and vegetation dynamics is expected of the reader.

The thesis document starts with an introduction of the thesis project. This is then followed by the used methodology, which consists of five steps. These are data pre-processing, land cover classification, determination of the time series components, forecasting and time series analysis. In the third part of this thesis, the generated land cover and time series results are described. In part four, the summary and conclusions are given and recommendations are made for future work. The parts and chapters contained in this document are largely independent, and interested readers can therefore choose to read only the parts or chapters that are of most interest to them.

I would like to thank Dr. Irene Dedoussi for her excellent supervision during the thesis project. Without her, the quality of the work as well as the pleasure I have taken in producing it would both have suffered greatly.

This work was carried out on the Dutch national e-infrastructure with the support of SURF Cooperative. Without the use of a supercomputer, it would have been impossible to generate results at such a high resolution.

V. B. Verhoeven
Delft, January 2020

The cover image shows filtered NDVI data of the European domain on the 1st of June 2018.

Abstract

Vegetation is a key part of the world's ecosystem and is affected by human activities both directly, in the form of changing land cover, as well as indirectly due to climate change. In turn, vegetation has an effect on the climate itself due to forming an integral part of several climatic processes, making it a complex system to analyse.

To better understand the resulting changes to vegetation, remotely sensed SPOT-VGT and PROBA-V normalised difference vegetation index (NDVI) data was analysed. The spatial resolution is one by one kilometres and it is produced every ten days. By investigating the data between 1999 and 2018 over the continent of Europe, long-term results have been determined for a large spatial domain.

To obtain these results, a methodology consisting of five core steps was applied. The first step is to reduce the noise inherent in the raw vegetation data. By substituting for invalid samples and applying the iterative Savitzky-Golay filter, a noise-reduced NDVI data set was created. The second step was then to split this data set into its various land cover types by applying bagged classification trees on spatially segmented data. The third step consisted of decomposing the time series into their trend and seasonal components using STL decomposition, which enabled an independent analysis of the vegetation trend and seasonal behaviour. Break points were detected of the trend components, in order to partition this data. These time series components were then forecast two years using an ARIMA model such that the expected future values were obtained. The final step involved analysing the trend and seasonal components for both the original twenty year time span, as well as the two year forecasts. To determine the trend properties, the Theil-Sen estimator and Mann-Kendall test were applied. For the growing season metrics, the $\text{midpoint}_{\text{pixel}}$ method was used.

With this methodology, results were obtained pertaining to the land cover, trend component and seasonal component. The over-all classification accuracy of the produced land cover data exceeds 80%. The highest accuracies were obtained for the common agricultural and grasslands and forests classes, and the classifier managed to capture all major land cover patterns in the domain. As annual results were determined, an analysis of the distribution in land cover and its evolution with time was performed for the countries and most populous cities within the domain. One example is the general increase in artificial land area covered by the analysed cities as well as most countries within the domain, particularly in South Eastern Europe.

Regarding the vegetation trend, an increase in NDVI of $2 \cdot 10^{-4}$ per year was found for the majority of Europe, with an equivalent decrease for the north-eastern part of the domain. The mountainous regions show a stronger decrease in NDVI of $-5 \cdot 10^{-4}$ per year. For the forecast two years, a transition from browning to greening is expected for these regions, with areas of Germany and France showing a transition from greening to browning.

The seasonal component was used to determine the growing season metrics. In general, a more postponed growing season was found for the northern latitudes compared to the southern latitudes. The mountainous regions are distinct from their surroundings, showing a noticeably shorter growing season.

By providing these classification, trend and growing season results, our understanding of Europe's landscape and vegetation is improved by providing insight into both the current state of land cover and vegetation, as well as existing trends.

Contents

List of Figures

List of Tables

List of Symbols

Symbol	Description	Unit
ρ	Reflectance	-
α	Level coefficient of the Holt and Holt-Winters models	-
β	Trend coefficient of the Holt and Holt-Winters models	-
γ	Seasonal component coefficient of the Holt-Winters model	-
ϕ	Coefficient of the autoregression polynomial in the (S)ARIMA model	-
θ	Coefficient of the moving average polynomial in the (S)ARIMA model	-
ϵ	Error term used in the (S)ARIMA model	-
λ	Box-Cox transformation parameter	-
C	Coefficient of the Savitzky-Golay filter	-
D	Degree of differencing for ARMA models	-
H	Hurst exponent	-
Q	Sum of the sign of change between all pairs of data points	-
R	Remainder component of the NDVI time series	-
S	Seasonal component of the NDVI time series	-
T	Trend component of the NDVI time series	-
T_L	Level of the Theil-Sen estimator	-
T_S	Slope of the Theil-Sen estimator	-
V	The variance of the pairwise sign of change	-
Y	NDVI time series	-
Y^*	Savitzky-Golay filtered time series value	-
Z_{MK}	Mann-Kendall test statistic value	-
a	Beginning of the time series	-
b	End of the time series	-
c	Constant in the (S)ARIMA model	-
d	Polynomial order of the Savitzky-Golay filter	-
d_1	Polynomial order of the first Savitzky-Golay filter	-
d_2	Polynomial order of the iterative Savitzky-Golay filter	-
e_t	Forecast error at time t	-
$f(t)$	Linear model of the trend component, dependent on time	-
h	Number of samples past the length of the original time series	-
k	Integer part of the division between the forecast samples and the seasonal period	-
l	Level of the Holt and Holt-Winters models	-
m	Single-sided window width of the Savitzky-Golay filter	-
m_1	Single-sided window width of the first Savitzky-Golay filter	dekad
m_2	Single-sided window width of the iterative Savitzky-Golay filter	dekad
m_n	Number of samples in the majority class	-
m_p	Number of samples in the minority class	-
n	Number of samples within the time series	-
o	Degree of sampling	-
p	Posterior class probability	-
p	Autoregressive order for ARMA models	-
p_s	Period of the seasonal cycle	-
q	Moving average order for ARMA models	-
q_t	Scaled forecast error at time t	-

Symbol	Description	Unit
r	Trend of the Holt and Holt-Winters models	-
s	Seasonal component of the Holt-Winters model	-
x	Box-Cox transformed time series	-
y	Time series used as input for the Box-Cox transformation	-

Abbreviations

AG	asymmetric Gaussian
AIC	Akaike information criterion
AIC_c	bias-corrected Akaike information Criterion
ARIMA	autoregressive integrated moving average
ARMA	autoregressive moving average
AVHRR	Advanced Very High Resolution Radiometer
BIC	Bayesian information criterion
CGLS	Copernicus Global Land Service
DL	double logistic
EEA	European Environment Agency
EoS	end of the growing season
FFT	fast Fourier transform
FOEN	Swiss Federal Office for the Environment
GCOS	Global Climate Observing Systems
GDP	gross domestic product
GIMMS	Global Inventory Modelling and Map Studies
HANTS	harmonic analysis of time series
LAI	leaf area index
loess	locally weighted regression and scatterplot smoothing
LoS	length of the growing season
MASE	mean absolute scaled error
MK	Mann-Kendall
MODIS	Moderate Resolution Imaging Spectroradiometer
MOSUM	moving sum
MVC	maximum value compositing
MVI	mean value iteration
NDVI	normalised difference vegetation index
NIR	near infra-red
PBA	percentage of built-up area
PROBA-V	Project for On Board Autonomy - V
RMSE	root mean square error
SARIMA	seasonal autoregressive integrated moving average
SG	Savitzky-Golay
SoS	start of the growing season
SPOT-VGT	Satellite Pour l'Observation de la Terre Vegetation
SSR	sum of squared residuals

STL	seasonal and trend decomposition using loess
STSG	spatio-temporal Savitzky-Golay
TS	time series
USD	United States dollar
VPD	vapour pressure deficit
WS	Whittaker smoother

I

Introduction

1 Introduction

The environment provides the basis for human livelihoods and well-being by supplying food, water and essential services to human societies and culture [1]. Furthermore, the annual value of the terrestrial ecosystems is estimated at 75 trillion USD, equivalent to the global GDP [2].

Vegetation is an integral part of the environment, and fulfils an essential role in multiple biogeochemical and biophysical processes [1, 3]. An example of an important biogeochemical function of vegetation is its absorption of carbon in the global carbon cycle through photosynthesis. This stored carbon can then be released into the atmosphere through consumption or environmental disturbances, such as wildfires. As the various actors of these processes have highly interdependent relationships, responses to changes are generally complex making these systems difficult to analyse [3].

For this thesis, a particularly relevant example of the role of vegetation in a biophysical process, is its absorption and reflectance of radiation. Earth receives radiation in a wide range of wavelengths from the sun. It is partially absorbed by the atmosphere and Earth's surface, and the remainder is emitted back into space [3]. Vegetation absorbs specific wavelengths of the received radiation, and commonly shows a maximum in the absorption of red radiation. As near-infrared radiation is absorbed to a far lesser extent, a strong contrast is present. Vegetation thus emits considerably more near-infrared radiation, compared to red radiation [4].

This radiation can be measured effectively using remote sensing. Data that is gathered remotely without interfering with the specimen is often acquired using an airplane or satellite, where satellites provide the capability to record data of a large geographical area spanning several decades [5, 6].

In order to quantify vegetation activity, use is therefore made of NDVI data acquired by satellite sensors, discussed in more detail in Section 2.1. It is the most widely used satellite observable vegetation index [7]. The definition of NDVI is given in Equation 1.1, with ρ being the reflectance value in the near infra-red (NIR) or red spectra. The aforementioned contrast in reflectance values is taken into account by the difference shown in the numerator. Meanwhile, the denominator normalises the values. The normalisation of NDVI makes it possible to perform meaningful comparisons and reduces noise [8]. Values thus range from zero to one, where zero denotes a complete lack of vegetation and one the maximum possible amount.

$$NDVI = \frac{\rho_{NIR} - \rho_{red}}{\rho_{NIR} + \rho_{red}} \tag{1.1}$$

Due to the strong correlation NDVI has with vegetation productivity [9, 10], NDVI time series can be used to determine a wide array of metrics with which vegetation dynamics can be studied. Metrics of interest that can be directly derived from the NDVI time series are the annual maximum, minimum and mean NDVI values and the time at which said maximum occurs. These are commonly used to gain insight into existing trends in vegetation activity. Other metrics that are commonly used in vegetation studies are the time of, NDVI value at, and rate of change at the start and end of the growing season. These metrics are closely related to the phenology of the vegetation itself [9, 11]. This allows NDVI data to be used for many purposes, such as the assessment of the ecological response to global warming, changes in phenology, agricultural crop statuses and changes in land cover [12–15].

The importance of vegetation to the environment and the ability to use remotely sensed data for its analysis have now been explained, however no motivation is yet given as to why it is important to perform a long-term analysis of vegetation. The introduction therefore continues by providing this motivation, by giving a brief overview of the effects caused by humans to vegetation.

1.1. Human-Induced Changes to Vegetation

Human-induced climate change is indirectly affecting vegetation by changing the earth's atmospheric composition and temperature [3]. As mentioned before, vegetation forms a key component of the global carbon cycle, and the increased CO_2 concentrations allow for greater carbon absorption. As a result, it has a positive effect on the yield of certain crops and primary production of forests [16–18].

Generally speaking however, the effects of climate change to vegetation are negative. Whilst increasing temperatures have lengthened the growing season in temperate and boreal regions [3, 19], they have also reduced crop yields and increased the possibility of extreme weather events reducing that yield even further [20, 21]. The 2019 European Environment Agency report regarding the agricultural sector in Europe further states that climate change has already negatively impacted European agriculture, and will continue to do so in the future [22]. On a global scale, world hunger has been on the rise since 2015 and this trend is expected to continue as a result of reduced crop yields [23]. Additionally, heat stress and drought have fundamentally changed forests in many regions [24, 25].

Humanity also changes the world's ecosystem directly by altering its land cover. Agricultural and urban areas have both increased in size to support the world's growing population. Since the 1960s, the global area used for cropland has increased by 15% and the area used for pastures by 8% [1]. Urban surface area has increased dramatically over the last few decades and is expected to triple between 2000 and 2030 [26, 27]. As a result of this expansion, since the year 1900 between 54 and 57% of the global wetlands area has been lost, with the majority having been converted to farmland. Furthermore, it is estimated that only 13% of the wetlands area that existed in 1700 remains today and the loss rate accelerates [28]. The global area covered by forests has also declined by 3% between 1990 and 2015 and continues to decrease, with the largest losses in Brazil, Indonesia and Nigeria [29]. Further environmental problems are presented by desertification, which, despite national and international efforts, is continuing at accelerating rates [30].

1.2. Previous Vegetation Studies

The previous section signified the importance of performing a vegetation analysis, by showing that vegetation is under-going rapid change, and remote sensing enables this analysis for a large data set. As a result, prior use of remotely sensed data for vegetation analysis is expected. The interest in this field is shown by the exponential increase in the number of studies published on remote sensing of vegetation data, indicating the great academic interest in the matter [7]. This section will therefore provide a brief overview of existing vegetation studies. Literature regarding more specific aspects of the project are cited in their respective sections, and the focus here lies on more generally relevant literature to provide context and motivation for the thesis research described in this document.

The used time frame for vegetation studies is commonly around twenty years [10, 13, 14, 31–35]. A few studies significantly deviate from this time span, exceeding thirty years [36, 37] or analysing only three years [38]. Much less consistency is found for the analysed geographical areas. There are studies which restrict themselves to a specific country [11, 33, 39–43], as well as studies which look into naturally defined regions [13, 14, 32, 37, 44, 45]. Finally, a great amount of research has also been conducted on a global scale [10, 31, 36, 46–48].

As this project analyses data in Europe, special attention is paid to research conducted on this region. Two studies analysed the data to determine the status of agriculture in Europe. S. Estel et al. studied the abandonment and recultivation of farmland and the intensity of cropland use [49, 50]. They found that abandonment of farmland continues, particularly in Eastern Europe, but the trend is declining. The agricultural intensity was also lower in Eastern Europe, compared to North-Western Europe. A classification study was performed by J.-L. Champeaux et al., who mapped different types of land cover in Europe using NDVI data in order to provide more information for weather forecasting [38]. As classification was not the primary goal, not all resulting classes are given representative names. Furthermore, the results are more than twenty years out of date. Y. Julien et al. studied the change in land surface temperature and NDVI between 1982 and 1999 to determine changes in vegetation. Their study found that dry regions in Southern Europe have become warmer

and have seen a decrease in vegetation, whilst the north became colder and greener [34]. One study analysing Europe that is most closely related to the project proposed here, is that by R. Stöckli and P. L. Vidale [35]. They used NDVI to determine various vegetation characteristics over a twenty year period, concluding that the growing season lengthened and NDVI increased for all regions, including arid ones. This is in contradiction with the findings from Julien et al. [34], and the authors do note that the data they used was of questionable accuracy. Furthermore, they used data from 1982 to 2001 and findings are outdated. Whilst the effects of climate change could undoubtedly already be observed, the trends in temperature and precipitation have continued at an accelerating rate. Furthermore, extreme weather events have increased in frequency and the short and medium term effects on the current status of vegetation would be unknown [51].

Of particular relevance to this project is the use of NDVI data to determine temporal trends. These trends have been used to determine the average rate of change of NDVI over a longer period, from which it can be readily derived whether the vegetation has seen an increase or decrease in activity with time, termed greening and browning respectively. No consistent conclusion is found as to the scale of greening and browning. Predominant greening was found in the recent study by Chen et al., caused to a large extent by land-use management [31]. These positive results aren't shared by other studies, which predict varying degrees of browning. De Jong et al. and Liu et al. concluded that while greening is more common than browning, few areas showed statistically significant trends and they find an insignificant or even browning trend in areas where Chen et al. detected greening [10, 46]. Furthermore, more recent studies by de Jong et al. and Pan et al. concluded that a significant increase in browning area is present [36, 47]. Finally, in a recent study Yuan et al. derived from NDVI, leaf area index (LAI) and vapour pressure deficit (VPD) data that trend reversals from greening to browning have occurred in most of the world around the change of the millennium, caused to a large extent by a decrease in oceanic evaporation [48]. Aside from vegetation activity, more detailed information such as changes within the trend, variability, volatility and suitability for forecasting can also be determined using NDVI trends [39, 40].

1.3. Research Questions

To summarise the previous sections, the world's ecosystems are essential and evolving rapidly due to changes to the climate and land cover. The importance of climate analysis is therefore evident, however is still fraught with uncertainty [1]. In order to reduce this uncertainty, the aim of this thesis project is to analyse vegetation data to calculate temporal trends and further the understanding of the way vegetation has evolved over the years. The main research question is therefore:

Which temporal trends can be found for vegetation in Europe between 1999 and 2018 using remotely sensed data?

In order to not only quantify the changes to vegetation over the current time period, the suitability for NDVI time series to forecasting is investigated and - if deemed suitable - the expected future behaviour of vegetation is determined.

Additionally, land cover is changing rapidly and has a direct effect on the state of Europe's vegetation. Furthermore, land cover is described as an essential climate variable by the World Meteorological Organization. This means that it is critically important for the understanding of Earth's climate and the ability to guide mitigation and adaption measures [52, 53]. As universal, annual land cover data is not freely available at high resolution, an additional aim is therefore to investigate the classification of European land cover using vegetation data.

These additional research aims will be completed, by answering the following three research questions:

- *Are NDVI time series suitable for forecasting?*
- *How is vegetation expected to evolve into the future?*
- *Can NDVI data be used to create a representative, continent-spanning land cover data set?*

To the best knowledge of the author, no study exists that has studied Europe's vegetation on the same level of detail as is proposed here. This thesis document will continue with a description of the methodology which will be used to answer these research questions in Part II. The results after applying this methodology to Europe can be found in Part III, and the conclusion and recommendations for future work in Part IV.

II

Methodology

Overview of the Methodology

This part of the thesis document presents the methodology used to obtain the desired results. It consists of five core part: data pre-processing, classification, determining the time series (TS) components, forecasting and finally an analysis of the time series components. To aid the reader, a flowchart of the methodology is shown in Figure 1.1. The five methodological steps are shown in the green boxes, with the numbers in the top-right of each box indicating the chapter in which they can be found.

Data pre-processing is performed in order to reduce the noise within the data set and is discussed in Chapter 2. This noise-reduced NDVI data set is then divided into its various land cover types, called classification, the topic of Chapter 3. In the following chapter, Chapter 4, the existing time series data is divided into its trend and seasonal components. The future behaviour of these time series components is then determined using the forecasting approach described in Chapter 5. Now that the current and future values have been computed, an analysis can be performed on the time series components. This is the topic of the final methodology chapter, Chapter 6.

The schematic overview shown below gives an overview of the methodology in its entirety, however lacks detailed information. More detailed flowcharts are therefore provided at the end of each of the aforementioned chapters, which show the specific steps taken for that part of the methodology.

It should be noted that the methodology was developed and tested on subsets of the domain, as the computational resources were not available to do so on the full data set.

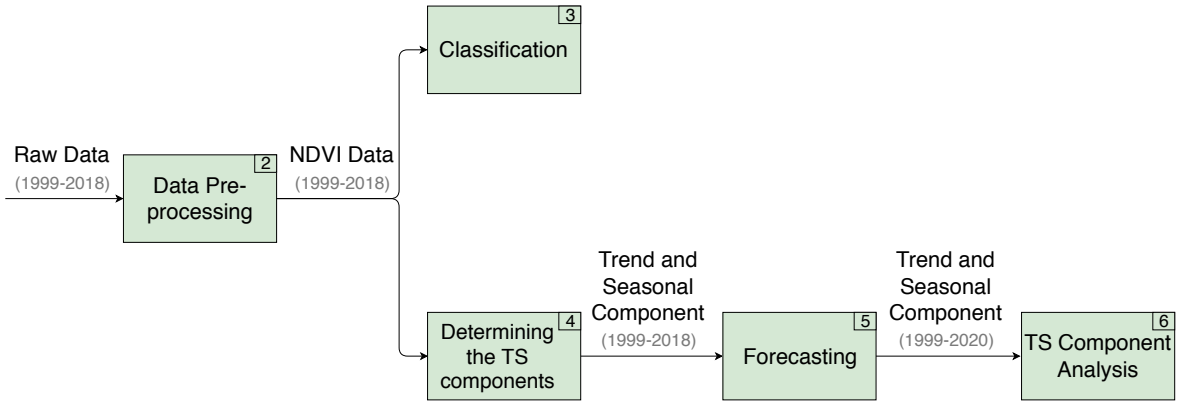


Figure 1.1: Schematic overview of the methodology. The green boxes show the five main methodological steps, with their respective chapter numbers in the top-right.

2

Data Pre-Processing

In this chapter the steps taken in data pre-processing are shown. Pre-processing the image data is essential to remove noise and thus increase the amount of information that can be retrieved from it and the confidence held therein. The chapter starts with a description of the raw data set itself, and the noise within this data set in Section 2.1 and Section 2.2 respectively. The chapter continues with the substitution of pixels already marked as being invalid in Section 2.3. After substituting these invalid pixels, noise still remains. Several algorithms exist to reduce this noise further, and the most promising ones are described in Section 2.4. To select the most suitable method, a comparison was performed as described in Section 2.5. The chapter concludes with the parameter estimation for the selected method in Section 2.6.

2.1. Vegetation Data Description

The project makes use of the Satellite Pour l'Observation de la Terre Vegetation (SPOT-VGT) and Project for On Board Autonomy - V (PROBA-V) NDVI data provided by the Copernicus Global Land Service (CGLS) [54]. Remotely sensed data makes it possible to analyse vegetation trends on large spatial as well as temporal scale and as such, the analysed data will span the continent of Europe between 1999 and 2018. The data is produced every ten days at a $1/112°$ resolution in both latitude and longitude, corresponding to roughly one by one kilometres [55]. This ten-day period will be referred to as *dekad*, from this point on, as this is the term used by the proprietors of the data [56, 57]. Compared to other NDVI data sets, the Advanced Very High Resolution Radiometer (AVHRR) data has the same spatial resolution, but is gathered daily [58]. Data at a resolution of 0.25 kilometres is offered by Moderate Resolution Imaging Spectroradiometer (MODIS) data, however this data is only produced every sixteen days [59]. Both Global Inventory Modelling and Map Studies (GIMMS) as well as the Pathfinder data use an eight kilometre resolution, produced in intervals of fifteen and ten days respectively [60].

The NDVI data provided by Copernicus takes values between 0 and 256, and can be converted to NDVI values using Equation 2.1.

$$NDVI = \frac{Raw\,value}{250} - 0.08 \tag{2.1}$$

As can be seen in this equation, the raw value is not simply divided by the maximum value of 256. Instead, the proprietors of the NDVI data have given several status flags to the data, to signify that these samples cannot be used to analyse vegetation. These status flags occupy values 251 through 255. The status types are given below and more details can be found in Section 2.3.

- 251 = Bad radiometric quality
- 252 = Obscured by clouds or shadows
- 253 = Snow or ice
- 254 = Water
- 255 = Missing input data

The European domain analysed in this project extends from -11 to 32 degrees longitude and 34.5 to 71.5 degrees latitude, shown in Figure 2.1. As a result, each image consists of approximately twenty million pixels. These images are produced every ten days, resulting in 720 images over the twenty year time span and approximately 14.4 billion pixels total.

Figure 2.1: NDVI Data of the European domain.

2.2. Description of the Noise in the NDVI Data Set

Before noise reduction methods can be discussed, the noise that should be expected is studied, as well as what steps have already been taken by the data proprietor.

Earth observation data contains noise and erroneous values due to interference from the earth surface and atmosphere. On the ground the data is affected by shadows, haze and scatter, and interaction with atmospheric substances such as ozone and water vapour further degrades the data quality. Additionally, optical remote sensing data cannot be gathered through cloud cover. These disturbances are also not constant in time, as regardless of atmospheric and surface flux the angle with which the satellite views a specific area changes continuously [5, 61, 62].

The SPOT-VGT and PROBA-V NDVI data provided by the CGLS has already seen corrections for the atmospheric conditions and surface reflectance values, as well as the different viewing angles, evolving orbital conditions and sensor degradation with time [55–57]. Additionally, to increase the data quality maximum value compositing (MVC) has been applied to the data [55]. MVC is a technique commonly applied to remotely sensed NDVI data, where the NDVI signal is maximised by processing multiple images within a given time frame. This reduces the effect of clouds, atmosphere and viewing angle [63]. The maximum signal is selected, as due to atmospheric noise the NDVI data is negatively biased [62, 64]. These measures however

do not mean that noise has been fully removed. In fact, CGLS acknowledges this as within the data set invalid pixels are still present and are given unique values to signify this, as was briefly explained in Section 2.1. Furthermore, noise reduction was still applied by CGLS in order to produce a land cover map from the NDVI data [65]. Furthermore, noise reduction is used almost ubiquitously in research, with two examples being those by R. Geerken et al. and J. M. B. Carreiras et al., where the data was used to classify land cover [66, 67].

As an example, an agricultural time series in South Holland is shown in Figure 2.2. In this figure the noise can be readily seen, as the time series contains several strong positive and negative outliers and lacks a smooth temporal profile. As discussed in Section 2.1, certain statuses are denoted by the higher pixel values. These strong positive outliers are therefore already marked, and the next section goes into more detail regarding these pixels and their substitution.

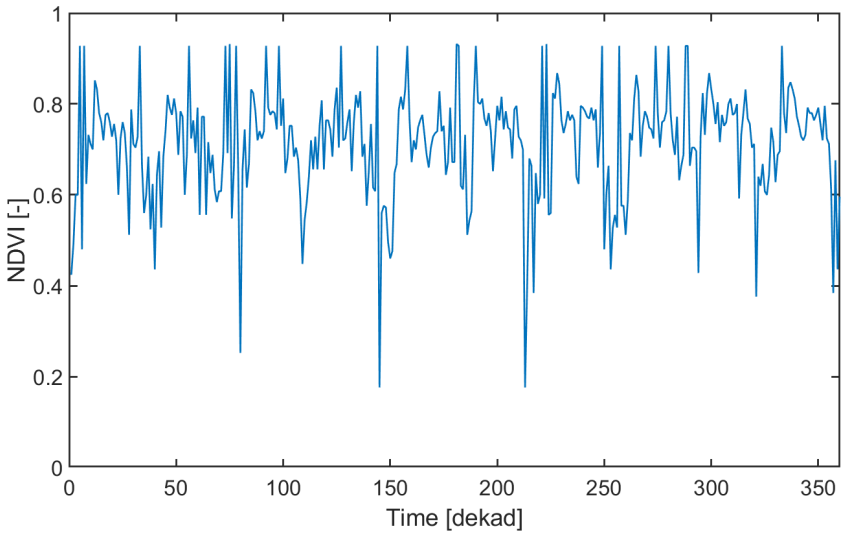


Figure 2.2: An agricultural NDVI time series before noise reduction. Invalid pixels are shown by the positive outliers.

2.3. Temporal Interpolation of Invalid Pixels

As described in Section 2.1 and shown by the positive outliers in Figure 2.2, the CGLS already identified several pixels as having specific statuses. Of these statuses, all aside from the 'sea' flag are seen as unusable for analysis purposes. Aside from these, pixels that show a highly unlikely increase are also considered invalid. More specifically, an increase greater than 0.4 NDVI in 20 days is used as the threshold, as suggested by Chen et al. [68]. The types of invalid pixels that are to be substituted are therefore the following:

- Bad radiometry
- Clouds and shadows
- Snow and ice
- Missing input data
- Unlikely increase in NDVI

In Figure 2.3, the percentage of invalid pixel for each time stamp are shown of the entire European data set. From this figure it can be deduced that the number of invalid pixels shows a seasonal cycle, with the maximum occurring each winter. Further, pixels with the missing and unlikely NDVI increase status are negligible in quantity compared to those with cloud and snow cover or bad radiometry. Data with poor radiometric quality is found exclusively for northern latitudes, where the ground receives insufficient light during winter months.

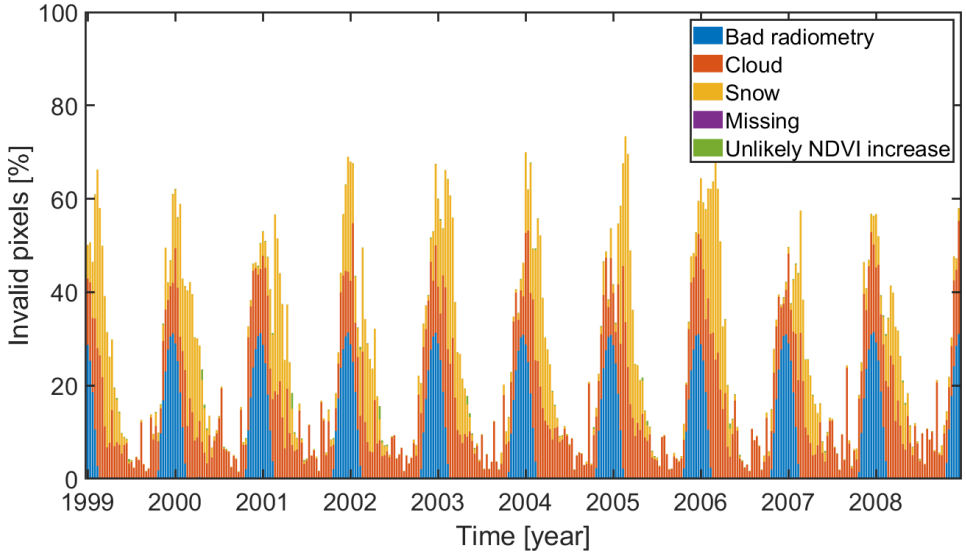


Figure 2.3: The temporal variation of the invalid pixel types in Europe.

To conclude, the quantity of invalid pixels that demand substitution varies both in time as well as space and whilst no significant temporal correlation is present for the invalid pixels in the vast majority of geographical areas, during winter months in northern regions the correlation is statistically significant. For further details, the reader is referred to Appendix A.

As these pixels are not temporally correlated, they can be substituted for using temporal interpolation. In this procedure, the weighted average of valid pixels in a certain temporal window width is used. The weight is determined by the difference in time, i.e. the next time stamp has twice the weight of the one after that and thrice the weight of the time stamp succeeding that. A symmetrical window width of five was used in the project, as a compromise between computational speed and accuracy. Figure 2.4 shows that the positive outliers have now been removed, but a significant amount of noise still remains. This noise is removed further, using the noise algorithm described in the next section.

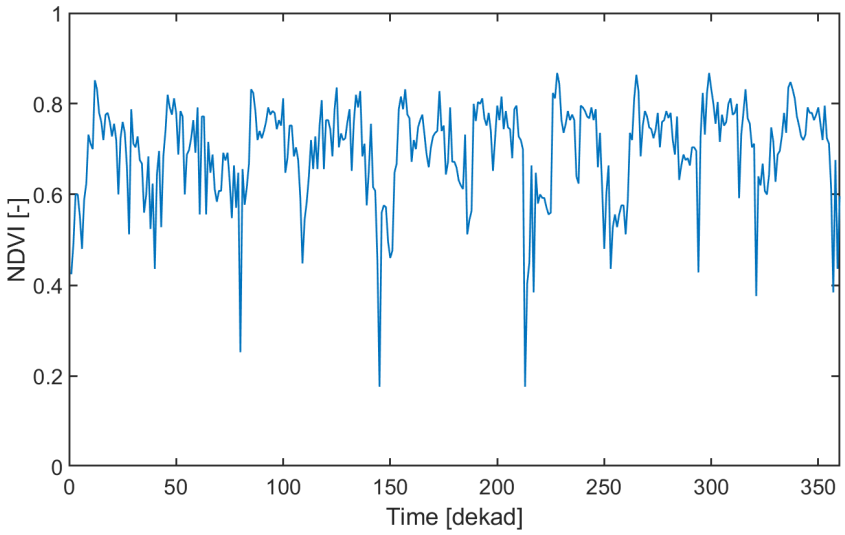


Figure 2.4: An agricultural NDVI time series after temporal interpolation. The positive outliers have been removed.

2.4. Overview of Existing Noise Reduction Methods

A wide variety of methods for noise reduction exist, with as many as sixteen and seventeen being listed by W. Wei et al. and L. Geng et al. respectively [60, 69]. Only the methods that are deemed most promising due to repeated successful applications in the field of remote sensing of vegetation are briefly described here, such that the most suitable method can be selected via a comparison in the next section.

- **Savitzky-Golay (SG):** The Savitzky-Golay filter is a moving average filter with weighting given as a polynomial. These weight coefficients when applied to a signal, perform a least-squares fit within the filter window. To apply the filter, the smoothing window width and the degree of the smoothing polynomial must be determined [70]. The weighting coefficients determined by Savitzky and Golay in 1964 contained numerical errors and were corrected by Steinier, Termonia and Deltour in 1972 [71]. One final correction and an extension of the existing coefficients was provided by H. H. Madden in 1978 [72].

  The existing Savitzky-Golay filter was modified in 2004 to better suit NDVI data and was tested using SPOT-VGT data. To compensate for the negative bias in NDVI data, an iterative Savitzky-Golay filter is applied where positive outliers are favoured over negative outliers [68]. This iterative Savitzky-Golay filter was used as the basis for the spatio-temporal Savitzky-Golay (spatio-temporal Savitzky-Golay (STSG)) filter in 2018. In this algorithm, values are estimated for invalid pixels by using the neighbouring pixels whose time series show a sufficiently high correlation with the time series of the invalid pixel. To estimate the value for the invalid pixel, the multi-year average of the valid NDVI values at that specific 10-day period is used. After substituting the invalid pixels, the iterative Savitzky-Golay filter is applied with slightly modified weights for the positive and negative outliers [73]. An issue with the estimation scheme for invalid pixels is that it is unsuitable when the NDVI time-series shows a strong trend or fluctuates with time, as the averaged values over all years are used regardless of where in the time-series the invalid pixel is located. One further issue is that when determining the performance of their algorithm they determined the mean absolute error (MAE) between the filtered time series and the averaged time series of each pixel, favouring their approach over others as this time series is used in the generation of the STSG filtered data set. No application of the STSG filter was found by the author, however the iterative SG filter has been successfully applied multiple times, one such example being the yield estimation of wheat in China by J. Ren et al. [74].

- **Mean value iteration (MVI):** The idea behind the MVI filter is that in an NDVI time-series, the value at one specific time point closely approximates the average between the values at the previous and next time steps. The first step is to determine the multi-year average NDVI value, which is used to replace invalid pixel values. Next the difference between the current pixel values and the average between the previous and next pixel values are calculated. If this difference exceeds a pre-defined threshold, the average between the previous and next pixel values is used instead [75]. From these steps it is evident that the filter is easy to implement, however it fails to take into account any existing biases within the data and by using multi-year averaged values it has the same issue as the STSG filter, where strong fluctuations with time make an average value unsuitable.

- **Whittaker smoother (WS):** The goal of the Whittaker smoother algorithm is to find the optimum of the similarity between the original and filtered data and the smoothness of the filtered data [76]. If the data is not filtered, the two data sets are fully similar, but the filtered data set is not smooth and thus no noise has been removed. On the other end, the data can be fully smoothed however has then lost almost all information within the original data set. This goal results in a simple linear set of equations. By setting the weight of invalid pixels to zero and valid pixels to one, the invalid values are automatically interpolated for by using a fitted polynomial, somewhat akin to the Savitzky-Golay filter. The used smoothing parameter is determined using cross-validation, where each element of the original data series is left out in turn and a prediction is obtained for said 'missing' data point. The error between the prediction and actual value can then be minimised to obtain the most optimal smoothing parameter. A downside of the WS algorithm is that the time series approaches a constant line when the smoothing parameter is increased. Consequently, the amplitude is always reduced regardless of whether the extrema can be considered to be noise or not.

- **Double logistic (DL):** This approach lends its name to the logistic functions that are fitted to the original data, one before and one after the maximum for each year. These curves are created using six parameters, namely the base level NDVI, maximum NDVI, two inflection points for the rise and drop

and the rates of increase and drop at the inflection points. The base level is determined using the data directly, the others using a least squares procedure. The base level NDVI is assumed constant and for long snow-cover it is determined by taking the geometric mean of the maximum NDVI values that were found over the years for each time step during the winter [77]. Downsides of using double logistics are that the NDVI seldom shows periods of negligible flux, therefore the use of a constant base NDVI is questionable. Additionally, the maximum is not always easily specified due to noise or for a vegetation type that shows little seasonality [68]. On the other hand, while the suggested method for substituting snow-covered NDVI values fails to take inter-annual variation into account, this method has shown to successfully approximate winter conditions [78].

- **Asymmetric Gaussian (AG):** The asymmetric Gaussian is a function fitting approach similar to DL, which uses two Gaussian functions rather than two logistic functions to model each year before and after the maximum. The data that is used for the algorithm is first smoothed using a running-window averaging procedure, where the number of points before and after the current point need to be specified. The two Gaussian functions use five parameters in total, namely the time at which NDVI is maximum and two parameters that determine the width and flatness for each of the two sides. Finally, the Gaussian function is multiplied with the amplitude and the base level NDVI is added resulting in the final local functions. A total of seven parameters are used to create the local functions. These parameters are determined by minimising a so-called merit function, which is the sum of the difference between the values obtained using the local functions and the original NDVI values. As a result, the time at which NDVI is maximum may be more accurately determined than when the maximum NDVI value is used in the double logistic algorithm. To take into account that the noise is negatively biased, the merit function is ran again where this time the weight of points below the function as a result of the first iteration are halved. To smoothly merge the local functions at the limbs, a global function is created by multiplying the local functions with two cutoff functions, one for the left half and one for the right half [79].

- **Harmonic analysis of time series (HANTS):** The HANTS algorithm removes faulty observations and reconstructs the time series by temporal interpolation using the remaining values. The original signal is divided into the frequencies containing the required information. Data points that are found to be invalid are removed from the series and the amplitude and phase coefficients of each frequency are determined again until the error is acceptable or too few points would remain. Invalid data points are those that have impossibly high or low values and those that show an unacceptable positive or negative deviation from the current curve [80]. Perhaps the most noteworthy application of the HANTS algorithm was its use in the generation of a land cover map by the Copernicus Global Land Service [65]. A reported issue of the HANTS algorithm is its poor performance at high latitudes. Zhou, Jia and Menenti created artificial gaps in the data to test the gap-filling properties of the HANTS algorithm and found that the error at latitudes greater than fifty degrees were significant [81]. De Jong et al. also reported issues when using HANTS for high latitude vegetation, as the algorithm returned an artificially gradual transition at the start and end of the growing season and therefore positively biased the data [10]. Another downside is that there are no objective rules to determine the five control parameters, meaning that for each application various combinations of these parameters must be tested and there is no guarantee that an optimal combination will be found [80, 81]. The number of frequencies used is crucial for the end result, however it depends on the application and study area and is not consistently applied [81]. To analyse vegetation in California, A. Moody and D. M. Johnson considered the first two harmonics, whilst Wen, Su and Ma used the first three when creating a cloud-free time series for the Tibetan Plateau [82, 83]. Geerken, Zaitchik and Evans reported that, despite two harmonics being commonly used, they are not sufficient when classifying rangeland and instead five harmonics should be used [84]. To evaluate the performance of HANTS, Zhou, Jia and Menenti used the first four harmonics [81].

2.5. Noise Reduction Methods Comparison

In the previous section the most promising noise reduction methods are listed, but with no mention yet of which one would provide the best results for this project. This section will answer that question, by first looking into existing studies that have compared the various noise reduction methods in Subsection 2.5.1. When looking at comparison studies however, it becomes clear that their test conditions, evaluation metrics and results differ making it difficult to definitively decide which method performs best. As such, an evaluation is performed to determine the best performing noise reduction algorithm for this application, as described in Subsection 2.5.2.

2.5.1. Previous Comparison Studies

Hird and McDermid tested the methods AG, DL, iterative SG and MVI as well as the less commonly used methods 4253H and ARMD3-ARMA5 in Alberta, Canada with long-term snow cover [78]. Performance was evaluated with the root mean square error (RMSE) and the season and NDVI metrics. The iterative Savitzky-Golay filter performed the best in terms of RMSE, however had poor accuracy when determining the vegetation metrics. The function fitting algorithms AG and DL performed the best over-all, however it should be noted that to calculate the error they used a modelled NDVI time series with constant winter NDVI, favouring AG and DL. A comprehensive study was performed by L. Geng et al, where they tested eight methods including all the ones listed in the previous section, aside from HANTS [60]. With SPOT-VGT data, the iterative Savitzky-Golay filter and Whittaker smoother both performed far better than the other methods. It should be noted that averaged time series were used for each land cover type, meaning that the data contained far less noise than would be present on a pixel-by-pixel basis. Y. Shao et al. tested the effect of the noise reduction method on classification performance, and as such they tested the variability within classes, separability between classes and classification results. The results were roughly equal for WS, iterative SG and the Fourier transform, with AG and DL performing worse [85]. The same methods, except the Savitzky-Golay filter, were compared by P. M. Atkinson et al. using the RMSE. Despite the fact that they did not use cross-validation to determine the smoothing parameter, WS performed the best closely followed by the Fourier transform. Again, the function fitting algorithms AG and DL performed the worst [86].

2.5.2. Quantitative Method Comparison

In the studies given in the previous subsection, the Whittaker smoother (WS) and Savitzky-Golay (SG) filters performed consistently well. As such, to make a final decision regarding the noise reduction algorithm, a comparison is drawn between the three different SG filters and WS. To put the results into context, a simple low-pass fast Fourier transform (FFT) filter is also applied, using a cut-off frequency corresponding to the parameters used by the Savitzky-Golay filter and Whittaker smoother. Further, the results are also shown for the case that no noise reduction is applied, or only the interpolation described in Section 2.3.

The effectiveness of the used noise reduction method can be determined in several ways. The most commonly used approach is to determine the error between one or multiple 'true' NDVI time series and the time series generated after filtering or function fitting. The difficulty lies in obtaining a data set that is truly representative of the real situation. In fact, the need for noise reduction is already indicative of the difficulty in obtaining this data. In literature, neither a consistent approach to this problem was found, nor one that could be applied to this project in a satisfactory manner [60, 73, 77, 78, 86]. Another way to measure the increase in data quality is by calculating the spatial and temporal autocorrelation. If the noise itself is not temporally or spatially correlated, correlation of the data will improve as the noise level is reduced. A measure for the spatial autocorrelation is the Moran's I value, which measures how disparate a pixel is to its neighbours. Noisy pixels have a pixel value that is strongly different from their neighbours, thus lowering the spatial autocorrelation and therefore lowering the Moran's I value [87]. The idea behind temporal autocorrelation is similar, only now the neighbours are in time rather than space.

The temporal noise of the time series is uncorrelated, as can for instance be seen in Figure 2.4. As such the temporal autocorrelation value is a suitable proxy for the noise level in the data. Similarly, the spatial autocorrelation requires that outliers occur individually in space. Otherwise, spatial autocorrelation can in fact be reduced when noise reduction is applied. Figure 2.5 shows several neighbouring time series. This figure acts as an example, in which it can be readily seen that these time series have the same outliers. Such an extreme example is however rare, and the fact that the Moran's I value increases when noise reduction is applied is indicative of this.

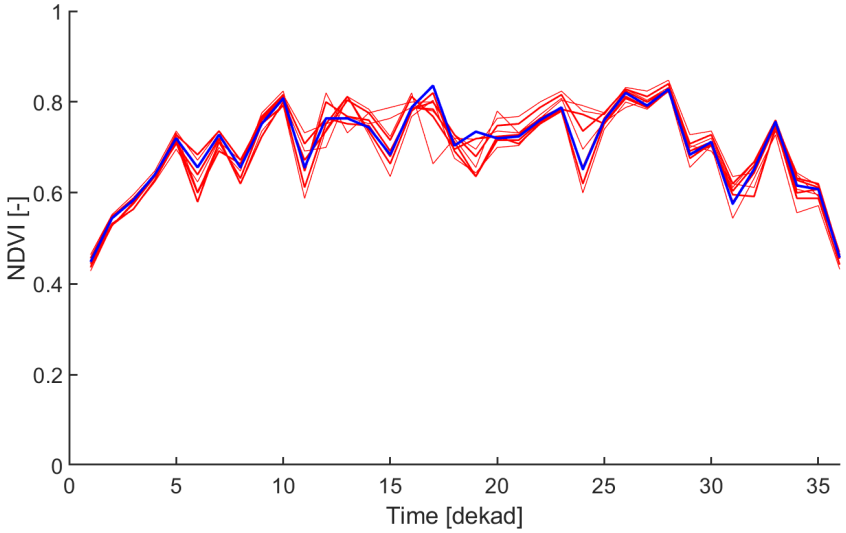


Figure 2.5: Example of neighbouring pixel time series (red) around the centre-pixel time series (blue).

The average Moran's I value and its statistical significance in the form of its z-score are presented in Figure 2.6 and the mean temporal autocorrelation and its standard deviation are given in Figure 2.7. The results shown are for South Holland, but similar results were obtained for other geographical regions. From these figures it is evident that all noise reduction algorithms improve the data quality and differences between the various methods are minor. Furthermore, the data is highly clustered even when noise reduction is not applied. Firstly, the iterative Savitzky-Golay filter consistently performed better than the other two Savitzky-Golay filters. Out of the remaining methods, the iter. SG filter ranks second in both tests, but is the only one that takes the negative bias described in Section 2.2 into account. As such, the iterative Savitzky-Golay filter was selected for noise reduction of the NDVI data in this thesis project.

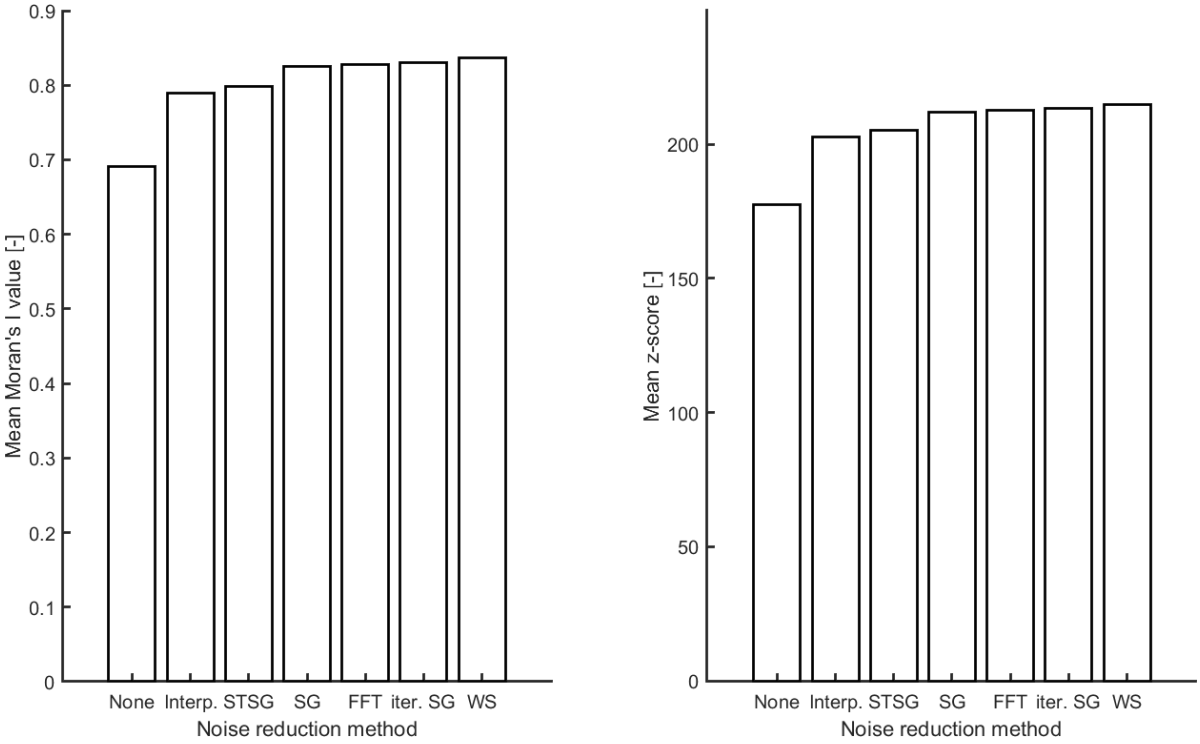


Figure 2.6: Spatial autocorrelation and its statistical significance of the various noise reduction algorithms.

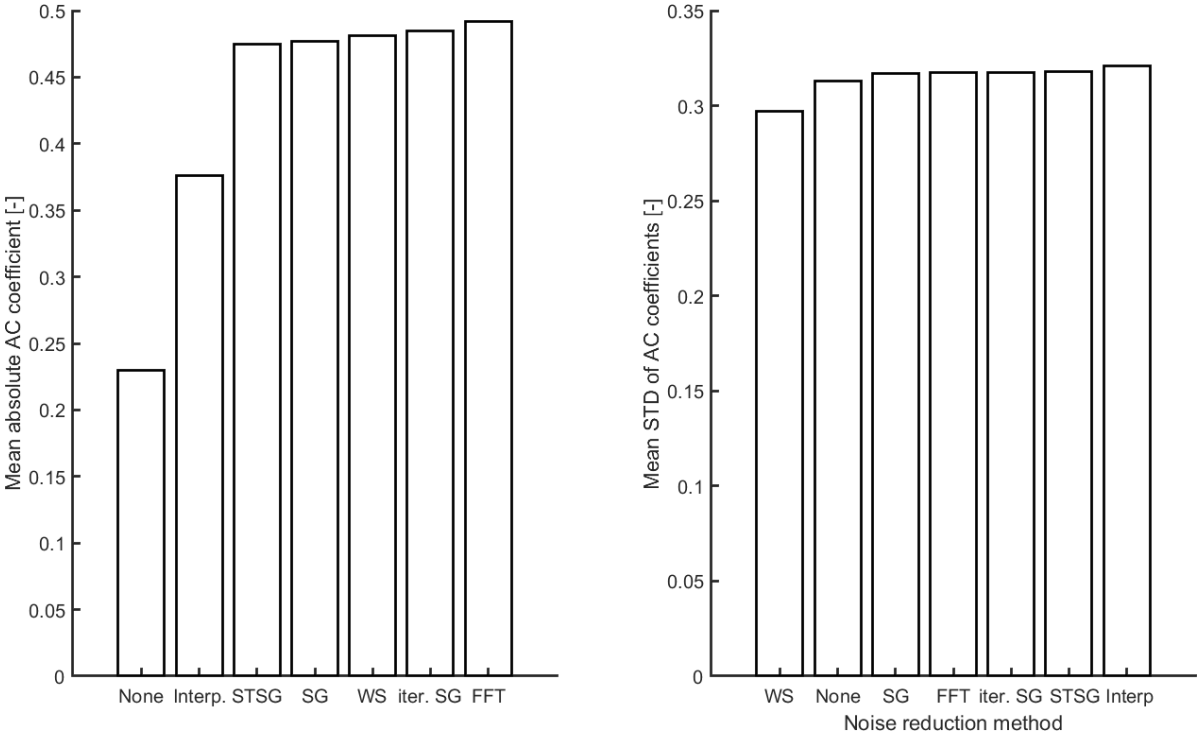


Figure 2.7: Temporal autocorrelation and its statistical significance of the various noise reduction algorithms.

2.6. Iterative Savitzky-Golay Filter Parameter Estimation

The iterative Savitzky-Golay filter has been determined in the previous section to be the most suitable noise reduction algorithm, however its performance depends greatly on the parameter values that have been selected. In this section, it will be explained which values are used for these parameters and why.

The Savitzky-Golay filter is shown in Equation 2.2 and uses two parameters, one for the degree of the used polynomial d, which dictates the used coefficient values C, and another for the single-sided window width m. The resulting filtered time series is Y^* [68]. For its iterative version, the data is passed through one initial filter and then iteratively passed through a second filter which can use a different polynomial degree and smoothing window width. As such, there are four degrees of freedom.

$$Y_j^* = \frac{\sum_{-m}^{m} C_i(d) Y_{j+i}}{2m+1} \tag{2.2}$$

The used window width can be viewed as the frequency of the low-pass filter. To determine the window widths of the first and second filters, the window width that succeeds in passing through useful information, and filters out high-frequency noise is selected. The frequencies that are included are crucial for the end result, however they depend on the application and study area [81]. To analyse vegetation in California, A. Moody and D. M. Johnson considered the first two harmonics, whilst Wen, Su and Ma used the first three when creating a cloud-free time series for the Tibetan Plateau [82, 83]. Geerken, Zaitchik and Evans reported that, despite two harmonics being commonly used, they are not sufficient when classifying rangeland vegetation and instead five harmonics should be used [84]. Finally, to evaluate the performance of HANTS, Zhou, Jia and Menenti used the first four harmonics [81]. As no consistent value is used in literature, the frequency spectra of the used NDVI time series were analysed in order to find a suitable cut-off frequency. As an example, the frequency spectra of an agricultural pixel are given in Figure 2.8. In this image it can be seen that the time series has two clear harmonics, the first one is annual and the second one semi-annual. Past this frequency, it becomes difficult to distinguish a higher order harmonic from noise. As such, and because an odd number of samples must be used, a window width corresponding to 19 samples is used. This means the window spans 190 days, as the samples are 10 days apart. It should be noted that the sensitivity to this parameter quickly diminishes when larger window widths are used.

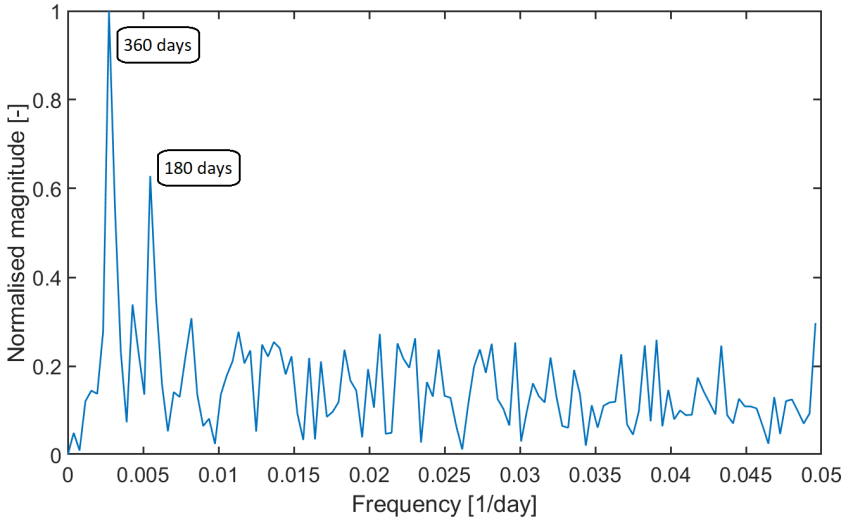


Figure 2.8: Frequency magnitudes of an example time series.

With regards to the polynomial orders, very little difference was found in the quantitative NDVI values. There is however a distinct difference between the obtained time series, as shown by Figure 2.9 in which d_1 is the degree of the polynomial used for the first Savitzky-Golay filter. In particular, there is a difference between a degree of one and higher degrees, which are in turn very similar. The latter time series correspond far more closely to the time series generated by using other noise reduction algorithms, and as such a degree of two is selected. Similar results are obtained for the second filter, and further details are given in Appendix B.

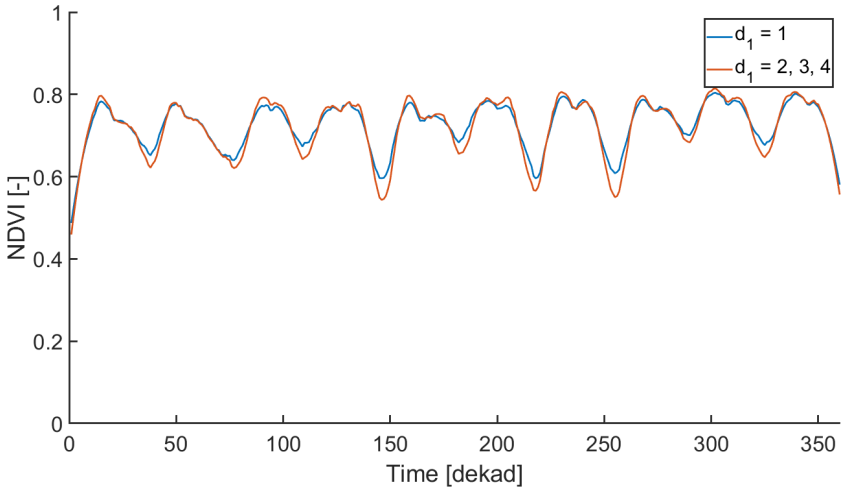


Figure 2.9: An agricultural NDVI time series for different values of the polynomial degree d_1 of the first Savitzky-Golay filter. Filters with d_1 greater than 1 produce time series that are indistinguishable.

To show the effect of the iterative Savitzky-Golay filter, the same time series as was shown before in Figure 2.4, now after having been processed by the iterative Savitzky-Golay filter, is shown in Figure 2.10 together with the raw data and interpolated time series. This figure shows the extent of the noise reduction, as a clear seasonal pattern has emerged without the outliers that were present before.

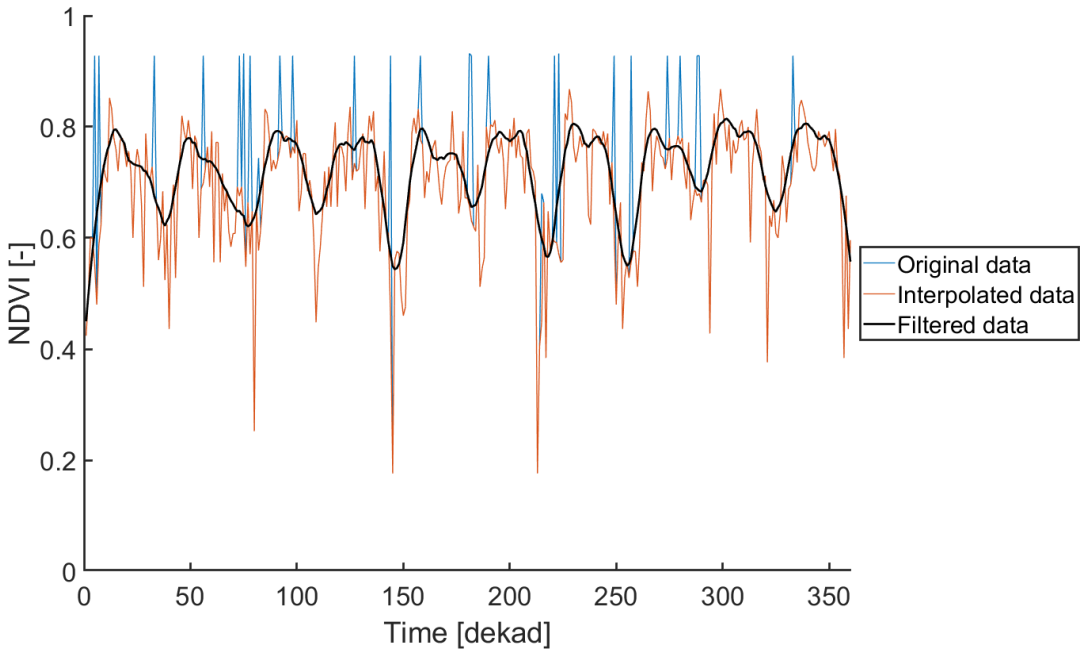


Figure 2.10: The interpolated and filtered agricultural time series, as well as the corresponding raw data and interpolated time series.

2.7. Data Pre-Processing Recap

The chapter is concluded with a brief overview of the data pre-processing approach. A schematic overview of the approach and its relation to the other aspects of the methodology is shown by the flowchart in Figure 2.11.

The raw data provided by Copernicus was not well suited for further analysis. To reduce the noise in the data, two steps have been taken. The first is the temporal interpolation of invalid pixels within the time series, using a symmetrical five-width interpolation window. As noise still exists after this substitution step, the iterative Savitzky-Golay filter devised by Chen et al. [68] is applied to the interpolated data and no clear signs of noise remain.

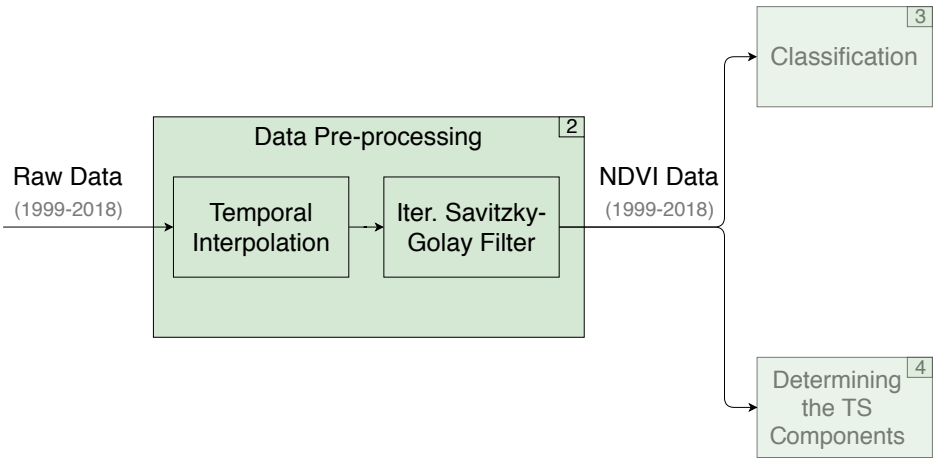


Figure 2.11: Data pre-processing flowchart.

3

Land Cover Classification

Land cover classification is the process of determining which class an NDVI pixel or group of NDVI pixels belongs to, for example urban or agricultural. Knowing which class a pixel belongs to greatly aids data analysis, as results can be analysed for each class individually. This chapter will start with a very brief overview of classification in Section 3.1. The method deemed most suitable is the classification tree. As this is a supervised method, the chapter continues with a description of the used ground truth data to train the classifier in Section 3.2. The types of land cover classes that will be distinguished between, as well as the profile that will be used to represent each class, are the topic of Section 3.3. The classification tree is discussed in detail in Section 3.4, along with methodological additions to make it more suitable for this specific application. The classification of a diverse geographic region is covered in Section 3.5. After this, the solution to inconsistent and discontinuous classification results is presented in Section 3.6, followed by temporal smoothing of the results in Section 3.7.

3.1. Overview of Image Classification

Two types of classification algorithms exist, supervised and unsupervised. Supervised algorithms rely on training data to teach the classifier, which unsupervised algorithms do not require. Unsupervised classification consists of two steps. First, clusters must be generated by grouping pixels based on their similarity, which can be determined by calculating the total Euclidean distance of the distinguishing metrics. The second step is then to classify these clusters, which can happen automatically, but is usually done manually [88].

Supervised classification does not require the data to be clustered and instead consists of the following three steps. First, a training data set is created which should be sufficiently large. Khorram et al. state that the training data should be as homogeneous as possible, whereas Foody and Mathur state that this is not effective. Removing outliers or boundary regions makes the classifier biased towards the purest class members and can therefore miss rarer cases [88, 89]. After this training data set is created, the second step is to develop profiles for each class that can be used to detect class members. In the final step, each pixel is compared to these profiles and it is determined which class the pixel belongs to [90].

Unsupervised methods are generally less accurate and have a higher computational cost. Additionally, one of the main advantages of unsupervised classification is that it can be performed without training data, however this is only the case if the generated classes can be identified manually. For a large data set, this is infeasible [6, 84, 90]. Furthermore, unsupervised methods are not frequently used in remote sensing of vegetation and Gómez, White and Wulder note that large-scale land cover mapping has evolved from unsupervised to supervised methods [91]. Taking these arguments into account, only supervised classification methods were looked into in more detail. Classification trees - also known as decision trees - are a supervised classification method that has several advantages that make it the most suitable choice for the classification of remotely sensed vegetation data. A detailed description of classification trees and their application in remote sensing is given in Section 3.4.

3.2. Ground Truth Data Description

In order to train the supervised classification methodology and assess its effectiveness, data of the true land cover types is required, commonly called the ground truth data. This data is often created manually with site visits or from high resolution aerial photography [66, 88, 90, 92]. An alternative is the use of a reference data set, such as the Corine Land Cover database covering most of Europe [14, 67, 80, 93, 94]. Like the NDVI data set, it is provided by the Copernicus Global Land Service and is openly available [95].

This land cover database was generated automatically using high resolution satellite imagery, with each of the data sets using different satellite sensors and thus data in different spectral ranges. NDVI data is not used, but the near infrared and red spectra that are used to calculate NDVI, are used to generate some of the data sets [96–99]. The results have been verified using expert assessment of aerial photographs, resulting in an over-all classification accuracy of 85% [100].

The Corine database consists of five data sets, four of which are within the time window analysed by this project. These are the years 2000, 2006, 2012 and 2018. They distinguish between 44 classes in total, in three hierarchical levels. On the highest level the classes are distinguished between artificial, agricultural and natural areas, as well as wetlands and water bodies.

The Corine data is at a constant 100 by 100 metre resolution, meaning it is both finer and has a different projection type as the NDVI data. The Corine data uses the elliptical Lambert-Azimuthal equal area projection, which has to be remapped to the Plate Carrée projection of the SPOT-VGT and PROBA-V data which is constant in longitude and latitude. To do this, the longitude and latitude of each pixel was calculated according to the methods described by J. P. Snyder [101]. The land cover type given to each pixel of the remapped land cover data set is determined by taking the mode of the Corine pixels that fall within the longitudinal and latitudinal confines of the vegetation pixel in question. To determine whether the conversion was accurate, coastlines of the remapped land cover classification and the vegetation data were overlaid.

The Corine land cover data for the year 2000 is shown in Figure 3.1, which when compared to Figure 2.1 shows not only that the projection type is different, but also that parts of Europe for which vegetation data is present do not have land cover data as a result of the way the database is created. The data of each country is produced by national groups and not all countries within the domain are involved with the creation of this database. This illustrates the lack of universal land cover data, which together with the fact that this data is produced only once every six years, motivates the use of NDVI data for land cover classification in this thesis project. As this ground truth data is generated every six years, the situation can arise that a six year old data set must be used to train the classifier. The effects of this on the classification results are minor, and shown in Appendix D.

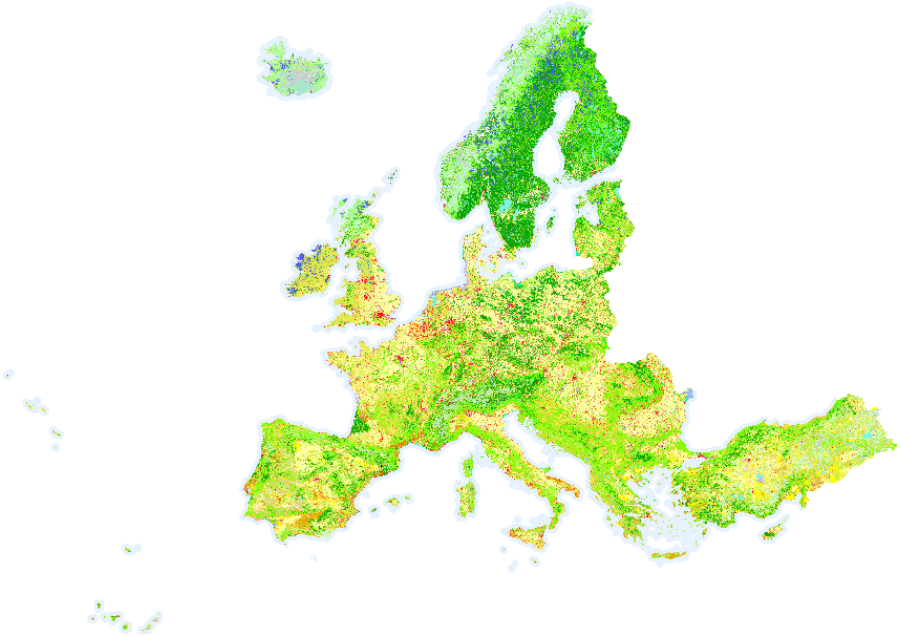


Figure 3.1: Corine land cover data of Europe, for the year 2000.

3.3. Types of Land Cover Classes and Their Classification Profiles

Determining which land cover classes are of interest is a trade-off between fidelity and accuracy. On the one hand, a useful portrayal of the ground cover diversity warrants a large number of classes. On the other hand, this makes the results more complex to analyse and reduces the accuracy of the produced land cover map. The Corine database that is used as ground truth data, distinguishes between 44 land cover classes in total. To have a level of complexity and a number of classes more in line with other classification studies, this number had to be reduced. Fortunately, many of these classes have very few members or are similar either logically, or in terms of their vegetation profiles. As such, they can be grouped together and the number of classes can be drastically reduced. Finding an optimal grouping is a difficult task, as an optimum must be found taking into account both the metrics that describe the class distribution - discussed in Subsection 3.5.1 - as well as the requirement for the created groups to be logically similar. Nevertheless, it is difficult to arrive at a more suitable grouping from the selected one, either logically or in terms of performance. In this list, the names for each group are given, as well as the corresponding Corine land cover codes [102].

- **Artificial:** Continuous urban (111), Discont. urban (112), Industrial and commercial (121), Road and rail networks (122), Ports (123), Airports (124), Mining (131), Dump sites (132), Construction (133), Green urban (141), Sports and leisure (142)
- **Agricultural:** Non-irrigated arable land (211), Permanently irrigated land (212), Rice fields (213), Vineyards (221), Fruit plantations (222), Olive groves (223), Pastures (231), Permanent crops (241), Complex cultivation (242), Agriculture with natural vegetation (243), Forestry (244)
- **Grasslands and forests:** Broad-leaved forest (311), Coniferous forest (312), Mixed forest (313), Grasslands (321), Moors (322), Sclerophyllous vegetation (323), Transitional woodland-shrub (324)
- **Non-vegetated:** Sand (331), Rocks (332), Sparsely vegetated (333), Burnt areas (334), Glaciers and perpetual snow (335)
- **Wetlands:** Inland marshes (411), Peat bogs (412), Salt marshes (421), Salines (422), Intertidal flats (423)
- **Water bodies:** Water courses (511), Water bodies (512), Lagoons (521), Estuaries (522), Sea and ocean (523)

Supervised classification methodologies require data with which they can be trained. In other words, classification profiles must be generated. Using the time series data directly is not efficient and greatly increases the computational cost when finding the optimal decision stage criterion for the classification trees. Instead, each time series is described by the following NDVI and growing season metrics. As annual land cover results are desired, these characteristics are determined on an annual basis. Examples of prior studies using these metrics are the studies performed by Martínez and Gilabert, and Reed et al. [9, 11].

- **NDVI metrics:** The mean, maximum and minimum, as well as the amplitude of the NDVI time series. The time at which the maximum occurs is also used, as well as the slopes at the start and end of the growing season. The latter are termed the rates of greenup and senescence respectively.
- **Growing season metrics:** The start, end and length of the growing season are also defined. Furthermore, the area under the NDVI time series during the growing season can be used as a proxy for the gross primary productivity [37].

This means that in total, eleven different parameters are used to describe each class. The values are normalised between 0 and 1 by dividing by the maximum of each metric. For instance, the end of the growing season is divided by the number of samples per year, in order to arrive at a normalised value. Whilst most of these parameters do not require further explanation, various methods exist to determine the growing season metrics and these values are therefore not self-evident. The method by which these are determined is given in Section 6.2.

3.4. Classification Trees

In Section 3.1 both supervised and unsupervised methods were discussed, with classification trees described as the most suitable classification algorithm. In this section, classification trees will be examined in more detail, starting with a description of the simple classification tree shown in Figure 3.2.

They start out at the 'root', which contains all the data. This data is then divided into 'branches', to create more homogeneous groups. Each division is determined with a decision stage, where the data either meets the given threshold for the specific feature or not. The 'leaves' are located at the top of the tree, and give the pixel its classification type. There have been numerous successful applications of classification trees in remote sensing of vegetation [67, 103–108].

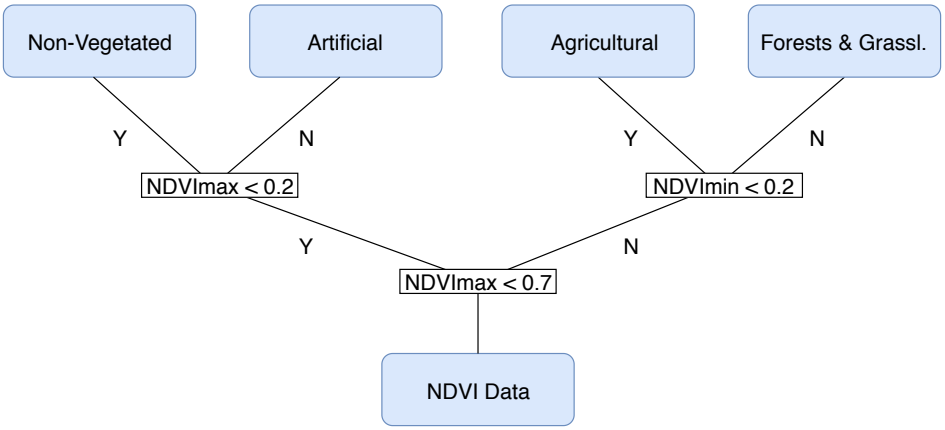


Figure 3.2: Example of a classification tree for vegetation data.

The performance of the tree depends strongly on its design. Furthermore, due to the construction of the tree, any errors accumulate upwards making the early decision stages critical. Designing a tree consists of three core tasks, namely the selection of a node splitting rule, decision rules to determine when nodes become terminal and finally the assignment of classes at the terminal nodes. Assigning classes is relatively straightforward, with the highest probability class being assigned. This is commonly, but not necessarily, the mode of the class samples remaining at the terminal node.

Deciding when a node is sufficiently pure and further splitting can stop, thus making the node terminal, is more complex. A simple solution would be to generate a classification tree until all terminal nodes are pure, however this is infeasible when the data set is large. A more suitable solution is to set a limit to the maximum number of splits, and the most optimal tree is the one with the greatest purity at the terminal nodes.

The decision stages are an optimisation problem where the design variables are the feature and threshold value and the design space is the number of features and zero to one (if normalised), respectively. The optimisation objective is then to maximise the decrease in impurity, or in other words, maximise the separability [109]. This optimisation process favours common classes over ones that are rare, as rarer classes have a small effect on the over-all purity of the data.

This is one of the issues that becomes clear when comparing the ground truth data in Figure 3.3 to the results obtained for the same region when using a single classification tree, given in Figure 3.4. The artificial class, which comprises roughly 1% of the data, is almost entirely ignored by the classifier. The section continues with a more detailed description of rarity. This is followed by two methodological techniques that enable the classifier to classify rare classes.

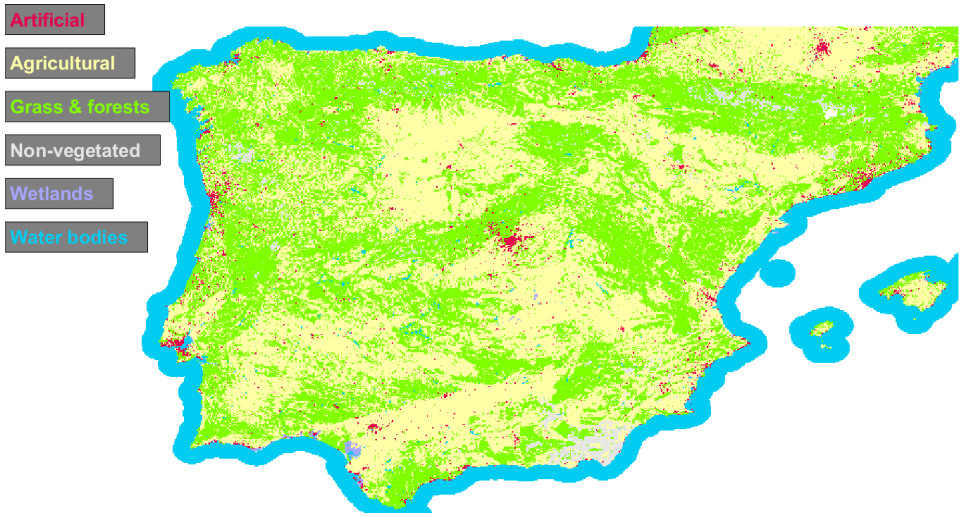


Figure 3.3: Corine land cover data for the Iberian peninsula.

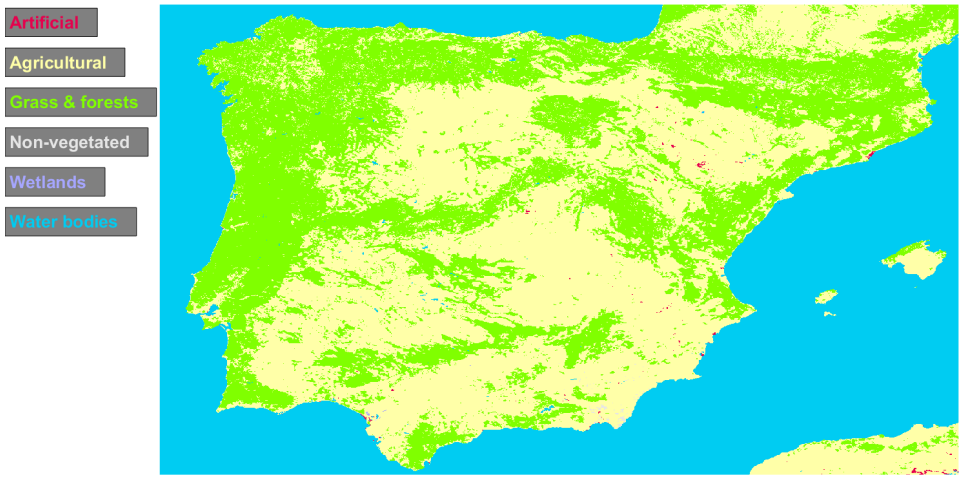


Figure 3.4: Classified land cover with a single classification tree.

3.4.1. Rare Classes

The initial results obtained when using a single classification tree were shown in Figure 3.4 for a subset of the domain. From this figure it is evident that the classifier fails for rare classes. For the European domain, the artificial, wetlands and non-vegetated classes are usually rare compared to the agricultural and grasslands and forests classes.

It should be noted that, whilst rare class is an easy term to use, it is not necessarily accurate. Rare classes are classes that simply have few members, whilst rare cases are rare subsets of the data. For instance, agricultural pixels may be common, but a specific type of agricultural pixel can be rare and distinct from the others. This type of agricultural pixel may then be missed when training the classifier. By their nature, rare classes contain a greater proportion of rare cases. These rare cases are difficult to classify, as representative classification profiles are difficult to generate. Classification trees allow for rare cases to be classified when they are grown to be very big, however rare cases have a small impact on the impurity and as such a limited effect on the creation of the early decision rules. An additional potential issue is that classification trees divide the data into ever-smaller parts. This reduces the issue of relative rarity, however it increases the issue of absolute rarity. In other words, the relative number of pixels may increase but the absolute number decreases and may not be sufficient to find a pattern [110].

In the following subsections, two techniques are presented that greatly improve the classification performance for these rare cases. These are sampling and ensemble techniques, described in Subsection 3.4.2 and Subsection 3.4.3 respectively.

3.4.2. Sampling Technques

The simplest solution to the issue of rarity, is to change the distribution of the classes such that rarity no longer exists. In other words, when all classes have the same number of samples no class can be considered rare anymore.

Three techniques exist to change the distribution of class samples, aptly called sampling techniques. Oversampling creates new samples that are copies of the existing minority class members, which can make the classifier vulnerable to overfitting. Undersampling is the opposite and removes samples of the majority class to create a more balanced distribution. As a result, information is lost [111]. A more complex sampling strategy is to synthetically generate new samples for the minority class. New samples are then generated at a random distance in the feature space between two neighbouring minority class members [112]. One should note that this is only viable if the minority class occupies a distinct region in the feature space, and for this project the classification performance was significantly worsened when using this technique.

A straightforward distribution would be a fully balanced data set, where all classes have the same number of samples. In reality however, the optimal distribution varies strongly with data set and classifier and may be found through iteration [113]. For this project, the fully balanced distribution was found to be the most suitable. Whilst other distributions performed better occasionally, the fully balanced one was the most consistently good. When the training data distribution is different from the actual data distribution the classifier will become biased. To prevent this, the prior probabilities must be changed accordingly [113, 114]. For this purpose, Laplace's law of succession is used, as it lacks the asymptotic behaviour for small sample sizes that the frequency based alternative has. The law is given in Equation 3.1, where p is the new probability estimate, also called the posterior probability, m_p is the number of samples in the minority class and m_n the number of samples in the majority class [113]. As in this project more than two classes are used, all classes aside from the current minority class (before any changes to the distribution are made) together are seen as the majority class. This is to ensure that the total probability sums to one. Finally, the degree of sampling is o. This is the ratio of minority to majority class members in the original distribution, relative to the new training distribution. When oversampling is used, it is greater than one. On the contrary, when undersampling is used, it is between zero and one.

$$p = \frac{m_p + 1}{m_p + o \cdot m_n + 2} \tag{3.1}$$

3.4.3. Ensemble Techniques

Aside from changing the training data, changes can also be made to the classifier itself to improve its performance. One example of this is the application of ensemble techniques. These ensemble techniques work by using the results from multiple classifiers, an ensemble, to arrive at the final result. The idea is that the error of a single classification is then outweighed by the group, for instance through majority voting where the mode of the classification results for a pixel is selected. Due to their quantity, they suffer less from over-fitting and thus enable the use of larger trees [114–116]. This enables the classifier to also consider rarer cases, which can otherwise be ignored.

Bagging is an ensemble technique where a group of classification trees is generated using different subsets of the training data. Random forests closely resemble this technique, with the difference being that in random forests a limited number of features are chosen for each tree making them less correlated. Finally, with boosting a group of weak classifiers - such as decision trees consisting of only one decision stage - are turned into a strong classifier. Unlike random forests and bagging, the full training set is used for each tree. The training samples are initially given equal weights and the weights of falsely classified samples are increased making the classifier focus on difficult samples [115].

To find the most optimal combination of sampling technique and ensemble technique, all options were tested. Of the sampling techniques, synthetic sampling was unable to provide adequate results, likely due to the poor separability within the classes in the feature space - a topic that will be covered in the next section. Similarly, boosting was unable to provide acceptable results for rare classes, either not classifying them at all or vastly overestimating their quantity. As such, only combinations between over- and undersampling and random forests and bagging remained. Random forests and bagging are very similar techniques, with the difference being the number of used features. The resulting classification accuracy of random forests was always worse than those obtained by using bagging, with the accuracy increasing with the number of features used. This leaves bagging as the best performing ensemble technique. Subsection 3.4.2 mentioned that the

main downside of oversampling is that it leads to overfitting. The majority voting of the bagging technique avoids this issue, and the classification results for a combination of bagging and oversampling are the best. As such, these two approaches are chosen for the classification methodology.

3.5. Classifying a Diverse Geographic Region

There are several more aspects which are important to consider when one wants to obtain high quality classification results for a large domain. The first aspect to consider is the suitability of the data to classification, given in Subsection 3.5.1. The other aspect one needs to consider is that a lack of ground truth data prohibits the classifier from having representative training data. The methodology used to address this particular issue is given in Subsection 3.5.2.

3.5.1. Class Separability and Variability

For a classifier to successfully distinguish between classes, the separability between them must be sufficient. To measure inter-class separability, the Jeffries-Matusita distance is used [117, 118]. This distance measures the similarity between two probability distributions, where a value of zero denotes complete overlap and two complete separability.

Another property of the data that is important to consider is the intra-class variability, which is an indication of how diverse the different class members are. Y. Shao et al. calculated the Euclidean distances in the feature space between each pair of data points and stated that the lower the mean Euclidean distance and its standard deviation, the more suitable the data is for classification [85].

Large heterogeneous landscapes are difficult to classify, as they have low inter-class separability and high intra-class variability [115]. This makes it difficult to accurately classify a large geographic area such as Europe, or even just Iberia. The Jeffries-Matusita distances between the various classes in Iberia are shown in Table 3.1. Aside from the water bodies class, the classes are difficult to separate and the values are considerably lower than those found by Y. Shao et al. for the great lakes basin in North America [85].

Table 3.1: Jeffries-Matusita distances between classes in Iberia when no segmenting is applied. The table is symmetric around the diagonal.

	Artificial	**Agri.**	**Grass & Forests**	**Non-vegetated**	**Wetlands**	**Water bodies**
Artificial	-	0.3247	0.2292	0.1102	0.2257	1.9259
Agri.	0.3247	-	0.2757	0.1825	0.3421	1.9683
Grass & Forests	0.2992	0.2757	-	0.2350	0.2874	1.9446
Non-vegetated	0.1102	0.1825	0.2350	-	0.2938	1.9658
Wetlands	0.2257	0.3421	0.2874	0.2938	-	1.4010
Water bodies	1.9259	1.9683	1.9446	1.9658	1.4010	-

To reduce the diversity encountered when training the classifier, only a limited part of the data set is assessed at a time by segmenting it in 60 by 60 pixel sections. The Jeffries-Matusita distances when using segmenting are given in Table 3.2, and the improvement compared to a lack of segmentation is applied in Table 3.3. The resulting values are significantly higher for all but the water bodies class. The values found for the water bodies class are lower, as it is impossible to calculate these distances in sections where no classes but the water bodies class exists, which are common. Therefore, this distance is only computed for coasts and inland lakes, which naturally show greater similarity to their surroundings due to the relatively large pixel sizes. The average Jeffries-Matusita distance of the land cover classes increased by 0.59, not taking the water bodies class into account due to the aforementioned computational issues.

The results for the class variability are given in Appendix F. The variability sees considerable improvement after segmenting has been applied for all classes, including the water bodies class, and an average reduction in Euclidean distance of 40.0% is found.

Whilst smaller section sizes showed improvements in both separability as well as variability, this resulted in obvious classification bias and sharp edges in land cover between sections, and a minimum allowed section size cannot be found by solely looking at these values. This is a topic that is discussed in more detail in Section 3.6.

Table 3.2: Jeffries-Matusita distances between classes in Iberia when segmenting is applied. The table is symmetric around the diagonal.

	Artificial	**Agri.**	**Grass & Forests**	**Non-vegetated**	**Wetlands**	**Water bodies**
Artificial	-	0.6895	0.8424	0.8055	0.6316	1.8987
Agri.	0.6895	-	0.6560	0.9072	0.9492	1.8737
Grass & Forests	0.8424	0.6560	-	0.7035	1.1934	1.8782
Non-vegetated	0.8054	0.9072	0.7035	-	1.0252	1.9304
Wetlands	0.6316	0.9492	1.1934	1.0252	-	1.5402
Water bodies	1.8987	1.8737	1.8782	1.9304	1.5402	-

Table 3.3: Improvement in Jeffries-Matusita distance when segmenting is applied. The table is symmetric around the diagonal.

	Artificial	**Agri.**	**Grass & Forests**	**Non-vegetated**	**Wetlands**	**Water bodies**
Artificial	-	0.3648	0.6132	0.6953	0.4059	-0.0272
Agri.	0.3649	-	0.3803	0.7247	0.6071	-0.0946
Grass & Forests	0.5432	0.3803	-	0.4685	0.9060	-0.0664
Non-vegetated	0.6952	0.7247	0.4685	-	0.7314	-0.0354
Wetlands	0.4059	0.6071	0.906	0.7314	-	0.1392
Water bodies	-0.0272	-0.0946	-0.0664	-0.0354	0.1392	-

3.5.2. Classifying Regions without Ground Truth Data

Aside from issues with diversity, further difficulties arise due to the ground truth data not covering the entirety of the NDVI data set. This can be seen most clearly in Figure 3.1, which shows the entire European data set, but one can also see that the North African part of Figure 3.3 is missing, yet classified by the classifier. To circumvent this issue, it is determined which sections of the data set contain an unacceptably large percentage of pixels that cannot be used for classifying land cover. Neither water bodies pixels nor missing pixels can be used to classify land cover, and if these pixels cover more than 50% of the section the ground truth data is seen as invalid. The exception to this rule is if the section contains solely sea pixels, in which case it can be classified normally.

In order to train the classifier for these sections, the closest valid region is used to provide the training data. In the case of this North African region, the south-eastern coast of Spain is used. It is important to note that neither the classification results for these invalid regions, nor for regions containing only sea pixels are used when calculating the classifier's performance. For the former, insufficient ground truth data is present to make a reliable estimate. For the latter, the results would be a hundred percent accurate regardless of the used classifier, obscuring any improvements in classification accuracy of other water bodies pixels. This is also the reason why the water bodies classification accuracy given by the confusion matrices is reduced, when segmenting is applied.

Appendix E presents results for a 900 by 900 pixels region. Comparisons are made with the original data set as well as when the eastern half is missing. As the classification accuracy is still acceptable even when such a large region is missing, this method can be considered successful.

3.6. Inconsistent and Discontinuous Classification Results

Consistent classification results are particularly important when the temporal change in land cover is of interest, as inconsistent results may present change that is in fact not present. The inconsistency of a single classification tree is readily shown, as the training data is randomly selected and the aforementioned sampling strategies introduce another degree of randomness. Using the ensemble techniques described in Section 3.4, the variance can be reduced as the results are now a function of a multitude of classification trees [119]. It should however be noted that the consistency of the ensemble is delicate, and that the variance may not approach zero [119, 120]. To the same end, Hastie, Tibshirani and Friedman state that there is a limit to the drop in variance that can be achieved using bagging, that depends on the correlation between the used trees [121]. They further state that random forests should have a lower variance than bagging due to the lower correlation between the assembled trees, however this result was not observed by the author in this domain.

Instead, more inconsistent results were obtained, further enforcing the decision to use the bagging technique over random forests.

To increase the consistency, a second layer of majority voting was introduced. Multiple iterations of the classification procedure described in this chapter are performed, with the most common classification for each pixel returned as the final result. The similarity between successive levels of majority voting for Iberia are shown in Figure 3.5, from which it can be observed that the similarity rapidly approaches the asymptote of 100%.

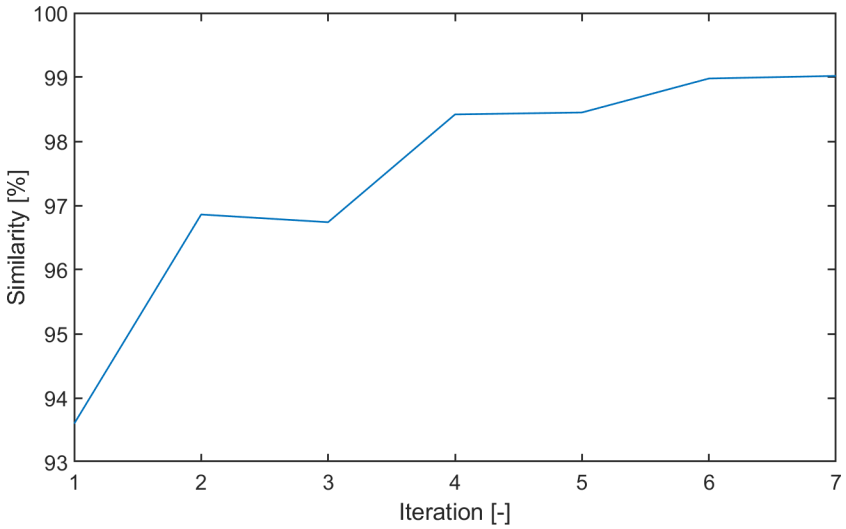


Figure 3.5: Similarity obtained between successive levels of majority voting for Iberia.

The previous section stated that discontinuous classification results are obtained when the data is segmented, as each segment is trained on entirely separate data. To solve this issue, as well as introduce the aforementioned second layer of majority voting, segmenting is applied in a sweeping manner horizontally as well as vertically. This is also one of the critical differences between the segmenting used for this research project and the segmenting (called tiling by the authors) used by J. Inglada et al. [122]. The other key difference is the use of segments specifically to create a less variable and more separable data set, whilst the aforementioned study includes combined use of a randomly selected group of tiles that are not otherwise related to one another, negatively affecting both of these metrics. As shown in Figure 3.6, segments have a degree of overlap. This smoothes out resulting edges between sections, and as these segments are layered on top of one another this inherently introduces a degree of iteration. An overlap of 80% in both dimensions is used, to provide a conservative ten iterations.

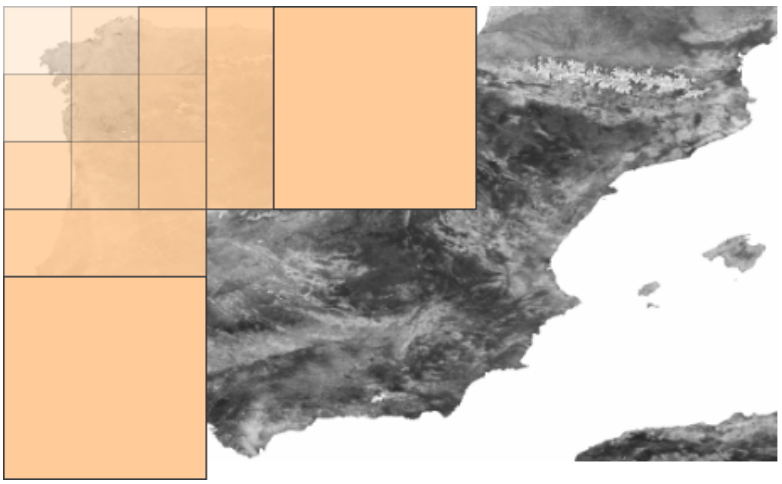


Figure 3.6: Segments are defined in a sweeping manner, with a degree of overlap both horizontally as well as vertically. The segment sizes are exaggerated for clarity.

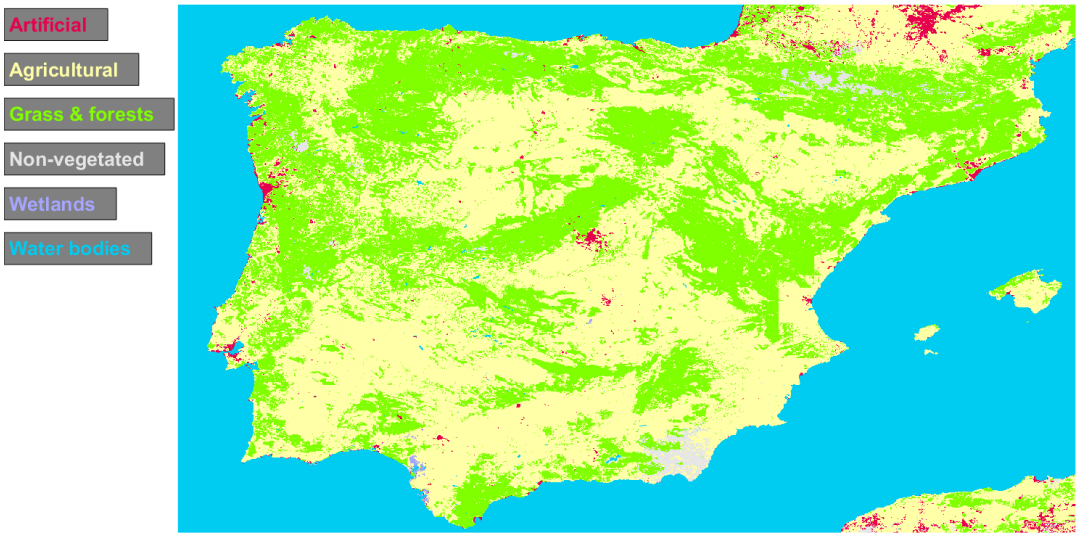
3.7. Temporally Smoothing the Classification Results

So far, this chapter has only discussed the results obtained for a single year, however annual land cover results are generated. As described in Section 3.2, the Corine land cover data used to train the classifier is only available for four years within the assessed time window, with six years in between. To obtain classification results for any of the other years, the classification trees generated using the nearest ground truth database are used. As the feature values evolve over the years, the classification algorithm can produce anomalous results.

By temporally smoothing the classification results, these misclassifications are removed almost entirely. A classification is seen as erroneous if the classification type does not cover at least three successive years. To find these anomalous values, a symmetrical substitution window of single-sided width two is used, where the outliers are substituted by the majority class within the respective window. The class that is the majority within the entire twenty year window is never substituted for. If after this procedure any outliers remain, they are substituted for by the majority class within the twenty year time span. These smoothing parameters were selected as smaller parameter values were insufficient for removing some of the most obvious noise seen in the data, whilst a greater degree of smoothing would reduce the temporal change in land cover beyond what would be expected of the data.

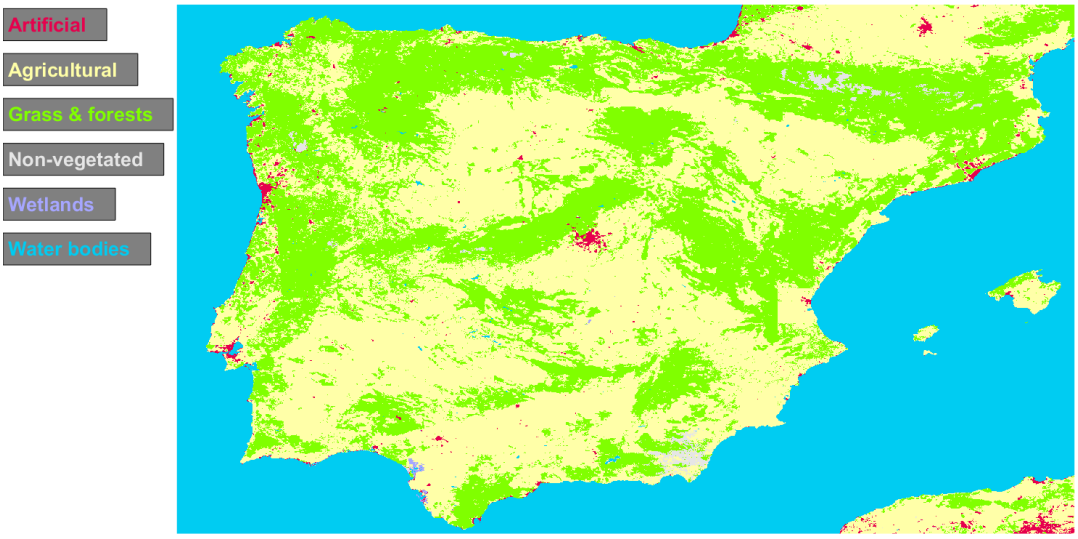
It should be noted that only some of the years will contain noise. Whilst results for the year 2000 were shown before, as ground truth data is available for that year, these results showed no significant amount of noise. As such, to present the improvement classified data from the year 2003 is used, as this year contained a large amount of falsely classified data. The results of this smoothing are shown in Figure 3.7, where the original data (a) greatly overestimated the urban land cover in southern France. After temporal smoothing is applied, a more realistic classification result is obtained and (b) acts an example of the land cover results obtained after all the steps in the classification approach have been completed.

From this figure it is apparent that the classification quality has improved considerably when compared to the initial classification data. It is also evident from the land cover map that the classifier manages to capture the wetlands in southern Spain, major urban areas and the non-vegetated areas in the Pyrenees and south-eastern Spain. For the final confusion matrix, as well as land cover maps and confusion matrices of results obtained when applying only part of the aforementioned methodology, the reader is referred to Appendix C.



(a) Original classified land cover

Figure 3.7: Comparison between smoothed and original classified land cover for the year 2003.



(b) Smoothed classified land cover

3.8. Classification Recap

This chapter is concluded with a brief overview, aided by the flowchart in Figure 3.9. In Section 3.4, classification trees were outlined and the results for a single classification tree are given. From these results it became clear that they were less than satisfactory, and improvements had to be made.

As the data shows poor separability between the classes, and high variability within each class, the data is segmented into smaller sections. To reduce variance, these segments partially overlap and a inherent degree of iteration is introduced. The use of oversampling (increasing the number of samples of rare classes) and the ensemble technique bagging (classifying each section with a multitude of classification trees) were found to be the most optimal. Due to the bagging approach, as well as the overlapping segments, multiple classification results are obtained for each pixel. By applying a majority vote, a single classification type is obtained for each pixel. Finally, misclassifications are occasionally introduced due to disparities between the training data and feature data of other years. To solve this issue, temporal smoothing is applied.

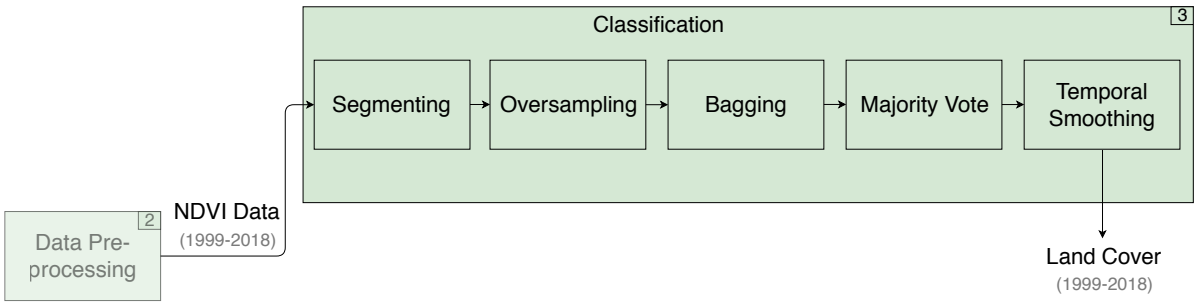


Figure 3.9: Classification flowchart.

4

Determining the Vegetation Trend and Seasonal Component

In this chapter, the methodology used to determine the trend and seasonal components of the NDVI time series is given, such that these can be forecast and analysed independently. First, the decomposition of a time series into its components is discussed in Section 4.1. This is followed by a discussion on how break points are detected in the trend component in Section 4.2. As the quality of the generated trend components and effectiveness of the break point detection algorithm greatly affect the forecasting accuracy, a statement to the sustainability of the generated trend components is provided in Section 4.3.

4.1. Time Series Decomposition

Vegetation time series contain information on different time scales, which can make it challenging to detect trends on the time scale of interest. Examples of these time scales are the annual time scale corresponding to the seasonal cycle of vegetation, or a multi-annual time scale such that long-term fluctuations in the vegetation activity can be determined [11]. A time series can be seen as a sum of time series at different time scales and obtaining these components is termed decomposition. A logical choice would be to use a Fourier or Wavelet transform, which has been applied previously to NDVI data from a monthly to a roughly three-year period [11]. Another solution is shown in Equation 4.1. Here, the time series Y is shown as a sum of its seasonal component S, trend component T and remainder component R. An example of this is shown in Figure 4.1. This removes the necessity to specify the time scales of interest and cleanly splits the seasonal variation, long-term trend and noise in the data set.

$$Y = T + S + R \tag{4.1}$$

De Jong et al. used a first degree polynomial to form the trend component, whilst Weatherhead, Stevermer and Schwartz used a time series of monthly means. Additionally, the seasonal component consisted of a sum of sines and a sum of sines and cosines respectively [47, 123]. A trend component has often been found by applying linear regression on the data set, however this method is rarely suitable. Linear regression requires the data to be uncorrelated and residuals to be random with zero mean and constant variance, which is generally not the case when applied to satellite image time series [124]. Instead, this project uses the statistically suitable and commonly applied method developed by Cleveland et al. in 1990, termed STL. STL stands for Seasonal and Trend decomposition using loess, which in turn is shorthand for locally weighted regression and scatterplot smoothing [125]. The loess filter applies polynomials generated in a window around each point to smooth the data.

The iterative approach of STL is shown in Figure 4.2. STL consists of a sequence of smoothing operations within both an inner and an outer loop. In the inner loop, the seasonal and trend components are determined through use of a low-pass filter as well as the loess filter on detrended and deseasonalised data, which means that either the trend or the seasonal component is subtracted from the original time series. The outer loop then uses these components to determine the remainder, which is used to reduce the influence of outliers. These weights, combined with the previous trend component estimation, are used in the next iteration of the inner loop. For the first iteration, the trend component and weights are set to zero.

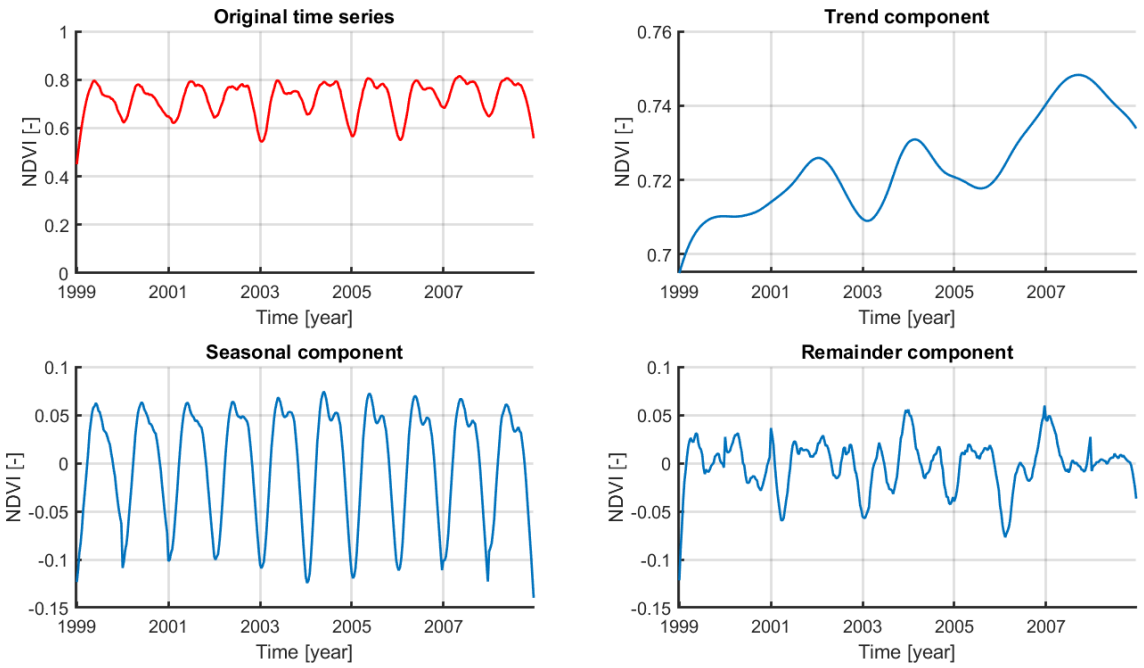


Figure 4.1: A decomposed agricultural NDVI time series.

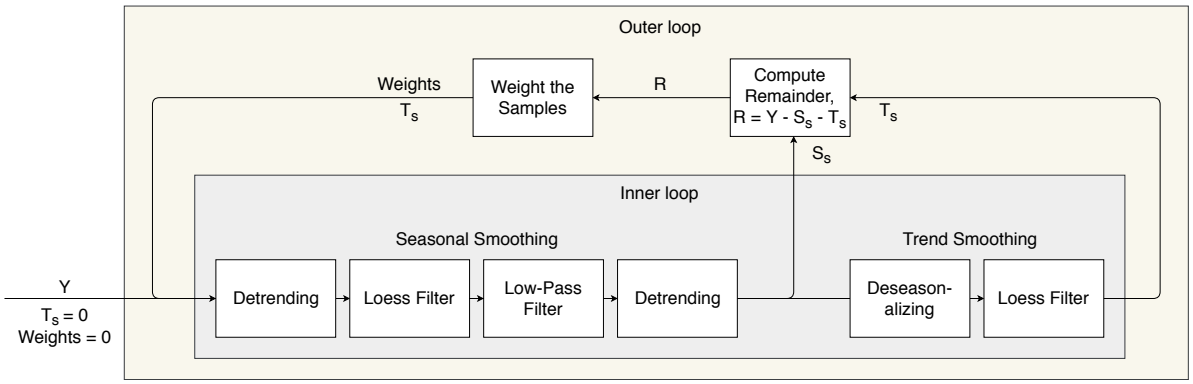


Figure 4.2: STL decomposition flowchart.

The time series are decomposed independently, making the process spatially independent. The filtered NDVI data was shown to be highly spatially correlated in Subsection 2.5.2, meaning that the resulting time series components should also be spatially correlated. In other words, if the input data is spatially correlated and the output data is spatially correlated as well without any algorithmic causality, that greatly increases the confidence held in the generated time series components.

The Moran's I values and the statistical significance in the form of the z-score are used to quantify the spatial autocorrelation, and are given in Table 4.1 for the trend and seasonal component of the same domain as was used previously in Subsection 2.5.2. Whilst the values are lower than those found for the filtered data that is used as input, they still indicate a high level of spatial autocorrelation.

Table 4.1: Spatial autocorrelation and its statistical significance for the decomposed time series.

	Mean Moran's I value [-]	**Mean z-score** [-]
Trend component	0.547	140.63
Seasonal component	0.659	169.33

4.2. Break Point Detection

Whilst the seasonal component in Figure 4.1 shows consistent behaviour throughout the observed time span, the trend component shows several turning points. To determine whether or not the trend component should be analysed as a whole, or whether it should be divided into partitions, break points are detected. These break points then indicate the boundaries between partitions. A break point is a sudden change in the time series, for instance a sudden positive or negative shift or a sharp change in trend direction from positive to negative or vice versa. When a trend is determined over a time series without taking its break points into account, the resulting trend may be false and forecasting results unreliable [126].

Break points can be caused by natural phenomena such as wildfires, floods and meteorological conditions like snow-cover. More gradual changes to the climate can also result in break points, such as a relatively wet period in a dry climate or a warm period at high latitudes [47].

Another potential cause can be issues with the data itself. Three different sensors were used to gather the NDVI data during the studied time frame, SPOT-VGT 1 and 2 as well as PROBA-V. Sudden shifts were detected in this NDVI data at the times a switch was made between different satellite systems [127, 128]. The data has been reprocessed in 2016 and again in 2017 to reduce these discontinuities and it is well within the Global Climate Observing Systems (GCOS) requirements, however minor differences may still be present [56, 57].

This section will continue by discussing how break points can be found for the trend component, as the used method precludes the use for a time series containing a seasonal component.

The presence of a break point can be statistically determined using a structural change test based on moving sums, shortened to MOSUM [39, 47, 129, 130]. At each point that is sufficiently far from the time series' start and end, ordinary least squares regression is applied on the data within a specified window width. Near a break point, the resulting sum of recursive residuals decreases rapidly. The MOSUM test value is independent of the mean value of the time series, as well as the absolute magnitude of the fluctuations therein [130]. The test is however sensitive to the aforementioned window width. If it is too small, the test may fail to detect persistent change and if it is too large, the test is insensitive and break points may be missed [130, 131]. A window width of 10% is used, or 2 years for a full time series. This is similar to the values used by Tian et al. [127]. The critical values from which the statistical significance value is obtained, is calculated using the equations determined by Kuan and Chen at a significance level of $\alpha = 0.05$ [132].

A similar idea was introduced by Bai and Perron in 2003 to determine the dates at which break points occur [133]. An upper-triangular matrix of dimension $n \cdot (n-1)$ is created, where n is the number of time stamps. The matrix contains the sum of squared residuals (SSR) when a linear model $f(t)$ is fitted to the trend component T data samples i through j, as shown by Equation 4.2.

$$SSR = \sum_{t=i}^{j} (T_t - f(t))^2 \tag{4.2}$$

As the computational cost rapidly increases with the length of the time series, the residuals can instead be calculated iteratively. After calculating the diagonal, entries in each row are calculated by adding to the previous value according to Lemma 2 as determined by Brown, Durbin and Evans [134]. Now that this matrix has been created, the optimal partition is found by minimising the total sums of squared residuals before and after said partition. After each break point is detected, the residual sums of squares affected by this break point are re-evaluated, making this method a partial structural change model.

Additional break points may not be too close to previously determined ones. Verbesselt et al. used a minimum of a single year between possible break points [129], Tian et al. used a minimum of two years [127] and Forkel et al. used a minimum distance of four years [39]. As no value is used consistently, the minimum time required for a trend to be statistically significant may be used instead. Weatherhead et al. stated that finding a trend relies on detecting it over the background noise and thus to calculate the minimum time span for a time series to be significant one must take the expected trend magnitude, variability and autocorrelation of the time series into account [123, 135]. When applying their analysis to the NDVI time series, times ranging between 9 and 15 years are found. As these lengths of time are far greater than those used by other studies and would prohibit the use of break points, instead a time of two years is used. Similarly, the maximum number of break points is not consistent. In fact, Tian et al. found different optima for the various land cover types, ranging between zero and five break points. As such, the average value of two is used, similar to the study by Forkel et al. [39].

The implemented algorithm presents the issue that whilst the sums of squared residuals are updated, any detected break point only takes the break points detected thus far into account. In other words, the first break point is determined such that it creates the optimal partition if only two partitions are allowed to be present. This phenomenon is illustrated in Figure 4.3, which shows the detected break points of a synthetically generated time series as vertical black lines. This time series has four clear break points, however only the fourth break point is detected correctly and the first break point is placed in the centre due to symmetry, despite a lack of proximity to an actual break point.

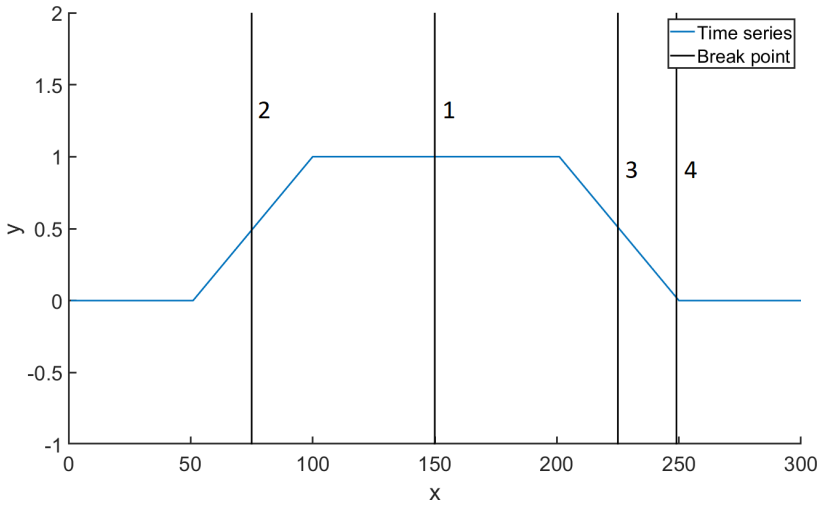


Figure 4.3: A synthetically generated time series, showing the detected break points as black vertical lines. The numbers indicate the order in which they are detected.

To ensure that all detected break points correspond to actual break points within the time series, all other break points must be taken into account when determining each break point. To ensure that only correct break points are detected, the partial structural change model is made iterative. First, the break points are determined according to the aforementioned model. Then, one detected break point is left out and a new break point is determined, taking the remaining break points into account. A break point is seen as valid, if it is found again after having been left out. When all break points are valid, the scheme has converged and iteration stops. The computational cost is increased by using this iterative scheme, but convergence is normally obtained within five iterations and thus the increase in computational cost remains acceptable. Figure 4.4 shows that the algorithm manages to converge to the correct break points even with the introduction of white noise.

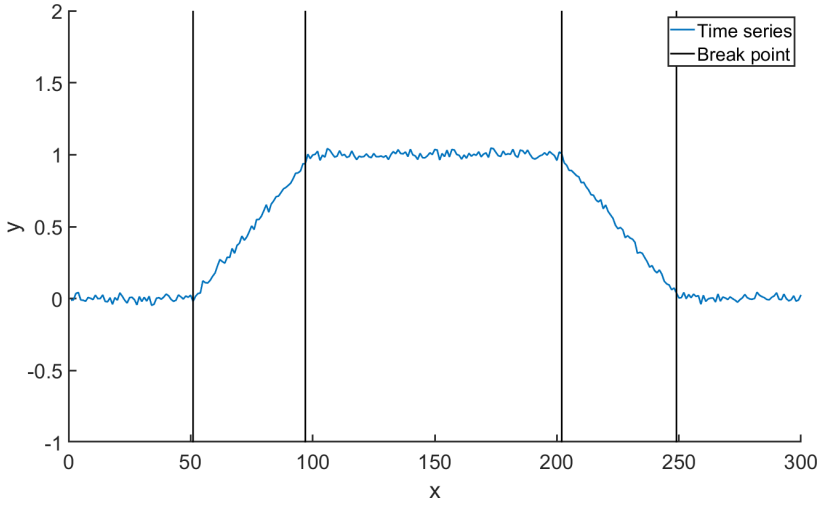


Figure 4.4: A synthetically generated time series, showing the detected break points as black vertical lines. The break points were detected iteratively.

To compare the performance of this iterative algorithm compared to the original, the sums of squared residuals (SSR) were computed after the time series were allowed to have a maximum of two break points, corresponding to the value set before. The mean value obtained over 8000 vegetation time series and the standard deviation are shown in Table 4.2, as well as the resulting improvement.

Table 4.2: Mean squared sums of residuals (SSR) and its standard deviation for the non-iterative and iterative break point detection models.

	Mean SSR [-]	**Standard deviation [-]**
Non-iterative model	0.1102	0.0776
Iterative model	0.1022	0.0713
Improvement	7.26%	8.12%

Once the sums of squared residuals corresponding to each break point have been determined, information criteria can be used to determine the correct number of break points within the time series [129, 133]. These information criteria function by penalising the likelihood, a measure of the suitability of a model to the data it aims to describe, for the model's complexity. The idea is that a highly complex model may fit the data well, however is not justifiable given the size of the sample used to produce it [136].

The Akaike information criterion (AIC) has been applied by Peng et al. [133], however it has been determined to have poor performance and overestimates the number of breaks [41, 137]. Instead, the Bayesian information criterion (BIC) is a better choice when the data is not serially correlated [133] and has been applied to vegetation data before [39, 127, 129]. The requirement against serial correlation means that it is suitable to use for the trend component, but not for a time series still containing its seasonal component. A choice is made for the number of break points by calculating the BIC value corresponding to a model split at each break point. As the BIC is partially constant, the absolute rather than relative change is used [136]. When the decrease in said value is greater than six, the change is considered strong [138] and the break point is seen as legitimate.

4.3. Trend Sustainability

The sustainability of a time series is indicative of its future behaviour. If the time series is sustainable, the existing trend is expected to continue as it is. If the trend is unsustainable, an increase is likely followed by a decrease and vice versa. A cyclical time series is therefore an example of an unsustainable time series, whilst the function $y = x$ is fully sustainable as its slope is constant and thus the trend will never change. The last remaining option is that the time series is stochastic and any future trend would be unrelated to the current trend, prohibiting forecasting of these time series.

A sustainable time series is thus needed for accurate forecasting results, the topic of the next chapter. To measure this sustainability, the Hurst exponent H is used. It is a measure of the long-term memory of a time series. H takes values between zero and one, and is divided into the following three ranges [40]:

$$\begin{aligned} 0 \leq H < 0.5 &\quad : \text{Unsustainable} \\ H = 0.5 &\quad : \text{Stochastic} \\ 0.5 > H \geq 1 &\quad : \text{Sustainable} \end{aligned}$$

4.4. Determining the Time Series Components Recap

In order to analyse the vegetation trend and seasonal activity independently, the time series are divided into a trend and seasonal component. This chapter is concluded with a brief overview of the methodology used to calculate these time series components, which is also given in the form of a flowchart in Figure 4.5.

The trend and seasonal components are obtained by applying STL decomposition, which consists of a sequence of filters and weighting to reduce the influence of outliers. Whilst each time series is decomposed independently, the resulting time series are shown to be spatially correlated. As the trend component is inconsistent throughout the time span, it is divided into partitions if more than one partition is present according to the OLS-MOSUM test. To detect the break points that limit these partitions, an iterative version of the partial structural change model described by Bai and Perron is used [133].

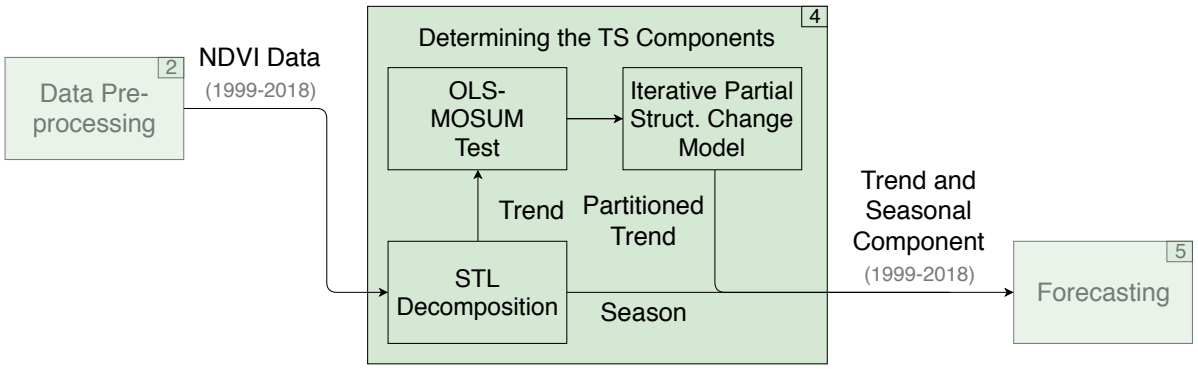


Figure 4.5: Time series component determination flowchart.

5
Forecasting

The previous chapter showed how the time series are decomposed into their trend and seasonal components, and discussed time series sustainability, a requirement for forecasting. This chapter investigates the continuation of the time series themselves beyond the time span for which it is available. The chapter will start with a brief section on prior studies on forecasting of vegetation data in Section 5.1. This is then followed by Section 5.2, which discusses the two most widely used forecasting approaches, ARMA and exponential smoothing models. The chapter continues with Section 5.3, which discusses the specific implementation of these models for the vegetation forecasting performed in this project. To determine which approach works optimally for this use-case, the methods are compared in Section 5.4.

5.1. Vegetation Forecasting

Forecasting of remotely sensed vegetation data has been used in literature before to predict crop yields, for example by Bolton and Friedl and Genovese et al. in the United States and Spain respectively [94, 139]. Both studies create a time series of the crop yield of each growing season using NDVI data. Linear regression is then applied to this time series, and extrapolated to predict the crop yield for the following growing season.

Other studies forecast the vegetation index itself, such as the study by Ji and Peters who applied an autoregressive distributed lag model using temperature and precipitation of Nebraska in the United States [140]. Autoregression indicates that the current time series value y is dependent on previous ones, i.e. $y_t = a \cdot y_{t-1} + b \cdot y_{t-2}$. Distributed lag means that the model uses explanatory variables x at the current time, as well as previous time steps, i.e. $y = a \cdot x_t + b \cdot x_{t-1} + c \cdot x_{t-2}$. A similar study was performed by Philippon et al., who used atmospheric dynamics data as well as NDVI data to forecast future NDVI values [141].

Neither model can be used for this thesis project, as both rely on climate data that is outside the project scope. The author was unable to find a study that forecast a vegetation index using only prior vegetation index data, and limited information is gained from these aforementioned studies. Instead, the chapter will continue with a section on the most commonly used general forecasting approaches and the selection of the most suitable one.

5.2. Forecasting Methods

In this section the two most common approaches to time series forecasting, exponential smoothing and autoregressive moving average (ARMA) models, are discussed [142]. Exponential smoothing methods use weighted averages of past observations, the simplest method of which is aptly called simple exponential smoothing. It uses a single variable that determines the weights for all past values, resulting in a declining influence with time. It is meant for forecasting time series without a clear trend or seasonal component, making it unsuitable for vegetation data. An exponential smoothing method that can be used for data containing a trend is Holt's linear trend method, which splits the forecasting equation into a level and trend component. The level acts as the constant, whilst the trend components is the slope [142]. As neither value changes regardless of how many samples are being forecast, the method is incapable of predicting anything more complex than a linear trend. An extended form of the method is called the Holt-Winters' seasonal method, given in Equation 5.1. Aside from the level l and trend component r, it also includes a seasonal component s, with coefficients α, β and γ respectively. In this system of equations, h is the number of steps ahead of the

given time series length, n, and p_s is the period of the seasonal cycle. Finally, k is the integer of $(h-1)/m$. Similarly to the level and trend component, the seasonal component is also incapable of changing with time.

$$\begin{aligned} Y_{n+h} &= l_n + h \cdot r_n + s_{n+h-(k+1)p_s} \\ l_n &= \alpha(Y_n - s_{n-p_s}) + (1-\alpha)(l_{n-1} + r_{n-1}) \\ r_n &= \beta(l_n - l_{n-1}) + (1-\beta) r_{n-1} \\ s_n &= \gamma(Y_n - l_{n-1} - r_{n-1}) + (1-\gamma) s_{n-p_s} \end{aligned} \tag{5.1}$$

An approach that is capable of predicting more complex behaviour is the ARMA method. The autoregressive part consists of a model using p previous time series values. The moving average part is similar, and uses q errors terms instead [142]. The autoregression order p can be found by determining when the autocorrelation coefficient drops below a specific threshold. The last lag above said threshold is an estimate of p. Similarly, the moving average order q can be determined by plotting the autocorrelation coefficients of the time series after a moving average process has been applied. If the used order exceeds the optimal one, the autocorrelation coefficients are distributed normally with a zero-mean [143]. A more systematic way of determining the coefficients is by using maximum likelihood estimation.

When fitting a high order model and using it for forecasting, errors will arise not only due to the white noise variance of the fitted model but also errors arising from parameter estimation, which are higher for higher order models. Using the Akaike information criterion (AIC) or Bayesian information criterion (BIC) helps prevent this as they include a penalty function for the used number of parameters [143]. Both R. J. Hyndman as well as Brockwell and Davis recommend the use of the bias-corrected Akaike information Criterion (AIC_c), which includes an addition that makes it more suitable for smaller sample sizes [142, 143].

ARMA models require stationarity, which means that the time series does not depend on the time at which it is observed and thus cannot contain a trend or seasonality. Stationary of a time series can be found by examining the autocorrelation coefficients. A non-stationary time series shows a gradual drop in the coefficients, whilst a stationary time series shows low values resembling white noise. A time series can be made stationary by differencing. The resulting time series consists of the difference between each consecutive pair of points. An ARMA model that applies differencing on the time series is called ARIMA, the I standing for Integrated. The ARIMA model is given by Equation 5.2, where Y^D denotes the time series differenced by degree D, ϕ are the coefficients of the autoregression polynomial and θ are the coefficients of the moving average polynomial, multiplied by the error terms ϵ. Finally, the constant c is added.

$$Y_t^D = \phi_1 Y_{t-1}^D + \ldots + \phi_p Y_{t-p}^D + \theta_1 \epsilon_{t-1} + \ldots + \theta_q \epsilon_{t-q} + \epsilon_t + c \tag{5.2}$$

The degree of differencing D cannot be determined by minimising the AIC_c as it changes the data on which it is computed, hence two values for different D cannot be compared. One way of determining the degree of differencing is then to apply differencing until the autocorrelation coefficients meet the aforementioned requirements [142]. A more systematic way is by finding the optimal p, q using the AIC_c for a range of D and comparing the resulting forecasting error measures to find the optimal combination. ARIMA models can also be made to include a seasonal component (SARIMA). This adds another order for the seasonal autoregressive and moving average functions as well as seasonal degree of differencing, which are determined similarly to the non-seasonal parameters [143].

5.3. Implementation of the Forecasting Models

In this section the implementation of the aforementioned models for this project is laid out, to explain any changes that are made to the way the models are used. As suggested by R. J. Hyndman, the trend and seasonal components obtained after applying STL decomposition, as described in Section 4.1, are forecast independently [142]. Therefore, whilst SARIMA and Holt-Winters' seasonal method can forecast time series containing both a seasonal as well as trend component, they are only used to forecast the seasonal component. In the case of SARIMA, this significantly reduces the computational cost of finding the optimal parameter values as the number of unknowns that need to be determined is reduced. When using the ARIMA functionality in MATLAB, the number of variables is further reduced to two, as the seasonal degree of differencing is a constant set to one when a SARIMA model is specified.

Anomalous trend and seasonal component forecasts were sometimes generated by the ARIMA and SARIMA models. To improve the suitability of the data to this model, the Box-Cox transformation is applied to increase the normality of the data and equalise variance. The transformation of a time series y into the transformed time series x is given by Equation 5.3, where λ is determined by maximising the log-likelihood function [144, 145].

$$x(\lambda) = \begin{cases} \frac{y^{\lambda}-1}{\lambda} & \text{if } \lambda \neq 0 \\ log(y) & \text{if } \lambda = 0 \end{cases} \tag{5.3}$$

Further, to forecast the trend component only the data after the final break point is used, as long as more than the number of forecast samples remains. As an example, if the forecasting window is two years, a minimum of two years of data is used to forecast the trend component. If the break point in question does not meet this criteria, the break point before it is used if any remain. If no break points meet this criterion or no break points have been detected for this time series, the full time series is used. This is done to balance the need for sufficient information for parameter estimation of the used model, with the lack of correlation between samples before and after a break point.

Finally, it should be noted that not all combinations of parameters result in an ARIMA model that can be used to forecast a particular time series, as the autoregressive polynomial can be unstable or the moving average polynomial non-invertible. Furthermore, for this project the MATLAB function *arima* of the Econometrics Toolbox is used. Whilst this greatly simplifies the creation of ARIMA and SARIMA models, impossible results are sometimes obtained when the polynomial coefficients are estimated in a parallel loop. If as a result of either of these issues the forecast is invalid, instead Holt's linear trend model or Holt-Winters' seasonal method is used.

5.4. Forecasting Method Intercomparison

To quantify the accuracy of generated forecasts, the mean absolute scaled error (MASE) is used. This error measure was suggested by Hyndman and Koehler for the comparison of forecasting errors [146]. The scale of the data is removed by dividing each forecast error e_t with the mean slope between each consecutive pair of data points. After this, the mean absolute value of q_t is taken to be the MASE, as shown in Equation 5.5. In order to compute the error e_t, the first 18 years of data are used to generate a forecast for the last two years. This allows the forecast to be compared to obtained results.

$$q_t = \frac{e_t}{\frac{1}{n-1}\sum_{i=2}^{n}|Y_i - Y_{i-1}|} \tag{5.4}$$

$$MASE = mean(|q_t|) \tag{5.5}$$

To find the best performing method, the average MASE and its standard deviation were computed of a group of randomly selected, non-constant time series. Furthermore, a combination of forecasts generated using different methods can result in an improved forecast accuracy [142]. As such, the averaged forecast produced by the exponential smoothing and (S)ARIMA models are also assessed.

The results for a hundred time series are shown in Table 5.1. From this table it is immediately clear that it is significantly easier to produce accurate forecasts of the seasonal component than the trend component, which is discussed in more detail later. The averaged forecast of the trend component is slightly more accurate than the ARIMA forecast, which in turn performs much better than Holt's linear trend model. As the difference is minor, and the standard deviation is in fact larger for the averaged forecast, the choice was made to forego the averaged approach in favour of the ARIMA model. The most optimal seasonal forecasting approach is clearer, as SARIMA outperforms the exponential smoothing method as well as the average. As such, for the trend and seasonal component the ARIMA and SARIMA models are used respectively.

Table 5.1: Mean MASE and its standard deviation for various forecasting methods.

	Mean MASE	**STD MASE**
Trend component		
Holt's linear trend model	35.43	27.03
ARIMA	28.75	22.14
Averaged forecast	27.94	22.96
Seasonal component		
Holt-Winters' seasonal model	1.54	0.79
SARIMA	0.95	0.33
Averaged forecast	1.18	0.45

The results shown in Table 5.1 have been generated using time series from the region of South Holland. Another comparison was also performed on time series located on the Iberian peninsula. Despite a difference in absolute values, the same conclusions can still be drawn showing that the relative performance of the method is insensitive to the used selection of time series. It should further be noted that the computational time required by the (S)ARIMA models is significantly higher than for the exponential smoothing methods. If computational cost is a major concern, the reader may still choose to use the exponential smoothing models.

As mentioned previously the seasonal forecasts are notably more accurate than the trend forecasts. The reason for this is that the seasonal component behaves more consistently than the trend component. The example in Figure 5.1 shows an abnormally inconsistent seasonal component. Despite this, the change is still less than that commonly seen in the trend components and the forecast is accurate to the actual time series.

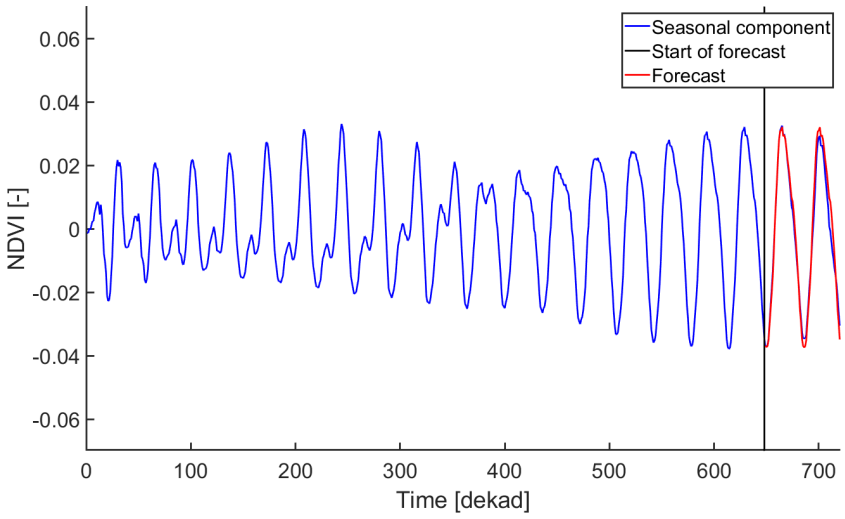


Figure 5.1: Example of an abnormally inconsistent seasonal component, with the generated forecast.

On the other hand, the trend component often contains break points, as was already discussed in Section 4.2. The forecasting methodology commonly fails to predict these changes, one such example being shown by Figure 5.2. Whilst methods exist that take the possibility of break points in the forecasting window into account, they are time consuming and difficult to implement [147] and good performance is not guaranteed [148]. As such, this aspect was not included in this project, but is one of the recommendations for future studies. As a result of this shortcoming, the forecasting error increases exponentially with the length of the forecasting window and forecasts are thus limited to two years.

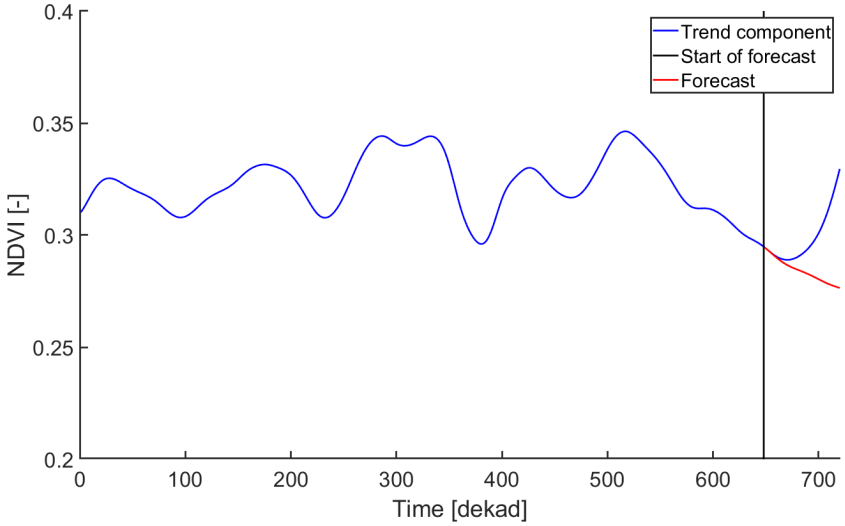


Figure 5.2: Example of a trend forecast showing a sudden change in the forecasting window.

5.5. Forecasting Recap

To conclude this chapter, an overview of the forecasting approach is given. A flowchart of the approach and its relation to the preceding and following chapters is shown below in Figure 5.3.

As the input data is not well suited to the used forecasting model, the Box-Cox transformation and its inverse are applied before and after forecasting is performed. Forecasting itself is done separately for the trend and seasonal components, where ARIMA is used for the former and SARIMA for the latter. These models use previous time series values, as well as error terms to create a polynomial that describes the future behaviour of the time series. To function correctly, the trend and seasonal components are differenced until stationarity is obtained. As the error for the trend component quickly grows due to unexpected changes in the forecasting window, the forecasting window is limited to two years.

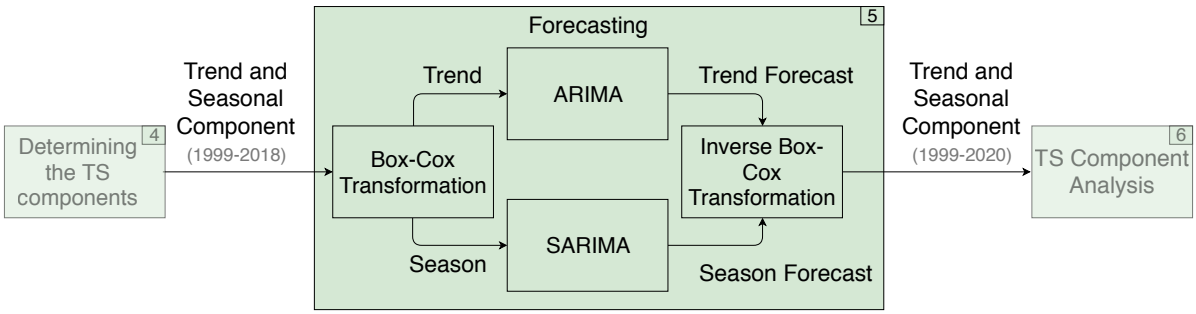


Figure 5.3: Forecasting flowchart.

6

Time Series Component Analysis

Now that the vegetation trend and seasonal activity have been calculated not only for the original data, but also for the two years after that, an analysis can be performed to determine the current and expected values for the trend properties as well as the growing season metrics. The analysis is split into two parts. The analysis performed on the trend component is described in Section 6.1, and the analysis performed on the seasonal component in Section 6.2.

6.1. Vegetation Greening and Browning

The increase and decrease in vegetation activity are termed greening and browning respectively. The introduction to this report, Chapter 1, showed that the degree of greening and browning is in contention. As such, these results are of interest to the project and can be determined with the generated trend components using the Theil-Sen estimator described in Subsection 6.1.1. To determine the statistical significance of the determined trends, the Mann-Kendall test described in Subsection 6.1.2 is applied.

6.1.1. Theil-Sen Estimator

The slope of the vegetation trend, in other words the degree of greening or browning, can be determined using the Theil-Sen estimator. This estimator can be used to compute the rate of change of a time series by calculating the median of the slopes between all pairs of data points. Once this slope T_S is obtained, the level T_L can be determined using the median of the trend component and its time stamps, as shown in Equation 6.1 [40]. As the time stamps are equally spaced, the median is equal to the average of the beginning a and end b of the time series. Unlike least-squares regression, it is non-parametric, circumventing the issues described previously in Section 4.1 [40, 46].

$$\begin{aligned} T_S &= \text{median}(\frac{T_j - T_i}{j - i}), \qquad a \leq i < j \leq b \\ T_L &= \text{median}(T) - T_S \cdot \frac{a+b}{2} \end{aligned} \tag{6.1}$$

To illustrate the results obtained after applying the Theil-Sen estimator, Figure 6.1 shows its application to the trend component of an agricultural time series. Unlike the synthetic time series used in Section 4.2 to illustrate the workings of the methodology used to date break points, this time series contains many more turning points than the maximum number of break points. In fact, this time series shows why this maximum has to be set as otherwise the time span of each partition would rapidly approach the minimum of two years. The detected break points are shown by the vertical black lines, and divide the data into three partitions. As can be seen by the green and red lines, the slope and level determined by the Theil-Sen estimator indicate a shift between the first and second partitions, and a shift as well as a trend change for the final partition.

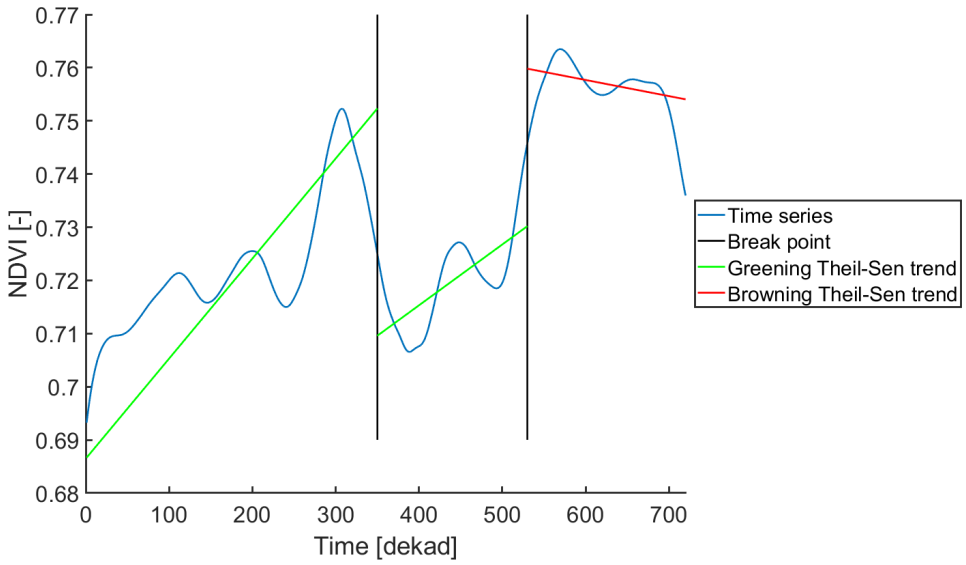


Figure 6.1: Trend component of an agricultural time series, showing the detected break points as black vertical lines and the slope and level determined by the Theil-Sen estimator in green and red.

6.1.2. Mann-Kendall Test

The Theil-Sen estimator indicates the rate of change of the time series, however it does not rate its significance. In order to find out whether a trend is consistently increasing or decreasing and whether this is statistically significant, the Mann-Kendall (MK) test is used according to Equations 6.2, 6.3 and 6.4 [40]. The statistic value is increased if a later value is higher and decreased if it is lower, as shown by Q which is determined by calculating the change between each data value and all subsequent data values. A high statistic value is therefore an indicator of a positive trend and vice versa. Before conclusions can be drawn however, the test statistic value Z_{MK} must be determined using the variance V, which increases exponentially with time series length n [10, 45, 124]. The absolute statistic value is indicative of the statistical significance. Unlike linear regression, it is a non-parametric test. It is however unsuited for serially correlated data such as the NDVI data, meaning it can only be applied to the trend component.

$$Z_{MK} = \begin{cases} \frac{Q-1}{\sqrt{V}} & \text{if } Q > 0 \\ 0 & \text{if } Q = 0 \\ \frac{Q+1}{\sqrt{V}} & \text{if } Q < 0 \end{cases} \tag{6.2}$$

$$Q = \sum_{j=a}^{b-1} \sum_{i=j+1}^{b} sign(T_i - T_j) \tag{6.3}$$

$$\begin{aligned} V &= \frac{n(n-1)(2n+5)}{18} \\ n &= b - a + 1 \end{aligned} \tag{6.4}$$

6.2. Growing Season Metrics

The previous sections of this chapter discussed the analysis of the trend component, with a methodology that cannot be applied to cyclical time series such as the seasonal component. The seasonal component does hold valuable information however, and in this section the method by which the start, end and length of the growing season are determined using the seasonal component is given.

The growing season metrics were briefly discussed in Section 3.3, as the start, end and length of the growing season are some of the parameters used to characterise each time series for classification. To calculate these metrics, the maximum and minimum NDVI are commonly used to determine the start and end of the growing season by using a threshold percentage of change between the two [35, 37, 69, 79, 87]. Additionally,

using this method on a pixel-by-pixel basis with a method called $\text{midpoint}_{\text{pixel}}$ was determined to be the most accurate, with the threshold being the average between the maximum and minimum NDVI [92]. To further ensure that the correct time stamps are found, for the start of the growing season the slope has to be positive and the point has to be before the maximum. For the end of the growing season, the opposite criteria are used. This helps to alleviate an issue that the growing season may be the inverse around the Mediterranean [49].

It should be noted that whilst the method deemed most accurate in the comparison drawn by M. A. White et al. was implemented [92], the study found that the results of the assessed methods varied considerably and contained errors on the order of days. As a result, M. A. White et al. caution against using these methods for the determination of trends in the growing season metrics [92]. Additionally, the rates of change present in literature are of the order of 0.1 - 1 day/year [10, 35, 37]. This exceeds both the expected magnitude of the error, as well as the temporal resolution of the NDVI data of ten days. As a result, neither the data, nor the applied methodology is well suited to determine the rates of change in the growing season metrics.

6.3. Time Series Component Analysis Recap

The final chapter of the methodology part of this thesis document is concluded with an overview of the analysis performed here, also provided in graphical form in Figure 6.2.

The trend and seasonal components are analysed separately, as the desired results differ. For the seasonal component, the $\text{midpoint}_{\text{pixel}}$ method is applied to determine the start and end of the growing season, from which the length can be readily determined. The start of the growing season is found when the average between maximum and minimum NDVI for the annual time series is first met, as well as the requirement for the slope to be positive and the point to be before the time at which NDVI is maximum. For the end of the growing season, the same threshold, but opposite criteria are used.

The trend component is used to determine the rate at which vegetation activity is increasing or decreasing, termed greening and browning. To do this, the statistically suitable Theil-Sen estimator is used which computes the median slope between all pairs of data points. Aside from these values, the statistical significance of the trend is also required to provide confidence in the obtained results. This statistical significance is computed via the Mann-Kendall test.

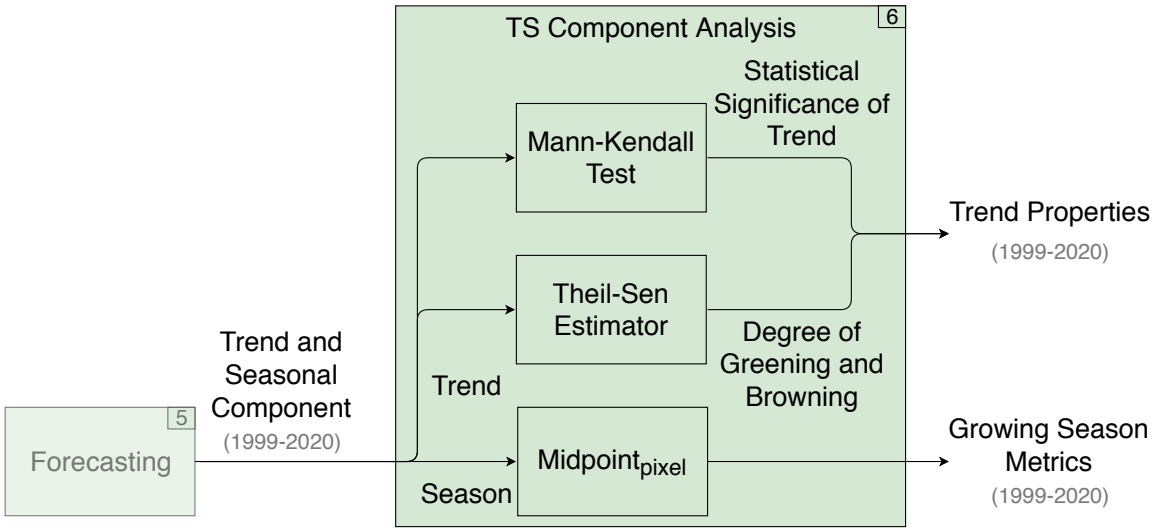


Figure 6.2: Time series analysis flowchart.

III

Results

Overview of the Results

In this part of the thesis the obtained results are given in three different chapters. The way by which these results have been generated using the methodology, as well as how they are interrelated, is shown by the flowchart in Figure 6.3. The green boxes contain the methodological steps that produced these results, with the white boxes showing the types of results that were generated. The chapters where each result can be found, are indicated by the numbers in the top-right of each white box.

The first chapter, Chapter 7, concerns itself with the annual land cover results that were produced for the full data set, as well as those for the countries and most populous cities within the domain. In the following two chapters, Chapter 8 and Chapter 9, the results obtained by analysing the trend and seasonal components of the time series are given. Key results obtained for the trend component are the degree of greening and browning, as well as the nationally aggregated trend components. Using the seasonal component, the growing season metrics were determined as well as similarly aggregated seasonal components.

Whilst the methodology could be developed using modest computational resources, this was insufficient to analyse the full data set. Instead, the Cartesius supercomputer owned and managed by SURFsara [149] was applied to the parallelised methodology, and more information regarding its use are given in Appendix I.

Finally, to make it easier to differentiate the various levels of the colormaps used for the results, ColorBrewer was used [150].

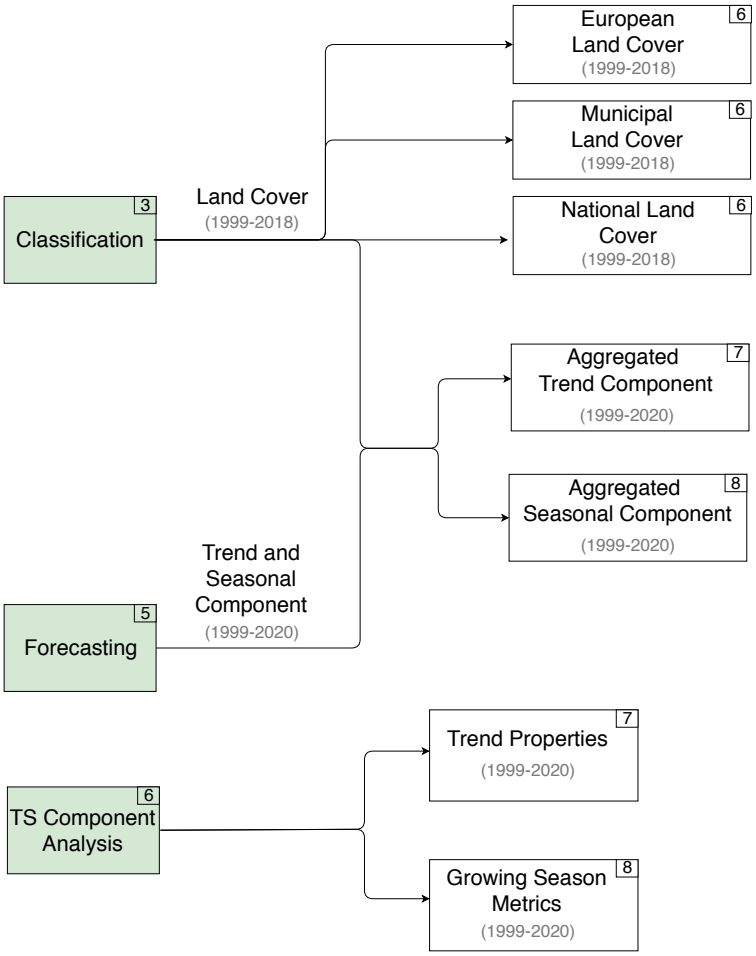


Figure 6.3: Schematic overview of the results (white boxes), and the methodological steps that created them (green boxes). The chapters where each result can be found, are indicated by the numbers in the top-right of each white box.

7

European Land Cover

In this chapter the annual land cover classification results are discussed for the European domain. The results for the full domain are given in Section 7.1, with the land cover map of 2018 shown in full which together with the confusion matrix demonstrates the effectiveness of the classification methodology. The subsequent sections, Section 7.2 and Section 7.3, provide more specific results, as the temporal change in land cover is given for the nations and largest cities within the domain.

7.1. Classification Results of the Entire Domain

In this section the results for the entire domain are presented. While existing land cover databases contain a data set of only one year, or have multiple years in between, this thesis has produced these results on an annual basis.

To assess whether the used smoothing parameters described in Section 3.7 are suitable, a comparison in the similarity between different years' data sets of the classified data and the Corine land cover data is shown in Table 7.1. These values are determined using the entire domain, minus sea pixels and land cover that is missing from the Corine data set. It can be observed that the similarity of the classified data is higher than the similarity between years 2000-2006 and 2006-2012 of the Corine data, but notably lower than the unusually high similarity between Corine data sets 2012 and 2018. It is therefore found to be an acceptable degree of smoothing.

Table 7.1: The similarity between different years of the ground truth data set and classified data sets.

	2000-2006	**2006-2012**	**2012-2018**
Corine land cover data	94.08%	94.34%	99.17%
Classified land cover data	95.65%	95.77%	95.80%

The table above showed the similarity over a six year gap, and the similarity between subsequent years is approximately 98%. Therefore, only the land cover result for the final year 2018 is presented in Figure 7.1, with results for other years available upon request. For a qualitative assessment of the results, the remapped Corine ground truth data is presented in Figure 7.2. The data is reduced to the same classes as the classified data, with missing data shown in white.

The classified land cover appears to capture all major patterns given by the Corine land cover data, with all large urban centra and natural areas accurately represented. Where the results differ, are the small areas covered by a land cover type that is distinct from its surroundings. These are often no bigger than a pixel, and most obvious for the urban class where the red 'noise' in the ground truth data is not present to the same extent in the classified data. The likely explanation for this is that these distinct areas can not provide the classifier with sufficient training data.

Figure 7.1: Classified European land cover for the year 2018.

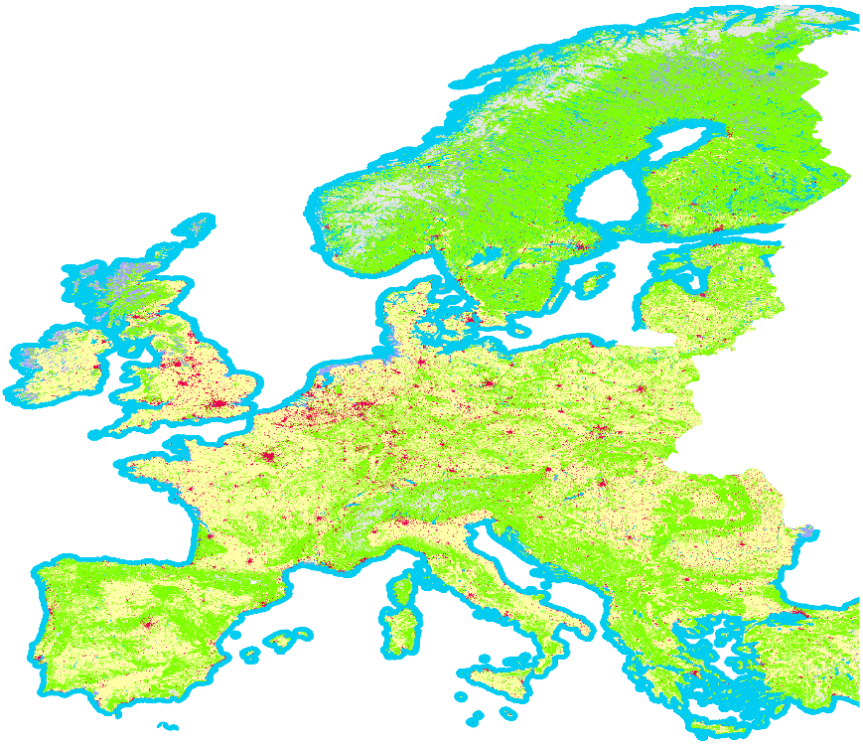
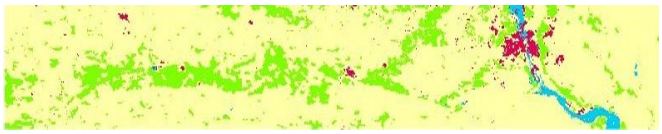
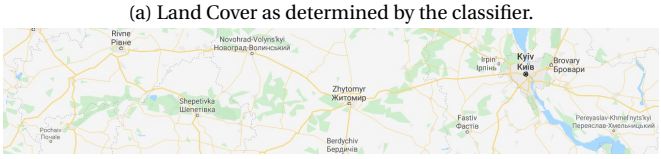


Figure 7.2: Corine land cover data for the year 2018. Data is missing for the areas covered in white.

For the regions where the ground truth data is missing, the classifier manages to classify major cities such as St. Petersburg, Minsk and Kiev, but also smaller cities such as Odessa on the coast of the Black Sea and Veliky Novgorod, south-east of St. Petersburg. The classifier also manages to capture the forests of eastern Europe well. Examples are the thin strip of forest cover west of Kiev, encompassing amongst others the Malovanka park, or the elongated forest on the border between the Berazino and Byalynichy districts of Belarus, between Minsk and Mogilev. The Ukrainian example is shown in Figure 7.3, and shows that even minor details in the natural and artificial landscape are captured despite the great distance to ground truth data.



(a) Land Cover as determined by the classifier.



(b) Land Cover as determined by Google Maps. [149]

Figure 7.3: Comparison between the classified data and Google Maps reference data. [149]

A more quantitative assessment of the results is the confusion matrix presented in Figure 7.4. This confusion matrix is the aggregated result of the test data used for the entire domain, for the years that ground truth data is available. The classification accuracy of each class is shown by the blue column to the right of the matrix. The over-all classification accuracy excluding sea pixels exceeds 80%, due to the very high classification accuracies obtained for the agricultural and grasslands and forests classes, as well as the high classification accuracy of the non-vegetated class. The high classification accuracies for the latter may not be as expected as that of the former two classes, as it is considerably rarer. Whilst the agricultural and grasslands and forests classes exist in great quantity throughout the domain, the non-vegetated class only exists in specific regions such as the mountain ranges of Europe. It is therefore a globally rare class, however not a locally rare class when applying segmenting. As expected, the classification accuracy for the water bodies class is nearly perfect as this class is most easily separable from the others and shows little variability.

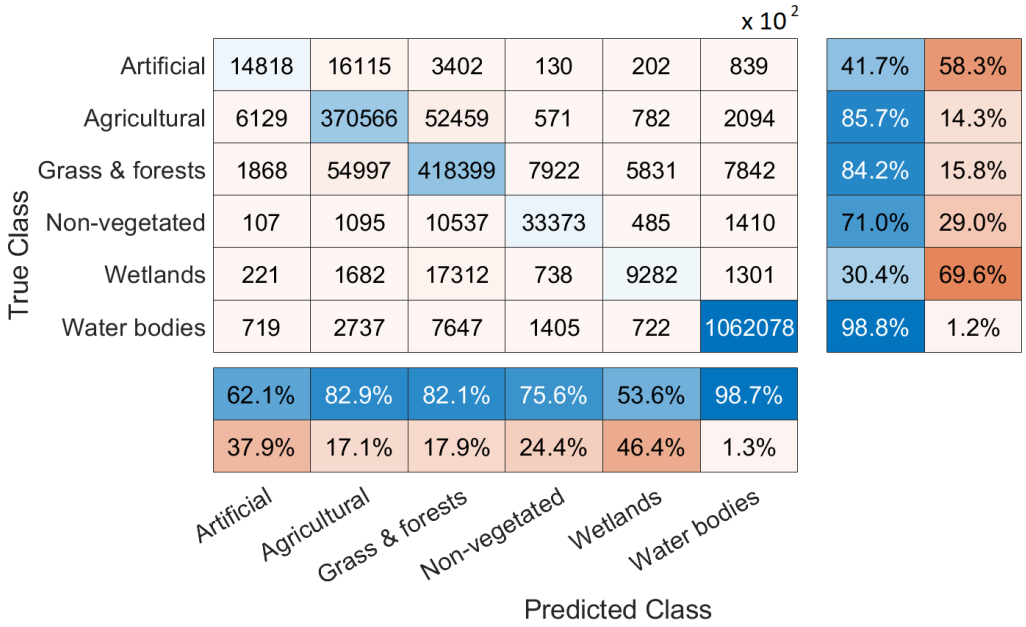


Figure 7.4: Aggregated confusion matrix for all used test data. For increased clarity, the number of samples is one hundredth the actual amount.

The classification accuracies of the wetlands and artificial classes are poorer, as these occur in smaller groups and have higher intra-class variability, as shown in Appendix F. Whilst these classification accuracies are at first sight disappointing, the reader is reminded that only one variable, NDVI, is used to generate these results and they are comparable with those found in other classification studies [67, 82, 85, 122].

The confusion matrix further shows that for the rarer classes, the precision of the classifier, shown in the bottom blue row, is considerably greater than its accuracy, shown in the blue row to the right. This indicates that the classifier is conservative when applying the rarer classes, and when it does so it is often correct. This is then a more quantitative representation of the lack of artificial 'noise' shown in Figure 7.1.

Finally, an absolutely true ground truth data set does not exist, and whilst the Corine land cover data base is highly accurate with a total classification accuracy of 85%, false classification labels reduce the quality of the training data used to grow the classification trees and result in reduced classification performance.

The distribution of land cover within the entire domain in 1999, as well as the change in distribution between 1999 and 2018, are given in Table 7.2. This table shows that changes in the land cover distribution are minor, and that the class distribution is highly imbalanced. A high relative increase in land covered by the wetlands class was found due to the Corine data set for the year 2000 not labelling the wetlands of northern Scotland, contrary to the data sets of 2006, 2012 and 2018. These were thus wrongly classified in the years 1999 to 2003, with the correct classification given in all succeeding years. As a result, in this chapter the classification results of Great Britain for the years 1999-2003 are substituted for those obtained in 2004.

The urban sprawl expected to be found in Europe [151, 152] can be seen through the high relative increase, something which will be discussed in more detail in Section 7.3. The non-vegetated class also increased in size, likely due to land degradation caused by the increase in global temperature [11]. Meanwhile the agricultural and grasslands and forests classes both decreased in size, albeit only by roughly 0.5%, and the water bodies class shows virtually no change in size.

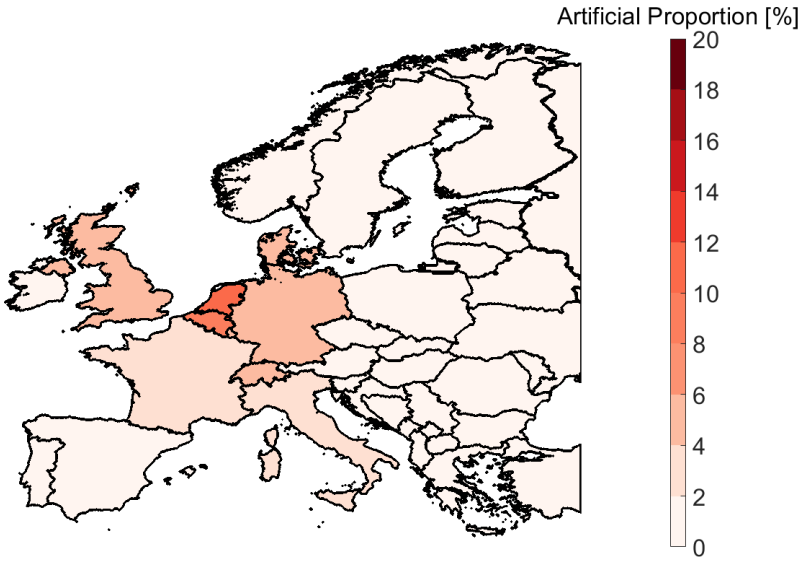
Table 7.2: The distribution of classes for the entire domain in 1999, including the absolute change w.r.t. to the total number of samples and the relative change w.r.t. the number of samples within that class, between 1999 and 2018.

	Distribution 1999 [%]	**Absolute Change [%] (1999-2018)**	**Relative Change [%] (1999-2018)**
Artificial [%]	0.82	0.14	14.39
Agricultural [%]	21.54	-0.14	-0.54
Grassl. & Forests [%]	24.28	-0.18	-0.62
Non-Vegetated [%]	1.64	0.04	1.99
Wetlands [%]	0.28	0.13	40.67
Water bodies [%]	51.44	0.00	0.01

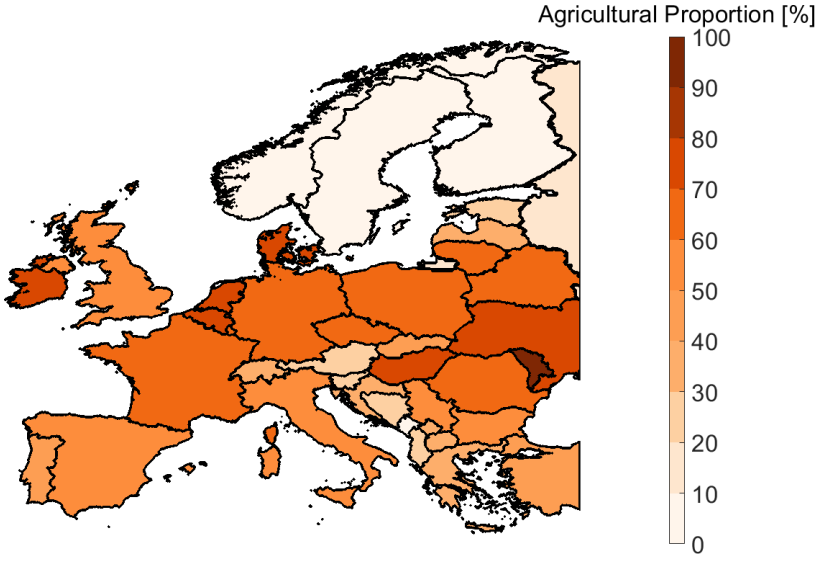
7.2. Changes in the Land Cover of European Nations

Classification results on a national level are presented in this section for the 39 largest countries within the analysed domain. The results are presented in graphical format in this chapter, with numerical results available in Appendix G.

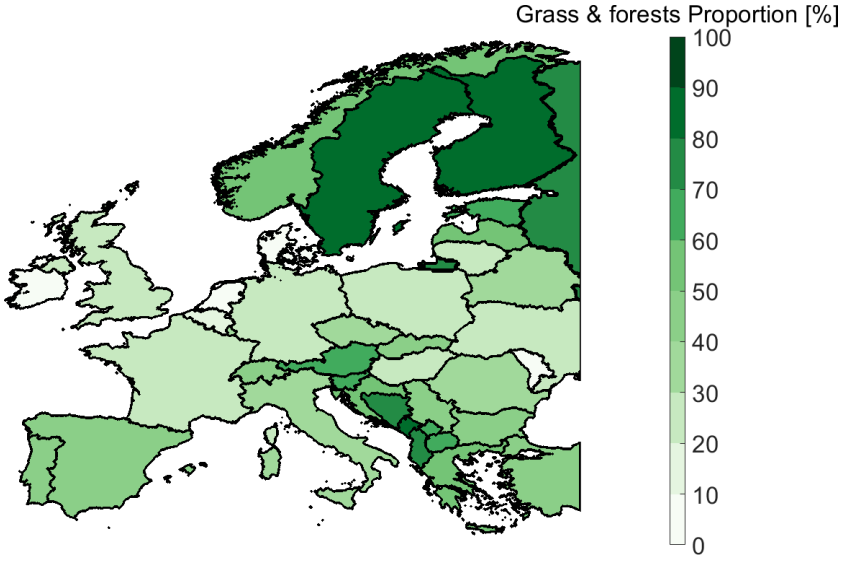
Figure 7.5 shows the class distribution for the first year of the data set 1999, as determined by the classification methodology, to provide context for the succeeding results. From this class distribution it can be deduced that the countries of north-western Europe have the greatest percentage of artificial land cover, something which is also apparent from the land cover map shown in Figure 7.1. On the contrary, the countries with the greatest percentage of agricultural land cover show no distinct geographical correlation, with Moldova showing a significantly greater percentage than any of the other nations. The land cover map presented earlier further shows that the forest class covers nearly all of Scandinavia, which is confirmed by the high fraction of Sweden and Finland covered by the grasslands and forests class. As expected, countries containing large mountain ranges have the greatest percentage of the non-vegetated class, with a negligible percentage for the vast majority of European nations. Only Ireland and Great Britain show a fraction of land covered by wetlands greater than 2%, and these countries contain a far greater area covered by wetlands than any of the other countries within the domain. For the water bodies class it is interesting to note that the four countries showing the largest proportion contain lakes found in the northern latitudes of Europe, with the Netherlands being the fifth due to its artificially enclosed lake, the *Ijsselmeer*.



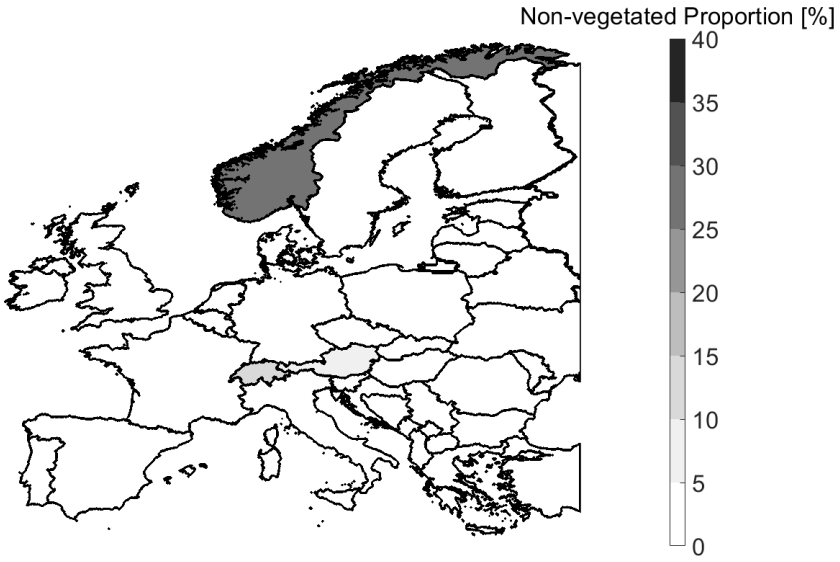
(a) Artificial



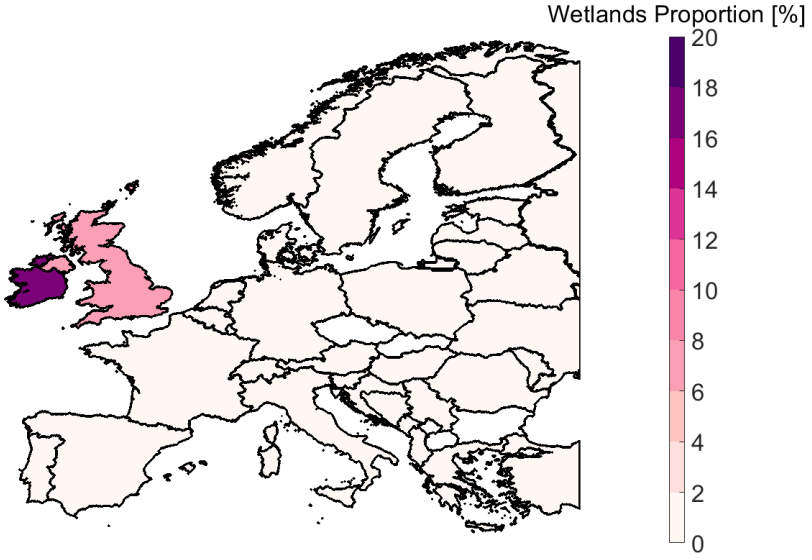
(b) Agricultural



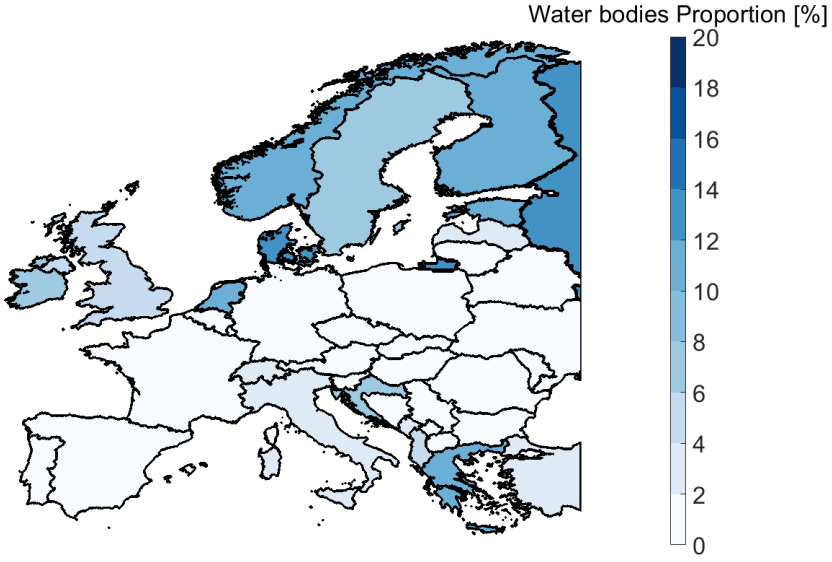
(c) Grasslands and forests



(d) Non-vegetated



(e) Wetlands



(f) Water Bodies

Figure 7.5: Proportion of the land cover types in 1999.

The difference between the national land cover distribution in 1999 and 2018, relative to the size of the class itself, is presented in Figure 7.6. Countries where the class forms less than 0.5% of the land cover, are shown in grey. Whilst the common classes show only moderate levels of change, for the rarer classes significant change is found. These percentages are indeed unrealistic, and show the challenge encountered when attempting to classify rare classes. It should be noted that similar results are obtained when computing these values using the Corine database, which are shown in Appendix H. This further illustrates the difficulty that lies in creating consistent classification results for rare classes, and as the Corine data is used to train the classifier, the magnitude of change is therefore as expected.

The countries in South-Eastern Europe show the greatest relative increase in artificial land cover. These results resemble the percentage of built-up area (PBA) findings reported by the European Environment Agency (EEA) and Swiss Federal Office for the Environment (FOEN) [151]. Furthermore, V. C. Radeloff and G. Gutman claim a high degree of urbanisation in many former Eastern Bloc countries [153]. While the results presented here quantify the spatial spread of artificial land cover, it is important to note that the classification data cannot be used to determine urban intensification, another important aspect of urban sprawl.

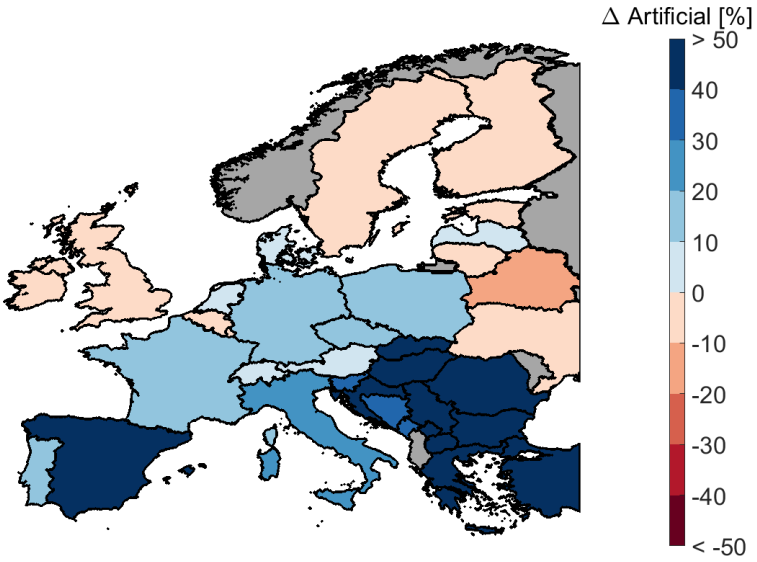
The agricultural class has changed less over the twenty year period, with an increase in most eastern countries and a slight decrease in western and central countries, although not consistently so for either region. S. Estel et al. state that abandonment of farmland occurred in Eastern Europe after the dissolution of the Soviet Union [49, 153]. Whilst farmland continues to be abandoned in these regions. The rate of recultivation in Eastern Europe is greater than the rate of abandonment and as such the increased proportion of farmland in Eastern Europe is to be expected. They further show that farmland abandonment has occurred in the mountainous regions of Europe, which is confirmed by the percentages shown by Norway, Switzerland and to a lesser extent Austria and Spain. They however also conclude that farmland in southern Scandinavia has been abandoned and not recultivated, which corresponds to the findings for Sweden, however not for Finland, which shows a large increase in agricultural land cover.

Many of the north-western countries of Europe show forestation, with deforestation in much of south-eastern Europe. The dissolution of the Soviet Union had not only a large impact on urban areas and farmland, but also on the forests of former Eastern Bloc countries. G. Gutman and V. Radeloff state that the logging industries of the Baltic countries and Romania grew to unsustainable levels, which agrees with the negative rates found for all but Lithuania [153].

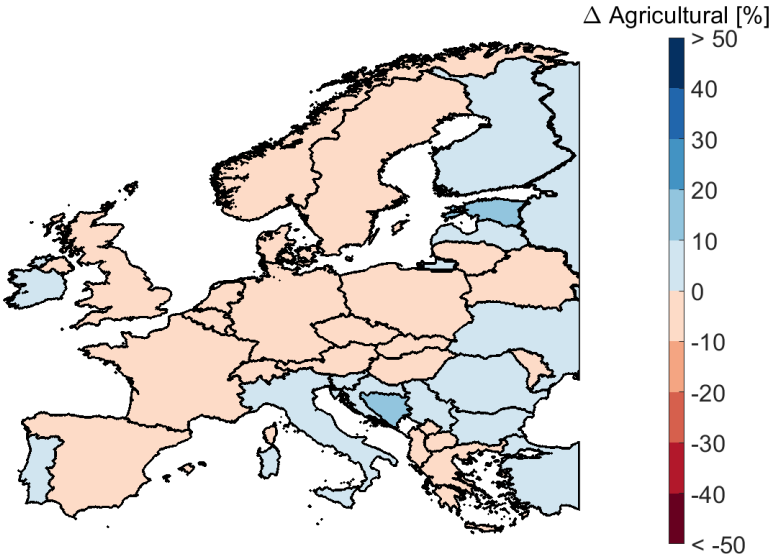
Whilst few countries contain a non-neglible amount of non-vegetated land cover, the ones that do almost ubiquitously show an increase with the most notable exception being Spain. The non-vegetated area in south-eastern Spain slowly transformed into forests and grasslands in the early 2000's according to the classifier.

The final vegetation class is that of wetlands, which only Ireland and Great Britain contain in great quantity. Ireland shows a moderate decrease of 6.4% over the time period, whilst Great Britain shows an increase of 8.9%. The former therefore corresponds with the negative rate of change reported by N. C. Davidson for European wetlands, whilst the latter negates it and for the domain as a whole a positive change was determined [28].

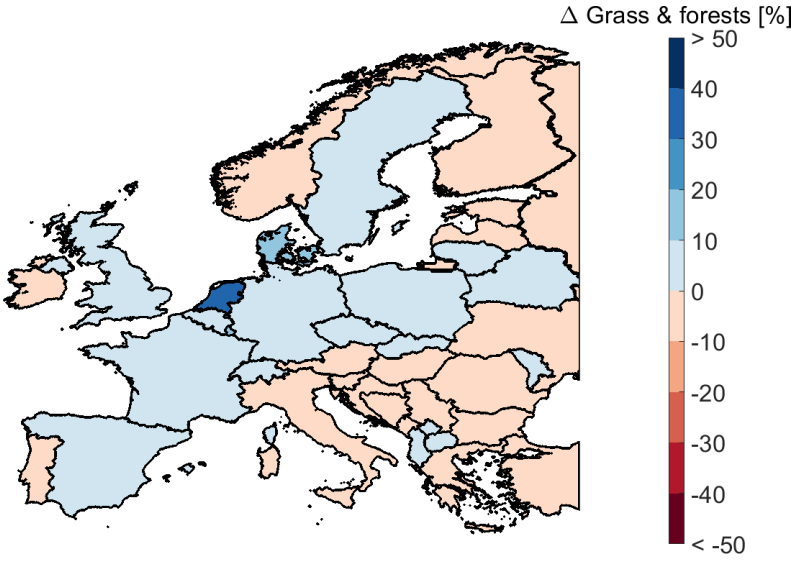
The temporal change for the water class distribution is smaller than that of the other common classes. As the changes are small in an absolute sense, the likeliest cause would be errors introduced due to classification. It is however interesting to note that an increase is found for nearly the entire domain, indicating a degree of spatial correlation.



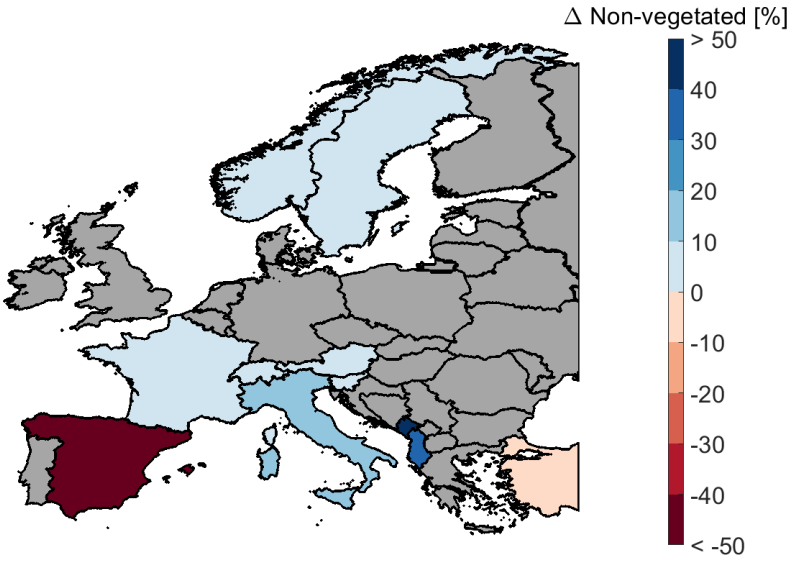
(a) Artificial



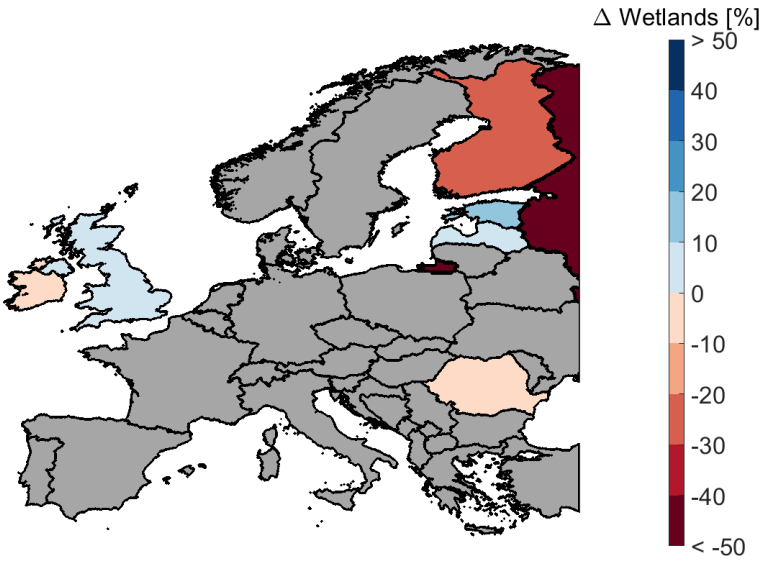
(b) Agricultural



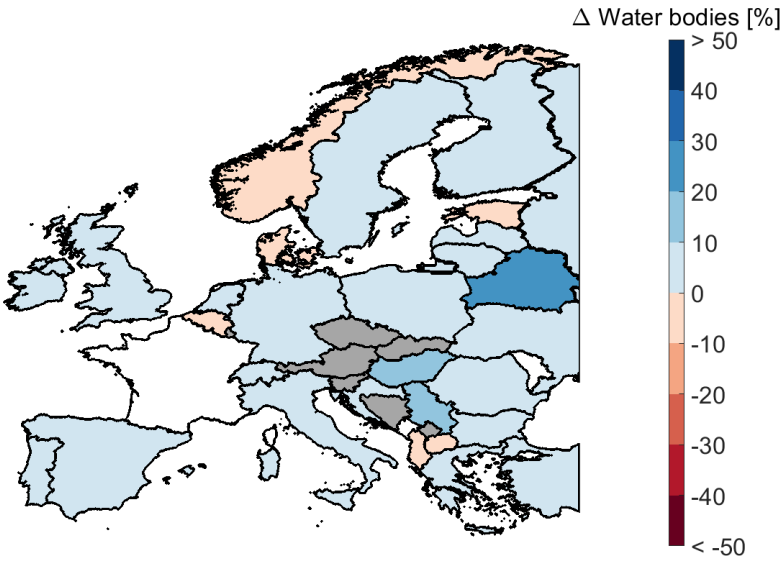
(c) Grasslands and forests



(d) Non-vegetated



(e) Wetlands



(f) Water Bodies

Figure 7.6: Relative change in land cover with respect to the number of samples within the class itself, between 1999 and 2018. Countries where the class covered less than 0.5% of the land are shown in grey.

7.3. Changes in the Land Cover of European Cities

Urban sprawl is a commonly recognised environmental problem, and could already be observed both for the full domain, as well as on the national level. This section looks into this phenomenon in more detail, by investigating the changes in land cover for the ten most populated cities within the domain. Again, the reader should be aware that the used classification method is only able to detect changes in area, and is thus not well suited to capture other urbanisation parameters such as utilisation intensity [151]. Furthermore, whilst the 1 by 1 kilometre resolution is adequate for results over a larger region, this results in a limited amount of pixels to draw conclusions from for these specific subsets of the database.

The city domains are chosen such that they encompass the city and its surrounding regions, and are shown in Figure 7.7. From these figures it is evident that the greater urban areas vary drastically in size and land cover distribution. As such, only the changes relative to the class itself are given in Table 7.3. As neither the non-vegetated nor the wetlands class was present in the considered urban areas, these results are omitted.

Of the ten most populated cities, St. Petersburg, Kiev and Minsk are located in regions where no ground truth data is available. Despite this, the classifier manages to classify these cities and the smaller urban regions in the greater St. Petersburg area. The classification results for St. Petersburg show that this is not without difficulty however, and a false transition from artificial to water is present for much of St. Petersburg, explaining the respective decrease and increase for these two classes. The only other city showing a significant decrease in artificial area is Kiev, where some of the artificial class has transitioned to agricultural. The remaining city outside the scope of Corine is Minsk, which shows highly consistent land cover over time.

Bucharest shows by far the strongest urbanisation. Whilst quantitative information regarding Bucharest's changing land cover could not be found in other sources, a strong degree of urbanisation to the east of Bucharest is described in the years 2000 to 2006 by J. Feranec et al., which agrees with the land cover differences shown in Figure 7.7q and Figure 7.7r [153].

Table 7.3: The relative difference between land cover distributions of 1999 and 2018.

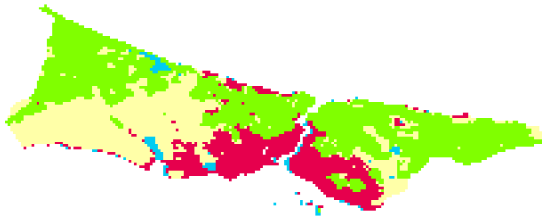
City	Artificial [%]	Agricultural [%]	Grassl. & Forests [%]	Water [%]
Istanbul	20.51	-9.45	-2.34	10.88
London	-0.69	-0.29	10.40	1.82
St. Petersburg	-16.35	25.00	-3.74	0.96
Berlin	11.87	-7.72	0.32	13.39
Madrid	9.77	2.87	-4.36	0.00
Kiev	-8.13	30.30	-11.31	0.00
Rome	11.24	-3.84	1.56	0.00
Paris	2.59	-2.48	4.43	0.00
Bucharest	48.78	-13.81	-4.40	0.00
Minsk	0.81	-1.66	6.33	20.69

The results found for London, Paris, Berlin, Madrid and Istanbul can be compared to the findings from the Urban Expansion Program from the New York University, who classified these cities using Landsat image and determined the annual rate of change between the early 2000's and early 2010's, with the specific time span varying per city [99]. The annual rates of change found by this study, and those found by the NYU, are presented in Table 7.4. From this table it is evident that the rates of change predicted by this study are consistently smaller in magnitude than those predicted by S. Angel et al [99], but also that increased urbanisation is determined for all cities but London, where this study predicts a negligible change.

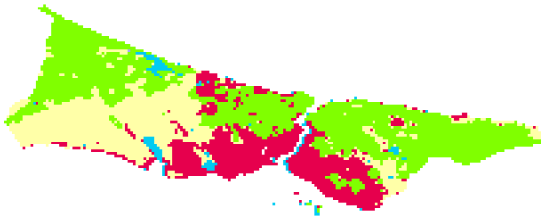
It should be noted that the reference study considers the inclusion of existing urban areas into a city's boundaries an increase in the city's size. This is not considered by this analysis and it forms a large part of the change found by the reference study, partially explaining the lower rates of change. Another important difference is that the artificial class includes land cover types that are not urban, such as the Berlin-Brandenburg and Istanbul airports, which began construction in 2006 and 2014 respectively [154, 155].

Table 7.4: Annual rates of change of the artificial class as determined in this thesis and the reference study [99].

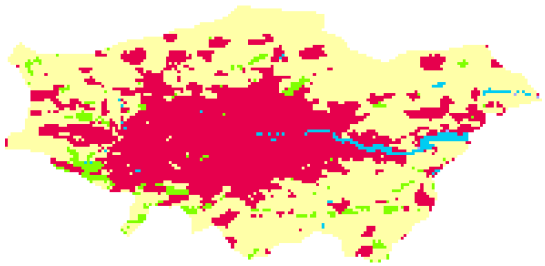
	London	Paris	Berlin	Madrid	Istanbul
NDVI Classification	-0.03%	0.13%	0.56%	0.47%	0.94%
Urban Expansion [99]	0.2%	1.3%	2.9%	6%	3.7%



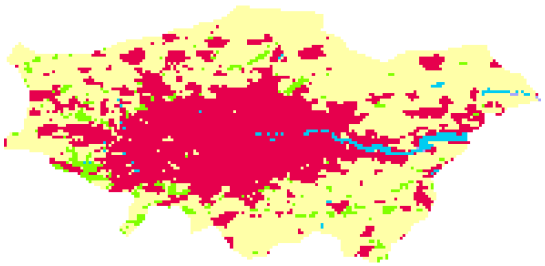
(a) Istanbul, 1999.



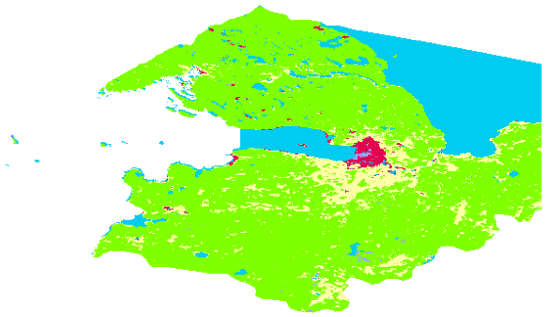
(b) Istanbul, 2018.



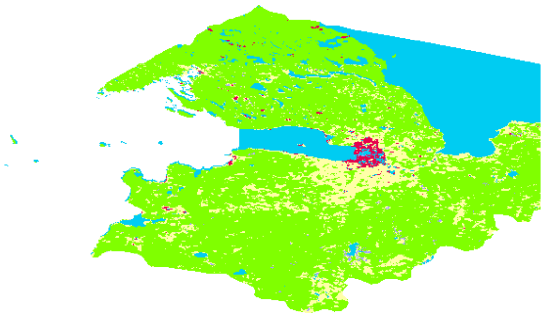
(c) London, 1999.



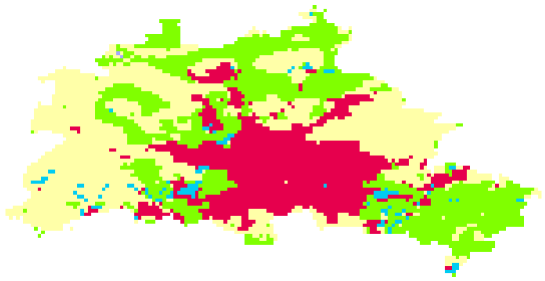
(d) London, 2018.



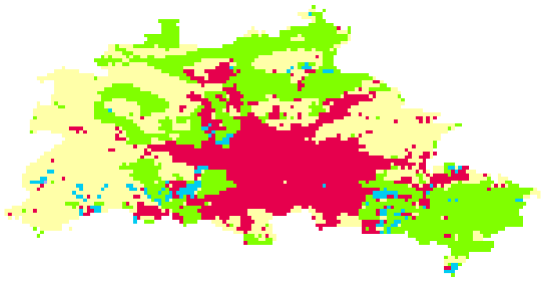
(e) St. Petersburg, 1999.



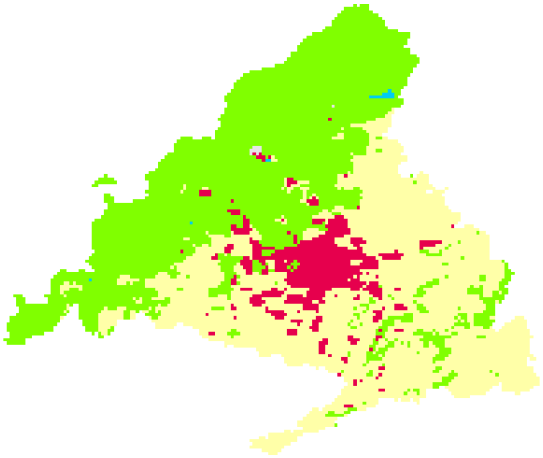
(f) St. Petersburg, 2018.



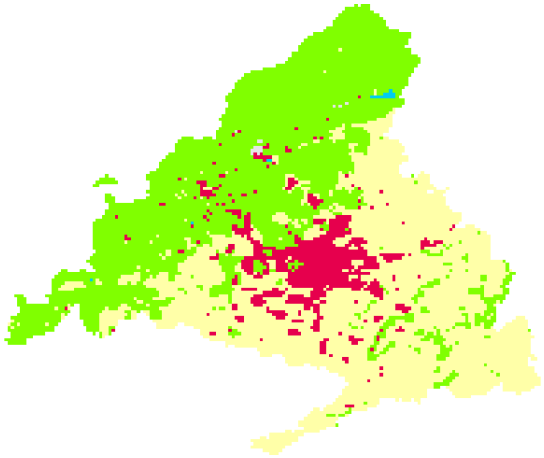
(g) Berlin, 1999.



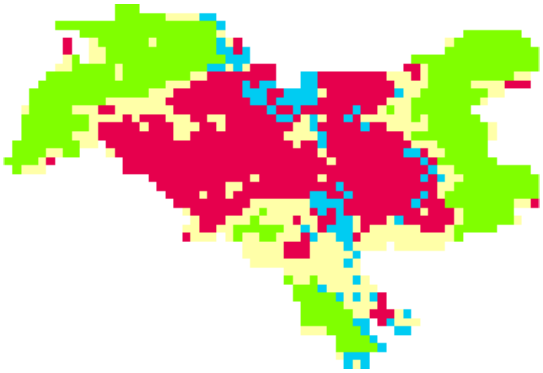
(h) Berlin, 2018.



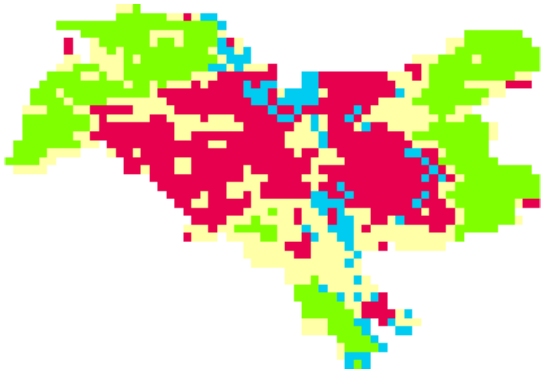
(i) Madrid, 1999.



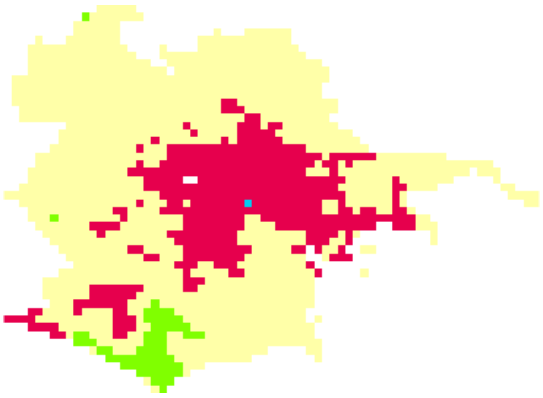
(j) Madrid, 2018.



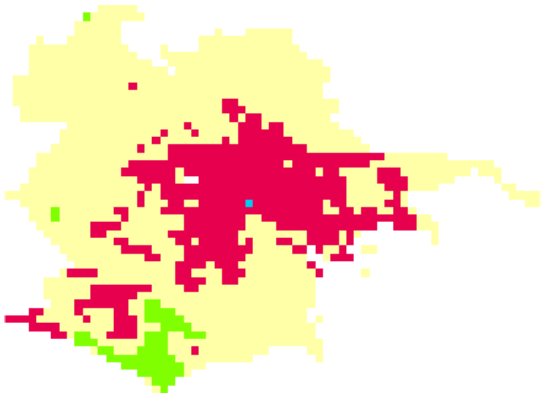
(k) Kiev, 1999.



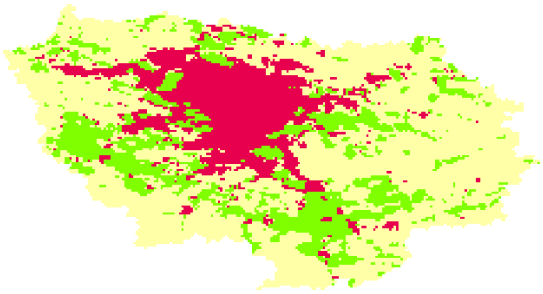
(l) Kiev, 2018.



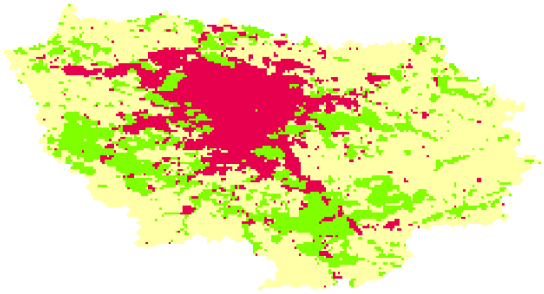
(m) Rome, 1999.



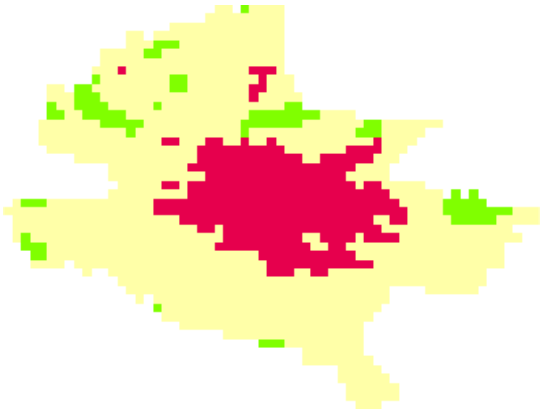
(n) Rome, 2018.



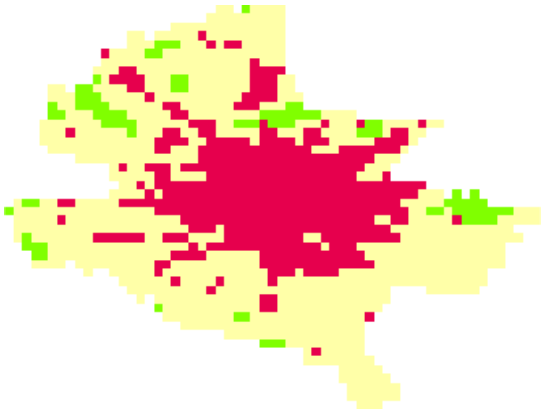
(o) Paris, 1999.



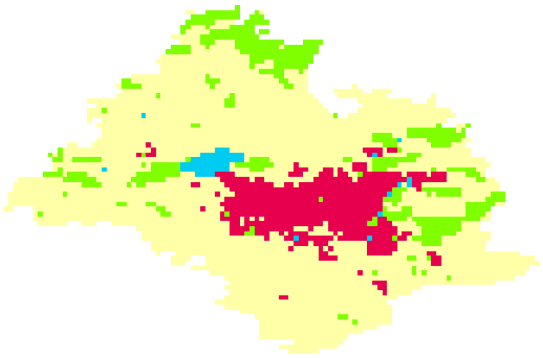
(p) Paris, 2018.



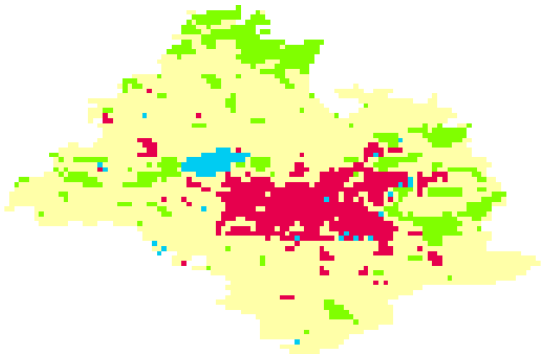
(q) Bucharest, 1999.



(r) Bucharest, 2018.



(s) Minsk, 1999.



(t) Minsk, 2018.

Figure 7.7: The land cover maps of the ten most populated cities in the European domain, with the results for 1999 on the left and 2018 on the right.

8

Changes in the Vegetation Activity over Time

In this chapter the results regarding the trend of the vegetation data are presented of the full European domain. The greening and browning of vegetation as well as its statistical significance are discussed in the first three sections. First for the years 1999 through 2018 in Section 8.1, followed by the last partitions of the data - the data after the last break points - in Section 8.2, and finally the results for the forecast trend component in Section 8.3. The chapter is concluded with the aggregated trend components of the various classes in the Netherlands in Section 8.4. Comparisons are drawn between the results found in literature and the results obtained by this study, however the driving factors behind these phenomena are not discussed. A detailed investigation into the phenological behaviour remains outside the scope of this research.

8.1. Greening and Browning over the Entire Time Span

In this section the rate of change of vegetation activity is given over the entire time span of the data, from 1999 to 2018. To determine the degree of greening and browning, the Theil-Sen estimator was applied to calculate the rate of change of each pixel's trend component. The results are given in Figure 8.1, expressed in the change in NDVI per year. This figure shows that Europe as a whole shows quite consistent greening, and NDVI increased with roughly $2 \cdot 10^{-4}$ per year. One exception to this are the mountainous regions, most prominently shown in the Scandinavian mountains and the Alps, but also visible in the Pyrenees and Balkans. These areas show browning at different degrees, and NDVI decreased on average on the order of $-5 \cdot 10^{-4}$ per year. Further decreases in NDVI of approximately $-2 \cdot 10^{-4}$ per year are also present on the Russian border with the Baltic states, as well as in Northern Scandinavia. It should be noted that a change in NDVI on the order of 10^{-4} corresponds to a change on the order of 10^{-2} % per year.

To determine the validity of these results, they are compared to those of reference studies. Whilst a detailed investigation into greening and browning of the European continent has not yet been performed to the best knowledge of the author, information can be retrieved from global studies. These were however often performed at lower resolution, and commonly only show statistical significance in parts of the European domain [10, 31, 36, 47, 156]. Furthermore, as was mentioned previously in Section 1.2 when discussing prior studies into vegetation activity, the findings of the various studies are inconsistent. A possible cause is the difference in data sets, resulting in different trends on regional and global scales [127]. This is further illustrated by the study performed by K. C. Guay et al., who obtained notable differences in greening and browning when analysing data-series GIMMS$_g$ and its successor GIMMS$_{3g}$, over the 1982 to 2008 time period [156]. Despite these issues, a brief comparison is still presented to provide a more detailed overview of previous studies, and it includes both sources that agree as well as disagree with the presented findings.

Greening in mainland Europe (Europe excluding Scandinavia and the United Kingdom) was determined by all of the examined studies that included this domain [10, 31, 36, 46, 47] except one, which bound browning throughout most of the European domain - including the Scandinavian mountain range [48]. Greening in mainland Europe is therefore a common finding, with a rate of change in NDVI of order 10^{-3} per year [36, 46, 47], which is therefore an order of magnitude greater than the one obtained in this study, however one order of magnitude smaller than the one found by R. de Jong et al. [10]. A slightly higher degree of greening was

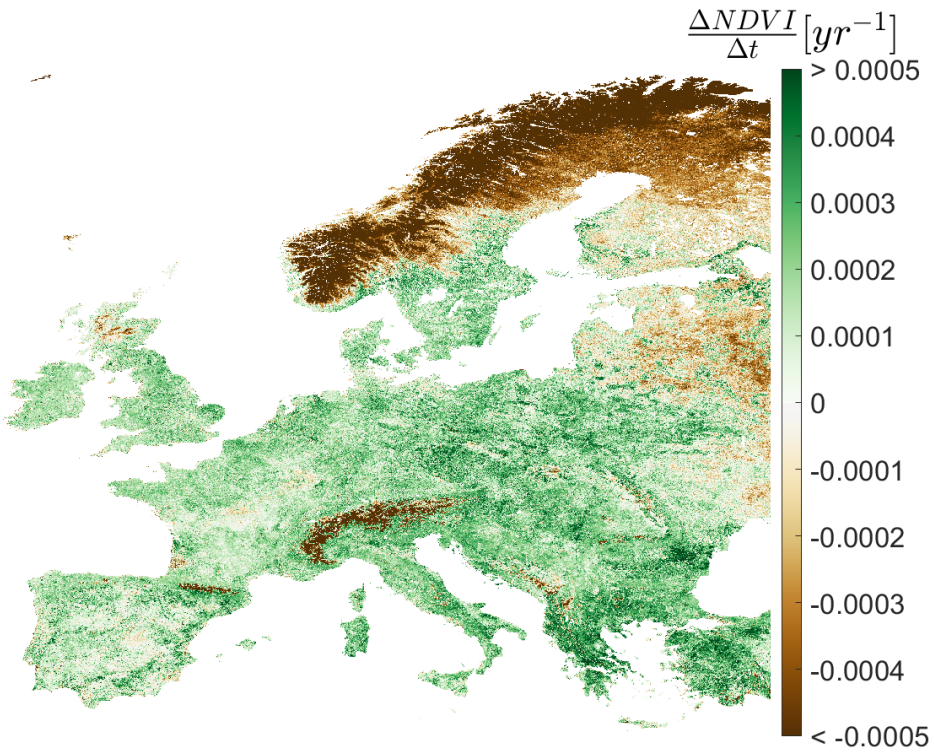


Figure 8.1: The rate of change of NDVI per year, determined over the complete trend component.

found for south-eastern Europe compared to its surroundings, in agreement with the findings of Chen et al. [31].

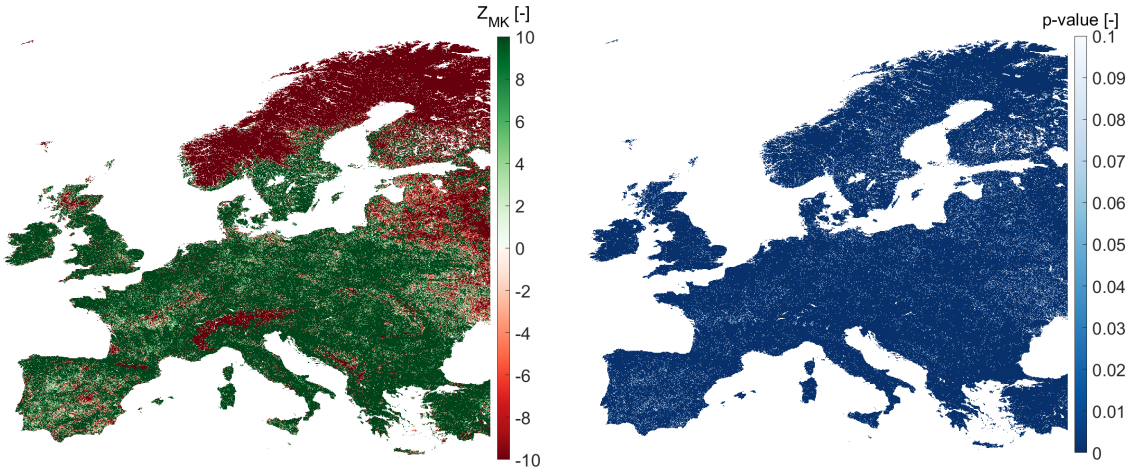
A predominantly greening domain is therefore well established, albeit at varying degrees. Another finding of the research is browning in Northern Scandinavia, for which less information exists in the global studies due to a lack of statistical significance for the trends obtained there [31, 46, 47]. The findings are supported by studies performed by R. de Jong et al., W. Yuan et al. and K. C. Guay [10, 48, 156]. Whilst the latter study shows disagreement between the two data sets for the entire time span after 1982, agreement that Scandinavia is browning is found for the years 2000 to 2008. A report by the US National Oceanic and Atmospheric Administration further states that areas at high latitude have transitioned from greening to browning in the 1990's, however no specific information for Scandinavia is provided [52].

Finally, the data clearly shows a browning trend for mountainous regions throughout Europe, regardless of their location. Whilst to the best of the author's knowledge no study exists that studied changes in vegetation activity for the Balkans, Pyrenees or Alps as a whole, B. Z. Carlson et al. studied a small region within the alps and found it to be greening over the entire 1984-2015 period, albeit at a decelerating pace [157]. Whilst the aforementioned global studies on vegetation greening and browning show these mountainous regions as well, a combination of low spatial resolution as well as low statistical significance makes drawing conclusions from these studies with regards to these mountainous regions difficult.

This section has now given the calculated rates of change in vegetation activity, and compared these findings to existing studies. Many of these studies showed low statistical significance for areas within the studied domain, which further necessitates answering whether the aforementioned trends are statistically relevant. To do this, the Mann-Kendall test is performed and the resulting test statistic values Z_{MK}, as well as the associated p-values, are determined. The reader is reminded that a positive test statistic value indicates a positive trend and vice versa, with the absolute value being indicative of statistical significance and not the magnitude of the present trend. On the contrary, a low p-value indicates a higher level of statistical significance.

The results are given in Figure 8.2, (a) closely mimicking the previously shown Theil-Sen slopes. As shown by (b), the statistical significance is high throughout the domain, and as such no results need to be withheld due to poor confidence. It should be noted that these results are determined on the trend component, and much lower statistical significance is found if the time series are not decomposed as the seasonal component

obfuscates present trends. Of the assessed studies, only those produced by Y. Liu and W. Yuan showed statistical significance throughout the domain [46, 48]. Both of these studies use GIMMS NDVI data, which has an 8 by 8 kilometres resolution and the produced results are thus at a coarser resolution than the 1 by 1 kilometre results produced by this study. Furthermore, neither study is as recent as this one as the used data sets end in 2012 and 2015 respectively.



(a) Mann-Kendall test statistic values.

(b) P-values.

Figure 8.2: Mann-Kendall test statistic and the associated p-values, determined over the trend component from 1999 to 2018.

8.2. Greening and Browning over the Last Partitions

In the previous section, the rate of change in vegetation activity between 1999 and 2018 was given. To provide further detail, this section shows the resulting degree of greening and browning, as well as its statistical significance, obtained over the last partition of each time series. Furthermore, these results are indicative for the forecasting results of the trend component that will be shown in the next section, as trend component data belonging to the last partition is used to train the forecasting algorithms. Similarly to before, comparisons between the obtained results and those published in literature are performed whenever possible, however limited information is available for the specific time spans assessed here.

These partitions are not equal in length, as the last partition begins at the last of the detected break points for each pixel individually. As such, the times at which the last break points are detected - and the last partitions thus start - are shown in Figure 8.3. No values are given beyond the year 2017 as a minimum length of two years was required. This figure shows that these dates vary quite strongly over the domain, but do show some degree of spatial correlation. Most noticeable are the late break points detected in Northern Scandinavia and the Russian border with the Baltic states, resembling the area that shows browning over the full time span. Late break points have also been detected for areas in Western Europe.

The rates of greening and browning for these last partitions are shown in Figure 8.4. Notable changes from the results over the entire time span are that mainland Europe does not show consistent greening anymore, and areas of France, Germany, Spain, Portugal and the Atlantic coast of Norway have transitioned from greening to browning. Additionally, the mountainous regions of Europe, as well as Northern Scandinavia and far Eastern Europe showed browning over the full time span, and the data from the last partitions shows a transition from browning to greening. As such, the majority of the areas that show late break points in Figure 8.3 have transitioned from greening to browning or vice versa. The study performed by N. Pan included their results obtained after the last turning points, and browning is visible in the same regions [36]. They however also detected browning for the British Isles, which is not present in these results.

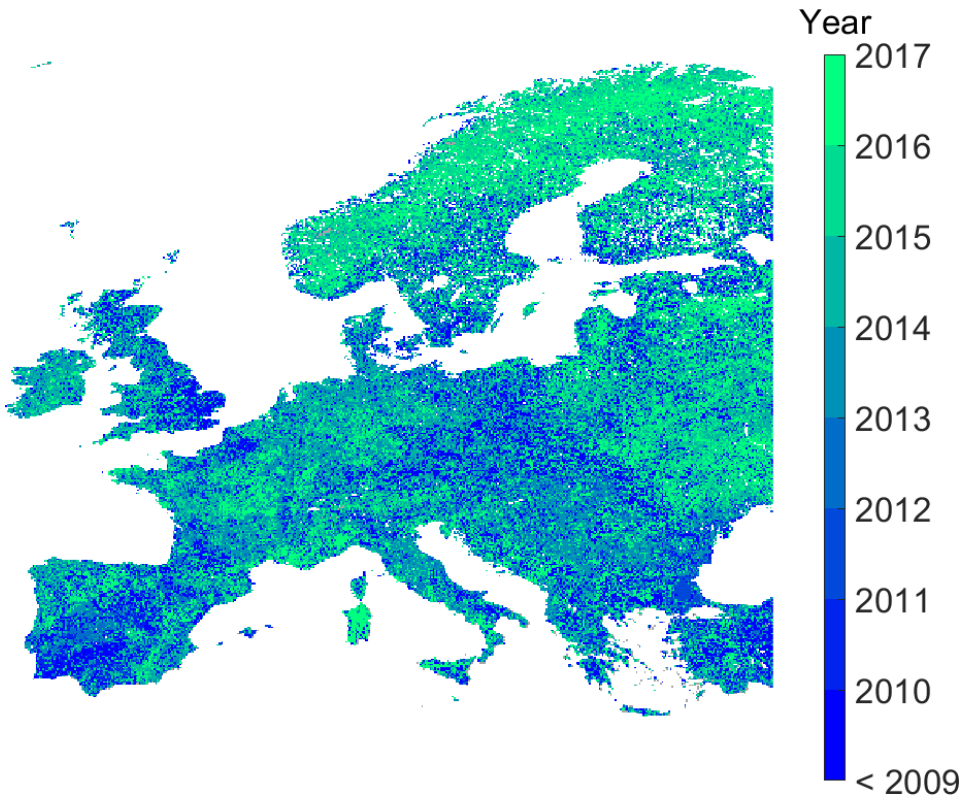


Figure 8.3: Dates of the last break points of each pixel, indicating the start of the last partition of each pixel's trend component.

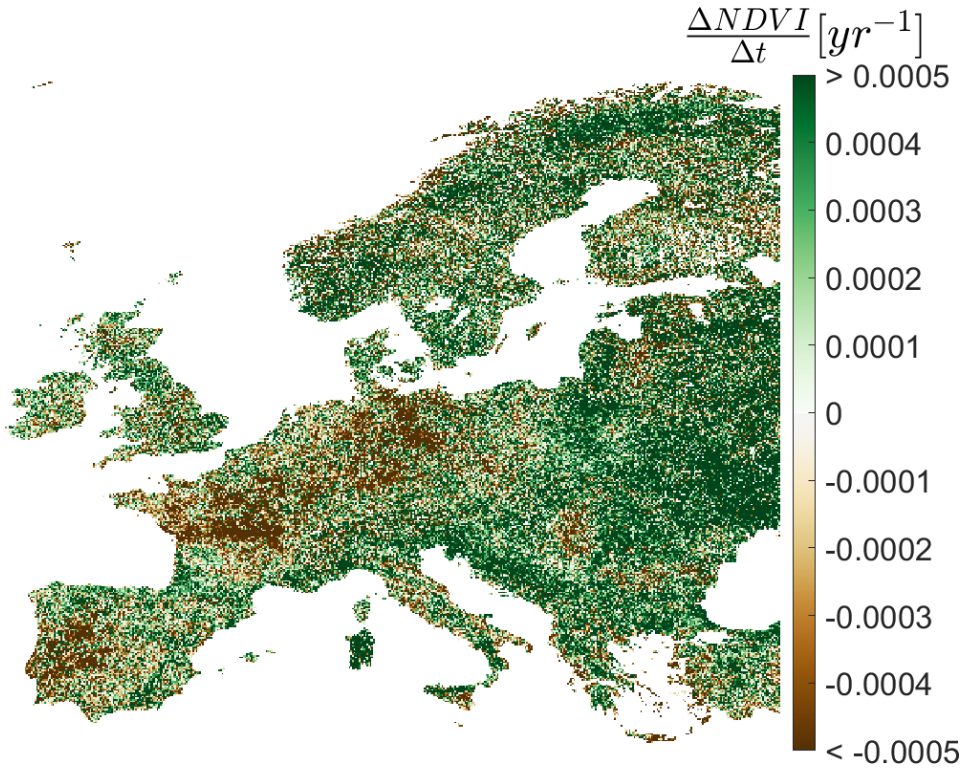
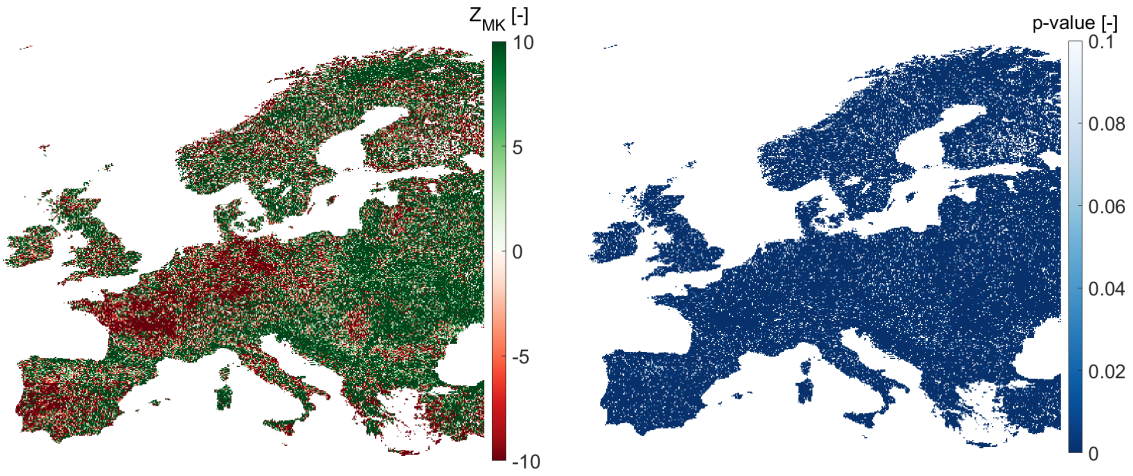


Figure 8.4: The rate of change of NDVI per year, determined over the last partition of the trend component.

Despite the shorter length of these partitioned trend components, statistical significance is high. This is likely due to the inherent nature of partitioning the data resulting in more uniformly increasing or decreasing time series, resulting in confidence in the results over the entire domain.



(a) Mann-Kendall test statistic values.

(b) P-values.

Figure 8.5: Mann-Kendall test statistic and the associated p-values, determined over the last partitions of the trend component.

The sustainability of the trend component, introduced in Section 4.3, is indicative of its future behaviour and is quantified by the Hurst exponent. A high degree of sustainability indicates that the current trend is expected to continue in a similar way and can be used to predict future behaviour. It is therefore critical that a high degree of sustainability is found for the last partitions, as this data is used for forecasting. Furthermore, a well-functioning break point detection algorithm should result in more sustainable partitions compared to the full trend component. The average Hurst exponent values for the full trend component, as well as for only the last partition, are shown in Table 8.1. The average values are calculated over the entire domain, without including the constant sea time series. As expected, sustainability is high for both due to the trend components being largely devoid of noise. An improvement of 5.4% is found as a result of partitioning the data.

Table 8.1: The average Hurst exponent values excluding sea pixels for the full trend component, as well as the last partitions.

	Full trend component	**Last partition**
Average Hurst exponent value [-]	0.9205	0.9702

8.3. Forecast Degrees of Greening and Browning

In the previous two sections the vegetation activity over the current time span of the data was given. This section expands upon these results, by providing the forecast degrees of greening and browning and its statistical significance for the forecast data. This section will therefore show how the vegetation activity is expected to change over the following two years, 2019 and 2020. Due to the large computational cost associated with forecasting, these results are generated at one tenth the resolution of the original data, and each pixel covers 10 by 10 kilometres.

As the forecasting algorithm uses the data for the last partition, provided sufficient data is available, the results shown here are similar to those shown in Section 8.2, however there are some important differences. The degree of browning shown in France, Germany and England has increased both in scope as well as magnitude. Similarly, the rate of greening has reduced in Eastern Europe. On the contrary, browning has transitioned into greening in areas of Iberia and Scandinavia. To the best knowledge of the author no study exists that allows for valid comparison of these results, and as such a discussion such as shown in the previous sections is not possible.

Whilst sporadic noise can be observed in the form of white pixels within the figure, the data is still spatially correlated with a Moran's I value of 0.26, assessed only over the land pixels. This increases the confidence held in the forecast results, as they are produced independently. It should be noted that this value cannot be directly compared to the Moran's I values given previously for the NDVI and time series component data in Subsection 2.5.2 and Section 4.1, as neither the domain nor the resolution are the same.

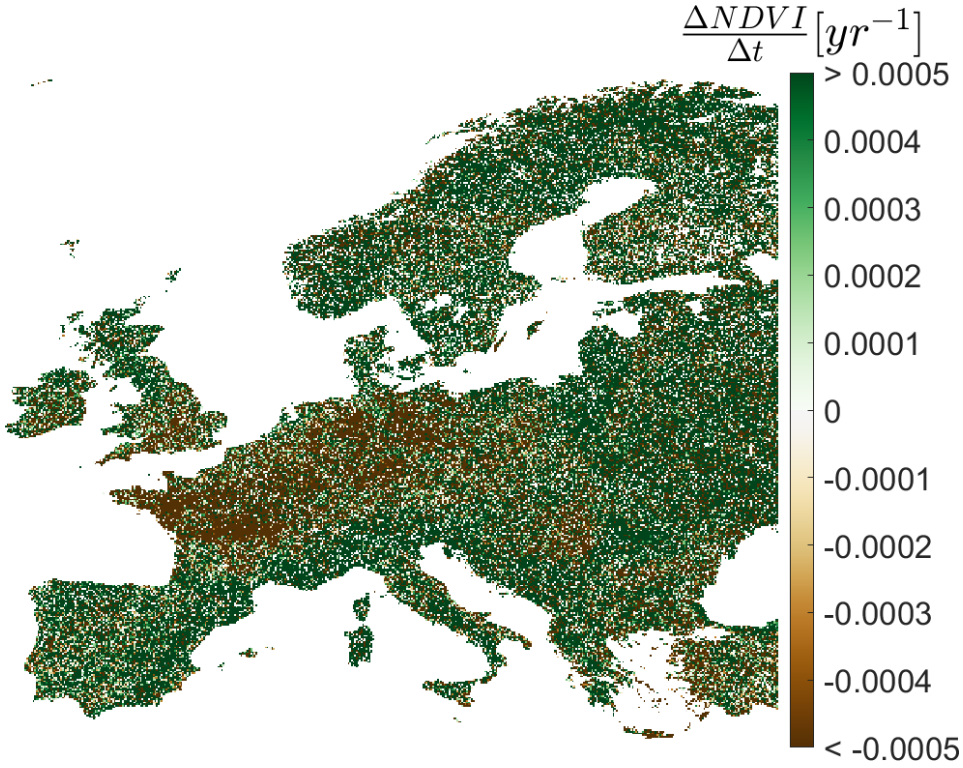
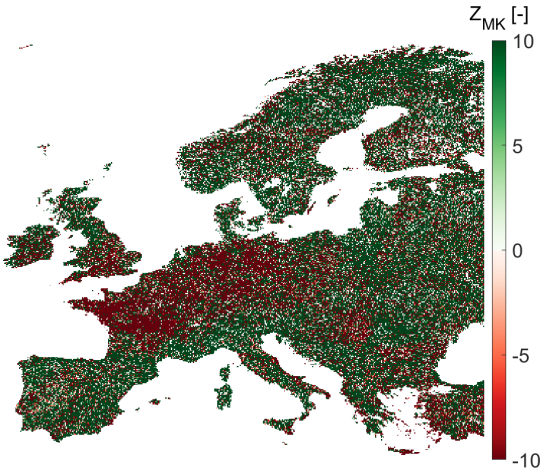
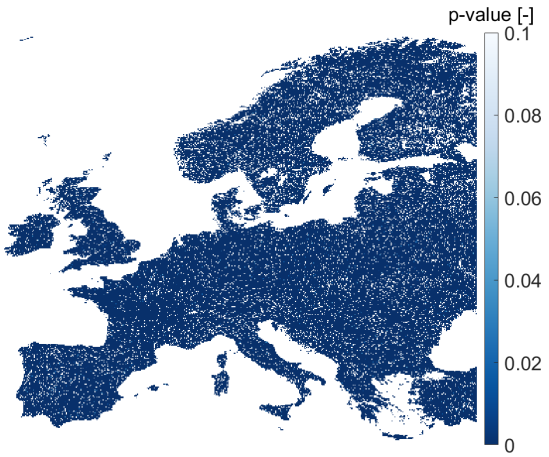


Figure 8.6: The rate of change of NDVI per year, determined over the forecast trend component.

The statistical significance of these results is again shown with the Mann-Kendall test results. Figure 8.7 shows that the statistical confidence held in the results shown above is high, with few time series showing a p-value exceeding 0.10. The percentage of insignificant trends is however greater than for the last partition, shown in Figure 8.5.



(a) Mann-Kendall test statistic values.



(b) P-values.

Figure 8.7: Mann-Kendall test statistic and the associated p-values, determined over the forecast trend component.

8.4. Vegetation Trends in the Netherlands

The preceding sections have shown an analysis of the trend component by showing values obtained over the entire time span of the given data, whether it be the period of 1999 to 2018, the last partitions or the forecasting window 2019 to 2020.

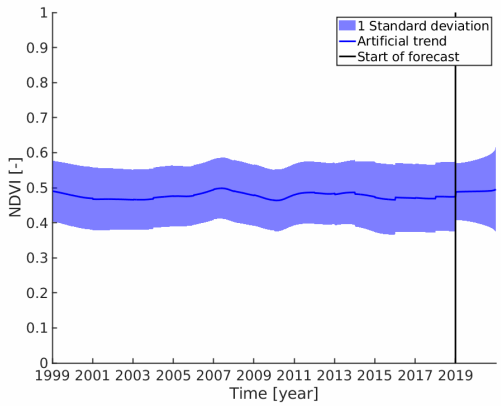
These results therefore do not demonstrate the level of temporal detail that is possible, as differences in vegetation activity on a smaller time scale are lost. As such, aggregated trend and seasonal components - including their standard deviation - have also been computed on a per country and per class basis, combining the original time window as well as the forecasting window. Furthermore, for the aggregated trend components break points were determined and the Theil-Sen estimator has been applied on each of the partitions, producing results similar to the example shown in Section 6.1. As this chapter only concerns itself with the trend component, the results for the seasonal component are given in Section 9.4.

The number of countries within the domain, when multiplied with the number of classes, leads to an amount of information that is impossible to show in this thesis document. These results are instead available upon request, and this chapter restricts itself to only showing the results pertaining to the Netherlands. Due to its small size relatively little data is available. As a result, if qualitative aggregated results are still available, this is a good testament to the results for larger nations.

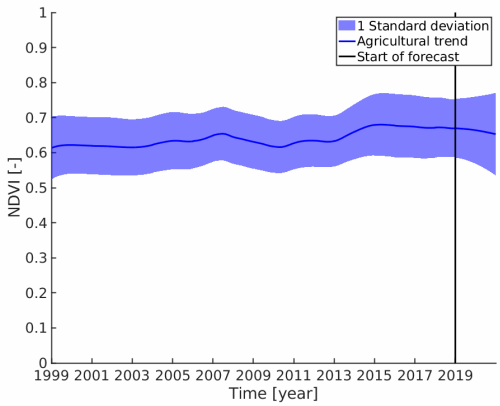
It should be noted however that, as the concentration of non-vegetated pixels within the Netherlands is approximately 0.02%, too little data was available to generate meaningful results for that class. Furthermore, results for the water bodies class are not given as vegetation data for this class is of little value. Whilst results are presented of the remaining artificial, agricultural, grasslands and forests and wetlands classes, the amount of data available still varies greatly. Detailed classification results themselves were presented in Chapter 7, so mention will only be made here that solely the agricultural class exceeds 11% of the data set, with the wetlands class containing the least data at only 0.4%. Finally, classification results are generated on an annual basis, and the aggregation is performed for each year individually. As such, due to changes in classification, discontinuities are present in the results. Furthermore, land cover data was not generated for the two forecast years. Instead, the land cover of 2018 is used for the years 2019 and 2020.

The average trend components and their standard deviations are shown in Figure 8.8. These graphs show that the standard deviation is not directly correlated with the number of samples. One might expect the standard deviation to increase with the number of samples, as one might expect the variability within the class to be greater. Additionally, it can be observed that the number of geographically separated regions is greater. Nevertheless, as the intra-class variability is not equal for all classes, certain classes show a higher standard deviation than others. The results for the class variability within Iberia shown in Appendix F further confirm this conclusion, and whilst results for other nations are omitted from this thesis document, similar conclusions can be drawn from their results. Another conclusion that can be drawn from the standard deviation of the forecast data, is that it increases in width as the forecast progresses. The standard deviation therefore mimics the behaviour that would be expected of the prediction interval, and both are closely related [142].

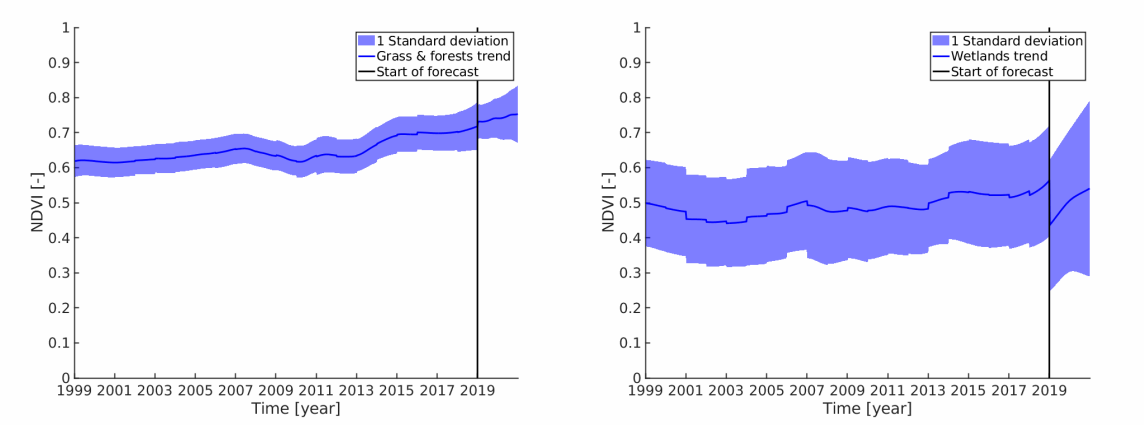
One further result that can be observed, is that the wetlands class shows a shift between the original time window and the forecast. This appears to happen when the number of samples is especially low, which is worsened by the fact that forecasting was performed at one tenth the resolution.



(a) Artificial.



(b) Agricultural.



(c) Grasslands and forests.

(d) Wetlands.

Figure 8.8: The aggregated trend components of the Netherlands for the full time span. The original data and forecast are divided by the vertical black line, and the standard deviation is shown by the blue shaded area.

To provide further information, the trend component was partitioned and the Theil-Sen estimator was applied on each partition. The reader is made aware of the fact that the graphs given in Figure 8.9 are zoomed in compared to the aggregated time series shown above, and changes therefore appear larger than they are in an absolute sense, and thus the magnitudes of change are not the same.

The artificial class appears to show the strongest discontinuities and turning points, whilst the relative amplitude remains within roughly 10%. The discontinuities can be explained by the relatively low classification accuracy of approximately 40%. Due to these discontinuities, it is difficult to distinguish the underlying trend from the noise.

A much cleaner trend can be seen for the agricultural class, which is also by far the most common within the Netherlands. The trend is increasing for the first two partitions, showing a negative shift between the two. The last partition shows a positive shift followed by a negative trend. As a negative shift is followed by a positive shift, the vegetation trend for the years 2009 to 2014 is anomalous compared to that of other years. An explanation for this anomaly could be the financial crisis, which resulted in a sharp fall in agricultural income in 2008 and 2009 and a decline in added benefit of the agricultural sector to the Dutch economy. This negative trend is reversed in 2010, and continued to rise until 2015 [158–160]. Whilst there are many other factors contributing to both the size of the Dutch agricultural sector and the vegetation intensity, this temporary drop in economic activity matches the behaviour of the vegetation time series well.

The grasslands and forests class is not as directly influenced by human activity as the artificial and agricultural classes, and it shows a more consistent greening trend over the entire time period. It should be noted though that this time series too, shows a slight negative shift for the central partition, however it is considerably smaller in relative magnitude. This could indicate that the aforementioned drop in economic activity is not, or only partially, the cause for this intermediate negative shift in vegetation activity for the agricultural trend. It should be noted that countries in the vicinity of the Netherlands do not show the same results.

Similar to the artificial trend, the wetlands time series shows many discontinuities as a result of the varying classification results, as well as the discontinuity between the original time window and forecast data. This makes drawing conclusions more difficult, however the observation can still be made that this time series shows a crest in the vegetation activity over the years 2010, 2011 and 2012, which is also shown in the agricultural and grasslands and forests trends and therefore likely the result of climatic factors that extend beyond any individual class or geographical region.

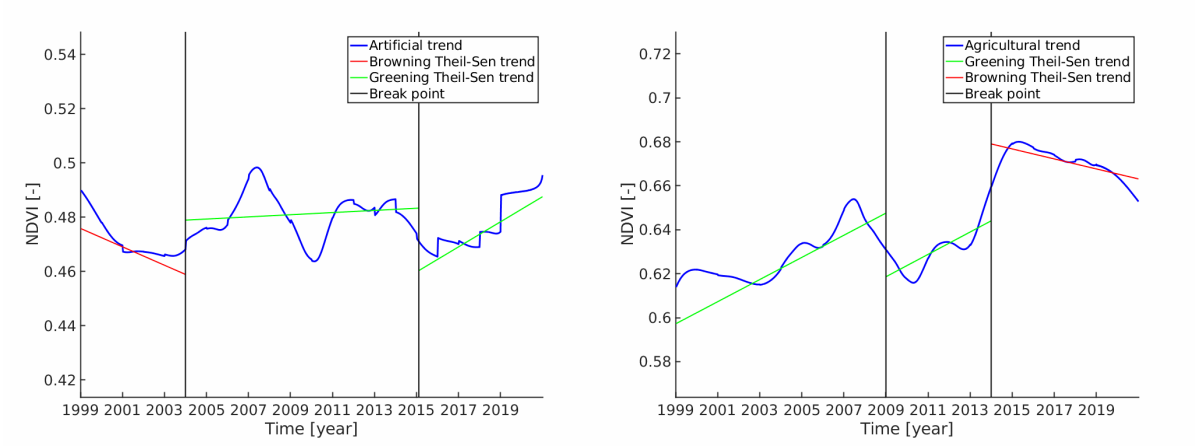
To further investigate the negative shift found for the agricultural and grasslands and forests classes, the correlation between these trends and precipitation and temperature is determined. To this end, annual precipitation and temperature data between 1999 and 2015 and 2017 respectively is used, provided by the Dutch government [161, 162]. As this data is annual, the averaged annual NDVI values within the respective time spans are used. The Pearson's correlation coefficients are given in Table 8.2, and show no clear correlation between either NDVI trends and precipitation. On the other hand, both NDVI trends show clear correlation with temperature. This is in line with prior findings by Schultz and Halpert, and K. Ichii et al. [163, 164].

The negative shift can thus be explained by the drop in temperature of 2.2° Celsius seen between 2007

and 2010, and a rise in temperature of 2.0° Celsius between 2013 and 2014. It should however be noted that the temperature dropped in 2015 and remained constant after that, which is not shown by the increasing grasslands and forests trend. Finally, the crest seen in 2011 can similarly be explained by a sudden increase in temperature for the same year.

Table 8.2: The Pearson's correlation coefficient between the agricultural and grasslands and forests NDVI trends, and precipitation and temperature data.

	Agricultural trend	**Grasslands and forests trend**
Precipitation	0.087	-0.017
Temperature	0.409	0.387



(a) Artificial.

(b) Agricultural.

(c) Grasslands and forests.

(d) Wetlands.

Figure 8.9: The aggregated trend components of the Netherlands for the full time span. The partitions are divided by vertical black lines, whilst the Theil-Sen estimator generated trends are shown in green and red.

9

The Vegetation Seasonal Cycle

In this chapter the growing season metrics as determined from the seasonal component, both for the original time span as well as the forecast time span, are given in Section 9.1 and Section 9.2 respectively. Next, the rate of change in the length of the growing season is discussed in Section 9.3. Similar to Section 8.4, the aggregated seasonal component for the classes within the Netherlands are given in Section 9.4.

As was already discussed in Section 6.2, the growing season metrics are likely to contain errors on the order of days, due to limitations of the underlying raw data and the applied methodology. To mitigate these issues, mean values determined over the respective time periods are given. This however does not extend to the resulting slopes, which is elaborated on in more detail in the pertaining section.

Finally, whenever possible, comparisons are drawn between the results found in literature and the results obtained by this study, however limited information was found for the assessed domain and time span. Furthermore, as the focus of this thesis remains a statistical and not a biological analysis, possible driving factors behind these results are not given and instead a few of the most noteworthy findings are discussed.

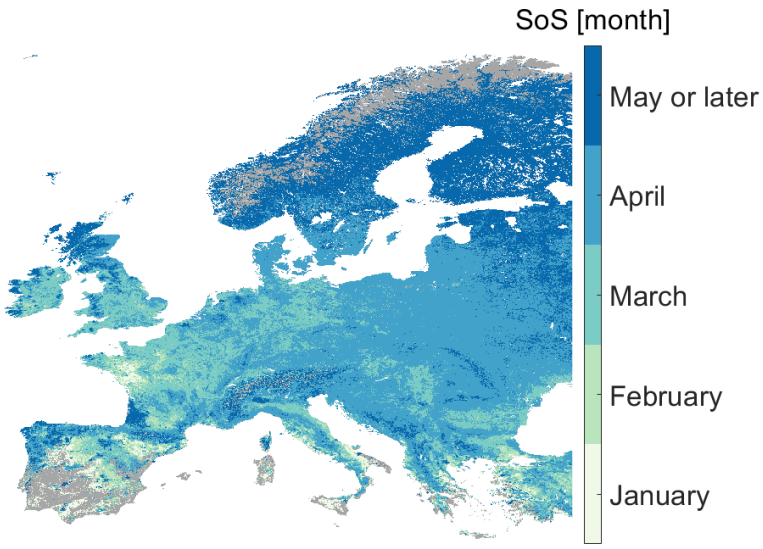
9.1. Growing Season Metrics

In this section the mean values of the annual start of the growing season (SoS), end of the growing season (EoS) and length of the growing season (LoS) are given in Figure 9.1. As these figures are highly interrelated, they are discussed as a whole rather than individually.

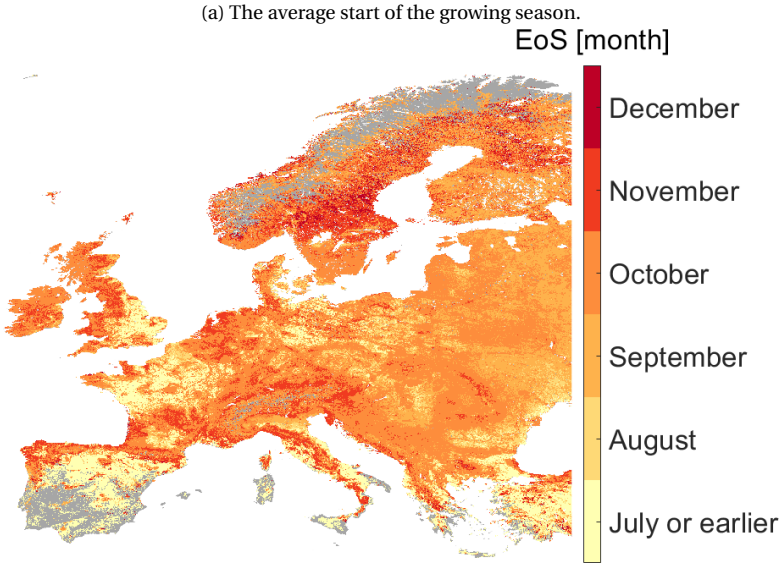
Overall, there is a relation between the latitude and the start and end of the growing season, where generally the higher the latitude the more postponed the growing season is. More regional phenomena can be seen around the coasts of Italy and Greece, where the SoS and EoS are earlier than the areas further inland. Similarly, but to a wider extent, the coastal regions around the English Channel show an earlier SoS and EoS.

The mountainous regions of Europe were already discussed in Chapter 8, as they showed a distinct difference to their surroundings in their trends. From the length of the growing season one can again see that it is distinctly lower than their surroundings, most especially noticeable for the Alps, Pyrenees and Scandinavian mountain range. The latter is however more difficult to distinguish from its surroundings, as fewer values could be determined for this region and the northern latitudes generally show a short growing season of roughly 4 months, similar to those expected for these regions [77].

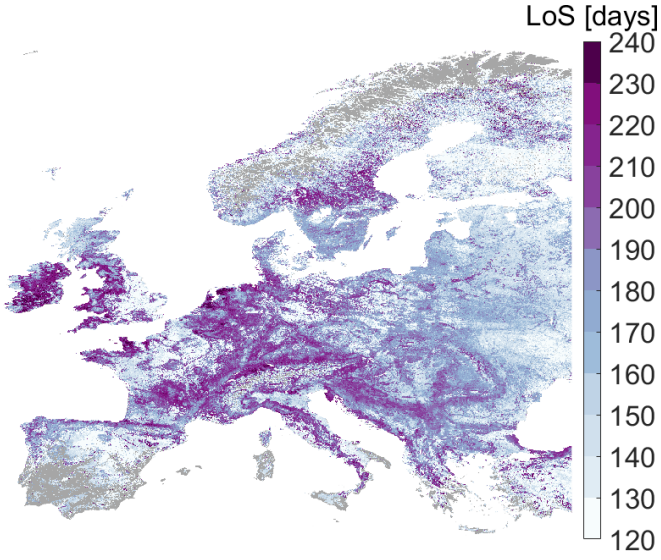
The class-averaged results over the entire European domain are shown in Table 9.1. As expected, the non-vegetated class has the shortest growing season, and the artificial class has the longest growing season. This is a result of the air temperature being higher in urban areas due to the heat island effect [165]. Furthermore, a positive trend has been determined between the degree of urban land cover and the end to the growing season. Urban areas thus experience a growing season that ends later than the surrounding areas, and is therefore longer [166].



(a) The average start of the growing season.



(b) The average end of the growing season.



(c) The average length of the growing season.

Figure 9.1: The average growing season metrics from 1999 to 2018. Regions where the metrics could not be determined due to a lack of seasonal vegetation cycles are shown in grey.

Table 9.1: Class-averaged growing season metrics for the European domain between 1999 and 2018.

	SoS [day of year]	EoS [day of year]	LoS [days]
Artificial	91	267	176
Agricultural	86	239	153
Grassl. & Forests	108	258	150
Non-vegetated	31	65	34
Wetlands	127	277	150

9.2. Forecast Growing Season Metrics

The last section showed the results obtained for the full time span of the used data. Here, the results are shown for the forecast seasonal component, with Figure 9.3 showing the average growing season metrics over the two forecast years. As there are no break points in the seasonal component, the data is not partitioned and the entire seasonal component could be used to train the forecasting algorithm.

The fact that the seasonal component is relatively constant compared to the trend component was already stated in Section 5.4, as the likely reason why higher quality forecasts were obtained. It is therefore unsurprising that the figures differ less than the figures shown in Chapter 8 for the trend component.

The end of season plot shows similar behaviour to the one obtained with the original time span, however important differences are present in the start and therefore length of season plots. Generally, the growing season has lengthened in the northern and far eastern parts of the domain. Despite the analysis being performed on different time series components, this lengthening coincides well with the regions that transitioned from greening to browning. Scotland also shows a longer growing season, and whilst a higher rate of greening was similarly found, it is less evident as no transition from browning to greening occurred there. In addition, the areas which already showed a longer growing season compared to their surroundings have become more distinct.

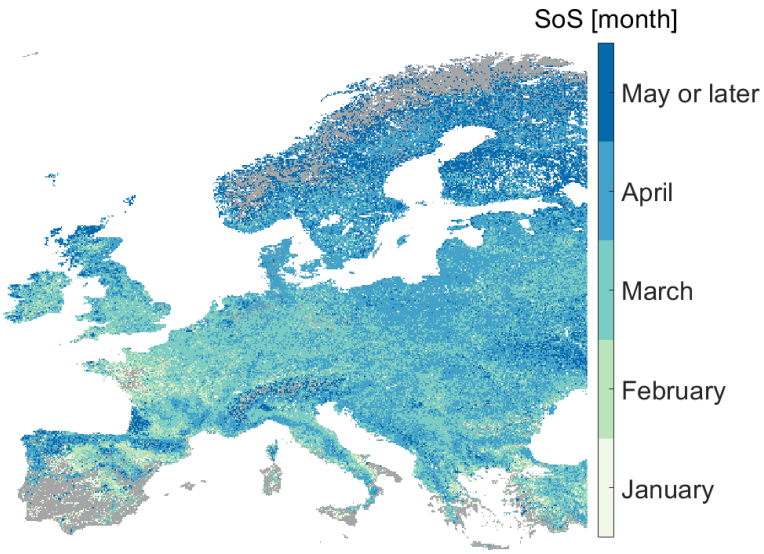
The class-averaged results are shown in Table 9.2, for which the land cover data of 2018 was used. The absolute and relative differences between the metric values for the forecasting window and regular time span are shown in Table 9.3. On average, the growing season starts and ends earlier, with a moderate change in the total length of the growing season. One hypothesis would be that the temperature required by vegetation to start the growing season is attained earlier due to global warming. Finally, whilst the figures for the EoS showed no apparent change, this change is still present in both an absolute as well as relative sense.

Table 9.2: Class-averaged growing season metrics for the European domain of the forecast data.

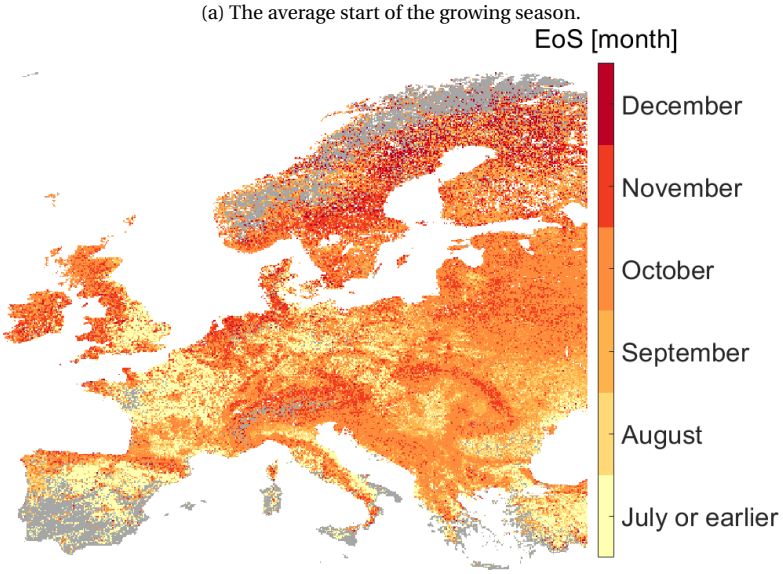
	SoS [day of year]	EoS [day of year]	LoS [days]
Artificial	77	246	169
Agricultural	76	228	152
Grassl. & Forests	90	249	159
Non-vegetated	26	59	33
Wetlands	105	264	159

Table 9.3: Difference in class-averaged growing season metrics for the European domain between the forecast data and the data obtained for the original time span, both relative as well as absolute.

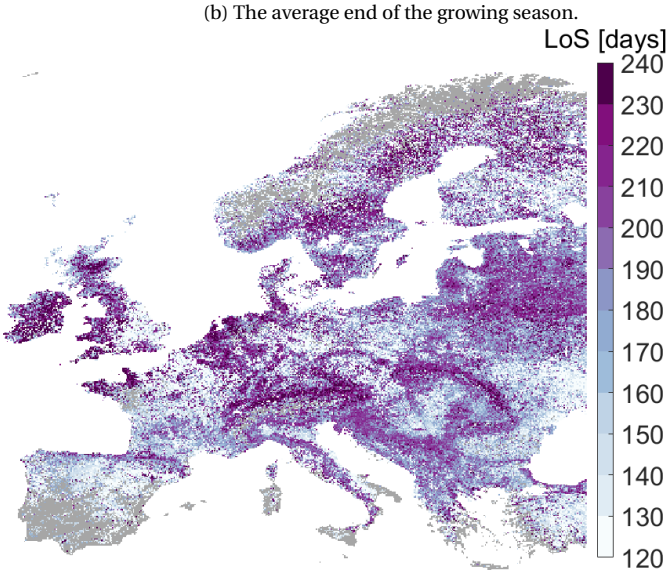
	Δ SoS [Days / %]	Δ EoS [Days / %]	Δ LoS [Days / %]
Artificial	-14.36 / -15.77	-20.71 / -7.76	-6.34 / -3.61
Agricultural	-10.14 / -11.76	-11.73 / -4.90	-1.61 / -1.05
Grassl. & Forests	-18.68 / -17.23	-8.82 / -3.42	9.87 / 6.60
Non-vegetated	-5.03 / -16.24	-6.83 / -10.45	-1.80 / -5.23
Wetlands	-22.03 / -17.32	-13.39 / -4.83	8.64 / 5.76



(a) The average start of the growing season.



(b) The average end of the growing season.



(c) The average length of the growing season.

Figure 9.3: The average growing season metrics for the forecasting window. Regions where the metrics could not be determined due to a lack of seasonal vegetation cycles are shown in grey.

9.3. Rate of Change of the Growing Season

In the previous sections the average results were given for the original time span, as well as for the forecasting window. This section displays the rate of change of the length of the growing season during the twenty year time span between 1999 and 2018. Section 6.2 stated that neither the methodology, nor the input data are sufficiently accurate to capture the slow rates of change that are expected, which are of the order of 0.1 - 1 day/year [10, 35, 37]. Nevertheless, the results are included as they show an interesting correlation with the results shown in Section 8.1 for the vegetation greening and browning.

To obtain the rate of change the Theil-Sen estimator is used again, as autocorrelation is also expected to be present in time series of the growing season metrics. The time series are interpolated to daily values, similar to the method used by T. Park et al. [37]. The obtained rate of change for the length of the growing season is shown in Figure 9.5. This figure shows that for the regions that transitioned from browning to greening, the growing season shows a strongly positive trend and the mountainous regions of Europe are clearly distinct in the plot. The rest of Europe shows more muted rates of change, both positive as well as negative. The lengthening in Scandinavia was also found in other studies - albeit for different periods of time [10, 19, 35].

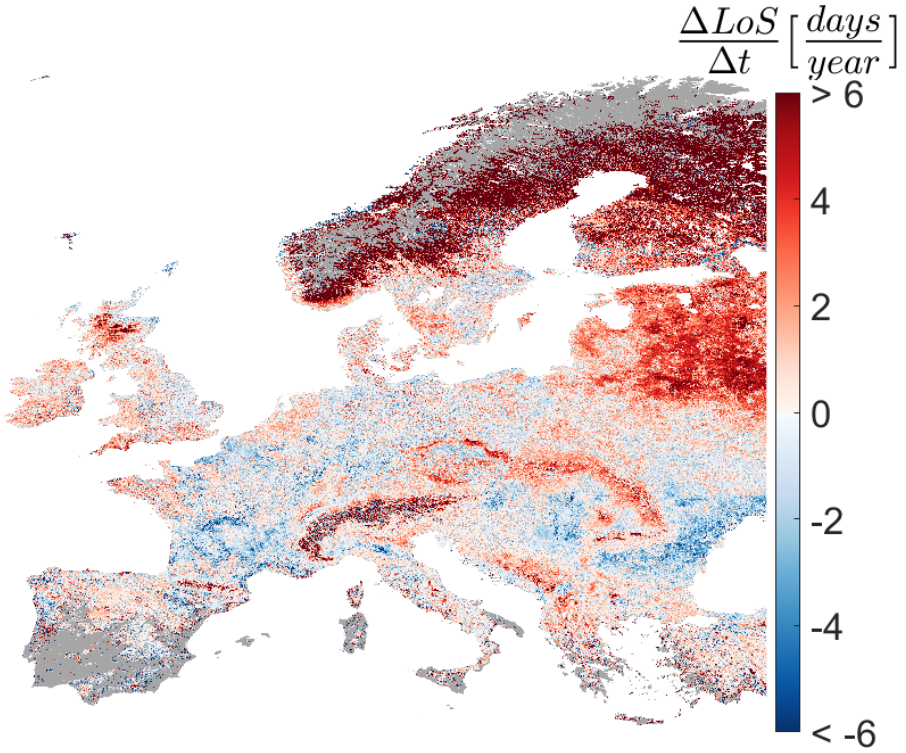


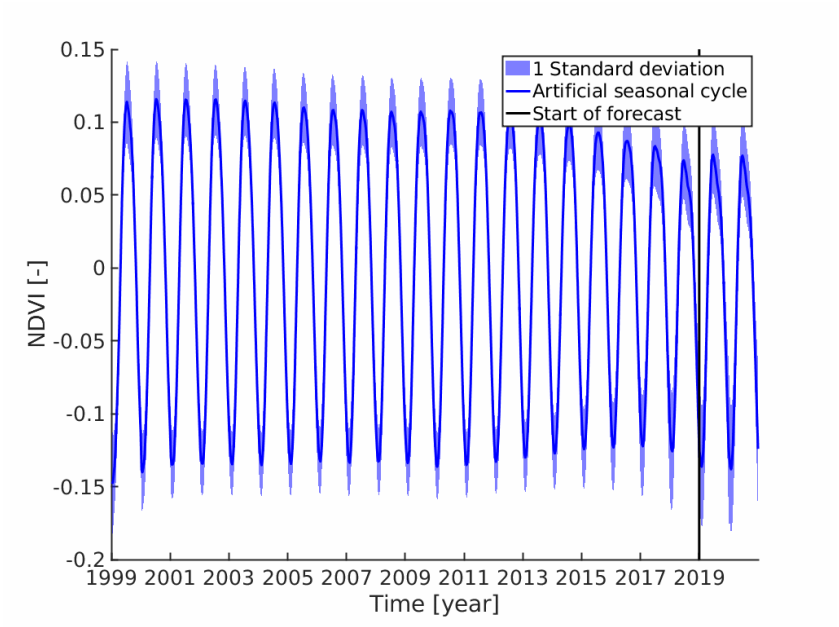
Figure 9.5: The rate of change of the length of the growing season per year, determined between 1999 and 2018. Regions where the metrics could not be determined due to a lack of seasonal vegetation cycles are shown in grey.

9.4. The Vegetation Seasonal Cycle in the Netherlands

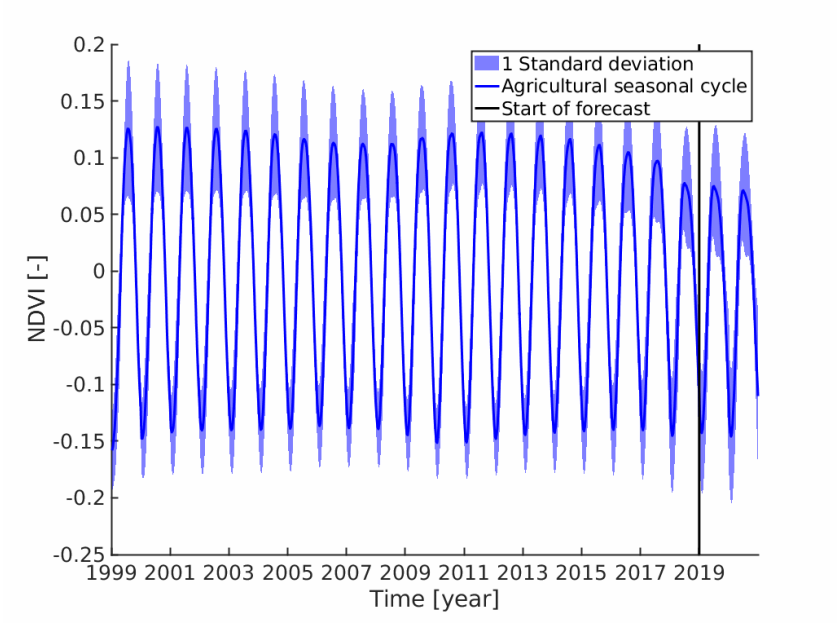
In this section the aggregated seasonal components for the artificial, agricultural, grasslands and forests and wetlands class of the Netherlands are shown. The choice to present only the results for this country and these classes is explained in Section 8.4.

Unlike the aggregated trend components, these seasonal components show little variation throughout the time series and the standard deviation is less affected by the forecasting window. The seasonal amplitude of the wetlands class is greater than that of the other classes, with small differences between the other three classes. Additionally, all classes show the same seasonality, with the maximum occurring on the 19tth or 20th time step of the year for all four classes. The average seasonal component shows a reduction in amplitude near the forecasting window, which is then carried through by the forecasting algorithm and is most obvious for the wetlands forecast. Belgium, Germany, France and to a lesser extent Luxembourg all show this decrease

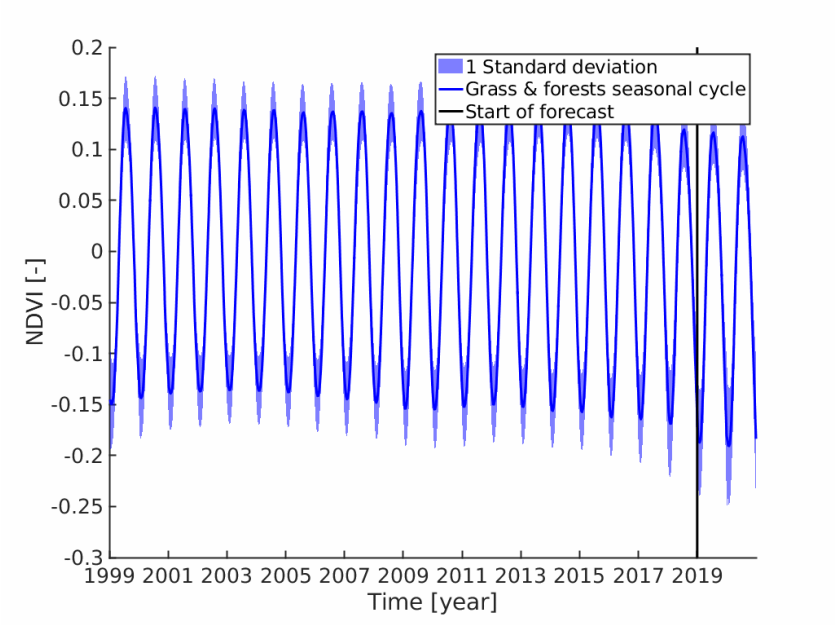
in amplitude. The same results are not obtained for many of the other countries, indicating that it is likely not a methodological artefact. L. Xu et al. studied the vegetation seasonality in the boreal region, which encompasses the Netherlands, and found it to be approximately constant between 1995 and 2007, after which it showed a negative trend until the end of their assessed time span in 2010 [167]. They do not provide statistically significant results for the Netherlands, making it difficult to draw direct comparisons. It is however interesting to note that they found a similar pattern, albeit roughly ten years sooner.



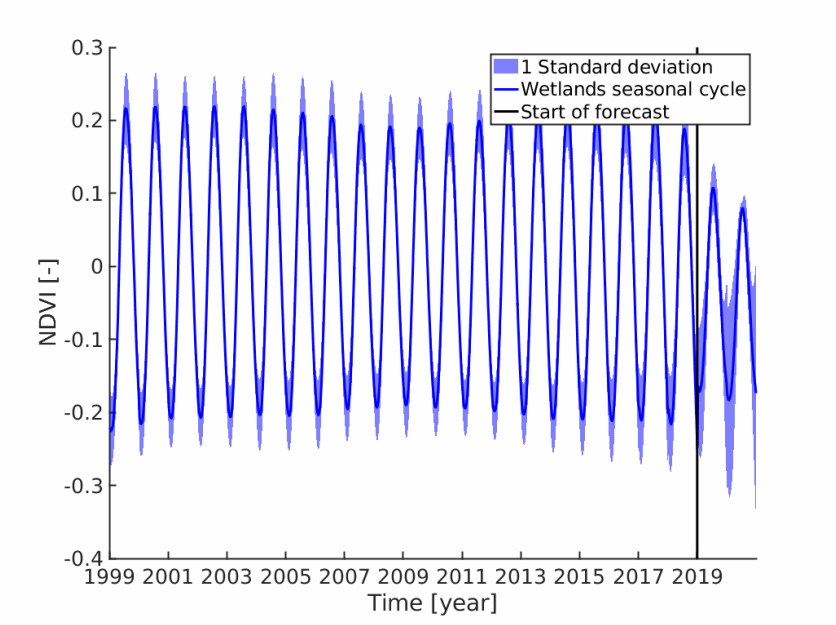
(a) Artificial.



(b) Agricultural.



(a) Grasslands and forests.



(b) Wetlands.

Figure 9.6: The aggregated seasonal components of the Netherlands for the full time span. The original data and forecast are divided by the vertical black line, and the standard deviation is shown by the blue shaded area.

IV

Conclusion

10

Summary and Conclusions

This MSc thesis project analysed SPOT-VGT and PROBA-V NDVI data [54] of Europe over the last twenty years in order to determine how vegetation metrics have changed over this time period and are expected to change in the future. A similar time span was used by most studies that have been conducted previously, but the geographical extent varies more significantly, from areas of countries to the entire globe. Several studies have already been conducted on Europe, however to the best knowledge of the author no study exists that studied Europe's vegetation in the same level of detail as this thesis project. As such, the main research question that the project aimed to answer is: *'Which temporal trends can be found for vegetation in Europe between 1999 and 2018 using remotely sensed data?'*. Other aspects that the research investigated are the use of NDVI data for land cover classification, determining whether NDVI trends are suitable for forecasting and what those forecasts indicate. To answer these research questions, a methodology was devised consisting of five main steps: data pre-processing, classification, determining the time series components, forecasting and time series analysis. This methodology was then applied to the full domain to provide the necessary results.

10.1. Summary of the Methodology

The first aspect of the methodology is that of data pre-processing, which is used to reduce the noise in the data, as well as fill in data that is missing, and therefore increase the information that can be gathered from it. Despite the measures the proprietors have taken to reduce the noise in the data, noise is still present and noise reduction is applied almost ubiquitously in research. The first step of the noise reduction algorithm is to substitute pixels marked as invalid using temporal interpolation. As this insufficiently reduced the noise in the data set, existing noise reduction methods were investigated and a comparison between the most promising methods was performed. The iterative Savitzky-Golay filter devised by J. Chen et al. [68] performed the best, and was thus used to reduce the noise in the data set further.

After the noise has been reduced, the data can be classified to determine which group a pixel belongs to. The classification tree was selected due to its applicability to the remotely sensed data set and prior use for vegetation classification. As this is a supervised methodology, training data is required. For this project, the Corine Land Cover database [95] was used which covers the majority of the analysed domain.

The basic classification tree was unable to classify rarer classes, necessitating a more elaborate classification methodology. To create a more balanced training data set, the data was oversampled. To mitigate the associated overfitting and improve the results further, the ensemble technique bagging was applied. With this technique, a multitude of classification trees are grown and majority voting is applied to arrive at the final result.

An investigation into the data itself led to the conclusion that it is not well suited for the classification of larger domains. By segmenting the data into smaller sections, both the intra-class variability and inter-class separability were improved drastically. The average Jeffries-Matusita distance was increased by 0.59 and the average Euclidean distance was reduced by 40.0%. Dividing the domain into smaller segments is therefore seen as an essential step in the classification algorithm.

These steps were sufficient to generate accurate spatial results, but for coherent annual classification results they were insufficient. To decrease the temporal variance, a second layer of majority voting in the form of overlapping segments was introduced, as well as temporal smoothing.

The noise-reduced data was not only used for land cover classification, but also to provide analytical results on its temporal behaviour. To this end, the time series were decomposed into trend, seasonal and remainder components using STL decomposition. The trend and seasonal components are analysed separately, with the remainder component receiving no further attention.

As the trend component cannot be expected to show consistent behaviour over the time period, its break points were determined by using a modified version of the partial structural change model as described by Bai and Perron [133]. By including iteration, the algorithm is capable of determining the most optimal combination of break points regardless of the order in which they were detected, and a modest improvement in performance of approximately 7% was obtained.

To determine how vegetation activity is expected to change in the future, the ARIMA and SARIMA forecasting methods were applied on the trend and seasonal component data respectively. As the component data proved ill-suited for the aforementioned models, the Box-Cox transformation is applied to improve the characteristics of the time series.

Despite these efforts, the forecasting window is limited to two years to limit the forecasting errors for the trend component. This is in no small part due to the forecasting methodology not taking the possibility of break points occurring in the forecasting window into account, due to the added complexity and only moderate prior success.

Finally, a methodology was described to analyse these time series components. As separate results are desired for the trend and seasonal components, a different method is applied to each.

To determine the degree of vegetation greening and browning and its statistical significance, the Theil-Sen estimator and Mann-Kendall test were applied to the trend component data. Further, to determine the sustainability - and thus suitability of the trend component for forecasting - the Hurst exponent was computed.

The seasonal component was used to determine the growing season metrics by applying the $\text{midpoint}_{\text{pixel}}$ method. As the seasonal component data varies strongly over the domain, this method was made more robust by including the necessity for the start and end to be before and after the annual maximum, as well as show a positive and negative rate of change respectively.

10.2. Summary of the Results

After developing this methodology, it was applied on the full European domain by using the Cartesius supercomputer. The results are split into three parts, the first being the generated land cover and the second and third the results obtained for the trend and seasonal component respectively.

The classification methodology was applied on the full data set, and annual land cover maps were generated for the years 1999 to 2018. These maps showed that the classifier manages to capture all the major patterns found in the Corine ground truth data. Furthermore, even for regions where the ground truth data is missing and suitable training data cannot be used, the classifier manages to capture even minor details accurately.

Quantitatively, the total classification accuracy exceeds 80%, due to the high accuracies obtained for the common agricultural and grasslands and forests classes. The accuracy of the rare artificial and wetlands classes is lower at 41.7% and 30.4% respectively, however at a level comparable to that of reference studies. Interestingly, for these rarer classes the precision is greater than its accuracy. This indicates that the classifier is conservative when classifying rarer classes.

Aside from these results for the full domain, results specific to the countries and largest cities within the domain are also given. The generated land cover was divided on a national basis, and a high urbanisation was found in Eastern Europe, a slight decrease and increase of farmland in Western Europe and Eastern Europe respectively, and forestation in north-western countries. The majority of the assessed cities grew in size during the twenty year time period.

To determine the rate of change of vegetation activity, the trend component was analysed. The results showed a consistent, low rate of greening over most of Europe and NDVI increased with approximately $2 \cdot 10^{-4}$ per year. Exceptions are the mountainous regions, which are browning more rapidly, and show a change in NDVI on the order of $-5 \cdot 10^{-4}$ per year, and North-Eastern Europe which is browning at a lesser rate and NDVI has decreased with $-2 \cdot 10^{-4}$ per year.

The rate of change in vegetation activity was also determined using the final partitions of each time series, to provide information that is both more recent in time and shows the basis used for forecasting. The areas that are browning over the full time period have nearly all transitioned to greening, with areas of France, Germany and Spain having transitioned from greening to browning. By applying partitioning, it was further shown that the Hurst exponent value was increased by 5.4% and is highly sustainable. This indicates that the trend component data can be used for forecasting, and its suitability is improved due to the break point detection algorithm.

As the forecasting algorithm uses the last partitions of the trend whenever possible, similar values are predicted for the future. Differences are that the browning in France and Germany has increased, and browning has transitioned to greening in Spain and parts of Scandinavia. The statistical significance of the trends for forecasting window, as well as for the twenty year time span and the last partition, is high throughout the domain.

The time series analysis discussed so far does not include the full level of temporal detail, and differences in vegetation activity on a smaller time scale are lost. To provide this detail, aggregated components were generated for each country and class individually. For both the agricultural and grasslands and forests classes, a notable negative shift in the trend is found for the central portion of the time series, stronger for the former than the latter. This is likely a result of the reduced temperature over this time period.

The final results that are given are those obtained for the seasonal component. To provide more information regarding the vegetation phenology in Europe, the start, end and length of the growing season for both the original as well as the forecast seasonal component were determined. Generally speaking, the growing season is more postponed the greater the latitude. Similar results were found for the mountainous regions of Europe, which display a later and shorter growing season compared to their surroundings.

A notable change shown by the forecast data is a lengthening of the growing season for the regions that previously showed a transition from greening to browning. Further, on average the growing season starts and ends earlier, with a moderate change in the length of the growing season.

Finally, the growing season lengthening shows correlation with the areas that transitioned from browning to greening. It should be noted however, that neither the applied methodology nor the used data is well-suited to capture trends in the growing season metrics.

Compared to the aggregated trend components, the seasonal components shows less variability, both within the original time span as well as the forecasting window. All of the assessed seasonal time series showed a repeating maximum just after the year's halfway point, as well as a decrease in amplitude for the later years.

10.3. Conclusions

Now that the thesis has been summarised, a conclusion can be drawn as to whether or not the used methodology and obtained results are sufficient to answer the research questions the thesis project aimed to answer. The main research question was:

Which temporal trends can be found for vegetation in Europe between 1999 and 2018 using remotely sensed data?

This question was answered by calculating the trend component of each time series using STL decomposition. Additionally, by detecting the break points of these trend components they could be divided into partitions. This allowed for the degree of greening and browning to be quantified not only for the full time span, but also after the last break points which is more representative of the current state of vegetation. As the generated trend component shows high spatial autocorrelation and statistical significance, the author is confident in the results.

Aside from the main research question, three sub questions were also answered. These questions are:

- *Are NDVI time series suitable for forecasting?*
- *How is vegetation expected to evolve into the future?*
- *Can NDVI data be used to create a representative, continent-spanning land cover data set?*

The first and second subquestion regard the forecasting aspect of this study. Testing showed the forecasting algorithm to be accurate and robust for the seasonal component, and the growing season metrics could be determined over the forecasting window without issue. Due to the turning points within the trend component, it is less suitable to be forecast. Nevertheless, the high sustainability determined for the trend component indicates that existing trends are expected to continue into the future, and the fact that sustainability is increased for the last partition increases the confidence held in the break point detection algorithm. Additionally, the observed spatial autocorrelation of the results and high statistical significance of the degree of greening and browning provide further confidence in the forecast rates of greening and browning.

The final research question is successfully answered by the obtained classification results that are accurate in both a quantitative as well as qualitative sense. The specific issue of classifying the full, continent-spanning domain was addressed by dividing the data into smaller sections which greatly increased the separability between classes, and decreased the variability within each class.

All the research questions have thus been successfully answered, and a number of scientific contributions have been made, as to the best knowledge of the author no prior study has generated these results or implemented these methodological steps.

The generation of annual land cover results of Europe, its countries and cities provides detailed information regarding land cover change, both spatially as well as temporally. In terms of the classification methodology, sweeping sectioning greatly increased the size of the domain that could be accurately classified, reduced the variance of the classified land cover and removed the discontinuous borders between the segments. Additionally, a supervised classification method was used successfully to classify regions where the required training data was not available, and the results obtained for these regions were shown to be of high quality.

Previous studies exist that determine the degree of greening and browning over the European domain, but they generally lack statistical significance for large parts of the domain, or provide outdated results at a significantly lower resolution. The decomposition of the time series into their trend and seasonal components resulted in a high statistical significance. Additionally, by analysing only the European continent these results could be generated at a 1 by 1 kilometre resolution and smaller spatial details such as the browning of the Pyrenees could be determined.

Finally, whilst growing season metrics and forecasting results exist, they restrict themselves to specific types of vegetation or parts of the European domain. This study did not restrict itself, and these findings were produced for all assessed classes for the entire European domain.

11

Future Recommendations

At the end of this thesis project, aspects remained that could improve the methodology and results, but were not included and remain the object of future research. These led to the following recommendations, which are split into sections for the classification, time series analysis and forecasting methodology in Sections 11.1, 11.2 and 11.3.

Additionally, it is noted that apart from classification the methodology can be readily applied to other domains, as the SPOT-VGT and PROBA-V NDVI data are available globally for the entire assessed time span. The limited extent of the Corine database restricts the possible extension of the classification results, which is addressed in more detail in Section 11.1.

11.1. Recommendations for Land Cover Classification

Whilst the classification methodology produced more than adequate results, several improvements can still be envisioned to the methodology to increase the performance as a whole, and more specifically the accuracy obtained for rarer classes.

- Use higher resolution NDVI data. Currently, the Corine land cover data is at a tenfold finer resolution than the Copernicus NDVI data. During the creation of the classification methodology it became evident that the classification performance is increased drastically when a finer resolution is used, in particular for the rarer classes. As such, further improvements are expected when the methodology is applied to finer data.
- Investigate more complicated sampling methods, i.e. base it on samples that the method first fails to correctly classify or base it on the distribution in feature space. Currently, no special attention is paid to the rare cases that exist within the classes. By leveraging the clusters within the feature space that these cases inhabit, these can be classified more accurately.
- Include spatial parameters in classification, akin to object-based classification, as the distribution and size of objects is expected to hold valuable information, in particular for the artificial class. From the resulting land cover maps it can be deduced that a large fraction of the artificial class takes the form of 'noise' within the land cover map. These samples are both rare, as well as distinct from their neighbours. Both of these attributes are useful information that can be deduced by dividing the data into objects.
- Compare the results to the ones obtained when more parameters are included to teach the classifier with, for instance temperature and precipitation or another vegetation parameter such as LAI. This would indicate the change in performance when using only NDVI, and quantify the need for extra variables to be included.

Recommendations are also made to extend the land cover results obtained thus far - both spatially as well as temporally.

- Spatially extend the land cover results by classifying a larger domain. The classification methodology proved capable of classifying areas where training data was not present, but in order for this to be the case the available training data must be representative. To classify a considerably larger domain, a different ground truth data set must be used, such as the Copernicus Global Land Cover database [65]. It is at a 100 by 100 metre resolution, and distinguishes between 23 types of land cover - predominantly vegetation types - at a total accuracy of 80%. Currently the global data set is available only for the year 2015.
- The land cover results can be temporally extended in a straightforward way by simply classifying annual NDVI results of the analysed domain as they are made available, and as of the second of January 2020, the 2019 NDVI data is complete [54]. The features that are currently used to describe the time series preclude the production of classification results in real-time. Instead, a different set of features would have to be used. One such solution could be to fit polynomials or other functions to the time series obtained thus far, and use the coefficient values.

11.2. Time Series Analysis Recommendations

With regards to the time series analysis, two recommendations can be made. The first pertains to the break point algorithm, which deserves further attention, and the second to the growing season metrics which are at this point insufficiently accurate.

- Investigate the iterative break point algorithm in more detail. Whilst the current algorithm showed an improvement over the original model, its exact capabilities have not been investigated in detail and the local minima found in the sums of squared residuals data have not been exploited to their full extent. The latter can in theory not only be used to increase the accuracy of the algorithm, but in some circumstances also greatly reduce the computational time.
- The methodology used to determine the growing season metrics proved insufficiently accurate to determine the rates of change of said metrics. Currently, the existing methodologies show little sophistication and do not take detailed phenological behaviour into account. A more accurate method is thus needed, which uses the available information to a greater extent.

11.3. Recommendations for Forecasting

For the forecasting aspect of this thesis, two recommendations can be made. The first is related to the data used to produce the forecasts, and the second is an improvement to the used methodology.

- Produce NDVI forecasts based on multiple data types. Currently, the NDVI forecasts are created using only the NDVI time series. Other climatic data, such as precipitation and temperature data, is however expected to hold valuable information regarding the future state of vegetation. In fact, the NDVI trend and temperature data of the Netherlands was found to be correlated in Section 8.4, and prior studies on vegetation forecasting described in Section 5.1 included other data sets.
- An improvement that can be made to the forecasting algorithm, is to include the possibility of break points in the forecasting window. Unexpected turning points were the biggest source of error for the trend component forecasts, and with an effective algorithm the forecast quality can be improved considerably. Currently, insufficient time was available for attempts to integrate such an algorithm. Furthermore, to the best knowledge of the author even the most advanced forecasting methodologies that include this feature have achieved only limited success.

V

Appendices

A

Autocorrelation of Invalid Pixels

In this appendix the autocorrelation coefficients are given for the European data set, as well as the Swedish subset, to prove that temporal interpolation described in Section 2.3 is suited as the invalid pixels are almost entirely uncorrelated.

To determine the temporal autocorrelation of the invalid pixels, binary time series were generated for each pixel with a value of 1 if the pixel is invalid, and 0 if not. High autocorrelation values will be returned in the case that successive pixels are invalid and vice versa. Figure A.1 contains the mean absolute autocorrelation coefficient value for time lags up to a year, as well as the standard deviations. From this figure it is evident that the bounds for statistical significance represented by the blue lines are not breached by the mean values.

The alternative is represented by Figure A.2, where the same plot is created of the Swedish subset. Here, long-term snow cover and poor radiometric quality is expected to result in correlation. This suspicion is confirmed, as several time lags are now correlated. As this is only an issue for a limited part of the data set, and no viable vegetation data is present for these periods, this is seen as acceptable.

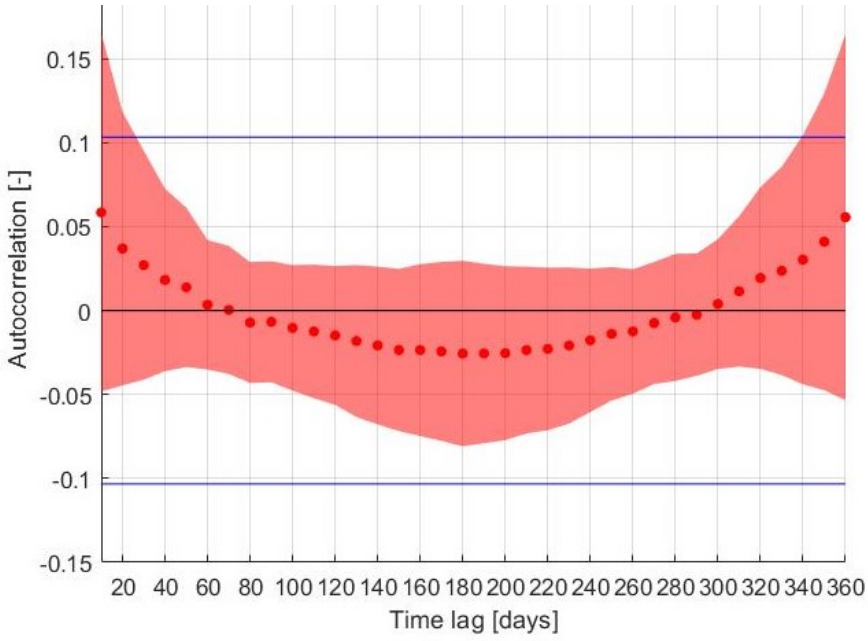


Figure A.1: Average autocorrelation coefficient, taking all types of invalid pixels into account in Europe.

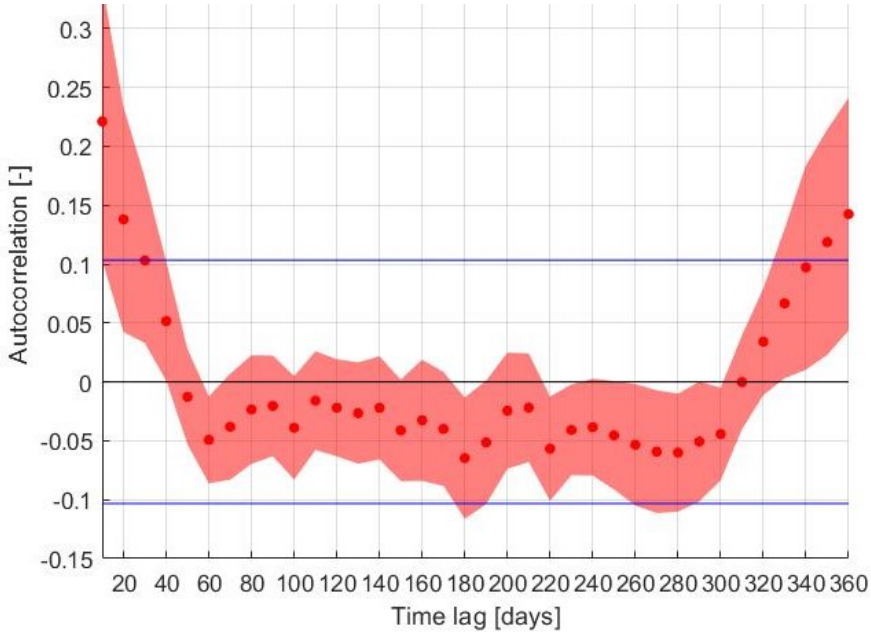


Figure A.2: Average autocorrelation coefficient, taking all types of invalid pixels into account in Sweden.

B

Iterative Savitzky-Golay Filter Sensitivity

This appendix contains the results from the sensitivity analysis of the iterative Savitzky-Golay filter, which was already described briefly in Section 2.6.

The iterative Savitzky-Golay filter uses four parameters in total, two polynomial degrees and two window widths which are used to form the polynomials. Whilst it may appear from Figure B.1 and Figure B.2 that the spatial and temporal autocorrelation differ greatly as a function of d_1 and d_2, the degrees of the polynomials used by the first and the iterative filter respectively. The difference is in fact less than a percent between the maximum and minimum for the spatial autocorrelation and only a few percent for the temporal autocorrelation. The selected value of 2 has nearly the highest temporal autocorrelation, with the lowest standard deviation of said value.

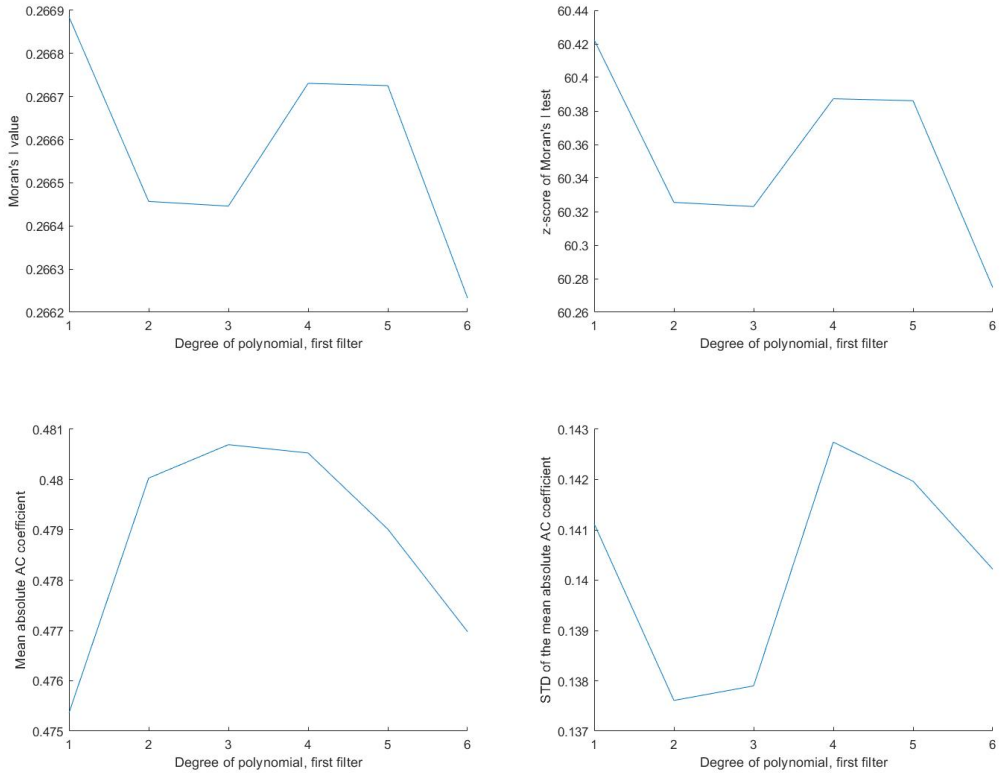


Figure B.1: Spatial and temporal autocorrelation for various degrees of the polynomial of the first filter.

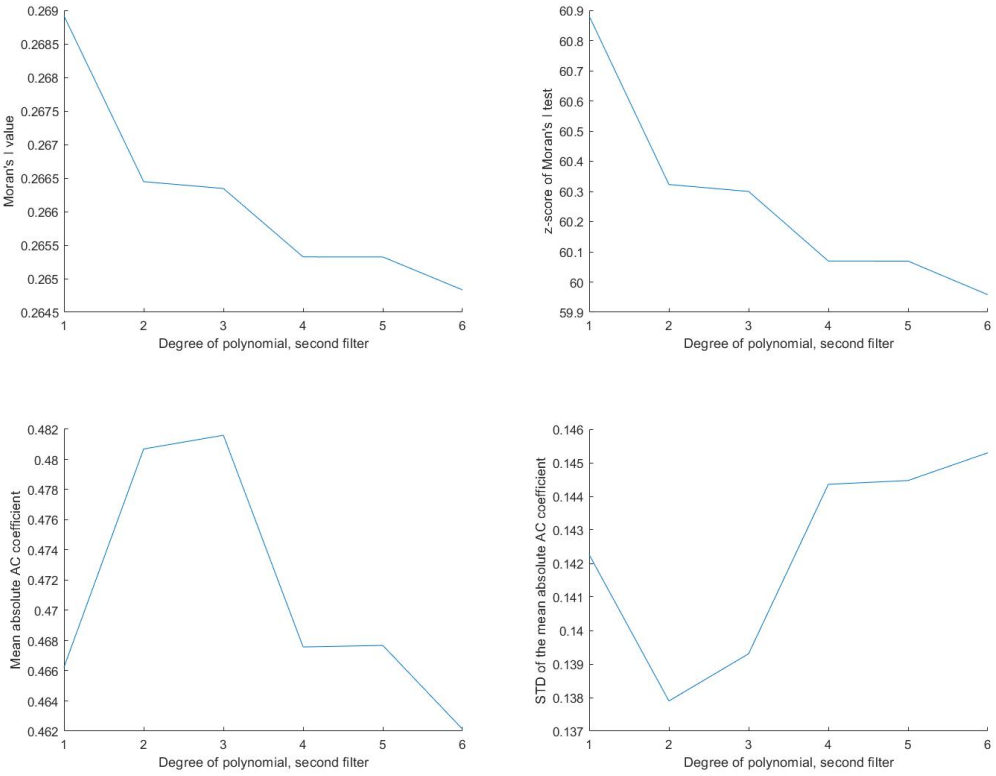


Figure B.2: Spatial and temporal autocorrelation for various degrees of the polynomial of the iterative filter.

Similarly, for the single-sided window widths m_1 and m_2, the spatial autocorrelation is again very insensitive, with the time series becoming increasingly temporally correlated with a greater width. Nonetheless, the differences are small and to ensure that the second harmonic is kept, a window width of 9 was selected. These results are shown for the first and iterative filters in Figure B.3 and Figure B.4 respectively.

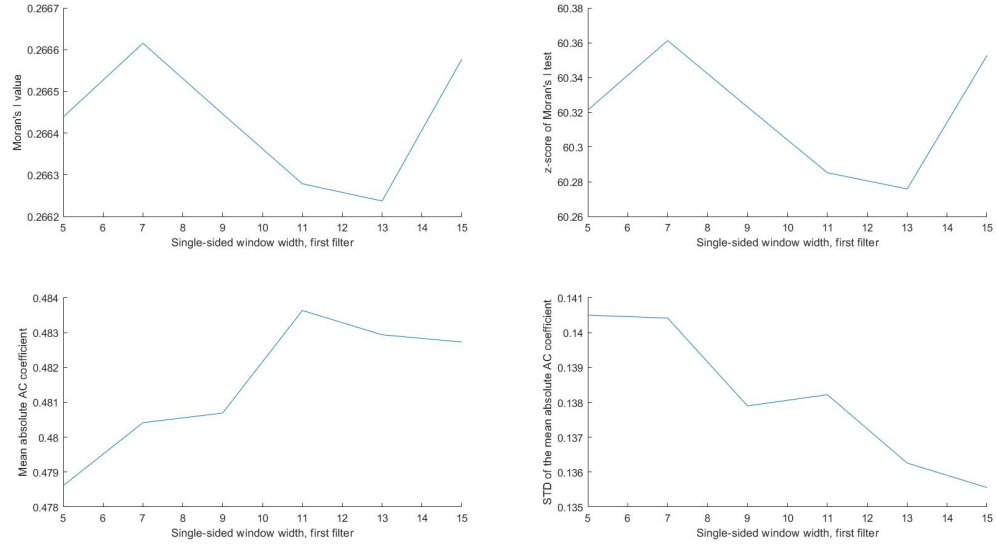


Figure B.3: Spatial and temporal autocorrelation for various single-sided window widths of the first filter.

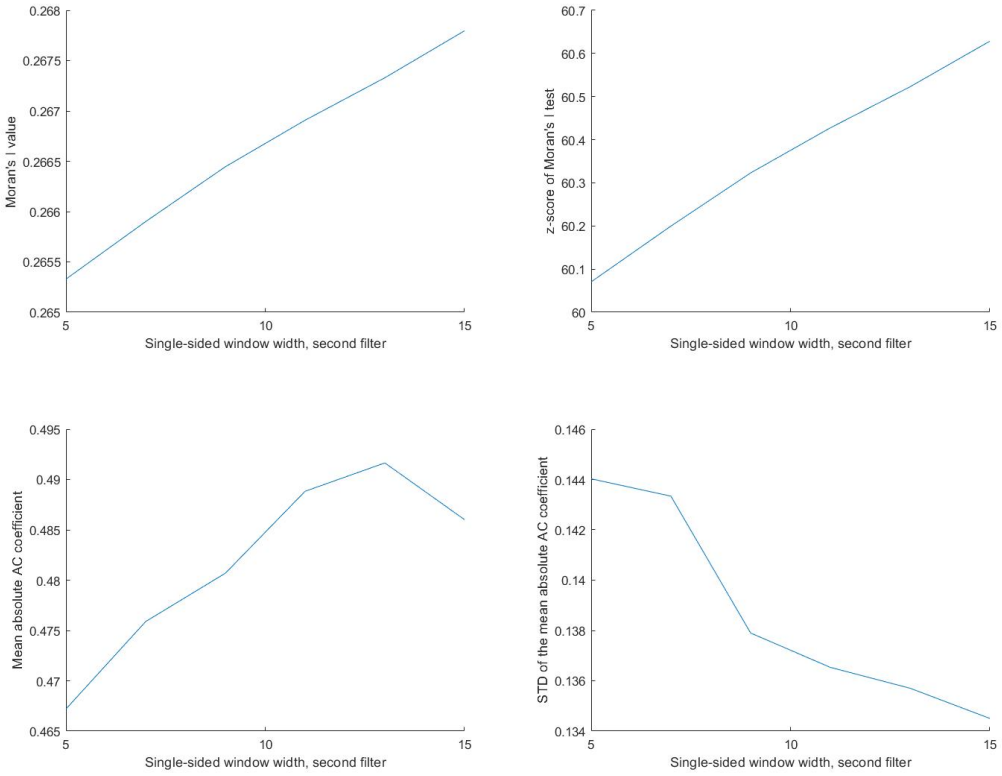


Figure B.4: Spatial and temporal autocorrelation for various single-sided window widths of the iterative filter.

C

Intermediate Classification Results

In this chapter, the intermediate classification results are given for the Iberian subset, such that more information is available than is given in Chapter 3. For reference, the ground truth data is given again in Figure C.1.

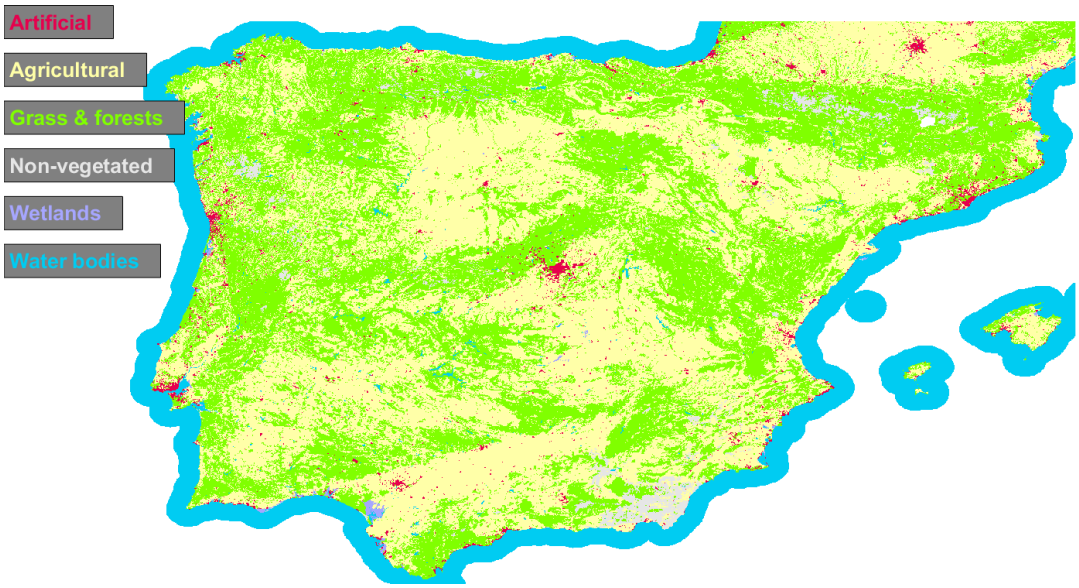


Figure C.1: Corine land cover data for the Iberian peninsula.

From Figure 3.4, it was already evident that the single classification tree did not provide acceptable results. Its confusion matrix is given by Figure C.2, which further shows that the classifier fails almost entirely to classify the rarer classes.

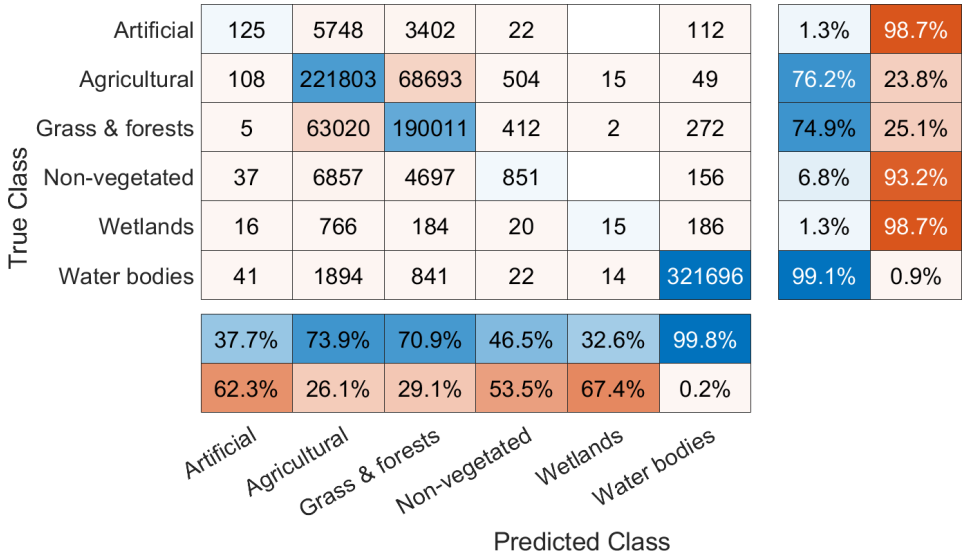


Figure C.2: The confusion matrix for the Iberian land cover data resulting from a single classification tree.

The classifier was changed by including oversampling and bagging, which by themselves did not provide a major effect on the results as can be seen in Figure C.3 and Figure C.4.

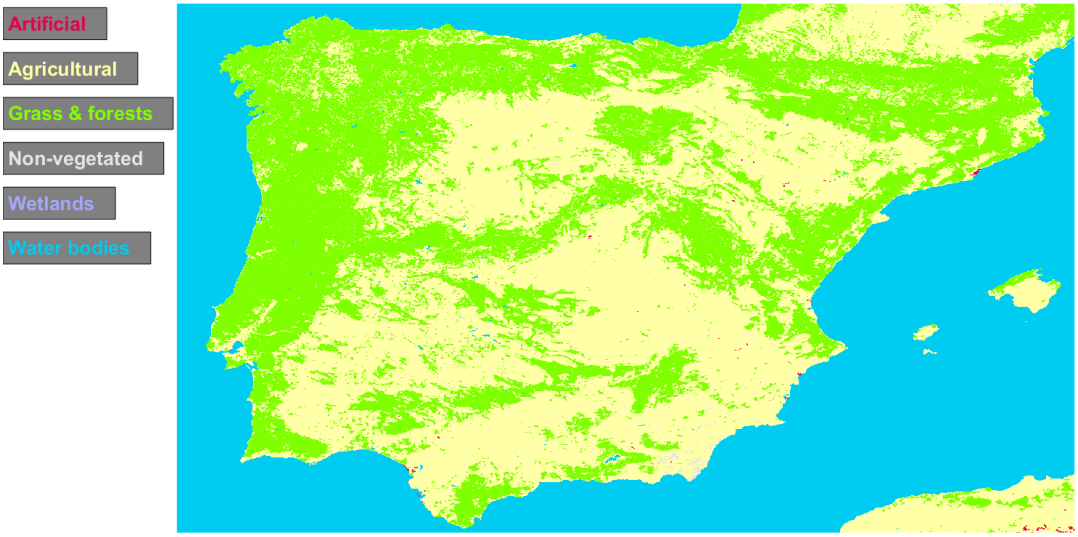


Figure C.3: Resulting land cover classification for the Iberian land cover data resulting from a classification tree using oversampling and bagging.

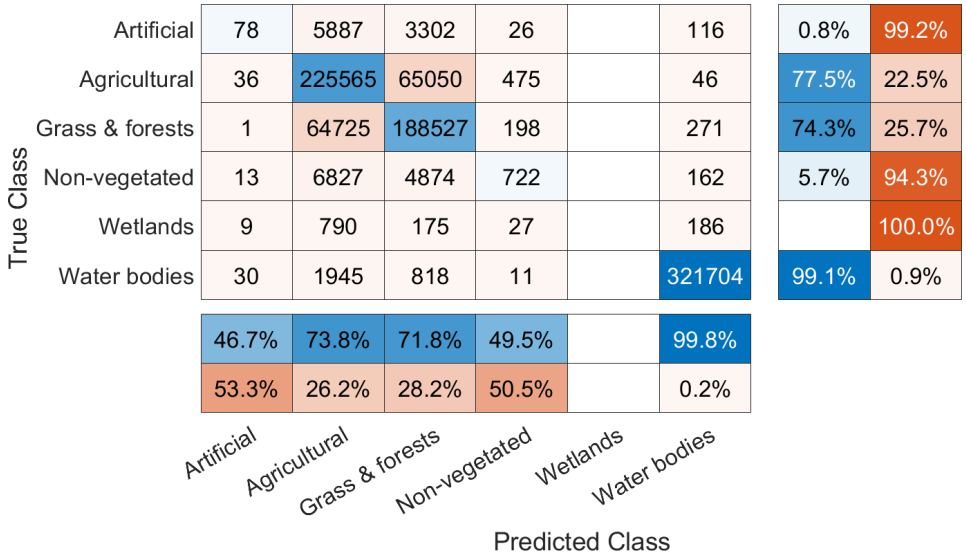


Figure C.4: The confusion matrix for the Iberian land cover data resulting from a classification tree using oversampling and bagging.

However, these methods combined with segmenting do have a positive effect on the classifier's performance. Figure C.5 shows the results obtained by using a single classification tree combined with segmenting, which when compared to Figure C.6 when oversampling, bagging and segmenting are used in conjunction shows that in particular the true positive rate for the final classification results has improved drastically. The final results for the year 2000 are shown in Figure C.7, and further show a large improvement in the classification quality.

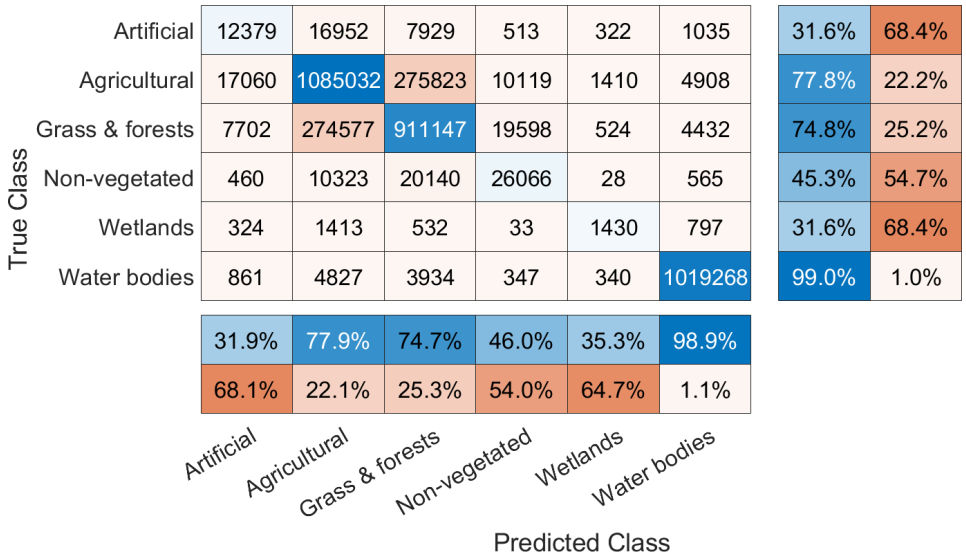


Figure C.5: The confusion matrix for the Iberian land cover data resulting from a single classification tree using segmenting.

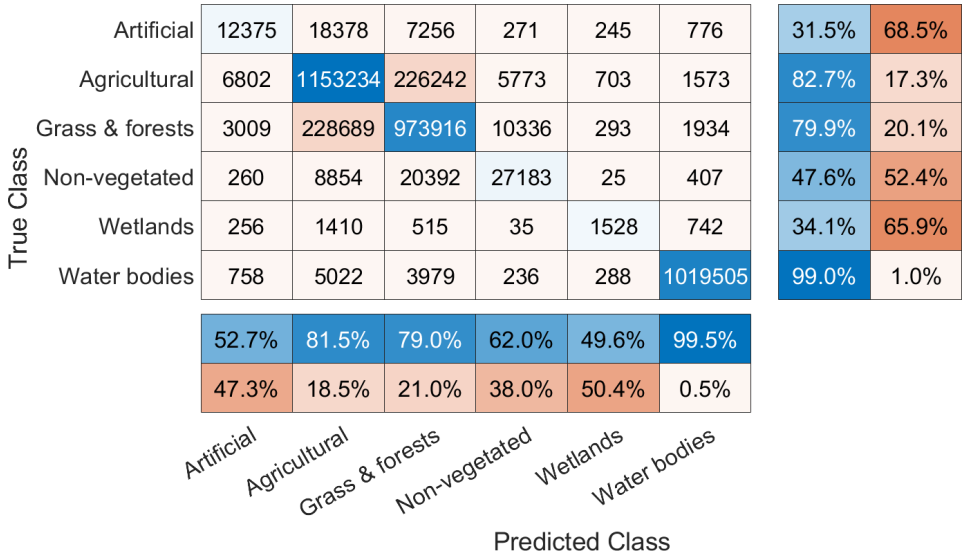


Figure C.6: The confusion matrix for the Iberian land cover data resulting from a classification tree using oversampling, bagging and segmenting.

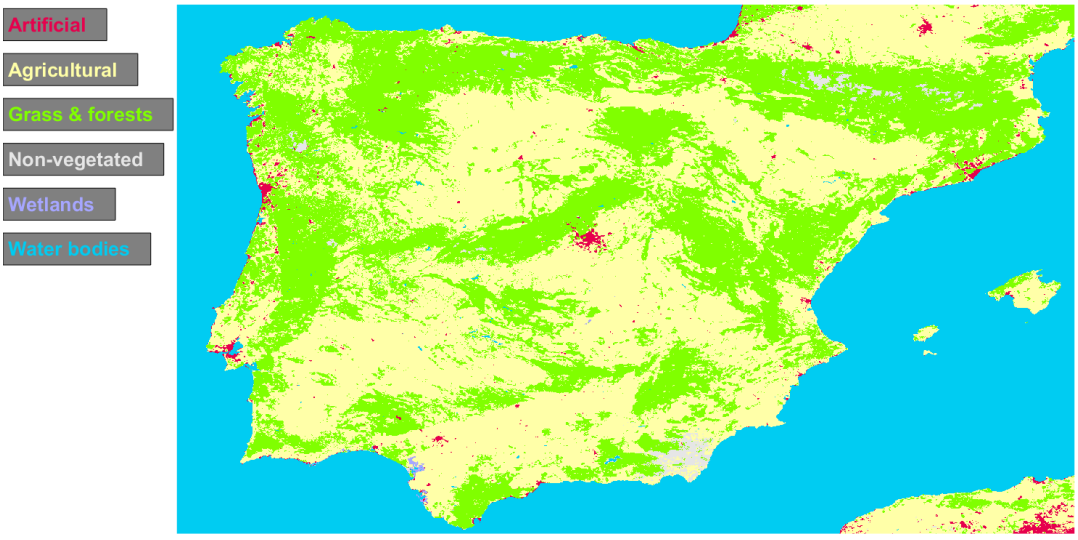


Figure C.7: The land cover map for the Iberian land cover data resulting from a classification tree using oversampling, bagging and segmenting.

D

Using Outdated Ground Truth Data for Classification

The Corine land cover data that is used as ground truth data to train the classifier is produced on a six-year basis. As a result, the situation can arise that a six year old data set must be used, compared to the current maximum of three years when the required data sets are available. This appendix shows that the effects this has on the classification performance, both quantitatively and qualitatively, are limited. The choice was made to classify NDVI data from 2012 using ground truth data from 2006, as the similarity between the 2012 and 2018 data sets is notably higher, as shown by Table 7.1.

First a quantitative comparison is made by comparing the confusion matrix for the year 2012 when the classifier is trained using the Corine data from 2006, shown in Figure D.1, and the confusion matrix when the correct Corine data from 2012 is used in Figure D.2. It is clear that whilst the latter outperforms the former on nearly all metrics, the differences are minor.

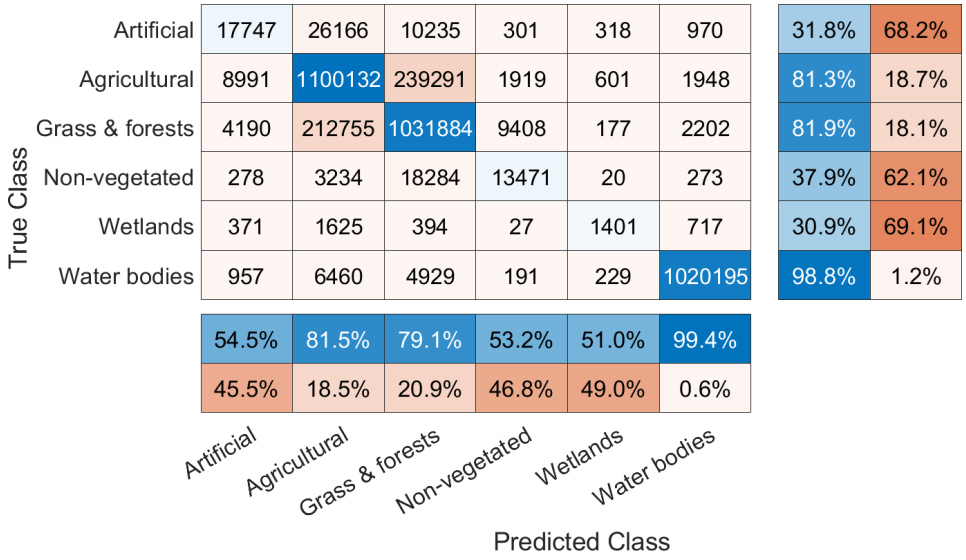


Figure D.1: Confusion matrix of the year 2012, when the classifier was trained using data from 2006.

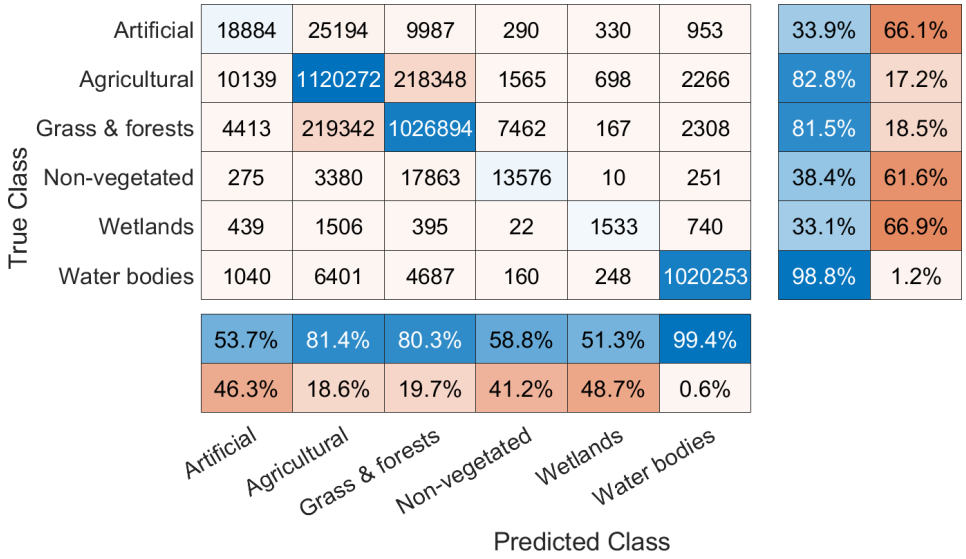


Figure D.2: Confusion matrix of the year 2012, when the classifier was trained using data from 2012.

This becomes clear when comparing the generated land cover maps (before smoothing has been applied). Whilst differences between the two exist, they are minor and no clear differences can be observed between the two data sets. The similarity between the generated land cover maps is therefore high, at 95.6%. As expected, this similarity is both higher than the similarity between the Corine 2012 and 2006 data for the same domain, which is 93.5%, as well as the 93.3% similarity obtained when classification is performed normally.

E

Classifying with Missing Ground Truth Data

To test whether the classifier is able to cope with missing ground truth data, the methodology described in Subsection 3.5.2 is tested on data of France. The eastern half is given missing values, such that the classification accuracy determined in that section when the data is present can be compared to when that data is missing. The Corine land cover is given in Figure E.1.

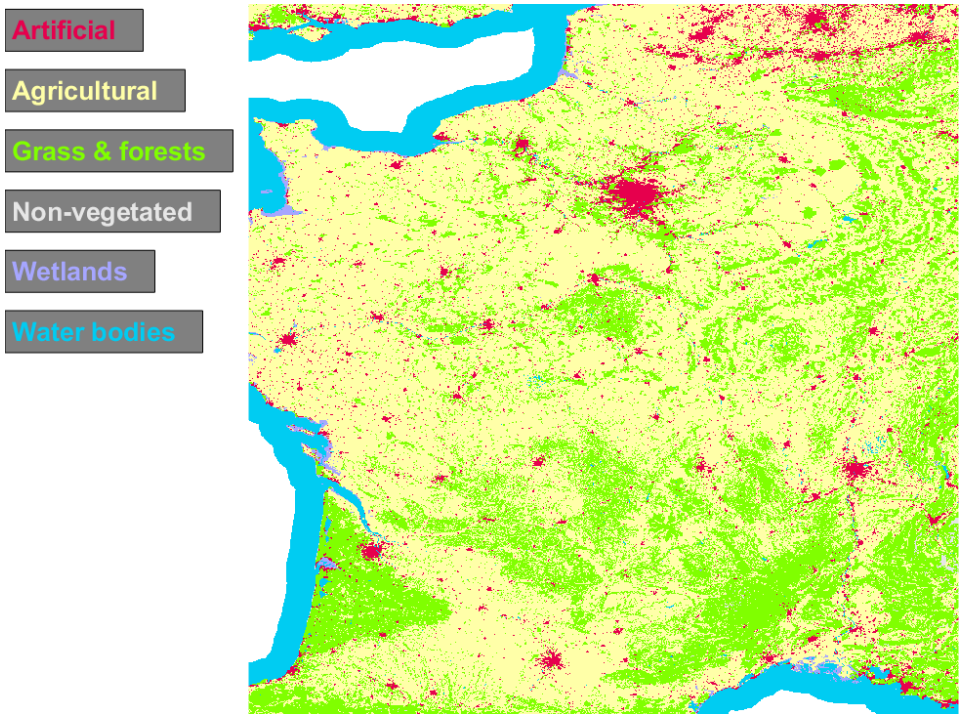


Figure E.1: Corine land cover data of France.

The confusion matrix for the eastern half is shown in Figure E.2. In this case, the classifier was able to use the data in this eastern half for training and the resulting classification accuracies are high. When this half is specified to be missing, the classifier must be trained using the data of the western half of France. Figure E.3 shows that the classification accuracy of the less common classes has dropped substantially. The likeliest explanation would be that balancing the data accurately is no longer possible, and the distribution of another section may not resemble that of one far away. Despite the fact that half the data was missing, and the eastern-most sections had to use ground truth data approximately 450 kilometres away, the classification accuracies of the most common classes are still high and the method is considered adequate.

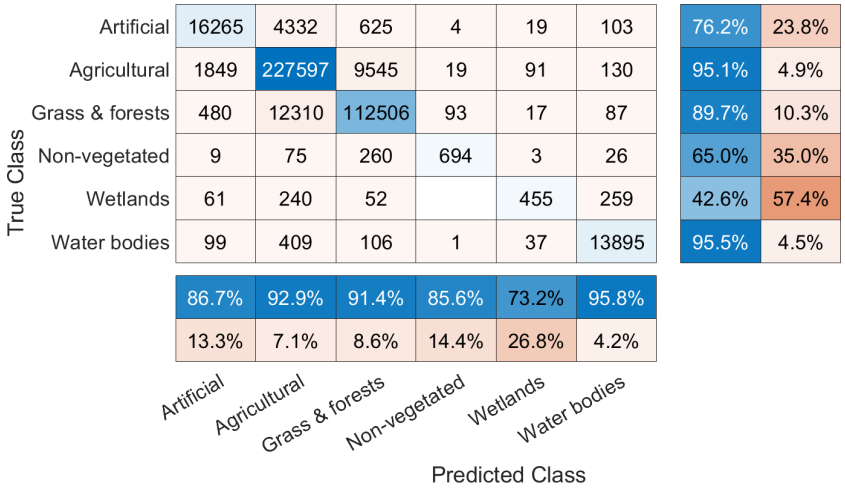


Figure E.2: Confusion matrix for the eastern half of the data set.

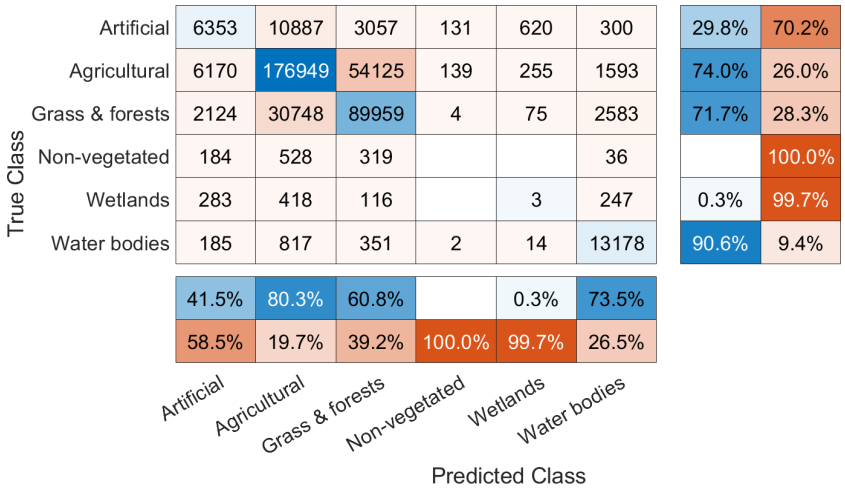


Figure E.3: Confusion matrix for the eastern half of the data set, when it is specified to be missing.

F

Variability within Each Class

In this appendix, the variability within each class in the Iberian data set is given by its average Euclidean distance and the standard deviation of these distances. The initial values before segmenting are given in Table F.1. As can be seen in Table F.2, the values are lower after segmenting has been applied. In other words, the variability within each class has decreased making it easier for the classifier to create an accurate profile. Contrary to the Jeffries-Matusita distances, these values are improved for every class. Finally, Table F.3 shows the relative improvement for each class via the relative decrease in the values obtained after segmenting. On average, the Euclidean distance is improved by 40.0% due to segmenting.

Table F.1: Average Euclidean distance and its STD in Iberia when no segmenting is applied.

	Artificial	**Agri.**	**Grass & forests**	**Non-vegetated**	**Wetlands**	**Water bodies**
Euclidean distance	0.1666	0.1562	0.1569	0.1577	0.2269	0.0307
STD	0.0724	0.0570	0.0631	0.0750	0.1439	0.0888

Table F.2: Average Euclidean distance and its standard deviation in Iberia when segmenting is applied.

	Artificial	**Agri.**	**Grass & forests**	**Non-vegetated**	**Wetlands**	**Water bodies**
Euclidean distance	0.1029	0.0870	0.0967	0.0916	0.1632	0.0156
STD	0.0557	0.0380	0.0431	0.0529	0.1249	0.0416

Table F.3: Relative improvement in average Euclidean distance and its standard deviation in Iberia when segmenting is applied.

	Artificial	**Agri.**	**Grass & forests**	**Non-vegetated**	**Wetlands**	**Water bodies**
Euclidean distance	38.24%	44.30%	38.37%	41.92%	28.07%	49.19%
STD	23.07%	33.33%	31.70%	29.47%	13.20%	53.15%

G

Quantitative Land Cover Results

This appendix contains the numeric values of the proportion of the land cover classes in Table G.1 and the absolute changes with respect to the total number of samples within the country, as well as the relative change with respect to the number of samples within the class itself, in Table G.2 and Table G.3 respectively. These tables then provide the numerical information used to create the plots in Section 7.2. The reader is reminded that due to an inconsistency in the Corine data of the year 2000, the classified data for 2004 is used for Great Britain instead of 1999.

Table G.1: The distribution of land cover classes for the European nations in 1999. Please note that for Great Britain the year 2004 is used instead.

Country	Artificial [%]	Agric. [%]	Grassl. & Forests [%]	Non-Veg. [%]	Wetlands [%]	Water [%]
Albania	0.30	24.91	71.36	1.10	0.02	2.32
Austria	1.81	29.57	62.53	5.56	0.11	0.41
Belarus	1.04	61.97	36.40	0.00	0.03	0.56
Belgium	9.29	70.56	19.79	0.00	0.08	0.29
Bosnia and H.	0.63	25.21	73.651	0.28	0.02	0.20
Bulgaria	0.90	52.95	45.42	0.08	0.01	0.64
Croatia	0.78	36.59	59.75	0.20	0.08	2.60
Czech Rep.	1.94	64.59	33.36	0.00	0.00	0.11
Denmark	4.59	84.15	6.91	0.01	0.18	4.15
Estonia	0.79	23.72	67.47	0.00	1.82	6.21
Finland	0.62	4.01	84.19	0.25	0.63	10.30
France	2.47	66.24	29.69	1.03	0.10	0.46
Germany	4.83	66.24	27.80	0.05	0.13	0.96
Great Britain	5.68	58.85	22.81	0.47	6.63	5.56
Greece	0.87	37.20	57.48	0.38	0.05	4.03
Hungary	1.59	75.61	21.83	0.00	0.14	0.83
Ireland	0.97	75.32	3.95	0.24	17.82	1.71
Italy	2.33	56.08	37.56	2.44	0.01	1.58
Kosovo	0.99	38.36	60.59	0.05	0.00	0.02
Latvia	0.64	37.10	59.11	0.01	1.05	2.09
Lithuania	0.97	68.20	29.08	0.00	0.09	1.66
Luxembourg	5.21	63.65	31.15	0.00	0.00	0.00
Macedonia	0.55	37.59	60.10	0.02	0.01	1.73
Moldova	0.06	93.59	5.55	0.00	0.03	0.77
Montenegro	0.68	4.94	90.81	1.20	0.37	1.99
Netherlands	10.48	76.24	4.92	0.02	0.41	7.93
Norway	0.27	1.65	59.70	29.82	0.24	8.32
Poland	1.52	68.80	28.13	0.01	0.09	1.45
Portugal	1.91	49.02	48.37	0.40	0.01	0.30
Romania	0.64	60.33	37.11	0.02	0.90	1.00
Russia	0.26	13.84	72.89	0.02	0.90	12.10
Serbia	1.13	53.43	45.02	0.02	0.01	0.39
Slovakia	1.02	49.26	49.30	0.19	0.00	0.23
Slovenia	1.20	29.04	68.53	1.12	0.06	0.04
Spain	0.62	55.11	42.73	1.04	0.07	0.43
Sweden	0.73	5.70	82.49	2.64	0.35	8.09
Switzerland	3.99	31.56	46.55	15.15	0.00	2.74
Turkey	1.13	46.28	47.09	3.51	0.10	1.89
Ukraine	0.62	75.91	21.89	0.00	0.18	1.41

Table G.2: The absolute difference between land cover distributions of 1999 and 2018, as percentage of total land cover for that country. Please note that for Great Britain, 2004 is used instead of 1999.

Country	Artificial [%]	Agric. [%]	Grassl. & Forests [%]	Non-Veg. [%]	Wetlands [%]	Water [%]
Albania	0.78	-1.28	0.22	0.30	0.00	-0.02
Austria	0.18	-0.22	-0.25	0.28	0.00	0.00
Belarus	-0.17	-0.87	0.94	0.00	-0.02	0.12
Belgium	-0.34	-0.85	1.15	0.00	0.04	0.01
Bosnia and H.	0.33	2.84	-3.46	0.27	0.00	0.02
Bulgaria	0.34	1.81	-2.20	0.02	0.00	0.03
Croatia	0.39	1.70	-2.14	0.06	-0.05	0.04
Czech Rep.	0.27	-2.77	2.50	0.00	0.00	0.00
Denmark	0.19	-1.46	1.45	-0.01	-0.04	-0.13
Estonia	-0.05	2.88	-3.05	0.00	0.30	-0.08
Finland	-0.01	0.28	-0.16	0.01	-0.15	0.03
France	0.46	-0.99	0.46	0.07	0.00	0.00
Germany	0.43	-3.10	2.63	-0.01	0.00	0.04
Great Britain	-0.10	-0.76	0.24	-0.08	0.59	0.11
Greece	0.74	-0.64	-0.37	0.21	0.01	0.05
Hungary	1.13	-0.79	-0.45	0.00	-0.01	0.13
Ireland	-0.01	1.05	-0.04	0.15	-1.14	-0.02
Italy	0.61	0.30	-1.34	0.42	.00	0.01
Kosovo	0.66	-1.34	0.62	0.05	0.00	0.01
Latvia	0.05	0.68	-0.84	0.02	0.03	0.06
Lithuania	-0.01	-1.43	1.39	0.01	0.01	0.04
Luxembourg	0.22	-3.46	3.23	0.00	0.00	0.00
Macedonia	0.23	-1.14	0.91	0.02	0.00	-0.02
Moldova	0.26	-0.34	0.09	0.00	-0.02	0.014
Montenegro	0.47	0.18	-2.46	1.76	0.00	0.05
Netherlands	0.57	-2.06	1.28	-0.01	0.09	0.12
Norway	0.00	-0.07	-0.73	0.88	0.08	-0.17
Poland	0.23	-0.89	0.61	0.00	0.02	0.02
Portugal	0.25	1.33	-1.46	-0.24	0.01	0.11
Romania	0.51	1.67	-2.22	0.01	-0.03	0.04
Russia	-0.07	0.46	-0.26	0.13	-0.39	0.09
Serbia	1.01	1.28	-2.36	0.01	-0.01	0.07
Slovakia	0.64	-0.93	0.27	0.00	0.00	0.02
Slovenia	0.42	0.14	-0.60	0.02	0.01	0.01
Spain	0.32	-0.13	0.33	-0.55	0.00	0.03
Sweden	-0.01	-0.38	0.10	0.13	0.14	0.03
Switzerland	0.16	-1.65	1.37	0.10	0.00	0.03
Turkey	0.56	0.25	-0.74	-0.20	-0.02	0.15
Ukraine	-0.00	1.28	-1.27	0.00	-0.03	0.03

Table G.3: The relative difference between land cover distributions of 1999 and 2018, as percentage of the number of samples within that class. N.A. is used when the number of samples in 1999 was zero and in 2018 non-zero, meaning that the relative change cannot be computed. Please note that for Great Britain, 2004 is used instead of 1999.

Country	Artificial [%]	Agric. [%]	Grassl. & Forests [%]	Non-Veg. [%]	Wetlands [%]	Water [%]
Albania	264.04	-5.15	0.31	27.29	14.29	-0.78
Austria	9.78	-0.74	-0.40	5.03	3.59	1.16
Belarus	-16.11	-1.41	2.58	N.A.	-54.78	21.42
Belgium	-3.67	-1.21	5.79	0.00	51.35	3.57
Bosnia and H.	51.42	11.28	-4.70	97.51	-11.11	12.33
Bulgaria	38.03	3.42	-4.84	29.03	0.00	3.94
Croatia	50.79	4.64	-3.57	28.49	-59.70	1.37
Czech Rep.	14.08	-4.29	7.49	0.00	0.00	-2.22
Denmark	4.08	-1.73	20.97	-80.00	-23.02	-3.13
Estonia	-5.76	12.14	-4.52	0.00	16.25	-1.22
Finland	-2.22	7.05	-0.19	5.15	-23.66	0.28
France	18.75	-1.49	1.54	6.44	-4.96	0.43
Germany	9.01	-4.68	9.44	-17.60	3.29	4.26
Great Britain	-1.75	-1.29	1.06	-16.88	8.94	1.89
Greece	85.69	-1.73	-0.65	54.84	25.00	1.32
Hungary	70.82	-1.05	-2.08	0.00	-7.07	16.09
Ireland	-0.61	1.40	-0.92	62.37	-6.40	-0.89
Italy	26.08	0.53	-3.56	17.21	-10.00	0.55
Kosovo	66.89	-3.49	1.02	114.29	0.00	33.33
Latvia	7.86	1.83	-1.43	185.71	3.19	2.72
Lithuania	-1.43	-2.10	4.78	N.A.	10.78	2.39
Luxembourg	4.27	-5.43	10.38	0.00	0.00	0.00
Macedonia	41.85	-3.05	1.52	116.67	0.00	-1.03
Moldova	406.25	-0.36	1.53	0.00	-60.00	1.81
Montenegro	68.25	3.63	-2.70	146.61	0.00	2.46
Netherlands	5.48	-2.71	26.02	-35.71	23.23	1.56
Norway	0.54	-4.37	-1.22	2.97	32.60	-2.01
Poland	15.44	-1.29	2.18	-8.06	23.21	1.35
Portugal	13.16	2.71	-3.01	-60.98	68.75	37.97
Romania	79.84	2.78	-5.97	65.15	-3.05	4.11
Russia	-25.31	3.60	-0.36	747.73	-43.23	0.75
Serbia	89.74	2.40	-5.25	47.37	-84.62	17.78
Slovakia	62.84	-1.89	0.55	2.11	0.00	7.06
Slovenia	35.33	0.48	-0.88	1.83	17.65	15.39
Spain	51.19	-0.25	0.78	-52.90	-0.22	7.19
Sweden	-1.47	-6.63	0.12	4.83	39.16	0.35
Switzerland	3.90	-5.22	2.94	0.63	0.00	1.08
Turkey	49.52	0.53	-1.57	-5.72	-17.56	8.01
Ukraine	-0.55	1.68	-5.78	N.A.	-18.35	1.83

H

Relative Changes on a National Basis in the Corine Database

In this appendix the change in the class distribution between 2000 and 2018 for the Corine database is shown in Table H.1 on a country-by-country basis, relative with respect to the number of samples found in the class itself. Note that some countries are omitted due to a lack of Corine land cover data.

Table H.1: Relative change w.r.t. the number of samples within the class itself from 2000 to 2018, according to the Corine land cover database.

Country	Artificial [%]	Agric. [%]	Grassl. & Forests [%]	Non-Veg. [%]	Wetlands [%]	Water [%]
Albania	75.66	-1.79	-2.13	16.53	21.74	6.02
Austria	21.23	-2.03	-0.26	-0.38	-12.46	8.14
Belgium	2.96	-0.92	-0.22	-44.44	14.60	2.11
Bosnia and H.	36.46	-9.71	3.66	79.23	19.05	-0.31
Bulgaria	-3.89	0.23	0.17	-9.08	-13.17	3.17
Croatia	25.63	-1.37	0.10	5.66	2.99	-0.85
Czech Rep.	10.93	-1.09	0.39	66.67	7.34	5.09
Denmark	16.85	-2.81	8.93	5.06	-0.08	8.87
Estonia	13.06	-3.12	0.81	-42.86	7.79	0.25
Finland	2.23	-3.35	0.62	38.18	-5.69	0.21
France	24.86	-1.60	-0.25	2.06	2.13	7.76
Germany	17.59	-3.54	3.45	-65.32	-9.95	11.82
Great Britain	14.28	-3.22	-22.57	-35.41	317.03	1.99
Greece	49.16	-3.22	-0.59	50.81	-1.67	4.75
Hungary	10.22	-3.80	9.27	12.90	2.65	1.79
Ireland	23.54	1.38	3.11	124.03	-13.31	-3.35
Italy	17.48	-0.81	-5.65	64.48	3.15	1.92
Kosovo	86.15	-4.36	-0.09	28.99	0.00	-8.70
Latvia	53.41	-11.28	8.66	270.00	2.63	9.07
Lithuania	8.20	-4.23	7.81	50.00	-3.81	2.84
Luxembourg	19.53	-5.55	4.92	0.00	0.00	-20.00
Macedonia	20.98	-3.68	0.92	202.47	16.00	4.04
Montenegro	81.62	-5.56	-3.97	60.00	-4.55	0.93
Netherlands	21.45	-4.79	6.71	-18.62	13.60	1.83
Norway	11.15	0.18	-0.57	0.17	2.62	-0.04
Poland	47.44	-5.17	6.32	-20.29	0.90	2.80
Portugal	21.93	1.14	-2.05	-24.89	4.98	53.21
Romania	-17.01	0.512	2.50	-33.40	-13.98	-0.35
Serbia	11.77	-3.20	3.70	-8.63	35.21	1.83
Slovakia	11.38	-2.91	2.01	5.36	37.50	-2.56
Slovenia	26.31	-1.15	-0.32	-0.73	2.56	27.78
Spain	49.41	-3.91	4.84	-40.88	2.32	13.74
Sweden	12.28	0.72	-0.33	1.84	2.51	-1.64
Switzerland	5.71	-3.55	2.43	-1.97	44.74	0.29
Turkey	23.89	2.67	1.78	-33.95	-4.21	5.73

I

The Cartesius Supercomputer

The Cartesius supercomputer hosted by SURFsara [149] provided the computational resources necessary to analyse the entire domain. In this appendix a very short overview of the system is given, as well as the computational cost associated with the analyses described in this thesis document. Should the reader desire to perform a similar analysis, the information given here will aid the selection of a suitable computational system as well as the planning for the required computational time.

Cartesius is a general purpose system that enables the use of a high quantity of cores with a large amount of memory. Two different types of nodes were used for this thesis, the fat nodes which contain 32 cores and have 256 GB of memory, and the thin nodes which have 24 cores and 64 GB of memory. As each parallel worker in MATLAB uses roughly 1 GB of memory even when idle, the amount of usable memory for parallelised tasks is limited when using thin nodes.

Nevertheless, as thin nodes are more frequently available, use was made of them for data pre-processing, time series decomposition, break point detection and forecasting. The data was divided into four roughly equal parts by latitude, to avoid exceeding the amount of available memory.

As the spatial domain affects the results obtained for classification, the aim was to have the sections be as large as possible and thus fat nodes were used. The data was divided into seven parts by latitude, which were roughly equal in size. A conscious effort was made to limit the effect of this procedure on the classification results. A qualitative analysis showed that the borders between these parts are not visible, and this segmentation of the data is deemed not to have affected the results in a meaningful way.

The required computational time is given in Table I.1, for both a single core and when either the 24 cores of the thin node are used or, in case of classification, the 32 cores of the fat node are used.

The reader is reminded that the full domain contains approximately 20 million pixels, of which roughly half are sea pixels. For all the steps listed in the table, the computational time required by these ten million sea pixels is negligible compared to that of the remaining land pixels, and the computational cost per land pixel is roughly twice the values given in the table.

Data pre-processing, classification and time series decomposition could be performed at native resolution due to the acceptable computational cost. As break point detection and forecasting proved costly even for Cartesius, these computations were performed for only one tenth of the domain.

Table I.1: Required computational time of the various tasks, assuming only one core is used in clock-seconds and clock-hours, as well as the total when the thin or fat nodes are used.

Task	**Computational time per pixel [clock-s]**	**Total computational time [clock-hr]**	**Total computational time [hr]**
Data pre-processing	0.12	672	28
Classification	0.60	3,360	105
Time series decomposition	0.20	1,080	45
Break point detection	8.38	46,560	1,940
Forecasting	54.0	300,000	12,500

References

[1] IPCC. *Climate Change and Land*. 2019. URL: https://www.ipcc.ch/srccl/ (visited on 12/31/2019).

[2] R. Costanza et al. "Changes in the Global Value of Ecosystem Services". In: *Global Environmental Change* 26 (2014), pp. 152–158.

[3] T. F. Stocker et al. *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Tech. rep. Cambridge, United Kingdom and New York, NY, USA: IPCC, 2013.

[4] R. B. Myneni et al. "The Interpretation of Spectral Vegetation Indexes". In: *IEEE Transactions on Geoscience and Remote Sensing* 33.2 (Mar. 1995).

[5] J. T. Kerr and M. Ostrovsky. "From Space to Species: Ecological Applications for Remote Sensing". In: *Trends in Ecology and Evolution* 18.6 (June 2003).

[6] Y. Xie, Z. Sha, and M. Yu. "Remote Sensing Imagery in Vegetation Mapping: A Review". In: *Journal of Plant Ecology* 1.1 (Mar. 2008), pp. 9–23.

[7] R. Houborg, J. B. Fisher, and A. K. Skidmore. "Advances in Remote Sensing of Vegetation Function and Traits". In: *International Journal of Applied Earth Observation and Geoinformation* 43 (2015), pp. 1–6.

[8] A. Huete et al. "Overview of the Radiometric and Biophysical Performance of the MODIS Vegetation Indices". In: *Remote Sensing of Environment* 83 (2002), pp. 195–213.

[9] B. C. Reed et al. "Measuring Phenological Variability from Satellite Imagery". In: *Journal of Vegetation Science* 5 (1994), pp. 703–714.

[10] R. de Jong et al. "Analysis of Monotonic Greening and Browning Trends from Global NDVI Time-Series". In: *Remote Sensing of Environment* 115 (2011), pp. 692–702.

[11] B. Martínez and M. A. Gilabert. "Vegetation Dynamics from NDVI Time Series Analysis Using the Wavelet Transform". In: *Remote Sensing of Environment* (2009), pp. 1823–1842.

[12] J. Verbesselt et al. "Phenological Change Detection while Accounting for Abrupt and Gradual Trends in Satellite Image Time Series". In: *Remote Sensing of Environment* 114 (2010), pp. 2970–2980.

[13] C. Tottrup and M. S. Rasmussen. "Mapping Long-Term Changes in Savannah Crop Productivity in Senegal Through Trend Analysis of Time Series of Remote Sensing Data". In: *Agriculture, Ecosystems and Environment* 103 (2004), pp. 545–560.

[14] X. Weixin and L. Xiaodong. "Response of Vegetation in the Qinghai-Tibet Plateau to Global Warming". In: *Chinese Geographical Science* 17.2 (2007), pp. 151–159.

[15] R. S. Lunette et al. "Land-Cover Change Detection Using Multi-Temporal MODIS NDVI Data". In: *Remote Sensing of Environment* 105 (2006), pp. 142–154.

[16] C. Müller et al. "Implications of Climate Mitigation for Future Agricultural Production". In: *Environmental Research Letters* 10 (2015).

[17] A. Nakamura et al. "Forests and Their Canopies: Achievements and Horizons in Canopy Science". In: *Trends in Ecology & Evolution* 32.6 (June 2017).

[18] B. A. Kimball. "Crop Responses to Elevated CO_2 and Interactions with H_2O, N, and Temperature". In: *Current Opinion in Plant Biology* 31 (2016), pp. 36–43.

[19] C. L. Quéré et al. "Global Carbon Budget 2015". In: *Earth Systems Science Data* 7 (2015), pp. 349–396.

[20] D. Deryng et al. "Global Crop Yield Response to Extreme Heat Stress Under Multiple Climate Change Futures". In: *Environmental Research Letters* 9 (2014).

[21] C. Lesk, P. Rowhani, and N. Ramankutty. "Influence of Extreme Weather Disasters on Global Crop Production". In: *Nature* 529 (Jan. 2016).

[22] G. Büttner, T. Soukup, and B. Kosztra. *Climate Change Adaptation in the Agricultural Sector in Europe*. Tech. rep. 4. Copenhagen, Denmark: European Environment Agency, Sept. 2019.

[23] *The State of Food Security and Nutrition in the World 2018*. Tech. rep. Rome, Italy: FAO, IFAD, UNICEF, WFP and WHO, 2018.

[24] C. D. Allen et al. "A Global Overview of Drought and Heat-Induced Tree Mortality Reveals Emerging Climate Change Risks for Forests". In: *Forest Ecology and Management* 259 (2010), pp. 660–684.

[25] W. R. L. Anderegg, J. M. Kane, and L. D. L. Anderegg. "Consequences of Widespread Tree Mortality Triggered by Drought and Temperature Stress". In: *Nature Climate Change* 3 (2013).

[26] F. Martellozzo et al. "Urbanization and the Loss of Prime Farmland: A Case Study in the Calgary-Edmonton Corridor of Alberta". In: *Regional Environmental Change* 15 (2015), pp. 881–893.

[27] C. B. d'Amour et al. "Future Urban Land Expansion and Implications for Global Croplands". In: *PNAS* 114.34 (Aug. 2017), pp. 8939–8944.

[28] N. C. Davidson. "How Much Wetland Has the World Lost? Long-Term and Recent Trends in Global Wetland Area". In: *Marine and Freshwater Research* 65 (2014), pp. 934–941.

[29] R. J. Keenan et al. "Dynamics of Global Forest Area: Results form the FAO Global Forest Resources Assessment 2015". In: *Forest Ecology and Management* 352 (2015), pp. 9–20.

[30] A. A. Abahussain et al. "Desertification in the Arab Region: Analysis of Current Status and Trends". In: *Journal of Arid Environments* 51 (2002), pp. 521–545.

[31] C. Chen et al. "China and India Lead in Greening of the World Through Land-Use Management". In: *Nature Sustainability* 2 (Feb. 2019), pp. 122–129.

[32] A. Anyamba and C. J. Tucker. "Analysis of Sahelian Vegetation Dynamics Using NOAA-AVHRR NDVI Data from 1981-2003". In: *Journal of Arid Environments* 63 (2005), pp. 596–614.

[33] L. Yang et al. "Thematic Accuracy of MRLC Land Cover for the Eastern United States". In: *Remote Sensing of Environment* 76 (2001), pp. 418–422.

[34] Y. Julien, J. A. Sobrino, and W. Verhoef. "Changes in Land Surface Temperature and NDVI Values over Europe between 1982 and 1999". In: *Remote Sensing of Environment* 103 (2006), pp. 43–55.

[35] R. Stöckli and P. L. Vidale. "European Plant Phenology and Climate as Seen in a 20-Year AVHRR Land-Surface Parameter Dataset". In: *International Journal of Remote Sensing* 25.17 (Sept. 2004), pp. 3303–3330.

[36] N. Pan et al. "Increasing Global Vegetation Browning Hidden in Overall Vegetation Greening: Insights from Time-Varying Trends". In: *Remote Sensing of Environment* 214 (2018), pp. 59–72.

[37] T. Park et al. "Changes in Growing Season Duration and Productivity of Northern Vegetation Inferred from Long-Term Remote Sensing Data". In: *Environmental Research Letters* 11 (2016).

[38] J.-L. Champeaux et al. "AVHRR-Derived Vegetation Mapping over Western Europe for Use in Numberical Weather Prediction Models". In: *International Journal of Remote Sensing* 21.6 (2000), pp. 1183–1199.

[39] M. Forkel et al. "Trend Change Detection in NDVI Time Series: Effects of Inter-Annual Variability and Methodology". In: *Remote Sensing* 5 (2013), pp. 2113–2144.

[40] W. Jiang et al. "Spatio-Temporal Analysis of Vegetation Variation". In: *Ecological Indicators* 51 (2015), pp. 117–126.

[41] S. Peng et al. "Recent Change of Vegetation Growth Trend in China". In: *Environmental Research Letters* 6 (2011).

[42] H. Lu et al. "Decomposition of Vegetation Cover into Woody and Herbaceous Components Using AVHRR NDVI Time Series". In: *Remote Sensing of Environment* 86 (2003), pp. 1–18.

[43] S. Eckert et al. "Trend Analysis of MODIS NDVI Time Series for Detecting Land Degradation and Regeneration in Mongolia". In: *Journal of Arid Environments* 113 (2015), pp. 16–28.

[44] A. Jacquin, D. Sheeren, and J.-P. Lacombe. "Vegetation Cover Degradation Assessment in Madagascar Savanna Based on Trend Analysis of MODIS NDVI Time Series". In: *International Journal of Applied Earth Observation and Geoinformation* 12 (2010), pp. 3–8.

[45] K. M. de Beurs and G. M. Henebry. “Trend Analysis of the Pathfinder AVHRR Land (PAL) NDVI Data for the Deserts of Central Asia”. In: *IEEE Geoscience and Remote Sensing Letters* 1.4 (Oct. 2004), pp. 282–286.

[46] Y. Liu et al. “Spatial and Temporal Patterns of Global NDVI Trends: Correlations with Climate and Human Factors”. In: *Remote Sensing* 7 (2015), pp. 13233–13250.

[47] R. de Jong et al. “Shifts in Global Vegetation Activity Trends”. In: *Remote Sensing* 5 (2013), pp. 1117–1133.

[48] W. Yuan et al. “Increased Atmospheric Vapor Pressure Deficit Reduces Global Vegetation Growth”. In: *Science Advances* 5 (Aug. 2019).

[49] S. Estel et al. “Mapping Farmland Abandonment and Recultivation Across Europe USING MODIS NDVI Time Series”. In: *Remote Sensing of Environment* 163 (2015), pp. 312–325.

[50] S. Estel et al. “Mapping Cropland Use-Intensity Across Europe Using MODIS NDVI Time Series”. In: *Environmental Research Letters* 11 (2016).

[51] A. Kelemen et al. *Regions 2020, The Climate Change Challenge for European Regions*. Tech. rep. Brussels, Belgium: European Commission Directorate-General Regional Policy: Conception, Forward Studies, Impact Assessment, Mar. 2009.

[52] M. O. Jeffries, J. Richter-Menge, and J. E. Overland. *Arctic Report Card 2015*. Tech. rep. NOAA, 2015.

[53] S. Bojinski et al. “The Concept of Essential Climate Variables in Support of Climate Research, Applications, and Policy”. In: *Bulletin of the American Metereological Society* 95.9 (Sept. 2014).

[54] Copernicus Global Land Service. *Normalized Difference Vegetation Index*. URL: `https://land.copernicus.eu/global/products/ndvi` (visited on 12/31/2019).

[55] B. Smets et al. *Gio Global Land Component - Lot I “Operation of the Global Land Component"*. Tech. rep. VITO NV, May 2016.

[56] C. Toté, E. Swinnen, and S. Sterckx. *Evaluation of the Re-processed VGT1 and VGT2 Archive*. Tech. rep. VITO NV, Apr. 2016.

[57] C. Toté et al. *Evaluation of the Re-processed PROBA-V Archive - Final Report*. Tech. rep. VITO NV, Nov. 2017.

[58] ESA. *NOAA AVHRR*. URL: `https://earth.esa.int/web/guest/missions/3rd-party-missions/current-missions/noaa-avhrr` (visited on 09/02/2019).

[59] NASA. *MODIS Vegetation Index Products (NDVI and EVI)*. URL: `https://modis.gsfc.nasa.gov/data/dataprod/mod13.php` (visited on 09/02/2019).

[60] L. Geng et al. “Comparison of Eight Techniques for Reconstructing Multi-Satellite Sensor Time-Series NDVI Data Sets in the Heihe River Basin, China”. In: *Remote Sensing* 6 (2014), pp. 2024–2049.

[61] R. B. Myneni et al. “Optical Remote Sensing of Vegetation: Modeling, Caveats and Algorithms”. In: *Remote Sensing of the Environment* 51 (1995), pp. 169–188.

[62] G. G. Gutman. “Vegetation Indices from AVHRR: an Update and Future Prospects”. In: *Remote Sensing of the Environment* 35 (1991), pp. 121–136.

[63] M. E. Brown et al. “SPOT-Vegetation, SeaWiFS, MODIS, and Landsat ETM+ Sensors”. In: *IEEE Transactions on Geoscience and Remote Sensing* 44.7 (July 2006), pp. 121–136.

[64] P. Jönsson and L. Eklundh. “TIMESAT - A Program for Analyzing Time-Series of Satellite Sensor Data”. In: *Computers & Geosciences* 30 (2004), pp. 833–845.

[65] M. Buchhorn et al. *Copernicus Global Land Operations “Vegetation and Energy"*. Tech. rep. VITO NV, Nov. 2019.

[66] R. Geerken, B. Zaitchik, and J. P. Evans. “Classifying Rangeland Vegetation Type and Coverage from NDVI Time Series using Fourier Filtered Cycle Similarity”. In: *International Journal of Remote Sensing* 26.24 (Dec. 2005), pp. 5535–5554.

[67] J. M. B. Carreiras, J. M. C. Pereira, and Y. E. Shimabukuro. “Land-cover Mapping in the Brazilian Amazon Using SPOT-4 Vegetation Data and Machine Learning Classification Methods”. In: *Photogrammetric Engineering & Remote Sensing* 72.8 (Aug. 2006), pp. 897–910.

[68] J. Chen et al. “A simple method for reconstructing a high-quality NDVI time-series data set based on the Savitzky-Golay filter”. In: *Remote Sensing of Environment* 91 (2004), pp. 332–344.

[69] W. Wei et al. “Selecting the Optimal NDVI Time-Series Reconstruction Technique for Crop Phenology Detection”. In: *Intelligent Automation & Soft Computing* 22.2 (2016), pp. 237–247.

[70] A. Savitzky and M. J. E. Golay. “Smoothing and Differentiation of Data by Simplified Least Squares Procedures”. In: *Analytical Chemistry* 36.8 (June 1964), pp. 1627–1639.

[71] J. Steinier, Y. Termonia, and J. Deltour. “Comments on Smoothing and Differentiation of Data by Simplified Least Square Procedure”. In: *Analytical Chemistry* 44.11 (Sept. 1972), pp. 1906–1909.

[72] H. H. Madden. “Comments on the Savitzky-Golay Convolution Method for Least-Squares Fit Smoothing and Differentiation of Digital Data”. In: *Analytical Chemistry* 50.9 (Aug. 1978), pp. 1383–1386.

[73] R. Cao et al. “A Simple Method to Improve the Quality of NDVI Time-Series Data by Integrating Spatiotemporal Information with the Savitzky-Golay Filter”. In: *Remote Sensing of Environment* 217 (2018), pp. 244–257.

[74] J. Ren et al. “Regional Yield Estimation for Winter Wheat with MODIS-NDVI Data in Shandong, China”. In: *International Journal of Applied Earth Observation and Geoinformation* 10 (2008), pp. 403–413.

[75] M. Ma and F. Veroustraete. “Reconstructing Pathfinder AVHRR Land NDVI Time-Series Data for the Northwest of China”. In: *Advances in Space Research* 37 (2006), pp. 835–840.

[76] P. H. C. Eilers. “A Perfect Smoother”. In: *Analytical Chemistry* 75 (2003), pp. 3631–3636.

[77] P. S. A. Beck et al. “Improved monitoring of vegetation dynamics at very high latitudes: A new method using MODIS NDVI”. In: *Remote Sensing of Environment* 100 (2006), pp. 321–334.

[78] J. N. Hird and G. J. McDermid. “Noise Reduction of NDVI Time Series: An Empirical Comparison of Selected Techniques”. In: *Remote Sensing of Environment* 113 (2009), pp. 248–258.

[79] P. Jöhnsson and L. Eklundh. “Seasonality extraction by function fitting to time-series of satellite sensor data”. In: *Transactions on Geoscience and Remote Sensing* 40.8 (2002).

[80] G. J. Roerink and M. Menenti. “Reconstructing Cloudfree NDVI Composite Using Fourier Analysis of Time Series”. In: *International Journal of Remote Sensing* 21.9 (2000), pp. 1911–1917.

[81] J. Zhou, L. Jia, and M. Menenti. “Reconstruction of Global MODIS NDVI Time Series: Performance of Harmonic ANalysis of Time Series (HANTS)”. In: *Remote Sensing of Environment* 163 (2015), pp. 217–228.

[82] A. Moody and . M. Johnson. “Land-Surface Phenologies from AVHRR Using the Discrete Fourier Transform”. In: *Remote Sensing of Environment* 75 (2001), pp. 305–323.

[83] W. Jun, S. Zhongbo, and M. Yaoming. “Reconstruction of a Cloud-Free Vegetation Index Time Series for the Tibetan Plateau”. In: *Mountain Research and Development* 24.4 (Nov. 2004), pp. 348–353.

[84] R. Geerken, B. Zaitchik, and J. P. Evans. “Classifying Rangeland Vegetation Type and Coverage from NDVI Time Series Using Fourier Filtered Cycle Similarity”. In: *International Journal of Remote Sensing* 26.24 (Dec. 2005), pp. 5535–5554.

[85] Y. Shao et al. “An Evaluation of Time-Series Smoothing Algorithms for Land-Cover Classifications using MODIS-NDVI Multi-Temporal Data”. In: *Remote Sensing of Environment* 174 (2016), pp. 258–265.

[86] P. M. Atkinson et al. “Inter-Comparison of Four Models for Smoothing Satellite Sensor Time-Series Data to Estimate Vegetation Phenology”. In: *Remote Sensing of Environment* 123 (2012), pp. 400–417.

[87] L. Zhou et al. “Variations in Northern Vegetation Activity Inferred from Satellite Data of Vegetation Index During 1981 to 1999”. In: *Journal of Geopysical Research* 106.17 (Sept. 2001), pp. 20069–20083.

[88] S. Khorram et al. *Remote Sensing*. 1st ed. Springer, 2012. ISBN: 978-1-4614-3102-2.

[89] G. M. Foody and A. Mathur. “The Use of Small Training Sets Containing Mixed Pixels for Accurate Hard Image Classification: Training on Mixed Spectral Responses for Classification by a SVM”. In: *Remote Sensing of Environment* 103 (2006), pp. 179–189.

[90] J. A. Richards and X. Jia. *Remote Sensing Digital Image Analysis*. 4th ed. Springer, 2005. ISBN: 978-3-540-25128-6.

[91] C. Gómez, J. C. White, and M. A. Wulder. “Optical Remotely Sensed Time Series Data for Land Cover”. In: *Journal of Photogrammetry and Remote Sensing* 116 (2016), pp. 55–72.

[92] M. A. White et al. “Intercomparison, Interpretation and Assessment of Spring Phenology in North America Estimated from Remote Sensing for 1982-2006”. In: *Global Change Biology* 15 (2009), pp. 2335–2359.

[93] R. R. Colditz et al. “TiSeG: A Flexible Software Tool for Time-Series Generation of MODIS Data Utilizing the Quality Assessment Science Data Set”. In: *IEEE Transactions on Geoscience and Remote Sensing* 46.10 (Oct. 2008), pp. 3296–3308.

[94] G. Genovese et al. “A Methodology for a Combined Use of Normalised Difference Vegetation Index and CORINE Land Cover Data for Crop Yield Monitoring and Forecasting. A Case Study on Spain”. In: *Agronomie* 21 (2001), pp. 91–111.

[95] Copernicus Global Land Service. *CORINE Land Cover*. URL: `https://land.copernicus.eu/pan-european/corine-land-cover` (visited on 12/31/2019).

[96] G. Büttner, J. Feranec, and G. Jaffrain. *Corine Land Cover Update 2000*. Tech. rep. Copenhagen, Denmark: European Environment Agency, 2002.

[97] *CLC2006 Technical Guidelines*. Tech. rep. Copenhagen, Denmark: European Environment Agency, 2007.

[98] G. Büttner, T. Soukup, and B. Kosztra. *CLC2012 Addendum to CLC2006 Technical Guidelines*. Tech. rep. European Environment Agency, 2014.

[99] G. Büttner et al. *CLC2018 Technical Guidelines*. Tech. rep. Vienna, Austria: European Environment Agency, 2017.

[100] *The Thematic Accuracy of Corine Land Cover 2000*. Tech. rep. 7. European Environment Agency, 2006.

[101] J. P. Snyder. *Map Projections - A Working Manual*. Tech. rep. Washington D. C., United States: U.S. Geological Survey, 1987.

[102] *CORINE Land Cover Nomenclature - Illustrated Guide*. Tech. rep. Copernicus Global Land Services.

[103] M. C. Hansen et al. “Global Land Cover Classification at 1 km Spatial Resolution Using a Classification Tree Approach”. In: *International Journal of Remote Sensing* 21.6 (2000), pp. 1331–1364.

[104] C. Lucas et al. “Identification of Linear Vegetation Elements in a Rural Landscape Using LiDAR Point Clouds”. In: *Remote Sensing* 11 (2019).

[105] M. A. Friedl, C. E. Brodley, and A. H. Strahler. “Maximizing Land Cover Classification Accuracies Produced by Decision Trees at Continental to Global Scales”. In: *IEEE Transactions on Geoscience and Remote Sensing* 37.2 (Mar. 1999), pp. 969–977.

[106] M. A. Friedl and C. E. Brodley. “Decision Tree Classification of Land Cover from Remotely Sensed Data”. In: *Remote Sensing of the Environment* 61 (1997), pp. 399–409.

[107] C. E. Pacheco, M. I. Aguado, and D. Mollicone. “Identification and Characterization of Deforestation Hot Spots in Venezuela Using MODIS Satellite Images”. In: *Acta Amazonica* 44.2 (2014), pp. 185–196.

[108] D. Stroppiana, J-M. Grégoire, and J. M. C. Pereira. “The Use of SPOT VEGETATION Data in a Classification Tree Approach for Burnt Area Mapping In Australian Savanna”. In: *International Journal of Remote Sensing* 24.10 (2003), pp. 2131–2151.

[109] S. Rasoul Safavian and D. Landgrebe. “A Survey of Decision Tree Classifier Methodology”. In: *IEEE Transactions on Systems, Man and Cybernetics* 21.3 (May 1991).

[110] G. M. Weiss. “Mining with Rarity: A Unifying Framework”. In: *SIGKDD Explorations* 6 (2004), pp. 7–19.

[111] G. M. Weiss, K. McCarthy, and B. Zabar. “Cost-Sensitive Learning vs. Sampling: Which is Best for Handling Unbalanced Classes with Unequal Error Costs?” In: *Proceedings of the International Conference on Data Mining* (2007), pp. 35–41.

[112] H. He et al. “ADASYN: Adaptive Synthetic Sampling Approach for Imbalanced Learning”. In: *International Joint Conference on Neural Networks* (2008), pp. 1322–1328.

[113] G. M. Weiss and F. Provost. “Learning When Training Data are Costly: The Effect of Class Distribution on Tree Induction”. In: *Journal of Artificial Intelligence Research* 19 (2003), pp. 315–354.

[114] D. K. McIver and M. A. Friedl. “Using Prior Probabilities in Decision-Tree Classification of Remotely Sensed Data”. In: *Remote Sensing of Environment* 81 (2002), pp. 253–261.

[115] B. Ghimire et al. “An Evaluation of Bagging, Boosting, and Random Forests for Land-Cover Classification in Cape Cod, Massachusetts, USA”. In: *GIScience & Remote Sensing* 49.5 (2012), pp. 623–643.

[116] L. Breiman. “Random Forests”. In: *Machine Learning* 45 (2001), pp. 5–32.

[117] S. Padma and S. Sanjeevi. “Jeffries Matusita Based Mixed-Measure for Improved Spectral Matching in Hyperspectral Image Analysis”. In: *International Journal of Applied Earth Observation and Geoinformation* 32 (2014), pp. 138–151.

[118] A. Ghiyamat et al. “Hyperspectral Discrimination of Tree Species with Different Classifications Using Single- and Multiple-Endmember”. In: *International Journal of Applied Earth Observation and Geoinformation* 23 (2013), pp. 177–191.

[119] P. Bühlmann and B. Yu. “Analyzing Bagging”. In: *The Annals of Statistics* 30.4 (2002), pp. 927–961.

[120] G. Biau, L. Devroye, and G. Lugosi. “Consistency of Random Forests and other Averaging Classifiers”. In: *Journal of Machine Learning Research* 9 (2008), pp. 2015–2033.

[121] T. Hastie, R. Tibshirani, and J. Friedman. *The Elements of Statistical Learning*. 2nd ed. Springer, 2009. ISBN: 978-0-387-84858-7.

[122] J. Inglada et al. “Operational High Resolution Land Cover Map Production at the Country Scale Using Satellite Image Time Series”. In: *Remote Sensing* 95.9 (2017).

[123] E. C. Weatherhead, A. J. Stevermer, and B. E. Schwartz. “Detecting Environmental Changes and Trends”. In: *Physics and Chemistry of the Earth* 27 (2002), pp. 399–403.

[124] K. M. de Beurs and G. M. Henebry. “A Statistical Framework for the Analysis of Long Image Time Series”. In: *International Journal of Remote Sensing* 26.8 (Apr. 2005), pp. 1551–1573.

[125] R. B. Cleveland et al. “STL: A Seasonal-Trend Decomposition Procedure Based on Loess”. In: *Journal of Official Statistics* 6.1 (1990), pp. 3–73.

[126] D. F. Hendry. *Econometric Modelling*. Oxford University, 2000.

[127] F. Tian et al. “Evaluating Temporal Consistency of Long-Term Global NDVI Datasets for Trend Analysis”. In: *Remote Sensing of Environment* 163 (2015), pp. 326–340.

[128] R. Fensholt et al. “Evaluation of Earth Observation Based Long Term Vegetation Trends - Intercomparing NDVI Time Series Trend Analysis Consistency of Sahel from AVHRR GIMMS, Terra MODIS and SPOT VGT Data”. In: *Remote Sensing of Environment* 113 (2009), pp. 1886–1898.

[129] J. Verbesselt et al. “Detecting Trend and Seasonal Changes in Satellite Image Time Series”. In: *Remote Sensing of Environment* 114 (2010), pp. 106–115.

[130] C.-S. J. Chu, K. Hornik, and C.-M. Kuan. “MOSUM Tests for Parameter Constancy”. In: *Biometrika* 82.3 (1995), pp. 603–617.

[131] J. Kim. “Comparison of Structural Change Tests in Linear Regression Models”. In: *The Korean Journal of Applied Statistics* 24.6 (2011), pp. 1197–1211.

[132] C.-M. Kuan and M.-Y. Chen. “Response Surfaces of MOSUM Critical Values”. In: *Applied Economics Letters* 9 (2002), pp. 133–136.

[133] J. Bai and P. Perron. “Computation and Analysis of Multiple Structural Change Models”. In: *Journal of Applied Econometrics* 18 (2003), pp. 1–22.

[134] R. L. Brown, J. Durbin, and J. M. Evans. “Techniques for Testing the Constancy of Regression Relationships over Time”. In: *Journal of the Royal Statistical Society* 37.2 (1975), pp. 149–163.

[135] E. C. Weatherhead et al. “Factors Affecting the Detection of Trends: Statistical Considerations and Applications to Environmental Data”. In: *Journal of Geophysical Research* 103.14 (July 1998), pp. 17, 149–17, 161.

[136] K. P. Burnham and D. R. Anderson. “Multimodel Inference”. In: *Sociological Methods & Research* 33.2 (Nov. 2004), pp. 261–304.

[137] S. Jamali et al. “Detecting Changes in Vegetation Trends Using Time Series Segmentation”. In: *Remote Sensing of Environment* (2015), pp. 182–195.

[138] R. E. Kass and A. E. Raftery. “Bayes Factors”. In: *Journal of the American Statistical Association* 90.430 (June 1995), pp. 773–795.

[139] D. K. Bolton and M. A. Friedl. “Forecasting Crop Yield Using Remotely Sensed Vegetation Indices and Crop Phenology Metrics”. In: *Agricultural and Forest Meteorology* 173 (2013), pp. 74–84.

[140] L. Ji and A. J. Peters. “Forecasting Vegetation Greenness with Satellite and Climate Date”. In: *IEEE Geoscience and Remote Sensing Letters* 1.1 (Jan. 2004).

[141] N. Philippon, N. Martiny, and P. Camberlin. “Forecasting the Vegetation Photosynthetic Activity over the Sahel: A Model Output Statistics Approach”. In: *International Journal of Climatology* 29 (2009), pp. 1463–1477.

[142] R. J. Hyndman and G. Athanasopoulos. *Forecasting: Principles and Practice*. 2nd ed. Accessed on 12.08.2019. Melbourne, Australia: OTexts, 2018.

[143] P. J. Brockwell and R. A. Davis. *Introduction to Time Series and Forecasting*. 2nd ed. Springer, 2001. ISBN: 0-387-95351-5.

[144] J. W. Osborne. “Improving Your Data Transformations: Applying the Box-Cox Transformation”. In: *Practical Assessment, Research & Evaluation* 15.12 (Oct. 2010).

[145] G. E. P. Box and D. R. Cox. “An Analysis of Transformations”. In: *Journal of the Royal Statistical Society* 26.2 (1964), pp. 211–252.

[146] R. J. Hyndman and A. B. Koehler. “Another Look at Forecast Accuracy”. In: *International Journal of Forecasting* 22 (2006), pp. 679–688.

[147] M. H. Pesaran, D. Pettenuzzo, and A. Timmermann. “Forecasting Time Series Subject to Multiple Structural Breaks”. In: *The Review of Economic Studies* 73.4 (Oct. 2006), pp. 1057–1084.

[148] M. H. Pesaran and A. Timmermann. “How Costly is it to Ignore Breaks when Forecasting the Direction of a Time Series?” In: *International Journal of Forecasting* 20 (2004), pp. 411–425.

[149] SURFsara. *Description of the Cartesius System*. URL: `https://userinfo.surfsara.nl/systems/cartesius/description` (visited on 12/18/2019).

[150] C. Brewer and M. Harrower. *ColorBrewer 2.0 Color Advice for Cartography*. URL: `http://colorbrewer2.org/#` (visited on 12/30/2019).

[151] *Urban Sprawl in Europe*. Tech. rep. 11. Copenhagen, Denmark and Bern, Switzerland: EEA and FOEN, Nov. 2016.

[152] E. I. Hennig et al. “Multi-Scale Analysis of Urban Sprawl in Europe: Towards a European De-Sprawling Strategy”. In: *Land Use Policy* 49 (2015), pp. 483–498.

[153] V. C. Radeloff and G. Gutman. *Land-Cover and Land-Use Changes in Eastern Europe after the Collapse of the Soviet Union in 1991*. 1st ed. Springer, 2017. ISBN: 978-3-319-42638-9.

[154] Google Maps. *Kyiv, Ukraine*. URL: `https://www.google.nl/maps/place/Kyiv,+Ukraine,+02000` (visited on 12/18/2019).

[155] *Istanbul Airport*. URL: `https://www.istairport.com/en` (visited on 12/17/2019).

[156] K. C. Guay et al. “Vegetation Productivity Patterns at High Northern Latitudes: A Multi-Sensor Satellite Data Assessment”. In: *Global Change Biology* 20 (2014), pp. 3147–3158.

[157] B. Z. Carlson et al. “Observed Long-Term Greening of Alpine Vegetation - A Case Study in the French Alps”. In: *Environmental Research Letters* 12 (2017).

[158] C. M. van der Heide, H. J. Silvis, and W. J. M. Heijman. “Agriculture in the Netherlands: Its Recent Past, Current State and Perspectives”. In: *Applied Studies in Agribusiness and Commerce* 5 (2011), pp. 23–28.

[159] M. van Leeuwen et al. *Het Nederlandse Agrocomplex 2013*. Tech. rep. Wageningen, the Netherlands: Wageningen University and Research, Jan. 2014.

[160] D. Verhoog. *Het Nederlandse Agrocomplex 2015*. Tech. rep. Wageningen, the Netherlands: Wageningen University and Research, Jan. 2016.

[161] CBS et al. *Temperatuur in Nederland en Mondiaal, 1906 - 2017*. URL: `https://www.clo.nl/indicatoren/nl0226-temperatuur-mondiaal-en-in-nederland` (visited on 01/06/2019).

[162] CBS et al. *Jaarlijkse Hoeveelheid Neerslag in Nederland, 1910 - 2015*. URL: https://www.clo.nl/indicatoren/nl050806-jaarlijkse-hoeveelheid-neerslag-in-nederland (visited on 01/06/2019).

[163] P. A. Schultz and M. S. Halpert. "Global Correlation of Temperature, NDVI and Precipitation". In: *Advanced Space Research* 13.5 (1993), pp. 277–280.

[164] K. Ichi, A. Kawabata, and Y. Yamaguchi. "Global Correlation Analysis for NDVI and Climatic Variables and NDVI Trends: 1982 - 1990". In: *International Journal of Remote Sensing* 23.18 (2002), pp. 3873–3878.

[165] K. George et al. In: *Atmospheric Environment* 41 (2007), pp. 7654–7665.

[166] A. J. Elmore et al. "Landscape Controls on the Timing of Spring, Autumn, and Growing Season Length in Mid-Atlantic Forests". In: *Global Change Biology* 18 (2012), pp. 656–674.

[167] L. Xu et al. "Global Correlation Analysis for NDVI and Climatic Variables and NDVI Trends: 1982 - 1990". In: *International Journal of Remote Sensing* 23.18 (2002), pp. 3873–3878.