

**A thesis submitted to the Department of Environmental Sciences and Policy of
Central European University in part fulfillment of the
Degree of Master of Science**

**Analysing Land Use Changes with Google Earth Engine:
A Case Study of the Azov Sea Basin**

Sara Margareta PRUCKNER

June, 2018

Budapest

**Erasmus Mundus Masters Course in
Environmental Sciences, Policy and
Management**

MESPOM



This thesis is submitted in fulfillment of the Master of Science degree awarded as a result of successful completion of the Erasmus Mundus Masters course in Environmental Sciences, Policy and Management (MESPOM) jointly operated by the University of the Aegean (Greece), Central European University (Hungary), Lund University (Sweden) and the University of Manchester (United Kingdom).

Notes on copyright and the ownership of intellectual property rights:

(1) Copyright in text of this thesis rests with the Author. Copies (by any process) either in full, or of extracts, may be made only in accordance with instructions given by the Author and lodged in the Central European University Library. Details may be obtained from the Librarian. This page must form part of any such copies made. Further copies (by any process) of copies made in accordance with such instructions may not be made without the permission (in writing) of the Author.

(2) The ownership of any intellectual property rights which may be described in this thesis is vested in the Central European University, subject to any prior agreement to the contrary, and may not be made available for use by third parties without the written permission of the University, which will prescribe the terms and conditions of any such agreement.

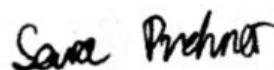
(3) For bibliographic and reference purposes this thesis should be referred to as:

Pruckner, S. 2018. *Analysing Land Use Changes with Google Earth Engine: A Case Study of the Azov Sea Basin*. Master of Science thesis, Central European University, Budapest.

Further information on the conditions under which disclosures and exploitation may take place is available from the Head of the Department of Environmental Sciences and Policy, Central European University.

Author's declaration

No portion of the work referred to in this thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

A handwritten signature in black ink, appearing to read "Sara Pruckner".

Sara Margareta PRUCKNER

CENTRAL EUROPEAN UNIVERSITY

ABSTRACT OF THESIS submitted by:

Sara Margareta PRUCKNER

for the degree of Master of Science and entitled:

Analysing Land Use Changes with Google Earth Engine: A Case Study of the Azov Sea Basin

Month and Year of submission: June, 2018.

Google Earth Engine revolutionizes land change analysis, as it provides possibilities to analyse satellite imagery and other datasets via cloud-computing technology and server-sided JavaScript programming. This saves resources on the client's side and can lead to vast improvements of results for scientists of all levels. The application of this new tool was tested on the land use classification of the Azov Sea basin and its subbasins. In total, a time series of seven land use maps from 1985 to 2015 was created, and urbanization and deforestation trends were detected. Additionally, the change within the region of the Donbas conflict was analysed from 2013 to 2017 to track the effect of the war on the environment. Subsequently, the collapse of the Soviet Union as a driver of agricultural land abandonment was discussed and examined. NASA population datasets were analysed to determine the relationship between the urbanization of the region and population change. Finally, the precipitation deficit of the Upper Don catchment basin was calculated with GLDAS data from 1984 to 2010 to draw conclusions on the effect of local land use changes on the water balance. It was concluded that the rising proportion of built-up land could have caused a rise in the local temperature in interaction with climate change, which caused the local evapotranspiration rates to rise and the water level of the Tsimlyansk reservoir to decrease.

The developed geospatial database was made available online on:

<https://syslab.ceu.edu/projects/GEE-land-use-azov-basin>

Acknowledgements

I would, first and foremost, like to thank my friends who helped me to not lose my mind in the past few months. You are what made these two years special.

Furthermore, I would like to thank Elizaveta Khazieva for her support with Google Earth Engine, as well as the Google Developer's Guide and Forum for their help and support. Also, thank you to Anastasia Kvasha for her help with finding and translating Russian datasets.

Finally, I would like to express my gratitude towards my supervisor, Viktor Lagutov, who gave me advice and feedback on my work and helped me in the making of this thesis.

Danke, köszönöm, thank you, ευχαριστώ.

Table of Contents

List of Tables.....	viii
List of Figures	viii
List of Abbreviations.....	xii
1 Introduction	1
1.1 Research Aim, Questions, and Objectives.....	4
1.2 Thesis Outline.....	5
2 Review of Literature and Theoretical Background	7
2.1 Land Change Science	7
2.2 DPSIR-Framework	11
2.3 Ecosystem Services	14
2.4 Conclusion of the Literature Review.....	16
3 Review of Google Earth Engine.....	18
3.1 GEE Application and Case Studies	18
3.2 GEE Land Use Classifiers	23
4 Methods and Approach	30
4.1 Land Use Mapping	30
4.2 Interpretation of Land Use Data	31
4.3 Collection and Analysis of Socio-Economic Data	32
4.4 Limitations.....	33
5 Case Study Area: Azov Sea Basin	35
5.1 Ecosystem.....	36
5.2 Agriculture.....	36
5.3 Industrial and Economic Development	39
5.4 Conclusion of the Case Study Area Review.....	41
6 Land Use Classification	42
6.1 Classification Process with GEE	42
6.1.1 Setting the Stage	42
6.1.2 Collecting Training Data	47
6.1.3 Classifying the Map	51
6.1.4 Accuracy Assessment.....	55
6.1.5 Mean Spectra.....	59
6.2 Land Use Areas of the Subbasins	61
6.3 Donbas Conflict.....	66
6.4 Validation of Results	71
6.4.1 Global Forest Watch.....	71
6.4.2 JRC Surface Water Data Set	75

7	Impacts and Drivers of Land Use Changes	78
7.1	Drivers: Transition to a Market Based Economy	78
7.1.1	Agricultural Productivity Changes.....	79
7.1.2	Abandoned Land	83
7.2	Drivers: Population Change and Urbanization	86
7.2.1	Population Development 1980 – 2000	88
7.2.2	Population Development 2000 – 2015	90
7.2.3	Correlation Population Development and Urbanization	92
7.3	Impact: Freshwater Ecosystem Services	94
7.3.1	Literature Review of LULCC in a Watershed	97
7.3.2	Calculating the Precipitation Deficit	99
7.3.3	Effects of the Precipitation Deficit on the Tsimlyansk Reservoir.....	103
8	Conclusion.....	106
	Bibliography.....	108
	Appendices	116

List of Tables

Table 1: Area of the Training data polygons for each class in km ²	51
Table 2: Overall validation accuracy and kappa values for all seven land use maps.....	56
Table 3: Confusion matrix for 1985.....	57
Table 4: Confusion matrix for 1990.....	57
Table 5: Confusion matrix for 1995.....	57
Table 6: Confusion matrix for 2000.....	58
Table 7: Confusion matrix for 2005.....	58
Table 8: Confusion matrix for 2010.....	58
Table 9: Confusion matrix for 2015.....	58

List of Figures

Figure 1: The DPSIR-Framework. Source: EEA 2016	12
Figure 2: Linking Actions (including, but not limited to LULCC) to a change in the value of ecosystem goods and services of a river system Source: Keeler et al. 2012, 1862	16
Figure 3: Set up and user interface of Google Earth Engine. Source: Google 2018a.....	20
Figure 4: A comparison of different Landsat satellites, their bands, and optimal usage Source: USGS 2015	24
Figure 5: Dataset visualization. Classifying water and vegetation with the dimensions red and NIR.....	25
Figure 6: Visualization of a Linear Classifier. The new point in red will be classified as vegetation in this fictional scenario	26
Figure 7: Visualization of the kNN classifier. The nearest neighbours all being vegetation, it will be classified as such with almost any k.	28
Figure 8: Workflow chart of the whole methodology.....	30
Figure 9: Workflow of the Land Use Classification Process with Google Earth Engine	31
Figure 10: The Azov Sea basin. Source: Lagutov and Lagutov 2012, 9.....	35
Figure 11: Top Producers of Corn, Barley and Winter Wheat in the Azov Sea basin and the rest of Russia. Not pictured areas of Russia are not considered top producers in any of the categories. Created by the author with ArcMap 10.4, Data Source: Gray 2017.....	37
Figure 12: "Changes in the Don River hydrological regime after completion of the Tsimlyansk reservoir" Source: Lagutov and Lagutov 2012, 40	39
Figure 13: Available images of Landsat 2 MSS in the case study region between 1975 and 1985	43
Figure 14: Choosing the image collection, filtering bounds and calendar range	44
Figure 15: Masking clouds out of the picture and taking the median of each band for each pixel of the remaining images.....	45
Figure 16: Calculating and adding the NDVI	46
Figure 17: The least cloudy pictures for 2015, visualized as band combination 6 - 5 - 4 of Landsat 8	47
Figure 18: Visualization of the NDVI of the least cloudy pictures in 2015, calculated from Landsat 8 images.....	47
Figure 19: Large industrial and mining areas that are included in the category "urban".....	48
Figure 20: Eutrophication visible in the Landsat 5 map	49
Figure 21: Eutrophic water falsely classified as forest (dark green patches).....	49
Figure 22: Data gaps due to cloudiness over an urban and agricultural area.....	50
Figure 23: Calculation of amount of pixels and total area of polygons of each class, example code for water polygons of 1985.....	50

Figure 24: Merging all five feature collections and transforming them into potential training data	52
Figure 25: Randomly splitting polygons into training and testing datasets, and using the training partition to train and classify the Landsat 5 image	53
Figure 26: Visualizing the classified land-use map in Google Earth Engine.....	53
Figure 27: Land use Maps for the years 1985 to 2015 of the Azov Sea Basin	54
Figure 28: Calculation of the different accuracy measures	55
Figure 29: Calculations of mean spectra for each class within the Azov Sea basin	59
Figure 30: Mean spectra of each land use class in the different Landsat 5 bands of the 1985 land use map	60
Figure 31: Extent and Land use in 2015 of the Kuban River basin	61
Figure 32: Extent and Land use in 2015 of the Don River basin	61
Figure 33: Extent and Land use in 2015 of the Upper Don River basin	61
Figure 34: Calculations of the area per class.....	62
Figure 35: Proportion of all agricultural area (split into the different sub-areas) in the Azov Sea basin between 1985 and 2015	63
Figure 36: Cropped versus bare agricultural land, in percent of the whole Azov Sea basin ...	63
Figure 37: Proportion of forest in the Azov Sea basin from 1985 to 2015	64
Figure 38: Proportion of urban areas in the Azov Sea Basin from 1985 to 2015	65
Figure 39: Proportion of water area in the Azov Sea basin from 1985 to 2015.....	65
Figure 40: Situation of the conflict in the Donbas region in October 2016. Used as a base map for all further analysis. Source: The Economist, 2016.....	67
Figure 41: Result of digitalizing the conflict zone in the Donbas region	68
Figure 42: Development of agriculture in the Donbas conflict region from 2013 to 2017, in percent of the whole buffer area	69
Figure 43: Development of urban area in the Donbas conflict region from 2013 to 2017, in percent of the whole buffer area	70
Figure 44: Development of natural vegetation and forests in the Donbas conflict region from 2013 to 2017, in percent of the whole buffer area	70
Figure 45: Preparation of data sets for the forest analysis	72
Figure 46: Using .expression() and mathematical operators to determine the masks for both, Hansen, and thesis only data sets.....	73
Figure 47: Forest classification of Hansen et al. 2013 versus this study	73
Figure 48: Area of water in km ² as according to the JRC, this thesis, or both data sets.....	76
Figure 49: Data gaps in the JRC water data sets of 1985. Areas with no data are marked as black	76
Figure 50: Areas identified as water in the thesis dataset (yellow), projected over the JRC data set	77
Figure 51: Box-Whisker Plots of grain yield in 100 kg per ha, averaged over the years 1986 – 1990 (a) and 1997 – 2000 (b).....	80
Figure 52: Box-Whisker Plots of the grain yield in 100kg per ha, averaged over the years 1986 – 1990 (a), and 1997 – 2000 (b).....	81
Figure 53: Box-Whisker-Plots of the milk yield in kg per cow and year, for 1990 (a) and 2000 (b).....	82
Figure 54: Box-Whisker Plots of the average milk yield in kg per cow and year for 1990 (a) and 2000 (b)	82
Figure 55: Development of agricultural areas in the Azov sea basin, in percent of the total area	84
Figure 56: Visualization of Abandoned agricultural land (red), added agricultural land (yellow), and land that was temporarily abandoned and re-developed (blue).....	85

Figure 57: Population development per grid cell (30 arc seconds resolution) from 1980 to 2000. Data Source: CIESIN 2017a	88
Figure 58: Population development from 1980 to 2000 in the different regions of the Azov Sea Basin, in No. of people per 30 arc seconds grid cell. Data Source: CIESIN 2017a	89
Figure 59: Population development from 1980 to 2000 of the different sub-basins within the Azov Sea Basin. Data Source: CIESIN 2017a.....	89
Figure 60: Population change from 2000 to 2015. Areas with a decrease in population are pictured in red, areas with an increase are pictured in green. The darker the colour, the stronger the change has been. Data Source: CIESIN 2017b.....	90
Figure 61: Population development from 2000 to 2015 of the different sub-basins within the Azov Sea Basin Data Source: CIESIN 2017b	91
Figure 62: Population development from 2000 to 2015 in the different regions of the Azov Sea Basin, in No. of people per 30 arc seconds grid cell. Data Source: CIESIN 2017b	91
Figure 63: Urbanized land divided into the point of time of urbanization and relationship with population increase	93
Figure 64: Land switching from urbanized to other land use types, split into point of time of the transition, and relationship with population decrease	94
Figure 65: The water cycle and its most important components Source: NASA 2018	96
Figure 66: Visualization of the GLDAS Evapotranspiration dataset in GEE. Red = high ET, green = low ET.....	99
Figure 67: Area-averaged storm surface runoff of the Upper Don river basin. Values are monthly with a 1deg spatial resolution. Data Source: GLDAS, Rondell et al. 2004. 100	100
Figure 68: Change of ET and P in the Upper Don river basin over time. Data Source: GLDAS, Rondell et al. 2004	101
Figure 69: PD of the Upper Don river basin in kg2 m-2 s-1, from 1984 to 2010, summed up monthly values per year. Data Source: GLDAS; Rondell et al. 2004	101
Figure 70: Mean annual air temperature in Kelvin over the Upper Don river basin. Data Source: GLDAS, Rondell et al. 2004.....	102
Figure 71: Water level in meters of the Tsimlyansk reservoir from 1992 to 2018. Source: Hydroweb, Theia Land 2018	103
Figure 72: Water level of the Tsimlyansk reservoir, averaged per calendar year. Ten data points have been averaged for each year, and are evenly distributed from January to December. Data Source: Theia Land 2018.....	104

List of Abbreviations

ANN	Artificial Neural Networks
API	Application programming interface
CIESIN	Center for International Earth Science Information Network
DPSIR	Driving forces-Pressure-State-Impact-Response Framework
dS	Change of water stored in rivers, ponds, lakes, or underground
EC JRC	European Commission Joint Research Center
EEA	European Environment Agency
EGS	Ecosystem Goods and Services
ET	Evapotranspiration
EU	European Union
FAO	Food and Agriculture Organisation
FAWR	Federal Agency of Water Resources
GEE	Google Earth Engine
GES DISC	Goddard Earth Sciences Data and Information Services Center
GFW	Global Forest Watch
GLDAS	Global Land Data Assimilation System
kNN	k-Nearest Neighbour
Landsat ETM+	Landsat Enhanced Thematic Mapper +
Landsat MSS	Landsat Multispectral Scanner
Landsat OLI	Landsat Operational Land Imager
Landsat TM	Landsat Thematic Mapper
LCS	Land Change Science
LULCC	Land use/Land cover change
MEA	Millennium Ecosystem Assessment
MLC	Maximum Likelihood Classifier
NASA	National Aeronautics and Space Administration
NDVI	Normalized Difference Vegetation Index
NIR	Near-Infrared
OCHA	Office for the Coordination of Humanitarian Affairs
OECD	Organisation for Economic Co-Operation and Development
P	Precipitation
PD	Precipitation Deficit
PSR	Pressure-State-Response
R	Runoff
SEDAC	Socioeconomic Data and Applications Center
SVM	Support Vector Machines
T	Temperature
TMT	Thousand Metric Tonnes
UNWPP	United Nations World Population Prospects
USGS	United States Geological Survey
USSR	Union of Soviet Socialist Republics
WRI	World Resources Institute

1 Introduction

More than half of all land surface has been permanently altered by humans, more than one-fifth of it has been transformed to permanent cropland, and more than one-fourth of all forests have been cleared before the start of this millennium (Steffen *et al.* 2002).

Changes in the land use and land cover (LULCC) of our planet have been happening since the dawn of humanity. However, with an accelerated population and consumption growth in recent decades, it has become more important than ever to monitor these changes, and draw connections between the social and ecological processes connected to them. LULCC, as they are happening now in the present scope and speed, have immense implications for all parts of our environment, and have changed the characteristics of our atmosphere, our soils, and especially our water.

Water is, without a doubt, one of the most crucial resources for human life, with rivers and lakes being one of the most accessible forms of freshwater. Our water consumption is continuously rising, and the agricultural sector is using more than 70% of our world water resources today (Cosgrove and Rijsberman 2000). Analysing LULCC in sea and river catchment basins that are located within agriculturally active areas, is, therefore, of utmost importance to ensure an adequate policy response to the increasing pressure on the planet's freshwater ecosystems.

Additionally to the already complex question of what changes happen in any given time frame, it is, furthermore, essential to discuss what drives these changes, and what impacts LULCC have on the water cycle of a catchment basin, and on the streamflow of rivers. Land use must be seen as an integral factor in a wider socio-ecological system that cannot easily be divided into its parts. Different economic and social influences can lead to changes in specific regions around the planet, and their effects can have vast implications for the ecosystem services of a watershed that, in return, influence human well-being.

The Azov Sea basin makes an especially intriguing case study, as it is among the most populated areas in Ukraine and Russia, and is of great agricultural and economic significance to the wider region (Lagutov and Lagutov 2012). Politically, this region has gone through a lot of changes recently that have influenced the land use – with the fall of the Soviet Union, the agricultural system of the area had to change completely from a centrally planned to a market-based economy, which has affected the livelihoods of people in the region (Ioffe 2005). Furthermore, the Donbas conflict, which has been going on since 2014, has led to the displacement of almost two million people (Coman 2017), which affects the land use and therefore, the environment in the border region. However, also ecologically, the Azov Sea basin is of interest as it combines two major rivers – the Don and the Kuban – and is home to several key species, such as sturgeons.

Most importantly, however, there has been little attempt to monitor the land use changes within this research basin. While there are general land use maps that cover the basin, no study has been found that observes the changes in small time steps over the whole basin and over large time frames, which is necessary to reach a high-quality research base on drivers and impacts of LULCC in a watershed. Therefore, this research gap needs to be filled in order to ensure a better understanding of the socio-ecological processes within the Azov Sea basin.

Remote sensing is making LULCC analysis more accessible and efficient. High-quality satellite imagery such as the Landsat or Copernicus programs improve research in this area and reveal new possibilities. Using online, server-sided computer programming to process satellite imagery, as made possible by Google Earth Engine (GEE), therefore saves time and effort on the researcher's side. This means that even researchers with otherwise insufficient resources, and especially small computational power, can analyse large areas of land in frequent time steps, which makes land change science (LCS) more accessible for low-budget research projects such as this one. With this new technology that is in many ways a game changer, a

research project of only a few months and a single researcher, like this one, can monitor LULCC over large time frames, with frequent time steps, and high accuracy measures. This could have never been possible before; only through open data policies and cloud computing platforms, these advances can be made and can help us reach a higher level of knowledge.

While remote sensing and automated classification technologies are more complex on the inside and give us more precise pictures of reality than ever, they are also easier to understand, adjust, and apply. Recognising and utilizing the immense possibilities opened up by a combination of machine learning, cloud computing, and remote sensing techniques, can help us make more informed decision about our future. This is a combination of techniques that is relatively new and not well explored, despite its advantages regarding research quality and efficiency. While the techniques are getting more and more developed, there is still a vast potential that we have not yet reached and can only get to by collecting case studies of their application (Lary *et al.* 2016). Therefore, to bridge the gap between science and decision makers, these cloud computing resources must be explored more, and their effectiveness has to be shown with case studies like this one.

To sum up, this study is combining a crucially important research focus with a key case study area and a novel and exciting methodology, to hopefully reach conclusions that can add to the current state of knowledge in all its facets. The problem as defined above must be viewed from several different angles, and therefore, this study will hopefully serve as a model for similar studies utilizing Google Earth Engine. The developed datasets will be made available on the CEU Environmental Systems Laboratory to enable future projects to benefit from this study¹.

¹ <https://syslab.ceu.edu/projects/GEE-land-use-azov-basin>

1.1 Research Aim, Questions, and Objectives

Resulting from the problem definition above, the aim of this research is two-fold:

On the one hand, the aim of this research is to analyse how land use changed in the Azov Sea basin and the subbasins of its two main feeding rivers, the Don and the Kuban, and to discuss what drives these changes, and what effects they have on freshwater ecosystem services.

On the other hand, this study also has an explorative component, as it aims to discover the possibilities offered by Google Earth Engine, a free, open-source API that allows users to process remote sensing products through cloud computing quickly and easily.

These aims will, therefore, be split into two main research questions:

1. How can Google Earth Engine be applied for the analysis of land use changes on a watershed scale, and what are some of the limitations?
2. How have the Azov Sea catchment basin and the defined sub-basins developed in recent years, what drives these changes and what impacts do they have on the Tsimlyansk reservoir and its ecosystem services?

To answer the research questions outlined above and achieve the two research aims, several objectives need to be adequately addressed:

- 1- Analyse land use changes in the Azov Sea Basin and its subbasins with Google Earth Engine, and validate the classification results
- 2- Calculate the abandonment or addition of agricultural and urban land
- 3- Discuss potential drivers, collect and analyse socio-economic datasets of the area and relate them to the land use changes observed before
- 4- Draw connections between land use in the area, and the provision of freshwater ecosystem services in the Tsimlyansk reservoir

1.2 Thesis Outline

This thesis started by explaining its objectives and aims, and the structure of the following pages will strongly be linked to the different objectives as explained in this brief **Introduction (Section 1)**.

Thereafter, the thesis will continue with a **Literature Review (Section 2)** that will discuss the guiding theoretical background and assumptions. The theoretical backgrounds discussed are land change science, the DPSIR-framework, and the ecosystem services framework, that all play a crucial role in the remaining parts of the thesis. Key terms will be defined there as according to the literature.

Afterwards, a review of the **Application and Layout of GEE (Section 3)** will be explained to set the stage for the methodology and main parts of the thesis, and the few case studies that already applied GEE and gave it its reputation will be summarized with their strengths and weaknesses. Subsequently, the types of classifiers that exist and are applied in land change science with GEE will be discussed; this is necessary to understand the choice of the random forest classifier in the later chapters and establishes the theoretical knowledge of processes that generally lie beyond the control of the user of GEE.

Following this literature review, **Section 4** will summarize the **Methodology and Approach** of the thesis. This section will furthermore explain how the programming skills were acquired, and where the utilized data came from. Workflow charts will visualize the process.

Section 5 then gives a background on **the Case Study Area – the Azov Sea basin** and its characteristics will be reviewed to give a clearer picture of the unique struggles of land use in that area. The focus of this section will be the ecosystem and the goods and services it provides, as well as the use of the area for agricultural purposes. Furthermore, the political challenges of the area and its development will be outlined.

After that, the **Land Use Classification Process (Section 6)** will be described in detail to enable the scope and tools of this research to be as comprehensible as possible. In particular, this section will include the description of Google Earth Engine and how it was used as a tool, which will include screenshots and snippets of code to make the extensive programming that has been carried out more understandable and to answer the first research question. Moreover, the extraction of data and the accuracy assessment will be discussed; in this section, the results will furthermore be compared to other land use maps made with GEE.

In the following **Impacts and Drivers of Land Use Change (Section 7)** chapter, the results obtained in section 6 will be reviewed in more detail and theories about what changed and why will be elaborated to answer the second research question. This chapter contains a lot of data that goes beyond the land use change and combines land change science with the DPSIR framework, as well as ecosystem services.

Finally, **Conclusions (Section 8)** will be drawn and will summarize concisely the answers to the research questions.

2 Review of Literature and Theoretical Background

In the following pages, different aspects of this research will be discussed to set the stage for the further chapters. With it, guiding underlying theories will be explained and key-terms will be defined as according to the literature. It will establish the theoretical and conceptual frameworks employed by this thesis, how they have been used in the past, and how they interact with one another. Three main frameworks will be applied:

First, Land change science (LCS) as a framework in itself will be explained. How this study developed in the past, how it can be differentiated from other concepts, and what guiding assumptions and theories lay beneath it will be the focus.

Second, the DPSIR framework, developed by the European Environment Agency (EEA), will emphasize how land use change can be seen in a bigger picture of its drivers, impacts, and potential policy responses and how the existing literature can be applied in this study.

Third, Ecosystem Goods and Services (EGS) will be defined and discussed briefly, as it is a guiding theory behind the assumption that land change has a direct effect on human well-being.

2.1 Land Change Science

How humans alter the land cover of our planet has been a priority for scientists for a long time. Nevertheless, land change science (LCS) is still developing as its own contemporary discipline as land change was exaggerated in its speed and scope in recent decades (Gutman *et al.* 2004). Furthermore, modern satellite and earth observation technologies only relatively recently gave us the possibilities necessary to efficiently monitor land use changes over larger scales; since this has never been possible before, it revolutionized LCS. Through these new technologies, the field became so big so quickly that a full literature review of all studies considering land use change is impossible to carry out (Gutman *et al.* 2004), and despite the extreme importance of land use changes for many other issues, such as climate change (Pielke 2005), or biodiversity loss (Sala *et al.* 2000), purely exploratory and methodological studies exceed real case studies

of land use change dynamics (Turner 2002). Therefore, the following pages should just give a brief overview of the current state of knowledge and serves as a theoretical background. A later chapter will review the explicit technology – GEE – used in this paper and the studies that applied it thus far.

Land change is usually split into a change of land use - the social component of how humans use a specific land area - and land cover - the natural component of vegetation cover, soil types, et cetera. Those two categories influence one another heavily and have mostly been treated together to combine the strengths of the different categories into one consistent body of research on LULCC (Meyer and Turner 1992). Combining sustainability science (Kates *et al.* 2001) and research concerning global environmental change, LCS seeks to understand the different dynamics of human-environment systems (Rindfuss *et al.* 2004; Turner *et al.* 2007). Correspondingly, the objectives of LCS include both the monitoring and observation of land change, as well as the understanding of how human systems change and influence the natural component of the land cover (Turner *et al.* 2007). Additionally, it aims to understand the effects LULCC can have on ecosystems and their services that are directly related to human well-being (Rindfuss *et al.* 2004; Turner *et al.* 2007). Accordingly, it is important to realise that land use changes are non-linear and depend on many different factors that are not easily dissectible (Lambin and Meyfroidt 2010). As such, LCS is an interdisciplinary field demanding the skills of experts from different major research subjects - remote sensing specialists, natural scientists, as well as social scientists are needed to fulfill the demands of LCS systems-thinking (Turner *et al.* 2007).

LCS is happening on all scales; from local and detailed to global and generalized studies, all levels can be found and are necessary to compile a comprehensive body of knowledge. Unifying them to a consistent knowledge base is difficult; nevertheless, spatial explicitness is essential to LCS (Magliocca *et al.* 2015).

A paper by Allan (2004) is criticising studies that focus on the differences in ecological integrity over the course of a river basin without keeping in mind factors such as non-linearity, the common natural factors (such as underlying bedrock and soil types) that influence both the land use and the river ecosystem, and legacy effects. To avoid problems of confusion between cause and effect, more studies need to be conducted that look at the temporal change of land use and EGS, instead of looking solely at the spatial differences of these two factors in an area (Allan 2004).

While LCS shares a lot with other theoretical frameworks, there are significant differences. For example, it shares its interdisciplinary focus on socio-ecological systems that cannot be analysed separately with political ecology; however, while LCS tends to focus on the outcomes of specific processes in either the human or the environmental subsystem, political ecology usually retains a strong focus on the human component and the effects of environmental changes on societal systems (Turner and Robbins 2008).

The intersections between economic geography and LCS are also large; both share their interest for spatially explicit models on the interactions of the society, economy, and the environment. However, economic geographers are more likely to focus on the role of markets and see the economy as something inherent to society, as opposed to outside and linked to it. An integration of these frameworks is advised to produce research that is more relevant to policymakers and that does not see the economy as something separate (Munroe *et al.* 2014).

In the past, research about land use changes has often focused on the role of forests and their development only, such as the classical work by Mather (1990) that explains the development of forests as a U-shaped curve; as a nation develops, its forest-covered lands will first decline as it is replaced with industries and built-up land; then eventually, it will increase again with further development as society understands the ecosystem's value (Mather 1990). A study by Lambin *et al.* (2001) expanded this theory and found that endogenous reforestation, meaning

an increase in forest area due to the degradation of its ecosystem services which lead to humans withdrawing from the space and moving elsewhere, is rare in today's society. Much more often, technological advances drive the further exploitation of forests, and reforestation is often exogenous, taking place due to an institutional decision to grant more space to the natural environment (Lambin *et al.* 2001).

However, forests cannot be seen as a separate entity; the question always needs to entail what land use/land cover replaces the forests, and which ones are, on the other hand, replaced by forests. In LCS, the natural environment must be seen and pictured as a part of the socio-ecological system, and not as separately de- or increasing (Barbier *et al.* 2010).

A conceptual framework used in land change science that breaks this issue down to a simple relationship is, as according to Lambin and Meyfroidt. (2011, 3465):

$$\text{Natural land} = \text{Total land} - (\text{Agriculture} + \text{Settlements})$$

While this is, of course, a simplified view of reality, it opens up opportunities to see how the system operates and how different land use classes interact with one another. While "natural land" and "non-natural land" are, of course, artificial categories that ignore the spectrum of human interference that really exists (Lambin and Geist 2008), some kind of differentiation is still useful for the analysis of land use effects, even if the cut-off points have to be defined. This research paper will therefore mainly take changes among these three classes into consideration – Natural land, which is mostly forested area in the research basin, agricultural land (including pastures and cropland), and settlements/built-up land, which includes mining areas and industries. As a necessary addition, the extent of the basin covered by water will be analysed. The guiding assumption for this paper will be, therefore, that whatever part of land is not utilized as agricultural or urban/built-up land, will be free for the natural environment and ecosystems.

The assumption that Lambin and Meyfroidt (2011) take further, is that the amount of agricultural land is determined by the population, the consumption per capita, as well as the yield per area. This would be true in a global context, however, as this study is spatially explicit and focussed on only a limited area within Ukraine and Russia, this assumption is not considered suitable, as trade between this area and others is taking place, and its analysis would be beyond the scope of this thesis.

Numerous papers (DeFries and Eshleman 2004; Dwarakish and Ganassi 2015) highlight the importance of technological advancements of hydrological modeling and a river's dependency on land use and land cover change. They argue that an interdisciplinary focus is more and more important, considering ecological, economic, social and other factors when designing scenarios and studies about the future development of land use changes and rivers. Decision makers need to be provided with clear information and instructions that are backed-off with suitable data – this seemingly impossible task comes closer to reality through the exploration of different information technology tools, such as GEE.

2.2 DPSIR-Framework

The Driver-Pressure-State-Impact-Response (DPSIR) framework is a well-known conceptual framework that aims to identify causal relationships between different processes. It started as the Pressure-State-Response (PSR) Framework that was established by the OECD (Organisation for Economic Co-Operation and Development) (OECD 2003) and was then developed further by the EEA (European Environment Agency) to include the driving forces and impacts as steps within the framework (Smeets and Weterings 1999). The official graph showing the relationship between the different frames, as well as examples for each of them, is shown in Figure 1 (EEA 2016).

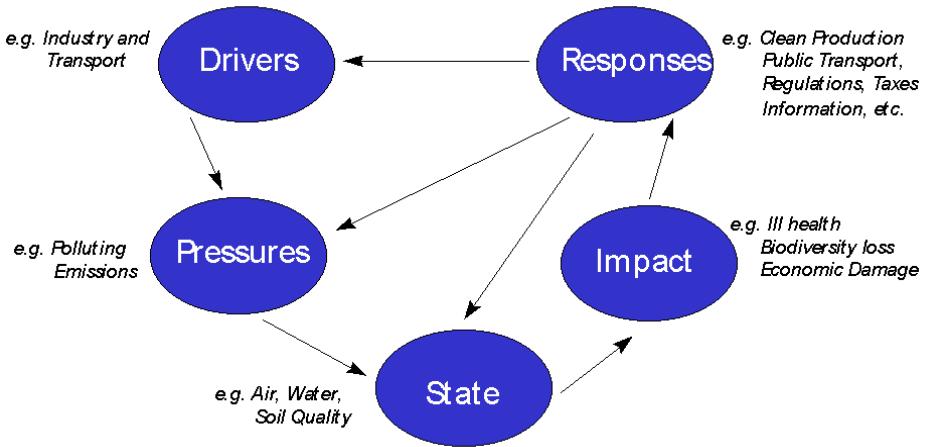


Figure 1: The DPSIR-Framework. Source: EEA 2016

These variables have been defined in various ways in the past when it comes to the role of LULCC. In some studies, land use change itself is seen as a driving force that e.g. leads to the release of pollutants (pressure) that then leads to a change in water quality (state) and furthermore to a change in the river quality (impact) (Benini *et al.* 2010). Others see LULCC as a pressure; then, the drivers would be migration or the transition of markets, the state can be defined as the change in the stock of a certain land cover, and impacts take the form of resulting changes in the water quality or streamflow (impact) (Potschin 2009).

The way Potschin (2009) defines the different stages adds the element of the human behaviour at the beginning of the process instead of just as a “Response”; this is more suitable for the further analysis of this research and will, therefore, be the applied framework hereafter. With LULCC being, in fact, a pressure, socio-economic drivers that lead to this pressure will be explained explicitly. The state will then be the absolute amount of each land use type at each point of time, and the impacts will especially be focussed on the ecosystem goods and services of the water bodies and the river system in the basin’s area and the Tsimlyansk reservoir. As policy responses have not yet taken place with respect to any of the mentioned aspects, this variable cannot be considered fully.

Since its development, the framework has been applied frequently for papers that aim to produce knowledge targeted at policy and decision makers. It wants to show how policy

responses can influence every single step of the way of environmental degradation and therefore aims to bridge the science-policy gap (Tscherning *et al.* 2012).

A recent review of existing scientific papers applying the DPSIR framework showed that while most studies mention it only in theory, it has indeed been successfully applied to make outcomes more understandable for professionals of different domains, and therefore has the potential to support decision making. Furthermore, it provides an ideal environment for interdisciplinary analysis of complex issues that are neither purely manifested in the natural world, nor purely in human societies (Tscherning *et al.* 2012).

Several papers have tried to embed the DPSIR framework within different other frameworks to establish a more holistic view of the processes. One example of this is how Hägerstrand's system of nested domains (Hägerstrand 2001) can bring the appropriate different levels of scale to the DPSIR framework that is otherwise not dealing with scale and can, therefore, be problematic in that area (Ness *et al.* 2010). This view, however, will not be established, as the "Responses" to land use change is not the focus of this paper, and the different levels of political scale are therefore less relevant to this research.

Several studies have furthermore treated changes of Ecosystem Services, as defined in the next chapter, as a potential impact within the DPSIR framework, which helps with relating social and ecological factors to one another (Xue *et al.* 2015).

Although helpful, the DPSIR framework has also been criticized in the literature for its discursive biases. The applicant can easily draw boundaries and exclude any social impacts or states; it is, therefore, important to consider as many aspects as possible and elaborate on any boundaries that have been drawn, to avoid excluding relevant parts of the socio-ecological system (Svarstad *et al.* 2008).

This thesis explicitly draws quite narrow boundaries due to time and resource constraints. The focus is going to be land use/land cover changes as a pressure that influences the state of land

covered by different types of vegetation or built-up material. The secondary focus is then going to be what drove these changes from a socio-economic point of view, and what impacts they have on the water cycle of the area and the provision of freshwater ecosystem services. Policy responses have not really taken place yet, and the scope of this research is too narrow to give recommendations; therefore, responses will not be analysed further.

2.3 Ecosystem Services

As mentioned before, the analysis of land use changes on the scale of a sea catchment basin is of importance due to the ecosystem services of the river system, the sea, and not least – the reservoir. Land use change is, therefore, directly linked to the well-being of humans, and can act as a “Pressure” on the “State” of the environment, as discussed above. For this purpose, the term Ecosystem Services needs to be defined.

The Millennium Ecosystem Assessment (MEA) defines ecosystem services as “*the benefits provided to people, both directly and indirectly, by ecosystems and biodiversity*” (Aylward *et al.* 2005, 216). This changed the way we see ecosystems forever by directly linking the well-being of an ecosystem to the immense benefits it can bring to humanity and our societies (Daily and Matson 2008). Nevertheless, the pressure on river ecosystems might have never been greater than now (Aylward *et al.* 2005). Ecosystem goods and services (EGS) of rivers are present in all four categories. Some examples for each of the four categories are (as according to Aylward *et al.* 2005, 216):

- Provisioning services: drinking water, power generation, transport, fish and other aquatic organisms for food
- Regulatory services: natural water filtration and treatment, disaster and flood prevention
- Cultural services: tourism, recreational activities
- Supporting services: nutrient cycling, ecosystem resilience

All of these services can only be provided by a watershed if the water quality is good enough, and if there is enough water at each point of storage or flow. A drying out reservoir can no longer be used efficiently for power generation, and with collapsing fish populations, humans can no longer gain benefits from them for food and leisure. Land use in that sense influences freshwater ecosystem services, as it can reduce or increase the evapotranspiration rates, the local temperature, and the amount of runoff (Dwarakish and Ganasri 2015; DeFries and Eshleman 2004). All of this will be discussed in greater detail in later chapters.

However, for this research, it is not only important to ask what ecosystem services the river system and the sea itself can bring, but also how the supporting ecosystem services gained from natural vegetation, especially forests, that are present in the river catchment basin, can help to sustain the river's balanced ecosystem and regulate the water quality and flow (Guo *et al.* 2000). Ecosystems and their services can therefore not be viewed separately, but the linkages and interactions between them must also be taken into account.

Defining the benefits our society can, directly and indirectly, gain from ecosystems, as per definition of EGS, can help us comprehend the importance of protecting them on a societal level. Furthermore, it is a framework for evaluating and assessing the ecosystem's status quantitatively, which helps to make interesting comparisons between different management practices to ensure the best long-term results. However, this should be handled with care as ecosystems are complex and diffuse; they are influenced by many different causal relationships that are not yet fully understood. Destroying an ecosystem beyond a certain point can, therefore, lead to the point of no return, or tipping point, that damages the system of biotic and abiotic components permanently (Huang *et al.* 2015; Dai *et al.* 2012; Veraart *et al.* 2012).

Linking human behaviour to a change in the value of ecosystem services is, therefore, a difficult yet important task. Figure 2 shows an example of visualizing the cause-and-effect relationship between "Action" (e.g. land use changes) and this change in the value of a river ecosystem.

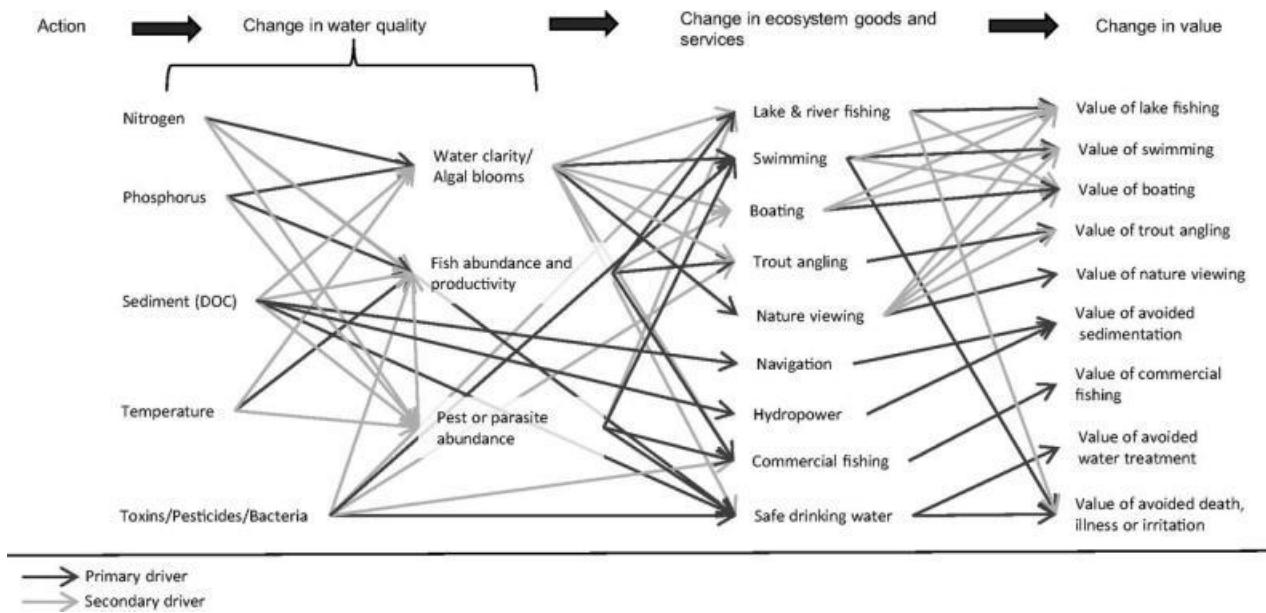


Figure 2: Linking Actions (including, but not limited to LULCC) to a change in the value of ecosystem goods and services of a river system Source: Keeler et al. 2012, 1862

What the graph does not show, however, are the concrete factors that determine the relationship between LULCC and a change in water flow; it just assumes that this link exists without further collaborating the ways of influence. While plenty of studies aim to show the relationship between land use in a watershed and the river's water quality and flow, there is still only limited understanding in terms of future development and adequate management practices and policy recommendations (Allan 2004), which is why the combination of the DPSIR framework with ecosystem goods and services is a welcome tool as discussed above (Potschin 2009).

2.4 Conclusion of the Literature Review

To sum up, LCS is an emerging, interdisciplinary subject that requires its scholars to think of systems as a whole that are made up of human-made and natural components that interact with one another. It requires spatial explicitness and time-series analyses rather than merely the comparison of different areas. While the possible levels of detail are uncountable, distinguishing natural, agricultural and built-up land needs to suffice for the purposes of this study, and the relationship between those categories needs to be examined.

The DPSIR-Framework further embeds land change in a wider array of driving forces, impacts and responses. LULCC can then be seen as pressures that are driven by societal processes and factors, and that influence, on the other hand, the provision of freshwater ecosystem services due to their effect on water quality and yield at different points of the water cycle. The combination of these three frameworks makes this study better embedded in the current level of knowledge and aims to make it as comprehensible as possible for decision-makers and scholars from different disciplines.

3 Review of Google Earth Engine

After the discussion of the theories above, the possibilities of remote sensing technology for water management and land use classification schemes need to be discussed in greater detail. This includes the importance of using different technologies and the novel possibilities to achieve more precise results due to Google Earth Engine. The few already existing case studies utilizing Google Earth Engine will be reviewed to emphasize the importance and opportunities offered by this new tool, and how this research study adds to the current knowledge.

Remote sensing is probably the easiest way to observe LULCC on large scales and has been applied and developed widely in recent years. Through remote earth observations, data can be collected over large time frames as well as large spatial areas, and it is nowadays used for plenty of research questions. Hereby, not only the remote sensing technology is important as it provides us with the necessary images for land use classifications; also, automated classifiers are needed to relieve human researchers from the burden of classifying maps by hand, which saves endless time and resources. The following chapters will first discuss the software used in this study, how it is setup and how it has been utilized in the past, and then discuss modern classification technologies. This analysis gives valuable insights into the automation of the classification process and justifies the choice of the classifier in this thesis.

3.1 GEE Application and Case Studies

Modern technology is evolving day by day and is expanding our possibilities to deal with pressing environmental issues quicker and more efficiently. However, in order to be able to further develop the various software available, more research and real-life applications of the tools are necessary to validate their functionality and reveal issues that can then be resolved.

Google Earth Engine (GEE) is offering scientists and researchers the possibility to utilize JavaScript programming in a web-browser based online interface to efficiently analyse satellite images and vector data from a variety of sources, mostly available on the global scale, with

high computational power. By utilizing cloud computing and an intrinsically parallel process, even scientists without the resources to do offline satellite imagery processing can take advantage of remote sensing technologies due to GEE (Moore and Hansen 2011). GEE combines all Landsat collections with over two million images, as well as Sentinel, MODIS, GlobCover, and other data, and through the online processing system, the user does not have to have access to large computational power, and there is no need to download large amounts of data. This opens up possibilities to lay researchers or students that often lack the adequate resources for offline processing of large areas of high-quality satellite imagery (Gorelick *et al.* 2017). There is also the possibility to use GEE with Python, however, as this will not be the approach of this case study, that possibility will not be discussed in detail.

The layout and setup of Earth Engine are pictured in Figure 3 and will briefly be explained here. At the top, GEE established a search bar where users can easily find the different data collections available via keywords, but also the geographic location of different places can be found.

On the left side, users can find their own as well as example *Scripts*, the *Docs* made available by GEE, and *Assets* users can upload and use later in their scripts.

In the centre, the main script area is located. How the code actually looks and works will be explained in more detail in the main area of this thesis via screenshots and code snippets.

On the right, there is space for three different outputs – the “*Inspector*” tab shows the band values of all images embedded to the *Map* at the bottom at exactly the location the cursor is placed. The “*Console*” shows any output the user called for with “*print(...)*” and can visualize results, tables or metadata of images and image collections this way. At “*Tasks*”, tables and images that were called by “*Export.toDrive()*”, “*Export.toAsset()*” or other export destinations will be shown and can be run on demand. This is useful for larger computations, as the *Console*

has a maximum computation time of five minutes, while *Tasks* have an unlimited computation time (Gorelick *et al.* 2017).

Finally, the *Map* at the bottom can be used to visualize results, draw polygons, or set geographic boundaries and locations.

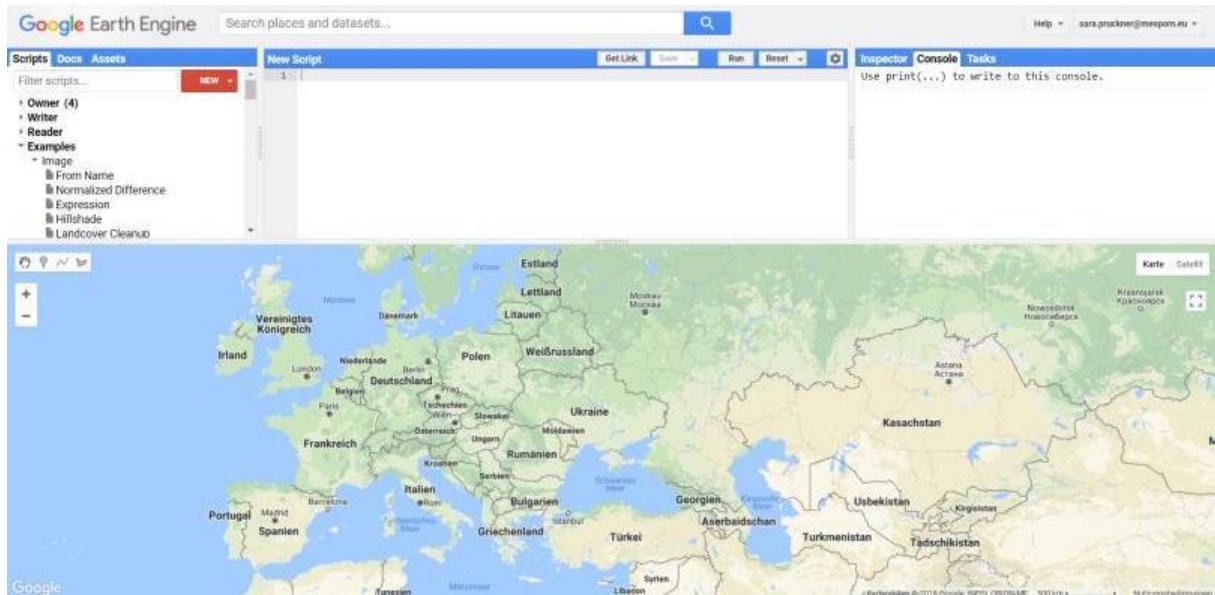


Figure 3: Set up and user interface of Google Earth Engine. Source: Google 2018a.

GEE is based on, and functions best with server-sided demands, as client-side methods often lead to timeouts, or result in wrong information. For this purpose, GEE has defined several objects, all found in the “*Docs*” tab, that start with *ee*. and are exclusively managed by the server. These objects can be images, image collections, feature collections, arrays, or others. If an object has to be managed on the client side, this can only be done by calling *.getInfo*, which should be avoided as the benefits of lazy server-side computing cannot be held up with this method, and other parts of the code need to be blocked temporarily until the client-sided computation is over (Gorelick *et al.* 2017). GEE utilizes lazy programming on its servers, thus it will not compute anything that is not necessary for an output in either the “*Console*”, the “*Map*”, or the “*Tasks*” (Navarro 2017). This increases computational speed and prevents the premature using up of the quota of concurrent requests for computation each user has. Due to the focus on server-sided computation and “objects”, using GEE is quite different from

traditional JavaScript programming, which increases the time it takes to fully comprehend GEE's analysis methods. However, after that initial period, GEE is a powerful tool and a true game changer.

Earth Engine is becoming more and more popular among researchers all around the world as more people recognise the benefits and wide array of applications of this tool. A few of the case studies that illustrate the possibilities of GEE will be summarized here.

One example of a global scale case study implemented with Earth Engine and the Copernicus Programme is the Global Surface Water Explorer, carried out by the European Commission Joint Research Center (EC JRC) (Pekel *et al.* 2016). This project monitors the extent of surface water on land, as well as its seasonality and changes over time, using Landsat 5, 7 and 8 imagery from 1984 until now. This analysis adds to the current knowledge on water seasonality, water stress and droughts on this planet. Where water is, and at what time of the year it will be available, is clearer and more understandable than ever before, which helps planning for sustainable development (Pekel *et al.* 2016).

Another prominent example using GEE is the Global Forest Watch (GFW) implemented by the World Resources Institute (WRI) in partnership with Google, Esri, and several other organisations and companies (Hansen *et al.* 2013). It aims to collect data about forest loss, forest gain and general canopy health and extend at a 30 meters resolution on a global scale. The GFW also uses Landsat 7 data just like the EC JRC, however, the temporal extent of the research is significantly smaller with initial data collection only reaching from 2000 to 2012, and in spite of new data being incorporated daily to keep the database up-to-date, historical data is not added. With this data, several publications were made possible that focus on key regions such as tropical rainforests, as well as broader analyses of drivers or impacts of deforestation (Goldman *et al.* 2015; Weisse *et al.* 2017).

A research team from different universities, among them Yale University and the University of Florida, has collaborated with the Google Earth Engine developers team to develop the “Map of Life”, as it is called. This map is supposed to show the places on earth with the highest biodiversity, and specifically, where certain endangered species have their habitats. This way, conservation efforts can be focussed better on the most important regions for the protection of specific species (Jetz *et al.* 2015).

In another study by the UC San Francisco, Google Earth Engine is utilized to map areas that have a high risk of future malaria outbreaks. In a later stage of this program, health professionals in risk areas will be able to add data about outbreaks themselves to feed an online database that will help to predict where the efforts to limit malaria should be concentrated to stop the spreading of the disease in due time. However, this project is still under development (Kurtzman 2014).

Just recently, a team of eight researchers from the U.S. Geological Survey (USGS) applied GEE to analyse and map cropland on the whole African continent from 2003 to 2014, collecting training data by citizen science, using a mobile phone application open to the civil society where people could enter different crop data. Thereby, they collected almost 3000 training samples. This took immense resources and while the results are impressive, a kappa value of around 0.72 means that the case study is still underdeveloped and in need of improvement (Xiong *et al.* 2017).

To sum up, Earth Engine is a powerful tool that is worth developing further. Despite the small number of case studies at the moment, it is getting more popular and can be expected to expand in the upcoming years. Utilizing it for a not yet explored, important case study such as the Azov Sea Basin, could make this research a model for further similar analysis. The fact that no images or Landsat tiles have to be actively downloaded for the application of GEE only made a research area of this spatial (a whole sea catchment basin) and temporal (thirty years of analysis in five

year steps) possible without a huge computational power. A land use analysis project has most likely not been performed on this scale by a single researcher within a few months.

3.2 GEE Land Use Classifiers

To classify basically any group of data, several different machine learning techniques can be used and should be considered. With artificial intelligence on the rise and technology more user-friendly, accessible and available than ever before, researchers of environmental sciences should recognise the immense possibilities of technology in order to improve our knowledge on the environment and test the use of machine learning on land use problems. This section is therefore also meant as a glossary of short descriptions of different techniques that must be considered, and will discuss the application of the different techniques in land use classification. Without a doubt, more case studies are needed to establish a good state-of-the-art methodology for land use change problems.

Beforehand, it is important to note that this chapter will only discuss pixel-based classification methods. While these pure pixel-based systems have encountered some headwind in the past few years, as more and more studies prefer using object-based or knowledge- and expert-based classification schemes, it is still the most widely used method (Aitkenhead and Aalders 2011; Chen *et al.* 2015; Richards and Jia 2013). Moreover, while these different schemes can be extremely useful and lead to a better accuracy and the ability to classify more types of classes, this goes beyond the scope of this research, and will therefore be left out for now.

To begin with, the different pixel-based categories of supervised and unsupervised classification must be made clear. Unsupervised classification clusters the available data points without knowing any labels yet. After the clustering, the user has to choose meaningful labels for the classes divided by the algorithm. The user can usually only determine the number of classes, but not much else (Lu and Weng 2007).

Supervised classification, on the other hand, needs a set of training data that has a set label attached by the producer of the map. The computer then tries to understand the differences between the categories (classes) in its different dimensions (when using remote sensing data, the bands are the dimensions) and therefore comes up with rules on how to classify a new point not contained in the training data (Lu and Weng 2007). This study will use exclusively supervised classification. Within the category of supervised classification, there are several different possible classification methods that are here discussed further.

To understand the different classifiers, the data needs to be imagined as a multi-dimensional field, where each dimension is a band of the satellite image, and each data point is a pixel in the image. The bands of the different Landsat satellites can be seen in Figure 4, as those will be the datasets used for the rest of the study.

Table 2. The bands of each Landsat satellite and descriptions of how each band is best used.

[--, not applicable]

Band name	Uses of Landsat bands					Description of use
	L8 OLI/TIRS	L7 ETM+	L4-5 TM	L4-5 MSS	L1-3 MSS	
Coastal/Aerosol	Band 1	--	--	--	--	Coastal areas and shallow water observations; aerosol, dust, smoke detection studies.
Blue (B)	Band 2	Band 1	Band 1	--	--	Bathymetric mapping; soil/vegetation discrimination, forest type mapping, and identifying manmade features.
Green (G)	Band 3	Band 2	Band 2	Band 1	Band 4	Peak vegetation; plant vigor assessments.
Red (R)	Band 4	Band 3	Band 3	Band 2	Band 5	Vegetation type identification; soils and urban features.
Near-Infrared (NIR)	Band 5	Band 4	Band 4	Band 3	Band 6	Vegetation detection and analysis; shoreline mapping and biomass content.
	--	--	--	Band 4	Band 7	
Shortwave Infrared-1 (SWIR-1)	Band 6	Band 5	Band 5	--	--	Vegetation moisture content/drought analysis; burned and fire-affected areas; detection of active fires.
Shortwave Infrared-2 (SWIR-2)	Band 7	Band 7	Band 7	--	--	Additional detection of active fires (especially at night); plant moisture/drought analysis.
Panchromatic (PAN)	Band 8	Band 8	--	--	--	Sharpening multispectral imagery to higher resolution.
Cirrus	Band 9	--	--	--	--	Cirrus cloud detection.
Thermal (T)	Band 10	Band 6	Band 6	--	--	Ground temperature mapping and soil moisture estimations.
	Band 11			--	--	

Figure 4: A comparison of different Landsat satellites, their bands, and optimal usage Source: USGS 2015

Depending on how much energy of a specific wavelength (corresponding to the bands, or dimensions) the pixel radiated, the data point is put on a specific position within the field (Richards and Jia 2013). To illustrate this, Figure 5 shows the theoretical scenario of only those two dimensions (bands) that are used for calculating the Normalized Difference Vegetation Index (NDVI), a well-known and widely-used parameter when detecting vegetation via remote sensing.

The NDVI is calculated as:

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

with NIR being near infrared (band 4 at Landsat 5 and band 5 at Landsat 8), and RED being the visible light perceived as red for the human eye (band 3 at Landsat 5, band 4 at Landsat 8). This parameter is useful as chlorophyll absorbs red well, while it disperses near infrared; therefore, the NDVI of green, chlorophyll-rich plants is high (around 0.8), while it is lower for unhealthy plants, and even lower for water or built-up land (usually around 0.2 to 0) (Pettorelli *et al.* 2005). Pixels containing vegetation and pixels containing waterbodies are, therefore, vastly different in their NIR and RED spectra and make a good example for explaining the use of classifiers in land change science.

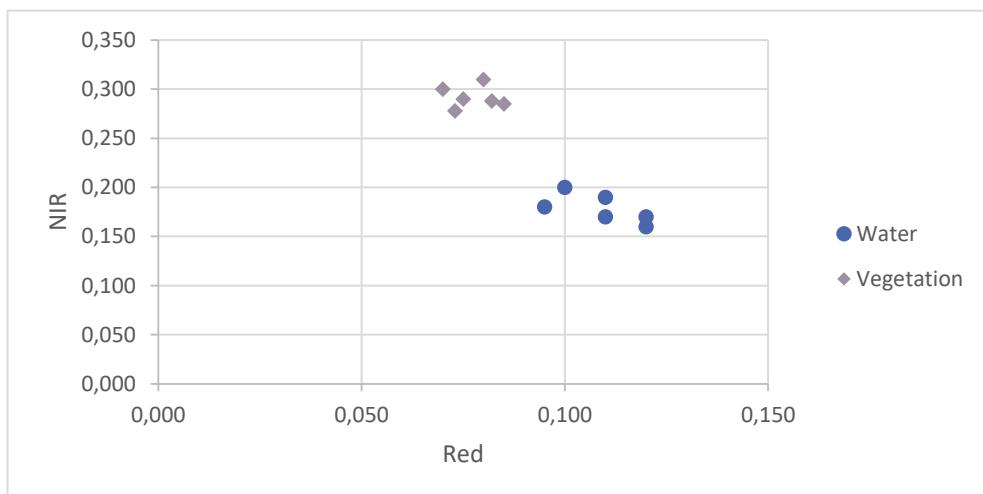


Figure 5: Dataset visualization. Classifying water and vegetation with the dimensions red and NIR

A machine learning classification system now takes all different dimensions into account (more than two in almost all cases) and tries to establish rules on how to put a new input data point in a category. The difference between the classifiers is how they try to establish those rules.

The simplest classifier is just a linear discriminant analysis (Figure 6). It tries out different lines drawn through the area and chooses the one that misclassifies the least training data, thus having the highest redistribution accuracy. New incoming data will only be classified as according to the established line (Lange 2016).

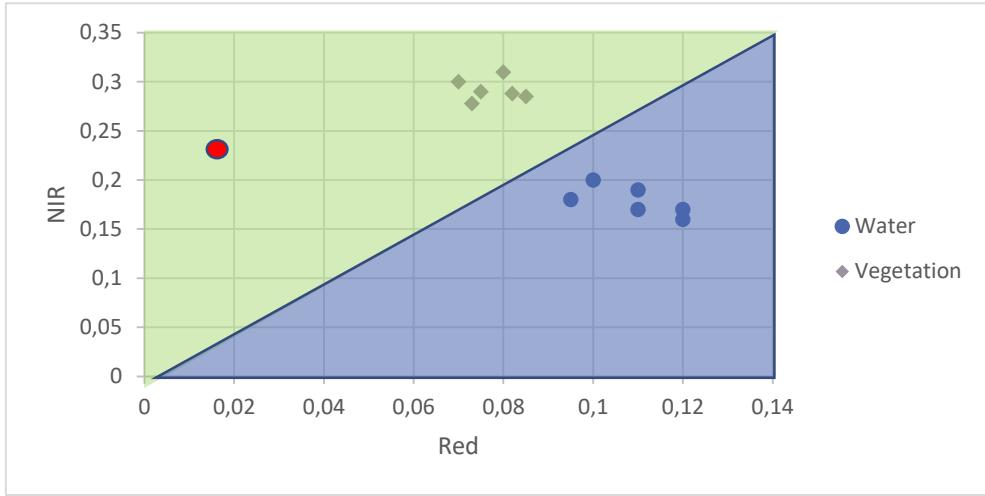


Figure 6: Visualization of a Linear Classifier. The new point in red will be classified as vegetation in this fictional scenario

Support Vector Machines (SVM) are one of the popular advancements to the mere “drawing of a line”. By establishing a kernel, a linear graph can be transformed to a polynomial line going through the graph and can therefore potentially accommodate outliers better (Lange 2016). Furthermore, an SVM usually has a “grey area”, as it is often a so-called “soft” classification method that allows new points to be put in between classes instead of forcing them into one class as a “hard” classifier does (Richards and Jia 2013). Even if used as a hard classifier, it improves the decision graph by drawing it where the distance to all training data points is largest, allowing for the biggest probability of classifying newcomers correctly (Lange 2016). Decision tree classifiers function in a similar way. Instead of drawing a single line, it always only cuts the area (or part of it) in two in the most optimal way, until most training data is classified correctly (Lange 2016). “Optimal” here means that it cuts off the most wrongly classified features possible in that move. This finally leads to a decision tree, where based on the specific value in one band, an element is either classified if that decision is distinctive, or moved further down to try out other criteria. This way, it is easy to visualize which bands have

the most influence on the classification, and in the case of restricted computational power, only the most important bands can be taken into account (Lange 2016).

A decision forest is a classifier combining several decision trees that have been built with only parts of the training dataset, which means that every single decision tree uses only parts of the available training pixels as well as only some of the available bands/dimensions. The classification is then chosen by a majority vote function for discretely numbered classification. A “Random Forest” classifier brings a random element to the decision forest which has been shown to reduce bias. It draws from the pool of training data pixels and returns the drawn element each time. To reduce even more bias, every decision tree only uses parts of the bands for its final decision, and randomly cuts out other bands. The random element is especially useful to avoid bias if the training data is not chosen as random points equally distributed over the area but just chosen strategically by the researcher in order to save resources and time (Lange 2016; Richards and Jia 2013). It is therefore well suited for a time-constrained project with limited resources, such as this one.

The k-nearest neighbour (kNN) classifier (Figure 7) is not drawing a line like SVM, decision trees, or linear discriminant analysis classifiers. It considers simply, for every new data point, the class of the nearest neighbour and classifies it this way. It can be improved by considering not one, but two or more (specified as “ k ”) nearest neighbours and choosing the class by a majority vote solution. KNN is a lazy learner, meaning that while the set-up of the model is extremely fast, the actual classification of the data takes more time as compared to other classifiers (Lange 2016).

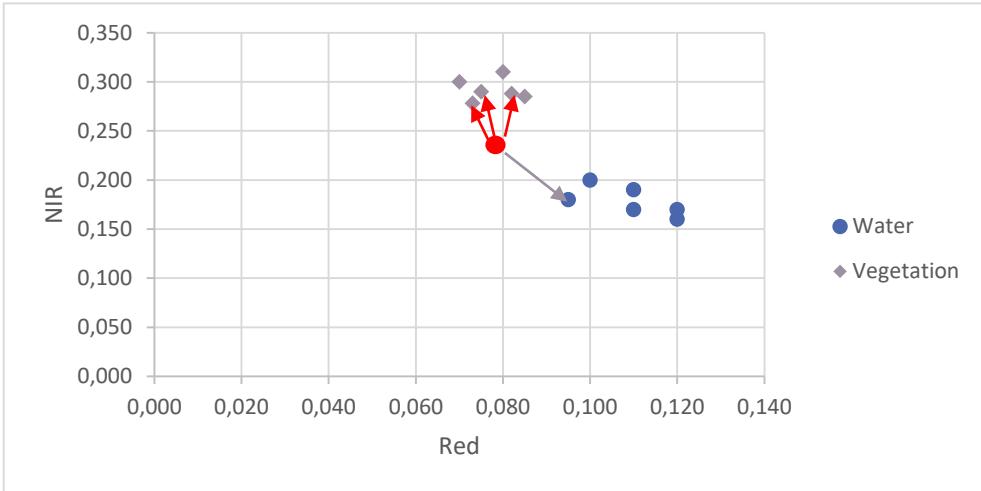


Figure 7: Visualization of the kNN classifier. The nearest neighbours all being vegetation, it will be classified as such with almost any k.

The Naïve Baye's law is another popular and simple tool. First, it calculates the probabilities of a point in a specific class having a specific feature. For example, the probability of a training pixel that is vegetation having a radiation value of above 0.25 in the Near infra-red wavelength, is in the above example 100%, as all training data points fulfill that criterion. The probability of it having a value of under 0.09 in the red spectrum is also 100%. Therefore, when classifying new data, the Maximum Likelihood Classifier (MLC) utilizing the Naïve Bayes algorithm will sum up and normalize the probabilities of each dimension of a pixel for each class, and select the class with the highest overall matching probability. For the maximum likelihood classification techniques, the number of training points being 10 to 100 times the number of bands in the dataset should be enough. (Richards and Jia 2013) This means that when considering 8 bands (dimensions) a minimum of 80 and a maximum of 800 pixels per class should be trained.

Just like random forests explained above, artificial neural networks (ANN) are combining different classifiers. Information is transmitted from each of the nodes to the next layer, with a unique weight and bias that is randomly chosen each time (Nielsen 2015). By combining the results of different classifiers and weighting them differently, this classifier can have very high accuracy rates (Nielsen 2015); however, it's training as well as classification time and power

need are very high, which is why it is unsuitable for a study of this scale. As of now, support vector machines (SVM) and artificial neural networks are the most explored techniques regarding the intersection of remote sensing and machine learning (Lary *et al.* 2016).

In conclusion, the variety of classifiers out there is large and for every case study, the ideal classifier needs to be detected. While all of them achieve good results with large enough training data sets, especially random forests, and other combined classifiers that decide by a majority vote tend to avoid biases that single classifiers might have towards specific classes. If training data is limited and not exactly equally distributed over the whole area, a random element should be introduced in the classification process, which is why the random forest classifier will be the classifier of choice for this study. As it is not a lazy learner, the overall classification process takes place faster, and the possibility to look at the different decision trees and the importance of the branches facilitates more transparency than other classifiers.

4 Methods and Approach

The main part of this thesis, especially Chapter 6 and 7, consists of several parts and smaller steps; the chronology and relationship of the workflow are pictured in Figure 8.



Figure 8: Workflow chart of the whole methodology

4.1 Land Use Mapping

First, the land use had to be analysed which is mainly explained in Chapter 6. For this purpose, skills in JavaScript programming and especially on the unique server-sided approach Google Earth Engine takes had to be acquired. This could mainly be achieved by the extensive Google Developer's Guide, online tutorials, and recordings of lectures of Google Earth Engine workshops that took place in the past (Google 2018b). All of these resources are freely available online and provided a good basis for this new skill. The workflow that followed can be seen in Figure 9.

After the basic skills were acquired, Google Earth Engine was utilized with a trial-and-error approach to analyse the longest possible time frame in as much detail as possible. It was concluded that Landsat 5 (for the years 1984 – 2011) and Landsat 8 (2013 – 2017) images are the best options for the case study at hand; this way, seven land use maps, representative for the years 1985, 1990, 1995, 2000, 2005, 2010 and 2015 were acquired that were each fed by data from August of three consecutive years (e.g. 1984, 1985 and 1986 for the data point labeled as “1985”). The training data was collected through a visual interpretation by the researcher of strategically selected data points of the Landsat images in the band combination 6-5-4 (Landsat 8) and 5-4-3 (Landsat 5), respectively. The visualization of the NDVI was used for decisions regarding vegetation. Hereby, it was necessary to ensure that the data point were well

distributed over the whole case study area. Subsequently, a random forest classifier was used for the supervised classification process which has been explained in the literature review, as the user has no influence on the process behind it. The accuracy was evaluated with separate validation data; training data was added bit by bit until the overall accuracy, the kappa value, as well as user's and producer's accuracy for each class, seemed appropriate for the purposes of this study. The cut-off values for appropriate accuracy were set to a kappa value and an overall accuracy value of minimum 90%, and user's and producer's accuracy of minimum 70% for each individual class. As a final step, the areas of the whole Azov Sea basin, as well as different sub-parts of the basin were calculated, to establish where in the basin change was happening.

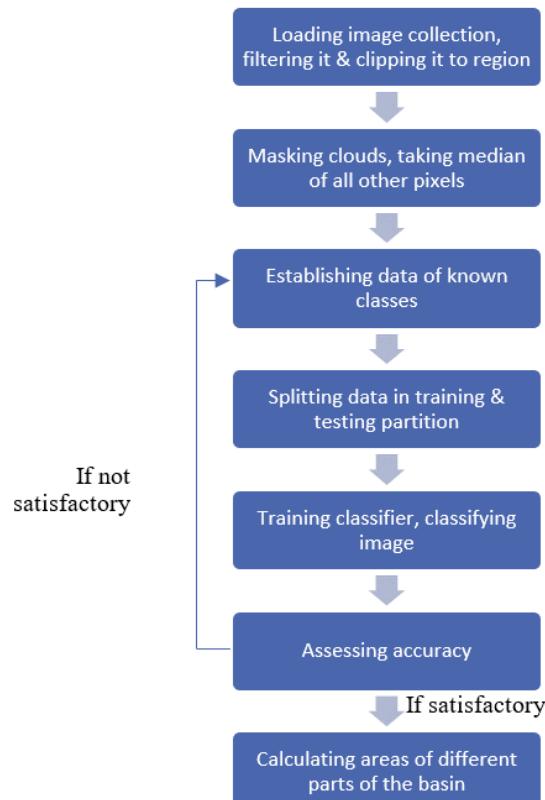


Figure 9: Workflow of the Land Use Classification Process with Google Earth Engine

4.2 Interpretation of Land Use Data

Additionally to the seven land use maps explained above, maps for the buffer region around the border of the Donbas conflict were created for the years 2013 to 2017, always taking the median

of all cloudless pixels available from May to September. The polygon of the buffer area was created with ArcGIS by geo-referencing and digitalizing the border of the war region by The Economist (2016), visible later in Chapter 6.3 in Figure 40. The digitized border was used to create a buffer zone of 20 kilometres on each side; the resulting polygon was imported to Google Earth Engine and further land change analysis was carried out there. For this part, the same training polygons as for the land use map of 2015 were used, minus those polygons situated directly within the conflict region to avoid confusion and false classifications.

To validate the overall results, the created land use maps of the whole Azov Sea basin were compared to data from the case studies mentioned above, namely the Global Forest Watch (Hansen *et al.* 2013), and the EC JRC Surface Water study (Pekel *et al.* 2016). This was done by calculating the areas of agreement and disagreement between the different datasets. This comparison gives further credit and validity to the results obtained in this study.

Furthermore, to interpret land use changes, the change in agricultural land was calculated in order to see which parts have always, never, or at certain points in time been agriculturally used. This is important as agriculture is the primary land use in the sea basin, and land abandonment is a well-known issue in the states of the former Soviet Union.

4.3 Collection and Analysis of Socio-Economic Data

After the finalization of the land use analysis, socio-economic data was collected to analyse the driving forces and impacts of the land use changes.

First, crop yield data acquired and made available by the EnviroGRIDS project (EnviroGRIDS 2012) for the yield of the Russian part of the basin has been analysed statistically with R to detect changes. As this was not a spatially explicit dataset, it was not fed to Google Earth Engine, but was only divided into the four subbasins.

Also, spatially explicit population data as acquired by the Center for International Earth Science Information Network (CIESIN) and made available from the NASA Socioeconomic Data and

Applications Center (SEDAC) (CIESIN 2011; CIESIN 2017a; CIESIN 2017b) has been used as an asset in GEE to calculate where the transformation to urban land was driven by an actual increase of population, versus where it has been driven by industrialisation and built-up without a population change.

Finally, to determine the effects of LULCC in the basin on the water cycle, data from NASA's Global Land Data Assimilation System (Rodell *et al.* 2004) has been used to calculate the area-average of different hydrological parameters of the Upper Don basin. While this data is available in Google Earth Engine, the author decided to use NASA's Giovanni application to download data and analyse it offline, as GEE's user memory limit was otherwise always exceeded. Nevertheless, the data quality and spatial and temporal resolution are the same, and the data has been visualized and verified with GEE. Among the analysed parameters are specifically evapotranspiration, precipitation, as well as temperature. This was, in further consequence, used to calculate the precipitation deficit and therefore, draw conclusions regarding the changing water level of the Tsimlyansk reservoir. Water level data was taken from the Theia Land Data Center (Theia Land 2018) that estimates their data through remote sensing. From the water level data, conclusions on the state of the ecosystem services of the Tsimlyansk reservoir can be drawn.

All of this data was set into the DPSIR-framework to analyse connections between the socio-ecological systems as a novel approach to the issue that has not yet been applied to the case study region.

4.4 Limitations

Limitations for the land use analysis definitely include the lack of training data from external sources. As the training data was, therefore, not randomly distributed over the whole case study region, biases towards specific land use types or regions are possible. This was addressed by choosing a random forest classifier, and by reviewing the maps and the underlying data several

times to increase accuracy and minimize errors. Nevertheless, an underlying bias (as with any land use classification) cannot be ruled out completely.

Furthermore, no different crop types or agricultural practices were analysed due to – again – limited data availability. Different vegetation types and agricultural practices, such as whether crops were rain-fed or irrigated and what the types of fertilizers were used, could potentially have a big independent effect that has a different influence than purely the amount of agricultural land. However, this type of data was not found despite serious efforts to collect it. Therefore, only the differentiation between “crops” and “natural vegetation” could be carried out, which still led to suitable results.

Another limitation is the imperfect Landsat data; several regions were cloudy which leads to data gaps and incorrect outcomes. Therefore, not the total area of each type was reviewed, but merely the percent-point changes between the land use types.

A limitation for the non-land use data was, once again, the unsuitable data availability in English for the region, and often the unsuitable time frames. For example, crop yield data was available for the Russian part of the case study region for 1990 and 2000 only, whereas, for the Ukrainian part, only data from 2006 to 2009 was covered. As these four years are of little interest to the overall research questions, the yield data only refers to the Russian part of the case study area and cannot speak for the whole Azov Sea basin.

Furthermore, the reliability and quality of water-related data in freshwater ecosystems in Russia is limited; several reviews of the publicly available data have shown that almost half of the water quality data gathered and analysed in the former USSR and in Russia until the early 2000s is not fit for scientific research (Zhulidov *et al.* 2003); therefore, an analysis of water quality could not be carried out.

In the following chapters, these issues will be addressed further and in greater detail.

5 Case Study Area: Azov Sea Basin

Figure 10 shows the location and extent of the research area – the Azov Sea basin – and visualizes the river system in the basin. It furthermore highlights the regional centres in the area, which all take advantage of the water the rivers Don and Kuban provide; the most important centres here being Kharkiv and Rostov-on-Don, which both have over one million inhabitants. Besides that, the location of the border between Ukraine and Russia, the two countries sharing the basin and managing the sea by a bilateral agreement, is pictured (Lagutov and Lagutov. 2012).



Figure 10: The Azov Sea basin. Source: Lagutov and Lagutov 2012, 9

In Russia, the Azov Sea basin covers the entire area of the provinces Rostov, Krasnodar Krai, Voronezh, Lipetsk, and Adygea, as well as parts of the provinces Karachay-Cherkessia, Stavropol Krai, Belgorod, Volgograd, Kursk, Oryol, Tula, Tambov, Penza, and Saratov. In Ukraine, parts of the oblasts Luhansk, Kharkiv, Dnipropetrovsk, Donetsk, Zaporizhia, Kherson and Crimea are covered.

5.1 Ecosystem

The Azov Sea itself has a unique ecosystem due to its relatively low salinity content of only 10,9‰ before the major river regulations and the establishment of the Tsimlyansk reservoir, and still only 14‰ thereafter. It is relatively shallow and the smallest sea in the world, but well supplied with nutrients and sediments by the two dominating feeding rivers, the Don and the Kuban. This makes it very favourable for a wide variety of fish and invertebrate species and causes its extremely high biological productivity as compared to other seas, which brings with it a high value for ecosystem services of all kinds (Lagutov and Lagutov 2012).

One of the most interesting and crucial fish of the Don are the different sturgeon species. They are widely used for their meat and are also of cultural importance, which makes them relevant for the provision of EGS (Dubinina and Kozlitina 2000). Due to the large alterations of the river flow in recent years, especially through the Tsimlyansk reservoir, spawning grounds for sturgeons become rarer, leading to a collapse of parts of the population. The increased salinity has furthermore damaged the bream and pikeperch population health and the overall ecosystem integrity (Dubinina and Kozlitina 2000).

In the catchment basin, most of the terrestrial ecosystem belongs to the biomes temperate forests and steppe; the steppe-zone of both countries being the most important agricultural area as explained below. Only in the South of the basin, in the Kuban catchment area, do mountainous forests dominate the natural ecosystems (Dupliak 1994; Trofimov *et al.* 2014).

5.2 Agriculture

The catchment basin of the Don is also known as the “Black Earth” or chernozem region of Russia, as it has one of the most fertile soils, making the region one of the most populated in both Ukraine and Russia (Rashleigh *et al.* 2012). However, to make up for unfavourable climate conditions in some parts of the catchment basin, the pressure to extract water for irrigation purposes is exaggerated (Lagutov and Lagutov 2012). Projects that tried to cultivate rice or

cotton – very water-intensive crops – have failed to do so due to unsuitable agricultural practices (Rashleigh *et al.* 2012). While agricultural productivity has decreased dramatically (up to 50%) in the 1990s, the water consumption has not decreased equally, and the pressure on the water ecosystems continue to increase (Dubinina and Kozlitina 2000).

All parts of the area use around 80% of the agricultural land for crop production, and the rest for hay production or pastures. The main crops produced are corn, sunflower, barley, wheat, rye, and vegetables, however, rice is also produced in the catchment basin of the Kuban (Dronin and Kirilenko 2012).

For several crops, the Russian provinces situated in the Azov Sea basin are the top-producers, as the data in Figure 11 shows. The data has been extracted from the Russian Federation Grain and Feed Annual Report of 2017, published by the FAS/Moscow Staff of the Global Agricultural Information Network (Gray 2017), and has then been visualized by the researcher with ArcGIS 10.4.

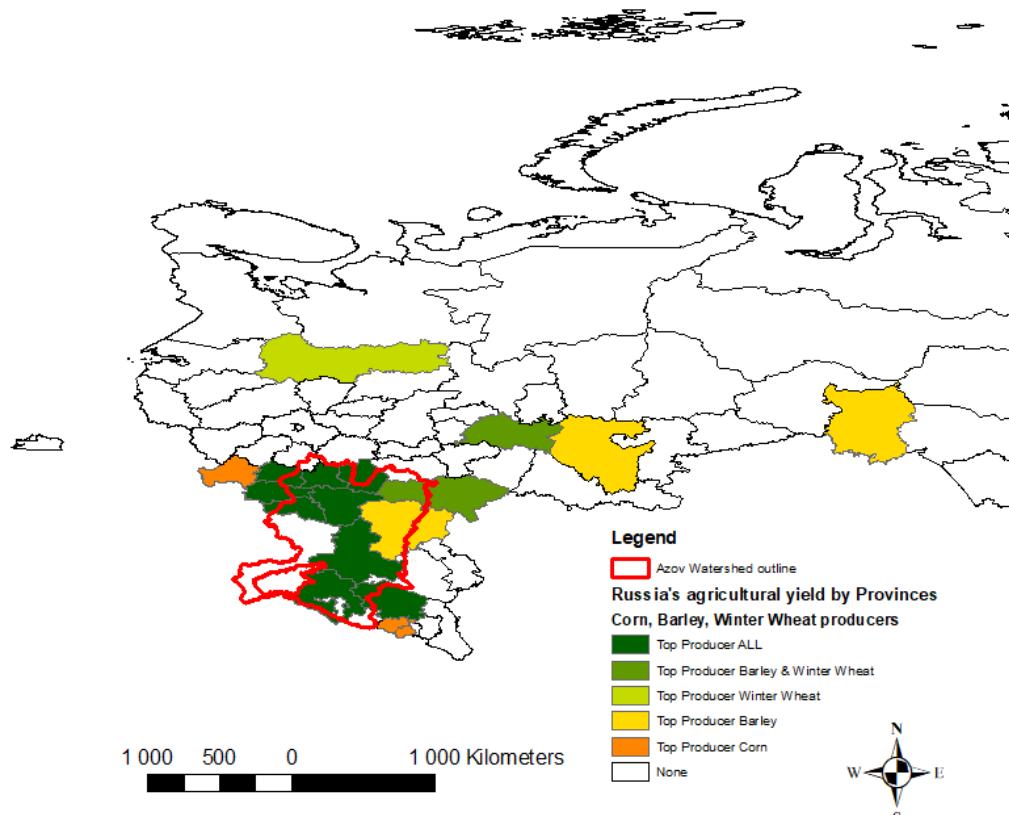


Figure 11: Top Producers of Corn, Barley and Winter Wheat in the Azov Sea basin and the rest of Russia. Not pictured areas of Russia are not considered top producers in any of the categories. Created by the author with ArcMap 10.4, Data Source: Gray 2017

A “top producer” in each category is defined as a province/oblast that produced over 1000 TMT (thousand metric tonnes) of winter wheat, over 500 TMT of barley, or over 400 TMT of corn per season. As visualized, several provinces that are within the Azov Sea basin are defined as top producers for all three crops, whereas no province outside of the basin can produce all three top priority commodities in those quantities. More specific numbers of the exact yield can be read in Grey 2017, page 16.

As grains are the most common crops in the region, it has been a tradition to have a field cropped one year while keeping it fallow the next, due to shortages in additional fertilizers and therefore the need to keep the natural fertility of the soil high. This has impacts on the overall productivity of the region and each farm (Ioffe *et al.* 2014).

The population of this region is extremely dependent on the agricultural yields and on the water that the two river basins provide, due to the high irrigation needs (Dronin and Kirilenko 2012). While this area is often called the “breadbasket” of Russia, its productivity could be much higher still with better mechanization, and even the Food and Agriculture Organization (FAO) has expressed high hopes in the agricultural intensification of this area (FAO 2010).

The agricultural market of the case study area experienced incredible changes during the timeframe of this study. Russia went through a shift from the USSR, a communist, centrally planned economy that heavily relied on state subsidies for fertilizer and machinery production, to a market economy with private farm owners and almost no subsidies for farmers. The massive drop in state subsidies led to a decrease in crop yields and animal yields at first, especially in the early 1990s, which in further consequence led to a lack of investments in agriculture and the abandonment of lands (Bobylev and Ligert 1994). More of this will be explained in chapter 7.1.

5.3 Industrial and Economic Development

The economic development in the region in the past century has altered the basin and especially the river Don, dramatically. The Tsimlyansk reservoir was built in 1951 to produce energy for the region and store water for irrigation purposes; Figure 12 shows the irrevocable change in the average discharge that happened as a result. The seasonal changes of the discharge have been smoothed drastically, which means that the natural annual spring flooding that occurred prior to these changes was widely stopped, which has vast implications for agriculture, fisheries and the connected ecosystem goods and services of the basin.

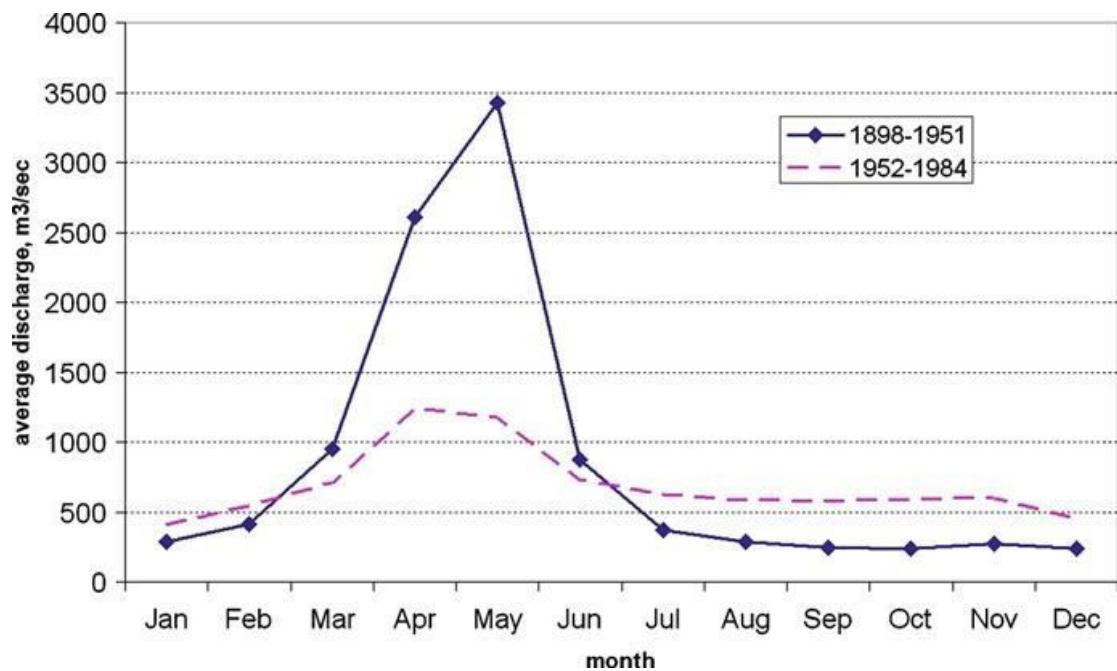


Figure 12: "Changes in the Don River hydrological regime after completion of the Tsimlyansk reservoir"
Source: Lagutov and Lagutov 2012, 40

Furthermore, while the whole basin has only scarcely available water resources (it accounts for less than 1% of Russia's water resources (Lagutov and Lagutov 2012)), the inhabitants of the region are vastly dependent on the ecosystem services it provides. Especially in the Don region, a lack of other smaller rivers leads to a dependency on freshwater from the Don for municipal purposes. The water and food security of the local population is therefore threatened by the further economic development of the region leading to surface water pollution and an irrevocably altered discharge cycle. A lack of regulatory ecosystem services that filtrate the

water naturally can therefore also increase water pollution-related diseases, and the lack of seasonal flooding due to river regulations and impoundment disturbs the natural system (Lagutov and Lagutov 2012). Nevertheless, not only the big regional centres that heavily depend on the Don are of importance; additionally, the traditional Cossack communities living dispersed all over the basin in ultimate proximity to the rivers are vastly dependent on the Don and the Kuban (Lagutov and Lagutov 2012).

Also in the light of the political situation of Russia and Ukraine, the past and future development of the Azov Sea basin seems worth studying. The Ukrainian crisis in 2014 was a new turning point for the region and the Donbas conflict will be discussed later on.

Beyond that, Russia articulated plans to significantly increase its Black Sea Fleet, and in general has mostly security related interests in the region, as opposed to environmental ones (Delanoe 2014). With a direct connection to the Black Sea, access to the Azov Sea is a high priority for both countries. Through the Azov Sea, the Don, the Volga-Don canal and the Volga River, there is a direct connection available from Moscow to the Mediterranean Sea and therefore the oceans (Sharvak *et al.* 2012). This route gains traffic every year and is of enormous importance to Russia's economic and security-related interests, which also makes train transportation to the ports of the Azov Sea more prevalent (Sharvak *et al.* 2012). The planned Eurasia canal that was proposed by Kazakhstan in 2007 would connect the Caspian to the Azov and the Black Sea, and would go through the Manych valley, a tributary of the Don. This would significantly alter the hydrological system once more on an unpredictable scale (Bekturganov *et al.* 2017). Coal is an important export good of the Azov Sea basin, as the Donetsk basin provides coal for Ukraine and Russia, and due to its proximity to the ports of the Azov Sea, it is easily transportable to Europe and different parts of Russia (Kurakov *et al.* 2010; Privalov *et al.* 2004). Productivity in the region has been massively reduced and several coal mines have

closed down from 1990 until today, however, it continues to be of importance to the energy security of the region (Kurakov *et al.* 2010; Privalov *et al.* 2004).

The rising economic importance of the region for all kinds of sectors (mining, agriculture, energy production, oil distraction) has increased the tension for both Ukraine and Russia, and Vladimir Putin articulated plans to increase the turnover on Russian ports as compared to Ukrainian and other foreign ones (Lagutov and Lagutov 2012). This could potentially go hand in hand with further regulations and therefore denaturalisation of the present ecosystem.

The Tsimlyansk reservoir, which is incredibly important for the above-mentioned transportation means, is endangered by increasing siltation processes. The water quality of the reservoir is, furthermore, continuously decreasing due to Manganese and other heavy metal pollution (Sharvak *et al.* 2012).

5.4 Conclusion of the Case Study Area Review

The combination of a unique ecosystem, the strong dependency of the population on the rivers and the sea, and the increasing economic development and important strategic location make the Azov Sea basin a remarkable case study that has been widely neglected by the international scientific community thus far. The changes that have occurred in the area in the past years have been dramatic, and just as much change is suspected to happen there in the future. Plenty of research gaps still need to be closed, and no time series analysis that monitors land use changes in the basin could be found. Future policy recommendations and management measures for this region could potentially be adapted to other similar ecosystems, which makes additional research on the Azov Sea basin even more crucial.

6 Land Use Classification

In the following section, the detailed process of land use classifications with Google Earth Engine will be discussed. Due to the novelty and unfamiliarity for most researchers of this tool to date, a detailed description of the possibilities of this technology, including screenshots of the code and outcomes, is important. However, despite showing code snippets in this section to give easily comprehensible examples, the full code can be found in Appendix A.

While the researcher had some previous experience with JavaScript programming, it is still important to mention here that this methodology was achieved by trial and error, as resources and expertise necessary for the handling of the program had not been available prior. The described version here is a description of the final version of the code. The methodology below will contain screenshots of the final code to ensure repeatability in future studies and optimal transparency of the achieved results, and to answer the first research question.

6.1 Classification Process with GEE

After the base map was chosen and visualized, the supervised classification process usually takes up several steps (Richards and Jia 2013):

- Choosing the type and amount of classes; collecting and classifying the training data; (both discussed in 6.1.2)
- Training the classifier; using the classifier to classify all pixels in the research area; producing a map that is fully classified; (discussed in 6.1.3)
- And assessing the accuracy of the process for user and producer (discussed in 6.1.4).

The results and their interpretation will finally be discussed thereafter.

6.1.1 Setting the Stage

The choice of the right data set has to take several aspects into account; the most important ones being availability, accessibility, spatial and temporal resolution, as well as data quality (Aitkenhead and Aalders 2011). For a case study of this kind that has a regional scope and

limited computational power, a medium spatial scale such as that of Landsat TM/ETM+/OLI is optimal and widely used by researchers (Lu and Weng 2007).

A time frame from 1985 (first Landsat 5 images) to 2015 (most recent available data for five-year-timesteps) was chosen. Extending the time frame with Landsat 2 MSS images to 40 years, from 1975 to 2015, was considered, however, the full collection of Landsat 2 MSS on Google Earth Engine is very limited; in Figure 13, all images recorded by Landsat 2 in August of the years 1975 to 1985 in the case study region can be seen; it is obvious, that the lack of data availability does not allow a bigger time frame.



Figure 13: Available images of Landsat 2 MSS in the case study region between 1975 and 1985

The land use images were always chosen for the specific year only for the month of August. All cloudy pixels were masked out; of the remaining data points for each pixel the median value was chosen for each band, as the reflectance over time is, of course, not normally distributed, which makes the median a more significant measurement. As a compromise between cloudless images and a narrow time frame, pictures for the August of three consecutive years were chosen (e.g. 1984, 1985 and 1986, thereafter labelled as “1985”; 1989, 1990, 1991, labelled as “1990”,

etc.), and are labelled in future sections only by the middle year the data came from. Only for the map standing representative for the year 2000 and being labelled as such, the data of 1998 has been included as well, as images from this time frame were extremely cloudy in the North-East of the research area, and the data loss would have been too large to ensure comparability between the years. This needs to be considered when interpreting the results.

For all six maps between 1985 and 2010, the USGS Landsat 5 Surface Reflectance Tier 1 collection was chosen in Google Earth Engine. The Image Collection ID at which this collection can be found in Google Earth Engine is LANDSAT/LT05/C01/T1_SR. The spatial resolution for all bands of the Landsat 5 ETM sensor is 30 metres per pixel, and the temporal resolution is 16 days. For 2015, due to a lack of data available from Landsat 5, the USGS Landsat 8 Surface Reflectance Tier 1 was used, available at Earth Engine with the Image Collection ID LANDSAT/LC08/C01/T1_SR, and also with a spatial resolution of 30 meters for most bands, and a temporal resolution of 16 days. For this year, again, images of August of 2014, 2015 and 2016 were used to fill up all cloud-covered areas.

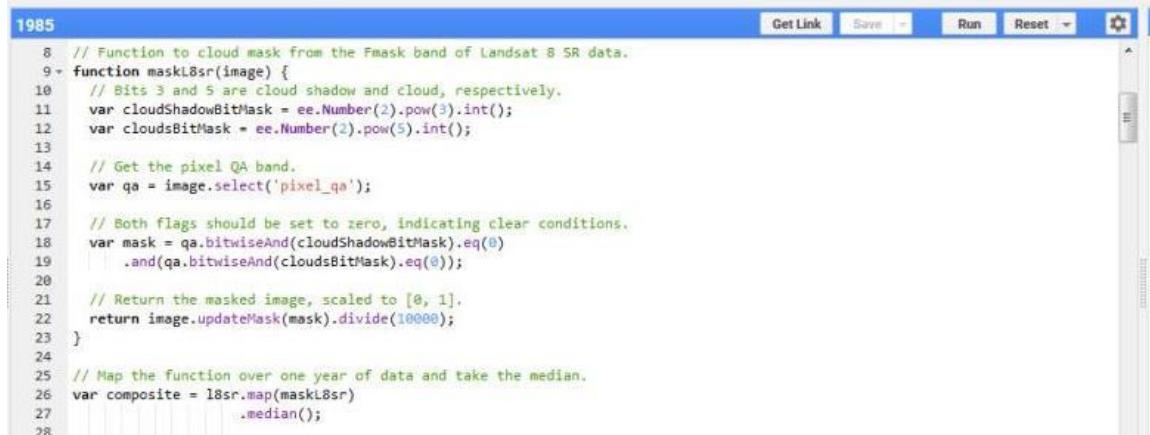
After choosing the image collection, it was filtered by the bounds of the basin, calendar range of the years, and the calendar range of August each time, and was sorted by least cloudy pictures, as can be seen in Figure 14.

```
1 // Load Landsat 5 surface reflectance data
2 // Look for adequate image filtered by date and cloud cover and region
3 var l5sr = ee.ImageCollection('LANDSAT/LT05/C01/T1_SR')
4   .filterBounds(shape_basin)
5   .filter(ee.Filter.calendarRange(1984,1986,'year'))
6   .filter(ee.Filter.calendarRange(8,8,'month')).sort('CLOUD_COVER',false);
7
```

Figure 14: Choosing the image collection, filtering bounds and calendar range

As a next step, all clouds had to be masked out. This was achieved by filtering the “pixel_qa” band the Landsat 5 Surface reflectance Collection 1 Tier 1 in Google Earth Engine already has embedded. This band gives a numeric value to each pixel; 1 being clear, 2 being water, 3 cloud shadow, 4 snow, 5 clouds, and 6-7 had an additional cloud confidence value split into low, medium, and high confidence. This band was then filtered to determine all pixels with the

values of 3 (cloud shadow) or 5 (cloud); these pixels were then masked out, and the median of each band of the remaining images was taken as the value for the whole map. The code therefore is visible in Figure 15.



```

1985
8 // Function to cloud mask from the Fmask band of Landsat 8 SR data.
9+ function maskL8sr(image) {
10   // Bits 3 and 5 are cloud shadow and cloud, respectively.
11   var cloudShadowBitMask = ee.Number(2).pow(3).int();
12   var cloudsBitMask = ee.Number(2).pow(5).int();
13
14   // Get the pixel QA band.
15   var qa = image.select('pixel_qa');
16
17   // Both flags should be set to zero, indicating clear conditions.
18   var mask = qa.bitwiseAnd(cloudShadowBitMask).eq(0)
19     .and(qa.bitwiseAnd(cloudsBitMask).eq(0));
20
21   // Return the masked image, scaled to [0, 1].
22   return image.updateMask(mask).divide(10000);
23 }
24
25 // Map the function over one year of data and take the median.
26 var composite = l8sr.map(maskL8sr)
27   .median();
28

```

Figure 15: Masking clouds out of the picture and taking the median of each band for each pixel of the remaining images

Furthermore, the NDVI for each pixel was calculated. Due to the common usage of this parameter in vegetation analysis, Google Earth Engine provides its own function named *.normalizedDifference()*, and the formula itself does not have to be calculated. For all images using Landsat 5, band 4 and band 3 were used for the NDVI calculation, whereas for Landsat 8, band 5 and band 4 had to be used, respectively. These bands represent the red and the near-infrared spectrum, as this normalized difference is normally high for healthy, green vegetation, lower for not healthy vegetation, and close to zero for water or built-up land. These differences can also be seen later in the mean spectra for the pixels in the different classes, which adds to the validation and discussion of the results. The NDVI was subsequently added as an additional band to the image, and the whole mosaic was clipped to the borders of the basin, as seen in Figure 16.

```

29 // NDVI band added to collection
30 var addNDVI = function(composite) {
31   var ndvi = composite.normalizedDifference(['B4', 'B3']).rename('NDVI');
32   return composite.addBands(ndvi);
33 };
34
35 // Test the addNDVI function on a single image.
36 var withndvi = addNDVI(composite);
37 var onlyndvi = addNDVI(composite).select('NDVI');
38
39 // Clip it to the exact region
40 var clipped = withndvi.clip(shape_basin);
41 onlyndvi = onlyndvi.clip(shape_basin);
42
43
44 // Display the results.
45 Map.centerObject(shape_basin,5);
46 Map.addLayer(clipped, {bands: ['B5', 'B4', 'B3'], min: 0, max: 0.3}, 'Landsat 5 image');
47 Map.addLayer(onlyndvi, {min:0, max:0.8, palette:['ffffff','ff0000']}, 'NDVI');
48

```

Figure 16: Calculating and adding the NDVI

The map was then centred to the area with a zoom level of 5 to ensure the whole basin is visible, but not unnecessarily much land around it. For the creation of the training data, mainly a visualization of bands 5 (as red), 4 (as green) and 3 (as blue) for Landsat 5 images, normalized to a minimum value of 0 and a maximum value of 0.3 was used. For 2015, using Landsat 8, the bands 6 (as red), 5 (as green) and 4 (as blue) were visualized, normalised to a minimum value of 0 and a maximum value of 0.5. These bands, namely Short-wave Infrared 1, Near Infrared, and Red were chosen since they offer the best combination for vegetation analysis (Butler 2013; Barsi *et al.* 2014). With this band combination, healthy vegetation can easily be seen as green, whereas water is almost black and bare land is brown/magenta coloured. The resulting map for 2015 in Google Earth Engine can be seen here in Figure 17 as an example visualization for the following methodology steps.

Additionally, a layer showing only the NDVI was created to compare the NDVI to the greenness of the picture, in order to be more certain about the vegetation status. It was normalised to a minimum value of 0 being shown as white (e.g. water is easily recognisable), and a value of 0.8 as red (very green forests). The example visualization for 2015 can be seen in Figure 18.



Figure 17: The least cloudy pictures for 2015, visualized as band combination 6 - 5 - 4 of Landsat 8

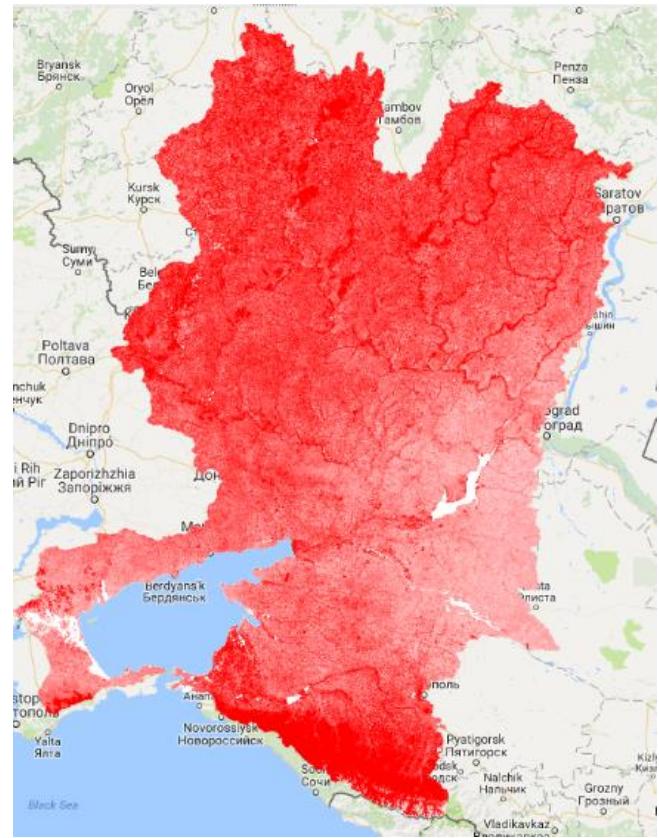


Figure 18: Visualization of the NDVI of the least cloudy pictures in 2015, calculated from Landsat 8 images

6.1.2 Collecting Training Data

Level 1 of the classic land use classification system established by the USGS for Landsat imagery (Anderson *et al.* 1983) has been adopted for this study as it was deemed detailed enough for the more exploratory purpose of the study and the theoretical framework explained in previous chapters. The categories evaluated and classified are:

- Urban or built-up land, which includes residential, commercial, and industrial land, as well as built-up land with transportation infrastructure. Most importantly, in this study, it contains open mining fields as well. The development of trade infrastructure and industry will be discussed in the later chapters, as this could potentially have immense implications on the urban percentage in the region. Figure 19 shows an example of an area with large mining fields and areas built up with train tracks and industrial buildings.



Figure 19: Large industrial and mining areas that are included in the category "urban"

- Agricultural land, which includes cropland, pastures, vineyards, horticulture, and others; here, this has been split up in land that was clearly vegetated in August, and land that was clearly bare, as this could have effects on the hydrological cycle and the overall yield and productivity. Bare land in this sense could mean that it has already been harvested at that point of time, or that it was kept fallowed, as this is common cultural practice in some provinces of the region.
- Forest land, which includes deciduous, evergreen, and mixed forest lands, as well as the mountainous forest of the Kuban river basin.
- Water, which includes streams, canals, lakes, reservoirs, and bays.

Not used, despite present in the report by Anderson *et al.* (1983), are rangeland, wetlands, barren land, tundra, and perennial snow or ice. These categories are not visibly present in the case study area and therefore uncalled-for. Categories going beyond this first level classification would need significantly more data which is not available and not possible to create due to the severe time constraints of this study.

After choosing those categories, polygons were drawn to create the training data. The land use classes and borders of the polygons were identified by zooming far into the 5-4-3 Landsat 5 map (aka 6-5-4 for the Landsat 8 map) and NDVI maps, and using the expert knowledge of the map producer about the region. Earth Engine's polygon drawing tool was used to set the polygons and classify them in one of the five different categories. Most polygons (especially

for water, urban, and natural vegetation) were initially drawn for 2015, and then revised and checked for the upcoming years; the classes Vegetated Agriculture and Bare Agriculture, however, were completely deleted and created from scratch for each map, as the location of vegetated versus bare agricultural land can differ extremely between the years due to crop rotation techniques applied in the area, and the different harvesting time frames for the different crops.

While assembling the training data, some issues emerged that will briefly be discussed here. A very common problem when assessing the water extent of an area is that the spectral reflectance of eutrophication is sometimes more similar to that of natural vegetation than that of water. This can lead to cases such as the Landsat image shown in Figure 20 that was misclassified as visible in Figure 21. Furthermore, too much of the surrounding areas have been classified as “urban”; these screenshots visualize the earlier process of back-and-forth of classifying and establishing training data.



Figure 20: Eutrophication visible in the Landsat 5 map



Figure 21: Eutrophic water falsely classified as forest (dark green patches)

Problematic were also clouded areas that made the underlying image invisible. While these are partly visible in Figure 20 and Figure 21 at the bank of the river, the following Figure 22 shows an even bigger area of clouds that stretch over the whole tile and over a large urban area that will consequently leave a data gap.



Figure 22: Data gaps due to cloudiness over an urban and agricultural area

The accumulated area of all training data polygons was calculated to show the comparability and draw conclusions about the accuracy measurements; the example of a code for the calculation of all “water” polygons can be seen in Figure 23. This was repeated in the exact same way for the remaining four classes.

```

1985
49 // calculating the amount of pixels I have in each polygon, and the area thereof
50 var mask = clipped.select(0).mask().rename('mask');
51 var area = ee.Image.pixelArea().multiply(mask).rename('area');
52
53 var sumDictionarypolygon1 = mask.addBands(area).reduceRegion({
54   reducer: ee.Reducer.sum(),
55   geometry: water.geometry(),
56   scale: 30,
57   maxPixels: 1e9
58 });
59 print('sum for water polygons, scale=30', sumDictionarypolygon1);

```

Figure 23: Calculation of amount of pixels and total area of polygons of each class, example code for water polygons of 1985.

The results of this calculation can be seen in Table 1.

Table 1: Area of the Training data polygons for each class in km²

	1985	1990	1995	2000	2005	2010	2015
Water	946	942	940	942	942	948	942
Crops	235	125	196	261	182	118	116
Bare	327	357	213	451	212	159	93
Urban	58	50	41	54	59	28	51
Forest	481	131	185	129	503	131	163
	2 046	1 605	1 575	1 836	1 898	1 384	1 365

These areas sum up to around 3.4 million pixels, depending on the year. There are plenty of reasons for the differences in the area of the polygons, the first and most important one being the accuracy of the training results as seen in the later accuracy assessment. As the accuracy for some maps, especially the map of 1985 and 2005, was initially very low, more training data was created than for the other maps (e.g. 2015), where the accuracy values were acceptable from the beginning and a minimum of training data had to be created. This could be explained by lower quality images, higher variability of the spectra of these years, or clouds covering up small crucial areas, such as regional centres or natural forests, which effectively makes certain crucially important polygons useless and requires additional polygons to be drawn. This process consisted of several steps back and forth between drawing new polygons and running the script again, until the accuracy was satisfactory.

To avoid just considering the redistribution accuracy of how well the classifier classified the pixels it was initially trained with, only 80% of the drawn polygons were used as a training partition, and the remaining 20% were saved for the accuracy assessment, as can be seen later.

6.1.3 Classifying the Map

After the training data was successfully established, the five different training feature collections (one for each class) had to be merged into one new feature collection and

transformed into a training data set. At this step, the scale had to be reduced from 30 meters per pixel to 300 meters per pixel, as Google Earth Engine would otherwise always time out in the later steps. The transformation from five different feature collections to one coherent training dataset can be seen in Figure 24.

```
94 //merging feature collections
95
96 var newfc = water.merge(croplandVEG).merge(croplandBARE).merge(urban).merge(naturalVEG);
97
98 // Get the values for all pixels in each polygon in the training.
99 var training = clipped.sampleRegions({
100   // Get the sample from the polygons FeatureCollection.
101   collection: newfc,
102   // Keep this list of properties from the polygons.
103   properties: ['class'],
104   // Set the scale to get Landsat pixels in the polygons.
105   scale: 300
106 });
...
```

Figure 24: Merging all five feature collections and transforming them into potential training data

Then, as already mentioned in the previous step, this dataset was split again randomly into two partitions – the first partition, using 80% of the initially classified pixels, will be used as a training partition, while the remaining 20% will be saved to build the testing partition for later on. For this purpose, a random column was created for each dataset, containing random numbers between 0 and 1. By splitting the data set according to its random column at a value of 0.8, a random choice of roughly 80% of pixels will be used for training.

Subsequently, the classifier was trained with this data, and the classification was carried out on the whole basin. The “random forest” classifier was chosen as it is considered the state-of-the-art classifier at the moment by several remote sensing specialists – a more detailed review of the different classifiers and their functionalities can be seen in the literature review. To ensure optimal comparability and not favour certain classes in certain time frames, the bands 1, 2, 3, 4, 5 and 7 were used for the training and classification of Landsat 5 images, and the bands 2, 3, 4, 5, 6 and 7 were used for those of Landsat 8 images. The code for all of the training and classification can be seen in Figure 25.

```

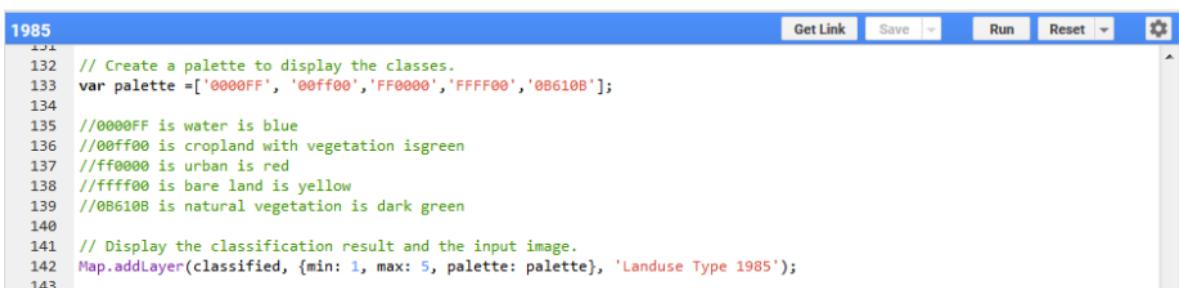
107
108 // adding a column of random numbers to the training dataset.
109 var withRandom = training.randomColumn('random');
110
111 // Reserving some of the data for testing, to avoid overfitting the model.
112 var split = 0.8; // Roughly 80% training, 20% testing.
113 var trainingPartition = withRandom.filter(ee.Filter.lt('random', split));
114 var testingPartition = withRandom.filter(ee.Filter.gte('random', split));
115
116 // Use these bands for prediction.
117 var bands = ['B1', 'B2', 'B3', 'B4', 'B5', 'B7', 'NDVI'];
118
119 // Create a random forest classifier with custom parameters.
120 var classifier = ee.Classifier.randomForest().train({
121   features: trainingPartition,
122   classProperty: 'class',
123   inputProperties: bands
124 });
125
126 // Train the classifier.
127 var trained = classifier.train(trainingPartition, 'class', bands);
128
129 // Classify the image.
130 var classified = clipped.classify(trained);
131

```

Figure 25: Randomly splitting polygons into training and testing datasets, and using the training partition to train and classify the Landsat 5 image

Following the classification, the results can be visualized in the *Map* of Google Earth Engine.

The code for the visualization is pictured in Figure 26.



The screenshot shows the Google Earth Engine code editor interface. The code is for visualizing land-use maps from 1985. It includes a color palette definition and a call to addLayer on the Map object. The code editor has a toolbar with 'Get Link', 'Save', 'Run', 'Reset', and a settings icon.

```

1985
131
132 // Create a palette to display the classes.
133 var palette =['0000FF', '00ff00','FF0000','FFFF00','0B610B'];
134
135 //0000FF is water is blue
136 //00ff00 is cropland with vegetation isgreen
137 //ff0000 is urban is red
138 //ffff00 is bare land is yellow
139 //0B610B is natural vegetation is dark green
140
141 // Display the classification result and the input image.
142 Map.addLayer(classified, {min: 1, max: 5, palette: palette}, 'Landuse Type 1985');
143

```

Figure 26: Visualizing the classified land-use map in Google Earth Engine

The seven classified maps are pictured here in Figure 27 for a quick visual comparison; a higher quality and larger image will be available in Appendix B, and the proportional development of each of the classes will be discussed in a later chapter.

Legend

- █ Water
- █ Agriculture with Crops
- █ Urban
- █ Agriculture without Crops
- █ Forest/Natural Vegetation

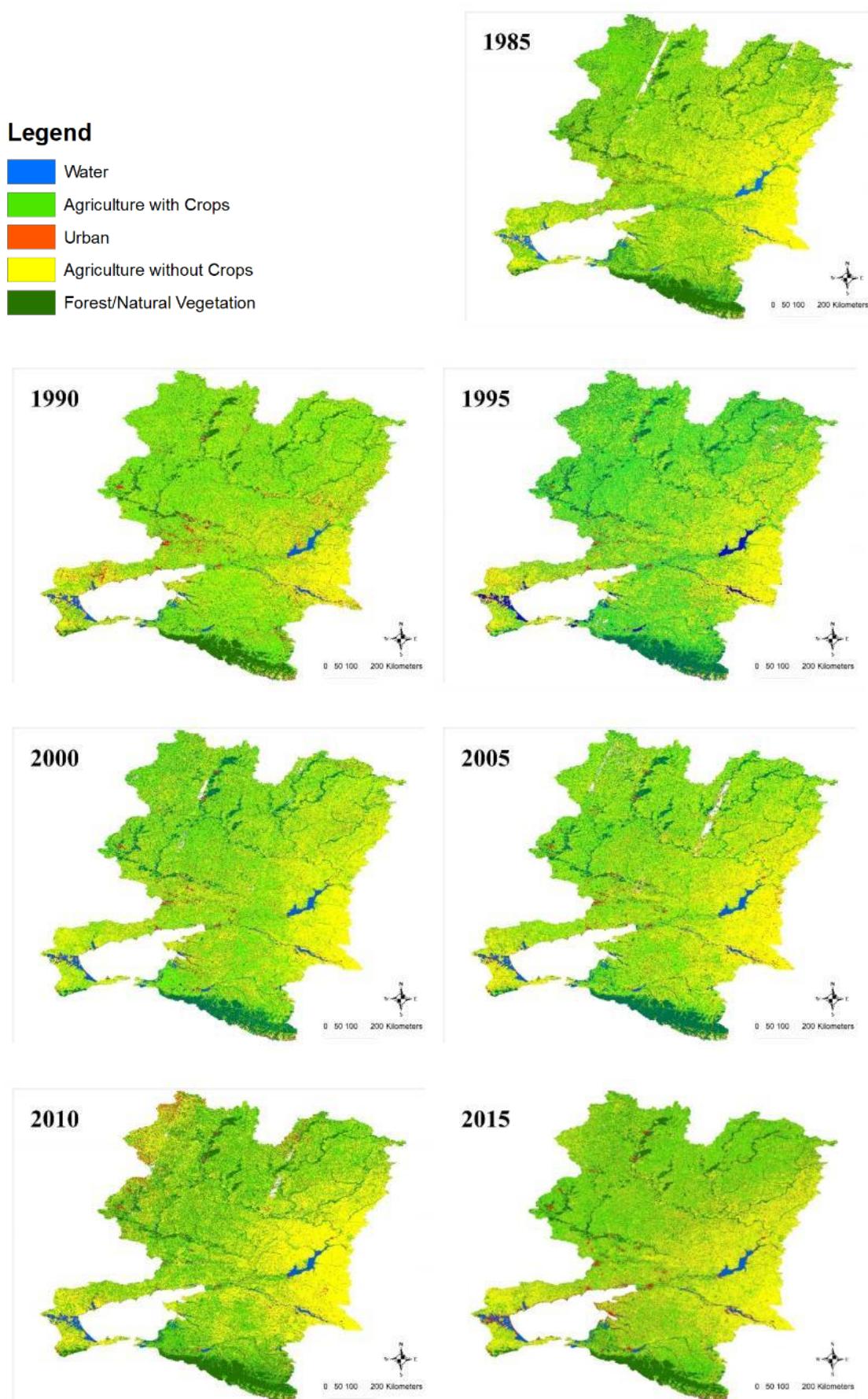


Figure 27: Land use Maps for the years 1985 to 2015 of the Azov Sea Basin

6.1.4 Accuracy Assessment

Immediately after the Landsat images have been classified, it is important to assess the accuracy – the code for which can be seen in Figure 28, where both the confusion matrix using the testing partition, as well as the overall accuracy and the kappa value of the confusion matrices have been calculated. To avoid a computation timeout, the confusion matrix had to be transformed into a feature collection and further exported as a CSV-file.

```
144 // Classify the test FeatureCollection to assess the accuracy!
145 var test = testingPartition.classify(trained);
146
147 // Print the confusion matrix.
148 var confusionMatrix = test.errorMatrix('class', 'classification');
149 print('Confusion Matrix', confusionMatrix);
150 print('Validation overall accuracy: ', confusionMatrix.accuracy());
151 print('Kappa value: ', confusionMatrix.kappa());
152
153 //Export table so it doesnt time out
154
155 var confusionMatrixExport = ee.Feature(null, {matrix: test.errorMatrix('class','classification').array()});
156
157 // Export the FeatureCollection.
158 Export.table.toDrive({
159   collection: ee.FeatureCollection(confusionMatrixExport),
160   description: 'exportAccuracy1985',
161   fileFormat: 'CSV'
162 });
163
164
```

Figure 28: Calculation of the different accuracy measures

For this study, threshold values of minimum 90% overall accuracy and kappa values, and minimum 70% user's and producer's accuracy for each of the five classes were chosen – if these thresholds were not reached initially, more strategically chosen training polygons were added to the existing base training data to increase accuracy values. These values are in line with what other studies deem relevant results (Goldman *et al.* 2015; Richards and Jia. 2013).

The overall accuracy is calculated by dividing the number of correctly classified pixels by the sum of all testing pixels.

The kappa value is a bit more complicated than the overall accuracy, and is therefore, believed to reduce bias, as it rules out any agreement that happened by chance (Richards and Jia 2013).

In words, simplified, it is defined by Richards and Jia (2013, 402) as:

$$K = \frac{(Probability\ of\ correct\ classification) - (Probability\ of\ chance\ agreement)}{1 - Probability\ of\ chance\ agreement}$$

As a formula utilizing the elements of a confusion matrix, it can be also expressed as

$$K = \frac{(N * d) - q}{N^2 - q}$$

, whereas N is the sum of all elements in the confusion matrix, d is all correctly classified pixels (the diagonal of a confusion matrix), and q is the product of the pixels classified for one class and the pixels that were supposed to be classified as a said class, summed up over all classes. A kappa value of above 0.81 is seen as “Almost perfect” (Richards and Jia 2013, 404) and therefore, an extremely good result was achieved by the classification of these maps.

The following Table 2 of kappa values and overall accuracy values gives an idea of the differences between the seven maps; subsequently, the confusion matrices are given to provide a more coherent picture of the quality of this land use classification.

Table 2: Overall validation accuracy and kappa values for all seven land use maps

Year	Validation accuracy	overall	Kappa value
1985	0.95		0.93
1990	0.95		0.91
1995	0.95		0.92
2000	0.96		0.93
2005	0.95		0.92
2010	0.97		0.94
2015	0.95		0.91

Table 3 to Table 9 show the full confusion matrices for each of the seven maps. All values are generally satisfactory. The highest accuracy is always achieved for water with 100% for each year except 1985, where the value is still 99%.

Second in place are usually forest/natural vegetation and bare land, with values over 90% for almost all maps. The only exceptions are here the bare land values of 2005 and 2015, which still over 85% and over 75%, respectively.

The categories that are posing the most problems, however, are urban land and cropland. The

lowest value present in all matrices is a 70% user's accuracy for urban areas of 1995; this is still, as mentioned, satisfactory for this type of research. These categories are often very close together in their average spectra as can be seen in the next chapter, making them easily confusable and therefore prone to error in any land use study.

Table 3: Confusion matrix for 1985

1985	Water	Crops	Urban	Bare	Forest	SUM	User's Accuracy
Water	3169	11	2	1	4	3187	99
Crops	8	693	20	60	44	825	84
Urban	5	27	149	19	3	203	73
Bare	2	63	33	1035	4	1137	91
Forest	1	44	5	1	1560	1611	97
SUM	3185	838	209	1116	1615		
Producer's Accuracy	99	83	71	93	97		

Table 4: Confusion matrix for 1990

1990	Water	Crops	Urban	Bare	Forest	SUM	User's Accuracy
Water	3090	0	4	5	2	3101	100
Crops	0	354	22	68	17	461	77
Urban	2	16	142	10	5	175	81
Bare	5	85	12	1126	3	1231	91
Forest	1	22	0	2	439	464	95
SUM	3098	477	180	1211	466		
Producer's Accuracy	100	74	79	93	94		

Table 5: Confusion matrix for 1995

1995	Water	Crops	Urban	Bare	Forest	SUM	User's Accuracy
Water	3041	3	2	2	0	3048	100
Crops	3	552	29	53	22	659	84
Urban	1	15	91	9	1	117	78
Bare	1	57	6	644	2	710	91
Forest	2	28	2	3	581	616	94
SUM	3048	655	130	711	606		
Producer's Accuracy	100	84	70	91	96		

Table 6: Confusion matrix for 2000

2000	Water	Crops	Urban	Bare	Forest	SUM	User's Accuracy
Water	3066	6	3	4	0	3079	100
Crops	2	774	19	80	11	886	87
Urban	2	27	150	11	0	190	79
Bare	0	76	6	1375	0	1457	94
Forest	0	17	0	1	431	449	96
SUM	3070	900	178	1471	442		
Producer's Accuracy	100	86	84	93	98		

Table 7: Confusion matrix for 2005

2005	Water	Crops	Urban	Bare	Forest	SUM	User's Accuracy
Water	3129	0	3	1	0	3133	100
Crops	0	474	34	88	34	630	75
Urban	4	27	152	18	6	207	73
Bare	1	69	10	618	2	700	88
Forest	1	40	6	0	1609	1656	97
SUM	3135	610	205	725	1651		
Producer's Accuracy	100	78	74	85	97		

Table 8: Confusion matrix for 2010

2010	Water	Crops	Urban	Bare	Forest	SUM	User's Accuracy
Water	3123	1	2	0	0	3126	100
Crops	2	425	17	17	18	479	89
Urban	2	24	167	15	2	210	80
Bare	0	24	20	516	1	561	92
Forest	1	32	0	0	841	874	96
SUM	3128	506	206	548	862		
Producer's Accuracy	100	84	81	94	98		

Table 9: Confusion matrix for 2015

2015	Water	Crops	Urban	Bare	Forest	SUM	User's Accuracy
Water	3133	1	3	0	0	3137	100
Crops	0	312	6	60	13	391	80
Urban	2	12	158	10	0	182	87
Bare	0	74	7	249	0	330	75
Forest	2	16	0	1	510	529	96
SUM	3137	415	174	320	523		
Producer's Accuracy	100	75	91	78	98		

6.1.5 Mean Spectra

With Google Earth Engine, it is fairly easy to visualize the mean reflectance of each of the classes in each of the bands. First, the computation and code thereof are explained in Figure 29.

```
189 // class band added to image
190 var addclass = function(clipped) {
191   var classes = classified.select('classification').rename('classband');
192   return clipped.addBands(classes);
193 };
194
195
196 // Test the addNDVI function on a single image.
197 var withclass = addclass(clipped);
198
199 //clip it to the exact region
200 var resultWithpixqa = withclass.clip(shape_basin);
201
202 var result = resultWithpixqa.select('B1','B2','B3','B4','B5','B7','classband');
203
204 print(result);
205
206 // Define a broad list of land cover categories.
207 var classNames = ee.List(['0','Water', 'Agriculture', 'Urban', 'Bare', 'Forest']);
208
209
210 // Define chart customization options.
211 var options = {
212   lineWidth: 1,
213   pointSize: 2,
214   hAxis: {title: 'Band'},
215   vAxis: {title: 'Reflectance'},
216   title: '1985 Mean Spectra of the Azov Sea Basin'
217 };
218
219 // Make the chart, set the options.
220 var chart = ui.Chart.image.byClass({
221   image: result,
222   classBand:'classband',
223   region:shape_basin,
224   reducer: ee.Reducer.mean(),
225   scale:2000,
226   xLabels: ['B1','B2','B3','B4','B5','B6','B7'],
227   classLabels: classNames}
228 .setOptions(options).setChartType('Linechart');
229
230 print(chart);
--
```

Figure 29: Calculations of mean spectra for each class within the Azov Sea basin

The mean spectra give an idea of why the accuracy between certain classes might be low; the following Figure 30 shows an example - the other 6 mean spectra graphs can be seen in Appendix C. Here it is visible how often, vegetated agricultural land and built up land have a similar reflectance over all bands, which explains the relatively high number of pixels confused between those two classes as shown in Chapter 6.1.4. The utter utility of the NDVI can, therefore, be easily explained – the most significant difference between the urban land and cropland spectra is that crops show a slightly lower value in band 3 (red), and a slightly higher value in band 4 (near infrared). Furthermore, it is visible that pixels in the class “natural

“vegetation/forest” have the highest reflectance in band 4, and the lowest in band 3.

This visualization can also be useful when deciding which bands to use for the classification process; in this case, since all classes have almost the same reflectance value in band 6 (temperature), this band can safely be ignored in the classification process, saving computational power and time. However, it is also evident why all six other available bands are important for classification, as the differences between the classes are variable and the highest accuracy and precision was always achieved when these six bands were utilized.

Comparing the seven different mean spectra of all classes can be interesting again in relation to the accuracy differences between the seven maps; for example, a higher average NDVI of the water class could indicate eutrophication.

When comparing the graphs in the appendix, it is important to note that the x-Axis is showing the bands as according to the Landsat 5 table shown above in Figure 4; in 2015, where Landsat 8 was used, the x-Axis was modified to show the Landsat 5 band titles to ensure easier visual comparability. This means that the actual B1 of Landsat 8 was ignored (for the classification as well as for this graph), B2 was labeled as B1, and so on and so forth.

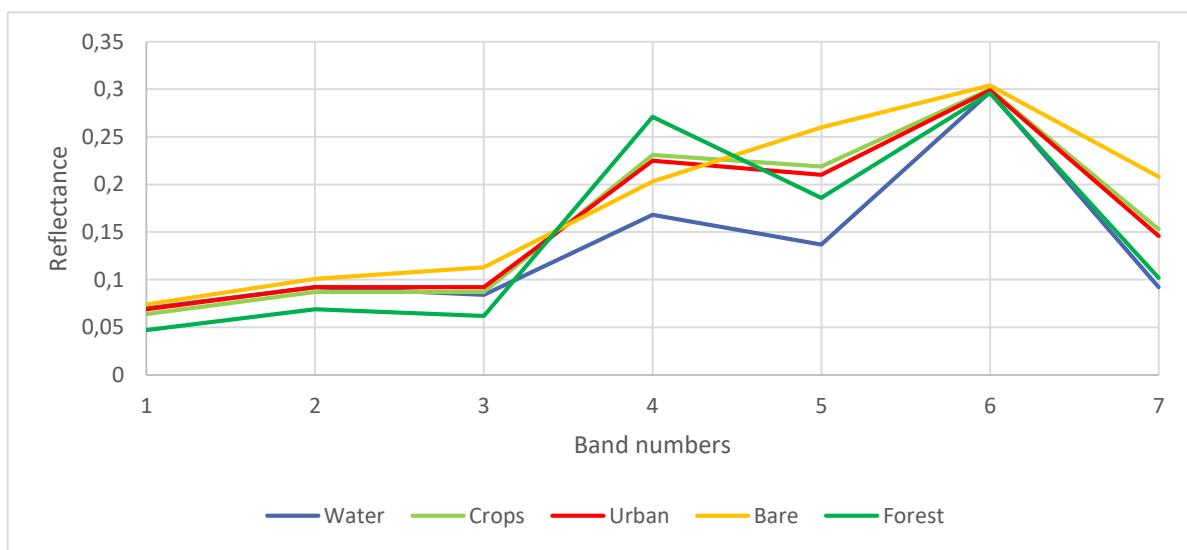


Figure 30: Mean spectra of each land use class in the different Landsat 5 bands of the 1985 land use map

6.2 Land Use Areas of the Subbasins

To make the maps provided in Chapter 6.1.3 more meaningful and analyse the land change differences in the regions, the overall area of each of the classes has been extracted five times: First, the areas of the classes for the whole Azov sea basin were calculated, as this is the basic case study area.

Second, calculation was carried out for the river Kuban, as this region is situated in the mountainous region in the South and therefore contains the majority of the forests in the region – the extent of this area and its most recent Land use map is pictured in Figure 31.

Third, the whole Don River basin has been extracted, as this is the biggest river of the basin and takes up the majority of the area – the extent and land use are pictured in Figure 32.

Fourth, only the Upper Don basin (Figure 33), containing everything upstream of the Tsimlyansk reservoir, has been calculated, as this reservoir is of the utmost importance for the region.

And lastly, the rest of the area was calculated, to see what is happening in the areas not directly surrounding the two main rivers. This area is home to more than 20 smaller tributaries (Lagutov and Lagutov 2012).

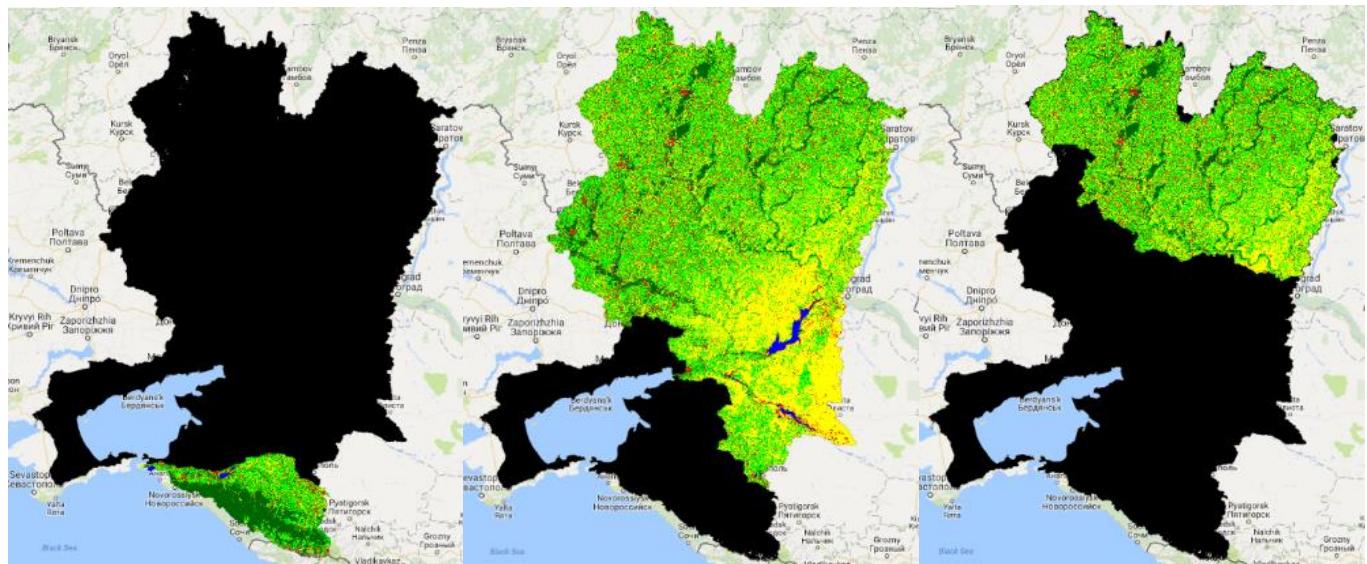


Figure 31: Extent and Land use in 2015 of the Kuban River basin

Figure 32: Extent and Land use in 2015 of the Don River basin

Figure 33: Extent and Land use in 2015 of the Upper Don River basin

The areas have been calculated with the following code (Figure 34); for the calculations of the sub-regions, the “results” and “classes” maps have been clipped accordingly – for the exact code Appendix A can be consulted:

```
-- 253 //forest
254 var forestMask = result.updateMask(classes.eq(5));
255 var forestarea = forestMask.multiply(area).select('classband').rename('forest');
256
257 //calculating the area of the waterclasses
258
259 var area_image = waterarea.addBands(croparea)
260     .addBands(urbanarea)
261     .addBands(barearea)
262     .addBands(forestarea);
263
264 var areas = area_image.reduceRegion({
265   reducer:ee.Reducer.sum(),
266   geometry: shape_basin,
267   scale: 300,
268   maxPixels:1e15
269 });
270 print(areas);
```

Figure 34: Calculations of the area per class

This code prints a matrix of all areas in the console; to make it more readable, this data has been exported into Microsoft Excel. The absolute area values can hardly be compared to the overall area of the whole basin differs to due data gaps resulting from cloudy areas. Therefore, the relative part of the area was calculated for each map.

As visible in Figure 35, the total area of agriculture in the whole basin has been slightly reduced, however, no big changes have happened from 1985 to 2015. In 1985, 74.6 % of the whole area were covered with agriculture, and while there have been ups and downs in between, the difference is negligible, as the total area was 73.9% in 2015. The same is more or less true for the analysed sub-areas as well. The whole Don region is the most agriculturally active of the region, in absolute numbers as well as percent. Agriculture in the Kuban basin has always been limited and there is no significant change detectable.

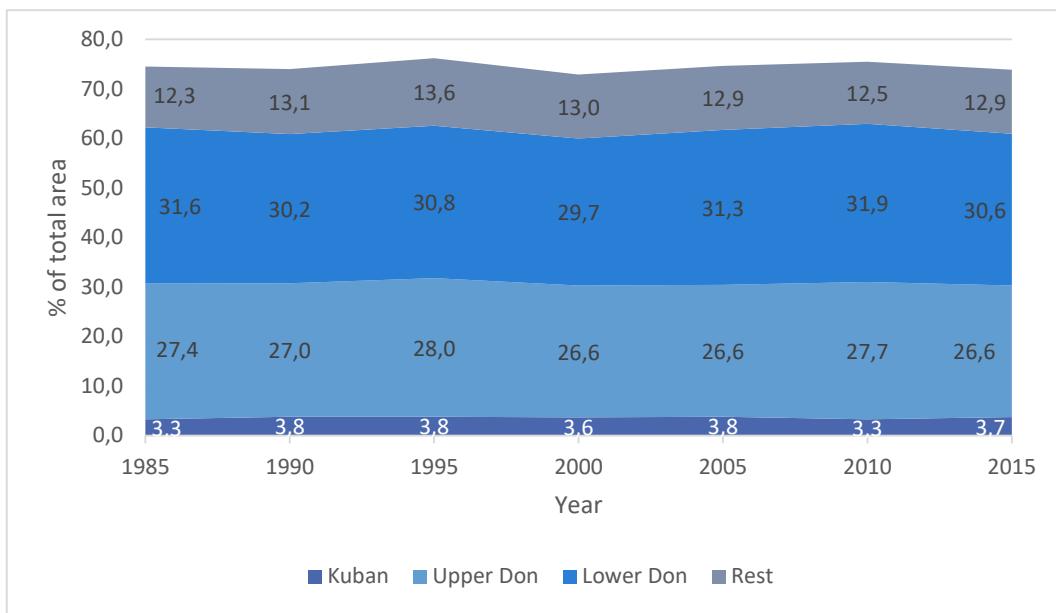


Figure 35: Proportion of all agricultural area (split into the different sub-areas) in the Azov Sea basin between 1985 and 2015

However, not only the development of the whole agricultural area, but also the partitions that are bare versus vegetated in August are interesting. This can help to make conclusions about the types of crops and connections to the water cycle. The two values are fluctuating strongly among one another, however, bare land is showing a downward trend especially since 2005, while cropland is increasing in the same time frame (Figure 36).

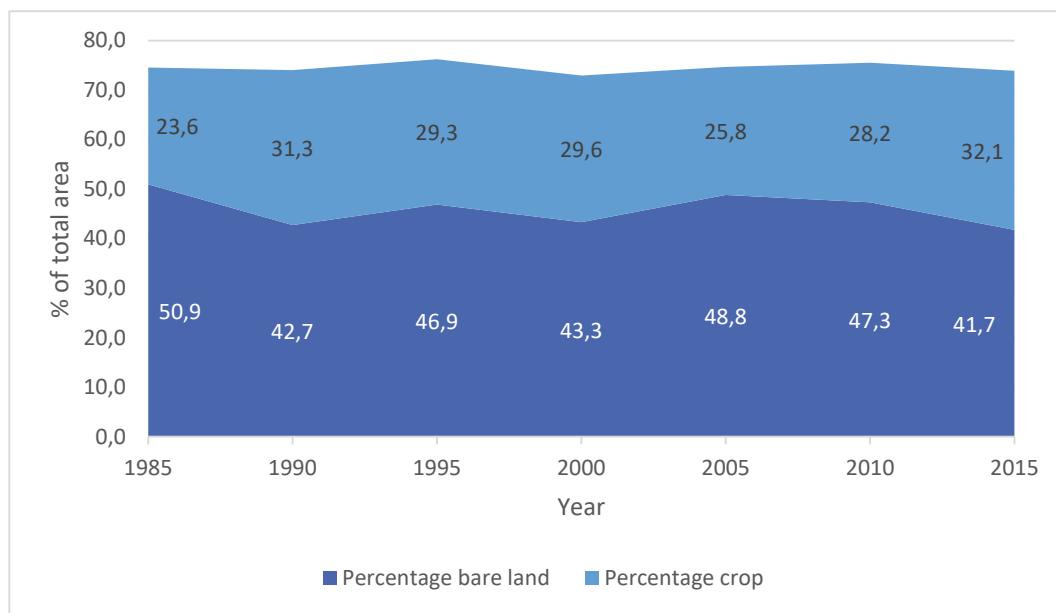


Figure 36: Cropped versus bare agricultural land, in percent of the whole Azov Sea basin

Figure 37 shows the change in the percent of forested area of the whole sea basin. In this graph, the area is visualized as a percentage of the total area. As with agriculture, while there are significant fluctuations and a decreasing trend from 2000 to 2010, no overall trend is visible. The majority of the Kuban's area – around 60% – is covered with forest, despite forested area in the whole basin making up less than 20%; this means that changes there are more significant than in any other sub-area.

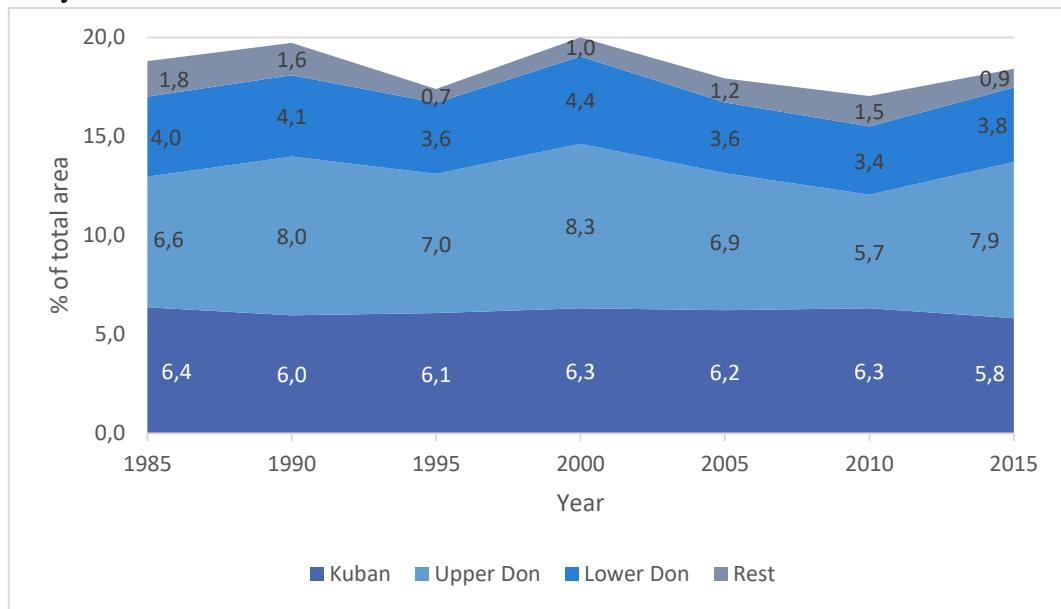


Figure 37: Proportion of forest in the Azov Sea basin from 1985 to 2015

When just looking at urban areas, however, the trend is clearly going towards more urbanization as visible in Figure 38. The whole Don area (upper and lower Don together), as well as the “Rest” area that is not directly surrounding one of the two main rivers, are clearly showing a rapid urbanization/industrialization ever since 1995. The Kuban, however, is showing a quite stable development from 1985 until 2010, with a more rapid urbanization happening in the most recent years. Still today, only just above 7% of the whole basin is urbanized – this means that while the change in urban areas seems drastic, only a – relative to the whole area – small percentage has changed in the other land use categories, despite them obviously having interactions. A trend line of the total urban area has shown to have a gradient of 0.3, and an R²-value of 0.89, showing a high compatibility.

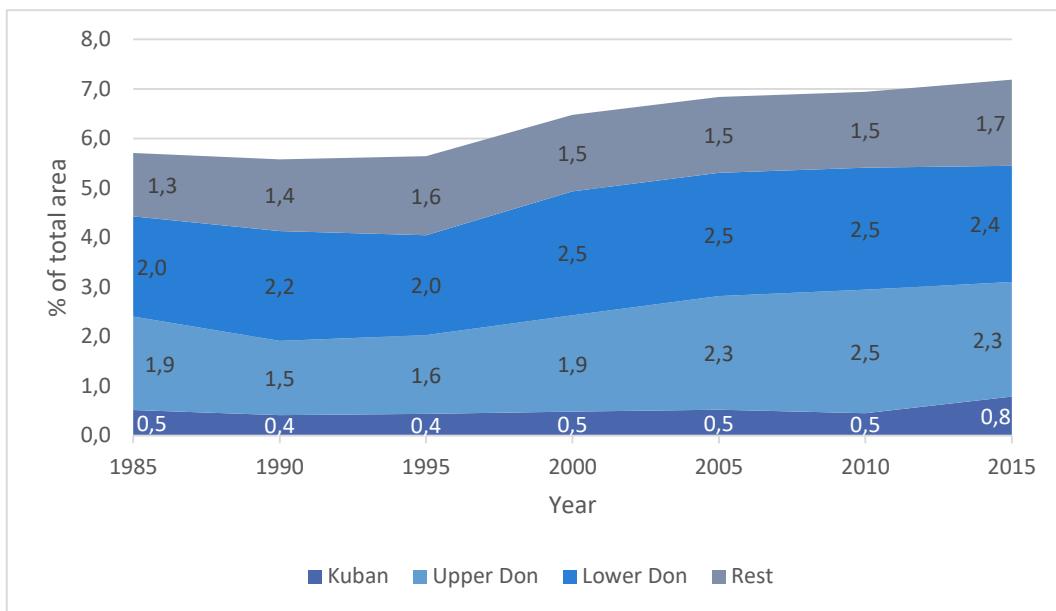


Figure 38: Proportion of urban areas in the Azov Sea Basin from 1985 to 2015

The percentage of area covered by water is small all over the sea basin, and heavily dependent on a variety of factors. Nevertheless, it has to be mentioned that this analysis showed a clear downward trend in water availability in all sub-areas of the basin, consistently over the past 30 years. A linear trend line of the total water covered area showed a gradient of 0.06, and an R^2 value of 0.82. The percentage of area covered by water in each of the subbasins can be seen in Figure 39.

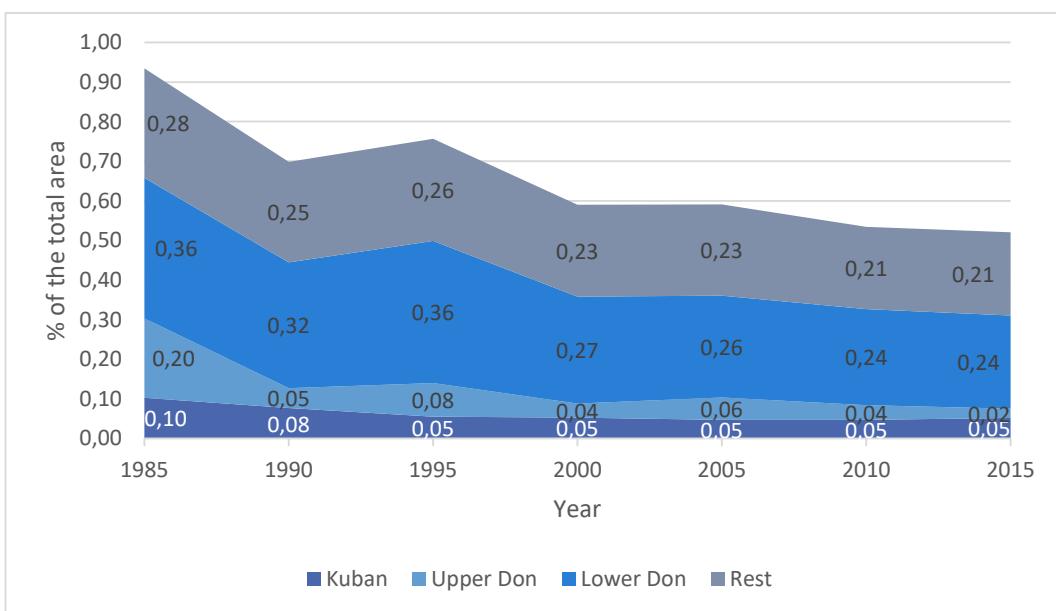


Figure 39: Proportion of water area in the Azov Sea basin from 1985 to 2015

6.3 Donbas Conflict

Since spring 2014, a violent conflict has been going on in the Donbas region as separatist rebels called out the Donetsk and Luhansk People's Republics (DNR and LNR, respectively), and have since taken control of large parts of the Donetsk and Luhansk oblasts in Ukraine (Katchanovski 2016; Kuzio 2015; The Economist 2016). This conflict has since then developed into a violent war that has received international attention and hurt thousands of civilians (Buckley *et al.* 2018). Furthermore, it brings enormous economic implications with it, not only for Ukraine, but for Russia and the European Union just as well, causing governmental leaders all over the EU to participate in the negotiations (Havlik and Astrov 2014; Katchanovski 2016). At the moment of writing, more than four years have passed since the conflict first started. By now, almost two million people have been displaced from the region, mostly internally to other parts of the Ukraine, and around 10.000 people have died in the conflict zones (Coman 2017). This abandonment of the region led to dramatic land use changes; quantifying those changes gives, on the one hand, a clearer picture of the potential influence of the disruptions on the land use, and, on the other hand, validates the local accuracy of the results discussed above, as it is quite obvious that agricultural areas and urban areas must have decreased in recent years, as so many inhabitants have left.

Figure 40 (The Economist 2016) is visualizing the situation and the border between rebel-held areas and the areas still in control of the Ukrainian government as it was in October 2016. Until now, this still remains the area of most conflicts and battles and lays in its majority within the Azov Sea basin.

The map provided in Figure 40 also served as a base map for the following analysis; it was geo-referenced with the aid of data from Humanitarian Data Exchange by the United Nations Office for the Coordination of Humanitarian Affairs Ukraine (OCHA Ukraine 2017) with ArcGIS, and then a buffer of 20km on each side of the border was established to receive a polygon including the most profoundly conflicted areas.



*Figure 40: Situation of the conflict in the Donbas region in October 2016.
Used as a base map for all further analysis. Source: The Economist, 2016*

In Figure 41, the result of this analysis can be seen, with the border in red, and the buffer zone in brown. For reference purposes, the borders of the case study region, as well as the country borders in the region are also visualized.

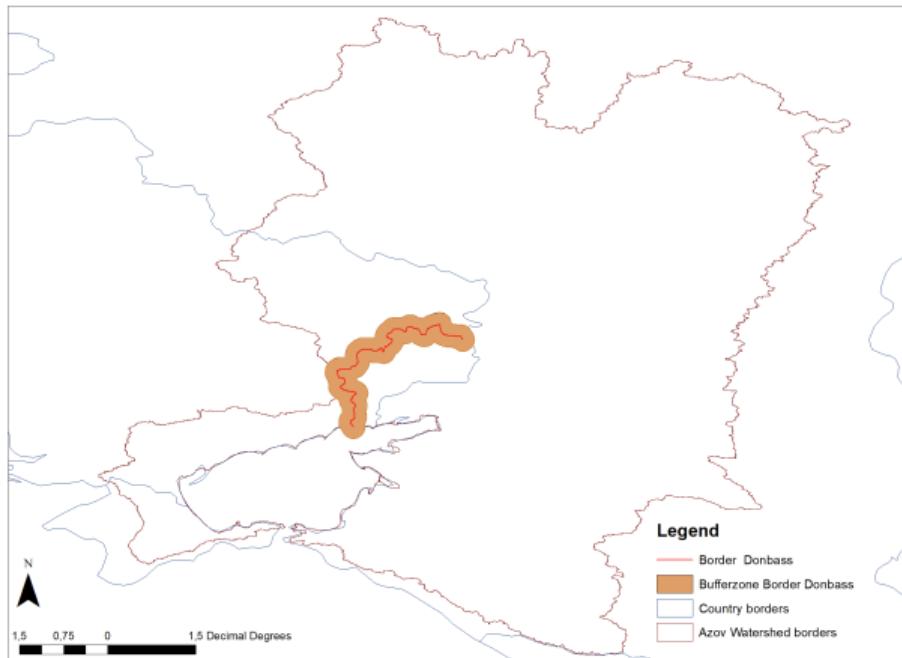


Figure 41: Result of digitalizing the conflict zone in the Donbas region

The analysis was then carried out with the same script as established above and used for the land use map of 2015, but focussing always on the summer months (May to September) of the years 2013 to 2017. As explained above, the median of all cloudless images for the time frame was chosen. The exact same training polygons were taken for all five years as established for the whole land use map of 2015. Polygons that were lying directly in the conflict zone have been eliminated to avoid falsely classifying those areas that have gone through large changes in the training polygons.

The results show clearly what was expected. The amount of agricultural area decreases from 73 % to 67% of the total area (Figure 42). Almost half of all urban landscapes were destroyed which caused a decrease from almost 10% built-up area in 2013 to merely 5.2% in 2017 (Figure 43), and natural vegetation took over the abandoned and destroyed landscapes, causing this category to enormously increase from 17% before the conflict to 27% in 2017 (Figure 44).

However, when interpreting these results, it is important to note that this analysis might be flawed as it does not distinguish between different vegetation types apart from “agriculture” and “forest”. Much more, if not all, of the agricultural fields, might have actually been

abandoned between the start of the conflict and now, but their change in reflectance in the different bands might not have been big enough for the simple differentiation carried out here. Moreover, bare soil has been classified as “agriculture” in this study, as it is mostly already harvested fields; the chernozem is so fertile, that any non-harvested area would be overgrown by wild plants quite quickly. However, as a lot has been destroyed in the region, some bare soils might just be destroyed areas where nothing can grow anymore, and those would have been falsely classified as agriculture in this study.

It is interesting to note at what time changes occurred. The most dramatic changes happened between 2013 and 2014, which is understandable as this is when the conflict started and was most intense. On the other hand, while agricultural areas dropped quite continuously from 2015 until the most recent map and are most likely still dropping now, urbanized areas seem to have risen slightly from 2015 to 2016. This could easily be a classification mistake, or it could be the temporary building of simple settlements by rebels or soldiers. Another possible explanation is the temporary cease-fire, called “Minsk II”, that lasted from fall 2015 to spring 2016 and led to a partial calming of the situation until the fights later became more extreme than ever before (Kramer 2017; Gibbons-Neff 2016; Zinets and Prentice 2016).

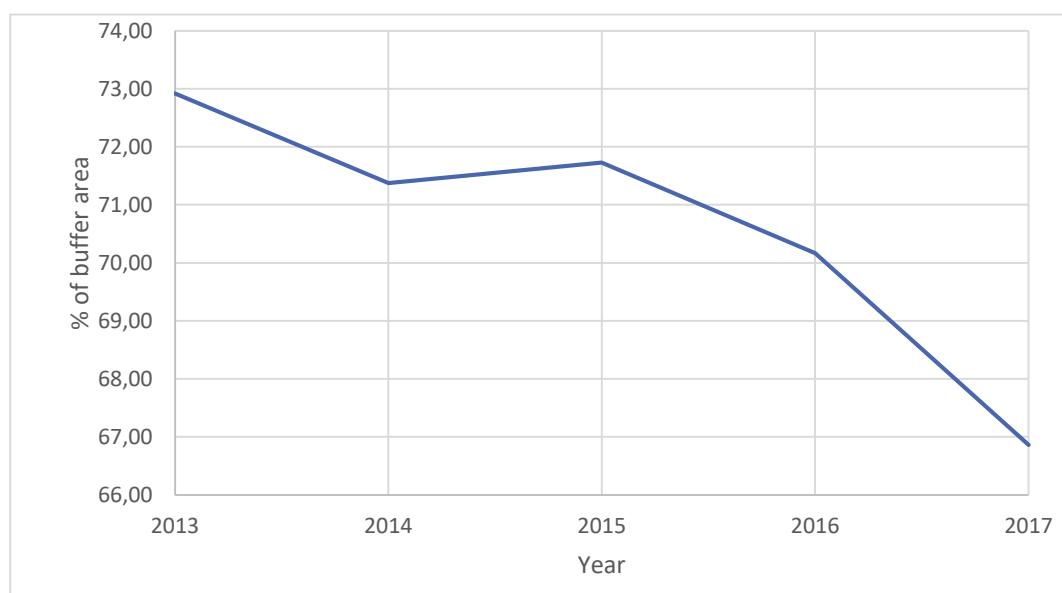


Figure 42: Development of agriculture in the Donbas conflict region from 2013 to 2017, in percent of the whole buffer area

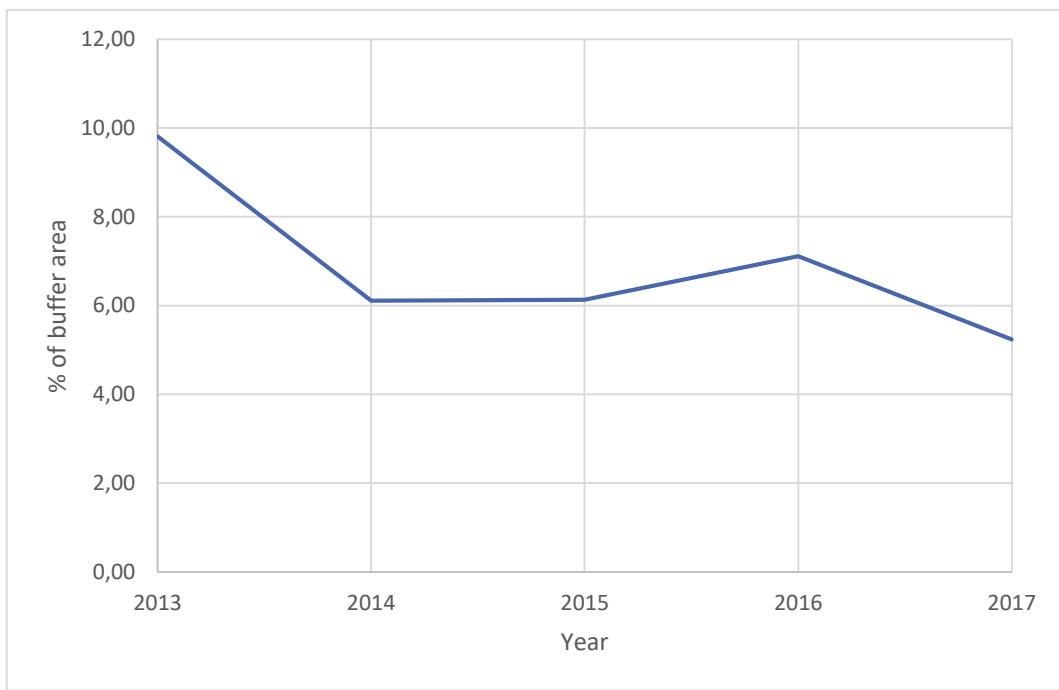


Figure 43: Development of urban area in the Donbas conflict region from 2013 to 2017, in percent of the whole buffer area

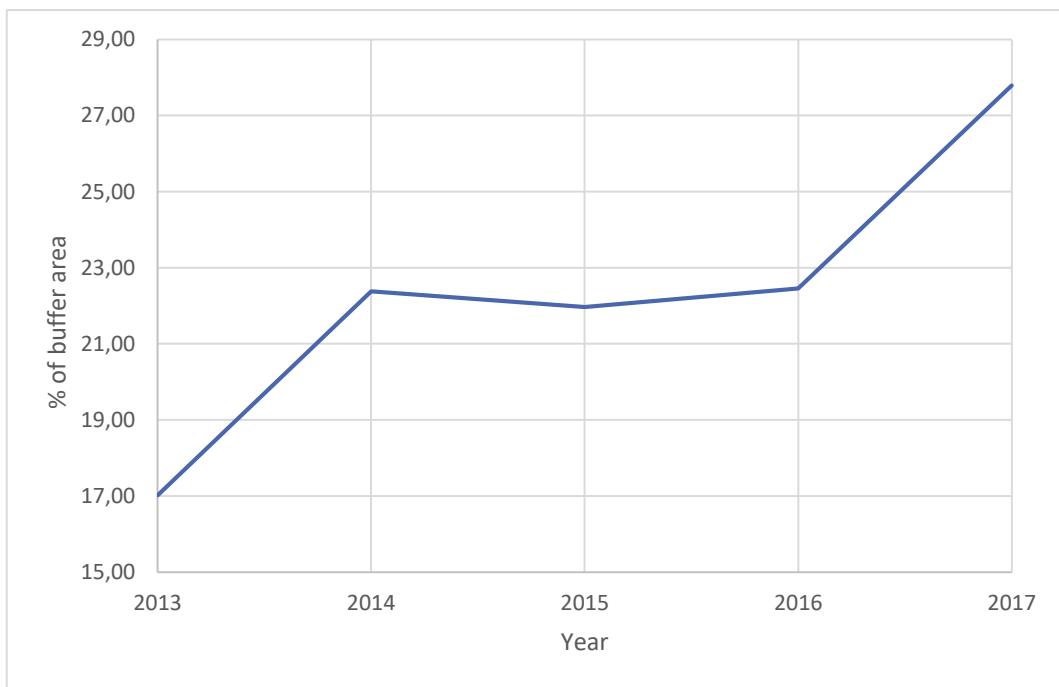


Figure 44: Development of natural vegetation and forests in the Donbas conflict region from 2013 to 2017, in percent of the whole buffer area

6.4 Validation of Results

As discussed in the literature review, Google Earth Engine has been utilized by several scientists and organisations in the last years. For some of the discussed results, such as those of the Global Forest Watch (Hansen *et al.* 2013) or the JRC Global Surface Water Explorer (Pekel *et al.* 2016), datasets are available online and embedded in GEE. This enables a form of validating the results obtained in this study by comparing them to results of other studies that used similar methods.

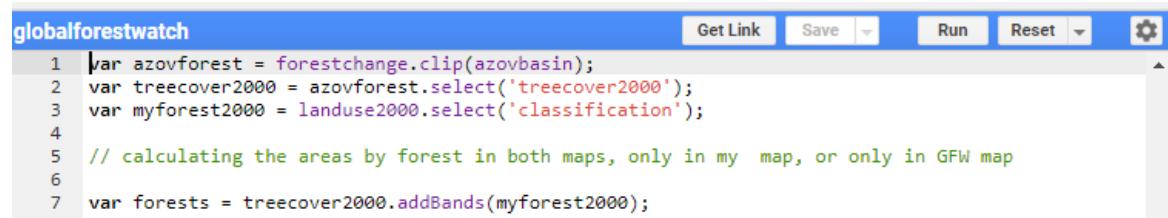
6.4.1 Global Forest Watch

The dataset of the Global Forest Watch is titled “Hansen Global Forest Change v1.3 (2000–2015), and can be imported into Earth Engine with the Image Collection ID “UMD/Hansen/global_forest_change_2015_v1_3”. It has several bands, such as “treecover2000” that contains data of the tree canopy cover for the year 2000 from 0 to 100 percent, which is the main band used for the following analysis. However, it also has bands about forest loss, that is only either marked as 0 (no loss from 2000 until 2015), or 1 (there was a loss from 2000 to 2015), as well as forest gain, again marked as 0 for no gain or 1 for gain between 2000 and 2012. The other bands are the same as those of the cloud-free Landsat 7 image composites that were used for the analysis. The reason why forest gain and loss has not been used for the comparison is the ambivalence of results; just “gain” or “loss” is too vague to draw clear conclusions as compared to the analysis of this study.

As a first step, the image has been imported and clipped to the area of the Azov Sea basin. Subsequently, the land use map of 2000 was imported as an asset.

For the comparison of the tree cover in 2000, a new image was created with the bands “treecover2000”, which is the corresponding map of the Hansen dataset, and “myforest2000”, which contained the classification bands of the 2000 land use map, and is therefore numbered

from 1 to 5, with class 5 being forests. The code for the merging of the data sets can be seen in Figure 45.



```
globalforestwatch
Get Link Save Run Reset
1 var azovforest = forestchange.clip(azovbasin);
2 var treecover2000 = azovforest.select('treecover2000');
3 var myforest2000 = landuse2000.select('classification');
4
5 // calculating the areas by forest in both maps, only in my map, or only in GFW map
6
7 var forests = treecover2000.addBands(myforest2000);
```

Figure 45: Preparation of data sets for the forest analysis

Afterwards, the image was masked three times in three different ways. While this research is only making a difference between forest and not-forest, Hansen *et al.*'s study actually determines the percentage of canopy cover as a number from 0 to 100 for each pixel. To still enable a meaningful comparison, different percentages were chosen as cut-off points of what still counts as "forest", to determine the percentage with the highest agreement between the maps.

The first mask called "both" contained all pixels that had a larger number in the "treecover2000"-band than the specified cut-off canopy cover, and a 5 as a classification ID number. The second masked, called "Hansen", had a larger canopy cover than the specified percentage, but was not classified as forest in the created map, thus was something other than 5. And the last mask, called "mine", was not reaching the specified cut-off canopy cover in the Hansen dataset, but it was classified as forest in the map created during this research. The classes were differentiated by using `.expression()` with the adequate comparison and ternary operators. The exact code for a canopy cover of over 90% can be seen in Figure 46.

```

9 //calculating where it was forest both times
10
11 var both = forests.expression(
12   "(b('treecover2000') >= 90) & (b('classification') == 5) ? 1 : 0"
13 ).clip(azovbasin);
14
15 var mask = both.eq(1);
16 var both = both.updateMask(mask).rename('both');
17 Map.addLayer(both, imageVisParam3, 'both');
18
19 //calculating where it was forest only for hansen
20
21 var hansen = forests.expression(
22   "(b('treecover2000') >= 90) & (b('classification') != 5) ? 1 : 0"
23 ).clip(azovbasin);
24
25 var mask = hansen.eq(1);
26 var hansen = hansen.updateMask(mask).rename('hansen');
27 Map.addLayer(hansen, {color: 'ff0000'}, 'hansen');
28
29
30 //calculating where it was forest only in my map
31
32 var mine = forests.expression(
33   "(b('treecover2000') < 90) & (b('classification') == 5) ? 1 : 0"
34 ).clip(azovbasin);
35
36 var mask = mine.eq(1);
37 var mine = mine.updateMask(mask).rename('mine');
38 Map.addLayer(mine, {color: '00ff00'}, 'mine');
39

```

Figure 46: Using `.expression()` and mathematical operators to determine the masks for both, Hansen, and thesis only data sets.

Consequently, the areas of all three classes were calculated as demonstrated in earlier chapters.

As this research does not clearly define how much the canopy cover has to be for a piece of land to be classified as covered by forest, the classes were calculated with several break-off points between 10% and 90% canopy cover, as visible in Figure 47.

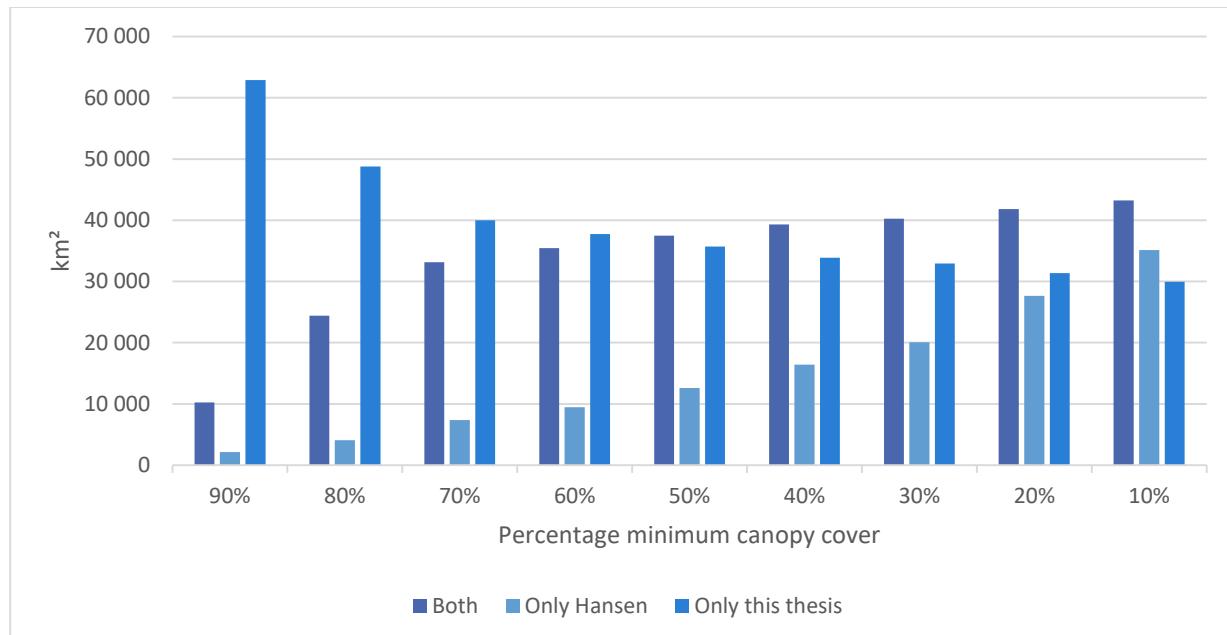


Figure 47: Forest classification of Hansen et al. 2013 versus this study

The results in Figure 47 show that the largest relative area has been classified as “both” for any canopy cover between 50% and 10%. For every value above 50%, the “thesis only” values exceed the “both” values, which is plausible, as a forest defined as only everything over 50% canopy cover is defined quite narrowly. The best ratio between forest classified in both maps, versus Hansen or thesis only, is reached at a minimum canopy cover of 50% or 40%. If only forest with a canopy cover of over 90% is counted as “forest”, the majority of forest area is only classified as such in this thesis, which is understandable; the small percentage that is still classified as “Only Hansen” then is most likely an inaccurately classified area or an area that was covered by clouds in the thesis data set.

The differences in forest classification can have plenty of reasons. Neither the GFW, nor this research is without errors, and it is difficult or near impossible to determine where the error lies in each specific location without a field visit.

Another reason can be the difference in base data used – while GFW used Landsat 7 images, this study used Landsat 5 images in order to avoid having to pre-process the faulty Landsat 7 images. Since the Scan Line Corrector failed in the early times of Landsat 7, around 22% of its pixels contain data gaps and faults (USGS 2018).

Furthermore, it has to be considered that this study is taking into account images of August only, but of four consecutive years (1998 – 2001), as the cloud cover was otherwise deemed too high for this specific year. GFW, on the other hand, uses all images of the year 2000, without any consecutive years or constraints to any specific months.

Equally important, as GFW has a global scale for its paper, clouds over specific areas are not taken into account when choosing the base map. Therefore, it is possible that indeed forested areas have been overshadowed by clouds in the GFW example.

As fifteen researchers have been co-authoring the mentioned paper and its database, and as they directly collaborated with Google engineers and plenty different companies and organisations

to achieve the best results, their resources have been completely different than the ones for this study. With 30 meters per pixel, they could achieve a ten times higher spatial resolution as they did not have to deal with timeouts. All of this needs to be taken into account when comparing the results. Nevertheless, the level of agreement between the datasets is satisfactory.

6.4.2 JRC Surface Water Data Set

Just like the forest data, the EC JRC's water dataset is also available and embedded in Google Earth Engine as an Image Collection, offering one image per year from 1984 to 2015 with a single band called "waterClass". The band is numbered from 0 to 3; 0 meaning no data, 1 equals no water, 2 seasonal water, and 3 permanent water cover of that year. The ID for this Image Collection is "JRC/GSW1_0/YearlyHistory" and is provided by EC JRC and Google as a joint project, and used images from Landsat 5, 7 and 8 (Pekel *et al.* 2016).

This data was embedded into the script and compared to all seven land use maps, similarly as explained in the previous chapter with the GFW data, just this time filtered by year instead of canopy cover. Only areas with a value of 3 (permanent water) were considered as water for this comparison, in order to retain simplicity.

Figure 48 shows the results of said analysis. The majority of water is classified as such in both data sets in all years. As the JRC dataset is using a scale of 30 meters per pixel for its classification, it usually detects more of the smaller water bodies and slim or partly overgrown parts of the rivers, as opposed to this study. A scale of 300 meters per pixel is too coarse to detect those, but was necessary due to timeout issues. As the JRC cooperated with Google, it could most likely use unlimited computational power which made a more precise scale possible; this thesis, on the other hand, could only operate with the normal Google user's computational power and time, which times out after five minutes per computation.

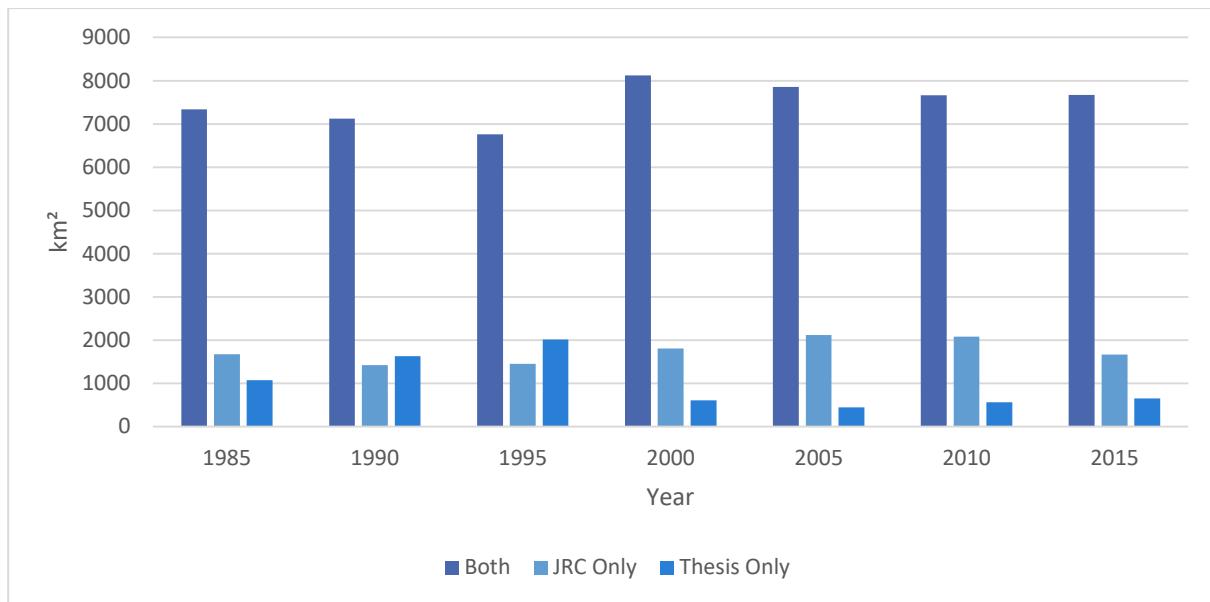


Figure 48: Area of water in km^2 as according to the JRC, this thesis, or both data sets.

From 1985 to 1995, the thesis only water detection is surprisingly high compared to the other years. This is the case due to data gaps in the JRC images; an example of this can be seen in Figure 49, where the black area means “no data” in the JRC set, but has been fully identified as water in the data set utilized for this thesis. The light blue in the graph means permanent water, green means no water, and red means seasonal water.

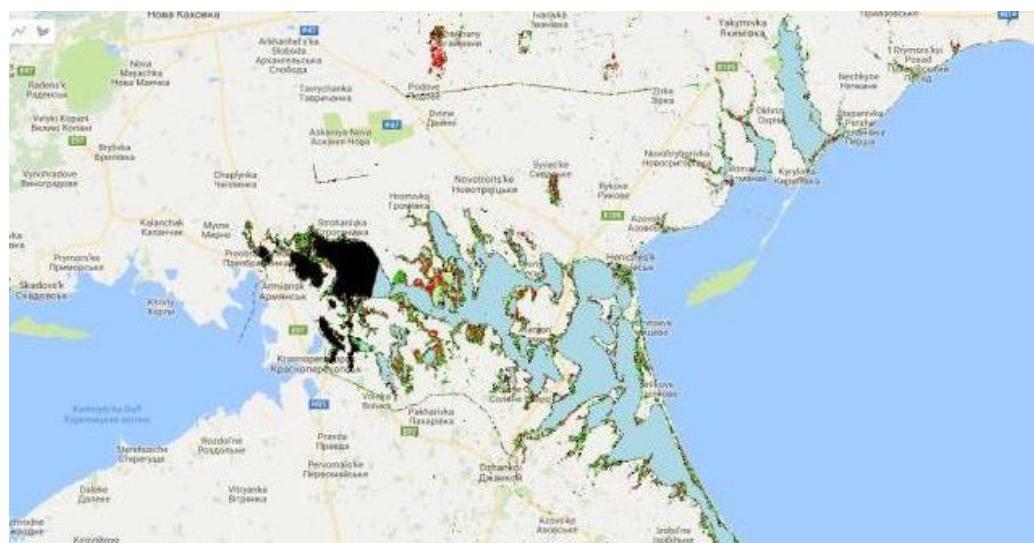


Figure 49: Data gaps in the JRC water data sets of 1985. Areas with no data are marked as black

Parts of the seasonal water of the JRC were in fact classified as water in this study. Figure 50 (map for 1985) shows in yellow which areas have only been identified in the thesis as opposed

to the JRC data for the same year. As visible, the data gaps make out the majority of the thesis only data.

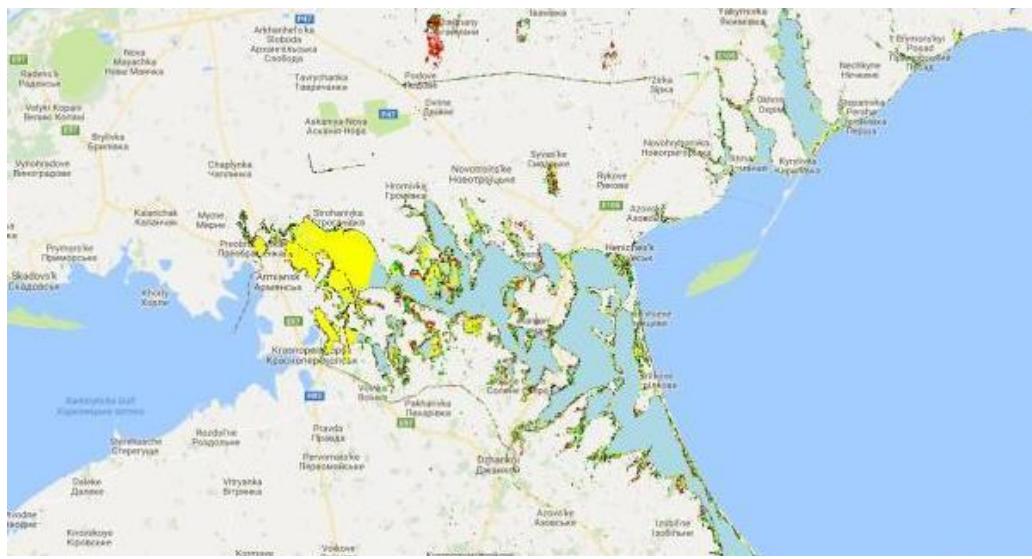


Figure 50: Areas identified as water in the thesis dataset (yellow), projected over the JRC data set

In general, therefore, the adequacy of the results obtained in this thesis can be called sufficient. Of course, it is not perfect, but no land use map that has purely been created with remote sensing techniques could ever be. Nevertheless, the quality of the maps achieved here can hold up to the quality of high-resource studies that were conducted over years and in cooperation with the Google engineers directly.

7 Impacts and Drivers of Land Use Changes

The results described in Chapter 6 will now be interpreted and set into the context of the DPSIR framework. Hereby, land use change will be seen as a pressure, and the amount of area occupied by each land use, especially natural vegetation, will be seen as the environmental state; the driving forces of the land use change, as well as the impacts of this state on the water cycle, will be evaluated. As no adequate policy responses tackle either of the discussed drivers, pressures, state or impacts, no responses will be discussed.

7.1 Drivers: Transition to a Market Based Economy

The most important socio-political driver of land use changes in the Azov basin in the time frame must have been the collapse of the Soviet Union in the early 1990s, as already briefly mentioned in Chapter 5.

In January 1992, the USSR's centrally-planned agricultural system suddenly collapsed and changed to a wild, market-driven economy that put pressure on Russia's peasants (Ioffe 2005). Consequently, the Russian agricultural landscape in the 1990s was shaped by chaos, as state subsidies decreased from around 60 Billion US Dollars a year in the late 1980s, to roughly one Billion Dollars a year in the late 1990s. Still in 2004, Russia was only providing its farmers with around 12.50\$ of state subsidies per hectare agricultural land on average, while the USA were spending around 200\$, and the EU 800\$ per hectare, in comparison (Ioffe 2005).

Prior to 1990, 98% of all agricultural land was managed by large collective farms that, instead of utilizing their economies of scale, were vastly inefficient and produced mostly unprofitably (Sedik *et al.* 1999). They were hardly able to adapt to the new rules of the free market, were not supported by the government anymore, but were broken down only slowly. Gradually, private family households took over large parts of the land during the chaotic transition as more and more people started to leave (Ioffe *et al.* 2004); however, a moratorium on trading land was established in the 1990s, which enabled local authorities to take away lands again that were

previously declared as private, which led to more land abandonment (Bobylev and Ligert 1994).

As a result, agriculture all over the country decreased in its inputs and outputs and did not resume growing until 1999, which caused the abandonment of land and a migratory trend away from the countryside (Ioffe 2005).

Overall, the state failed to establish a successful new agricultural system, due to a lack of technological advancement, not enough bank credits for farmers, and only an inadequate establishment of new market structures (Pallot and Nefedova. 2003). In parts of Ukraine, the abandonment of agricultural land as a result of the breakdown of the Soviet Union has been at least as dramatic as after the Chernobyl disaster in 1986, when parts of the country had to be evacuated (Hostert *et al.* 2011).

The environmental and social impacts of these agricultural changes have been immense. While the situation in the better-off and more fertile chernozem areas of the Azov Sea basin has been less extreme than in the rest of the country, poverty and desperation among the rural poor have still been tremendous (Ioffe 2005).

The following two sub-chapters analyse data on the two most important factors when considering land use change during the collapse of the Soviet Union – productivity changes per hectare land on the one hand, and agricultural land abandonment on the other hand.

7.1.1 Agricultural Productivity Changes

Policy changes as explained above led to, among other effects, a loss in the productivity of the Russian agriculture, especially since the decrease in state expenses also heavily resulted in a decrease in fertilizer subsidies (Prishchepov *et al.* 2013; Trueblood and Arnade 2001). This makes especially the analysis of pre-2000 data valuable, as all agricultural activity was decreasing during that time frame.

Data containing the grain yield per hectare averaged from 1986 to 1990 and from 1997 to 2000, respectively, was provided by the EnviroGRIDS Project (EnviroGRIDS 2012). The data set contains 203 data points that stand equivalent to rayons (districts) and that are situated within the boundaries of the Azov Sea basin. The analysis of this data set showed an overall reduction of the yield in its median as well as mean. A dependent, two-tailed t-test revealed that the difference between the means of the pre-1990 and pre-2000 data is significantly different, as the p-Value was $2.2 * 10^{-16}$, and therefore very significantly within the confidence level of 95%.

The Boxplot for these variables can be seen in Figure 51.

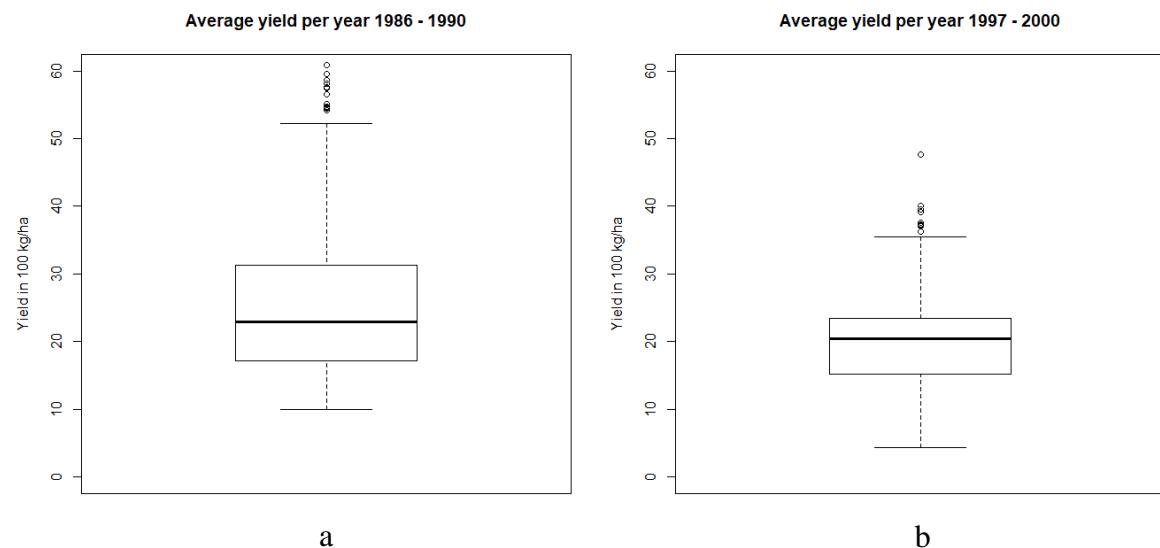


Figure 51: Box-Whisker Plots of grain yield in 100 kg per ha, averaged over the years 1986 – 1990 (a) and 1997 – 2000 (b)

For the purposes of reviewing the productivity of the four sub-basins, additional analysis was carried out that split the data into the sub-categories Kuban river basin, Upper Don river basin, Lower Don river basin, and the remaining parts of the sea basin labelled as “Rest”, most of which directly surround the sea and are made up of approximately 20 smaller streams. Of all the reviewed rayons where yield data was available, 22 were located in the Kuban, 93 in the Upper Don basin, 67 in the Lower Don basin, and 21 in the remaining parts. The Boxplots in Figure 52 make the changes of agricultural productivity in the sub-basins visible. The means in

all four areas have decreased dramatically, however, by far the lowest change has been observed in the Upper Don area, where the mean has only decreased by 7% (from 1.91 t/ha to 1.77 t/ha), whereas in all three other sub-basins, the change of the means have been over 30% (Lower Don 31% change, Kuban 33%, Rest 38%). It is interesting to note, additionally, that while the Upper Don region is occupied by more than 40% of agriculture, its grain yields even before the big structural changes in the 1990s have been fairly low compared to the Kuban and Rest, which are made up of agriculture by approximately 25% (Kuban) and 80% (rest), respectively.

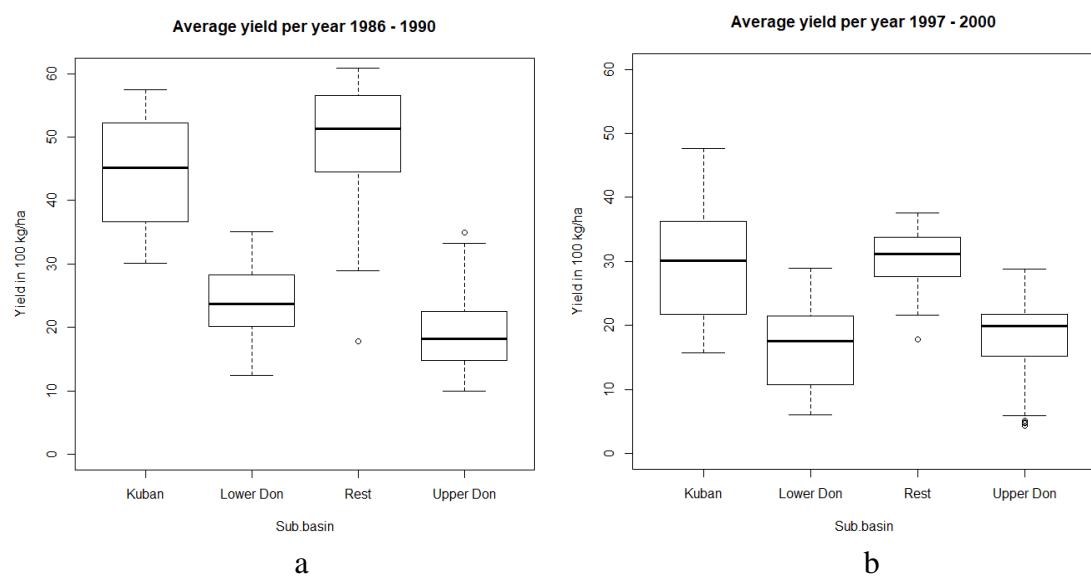


Figure 52: Box-Whisker Plots of the grain yield in 100kg per ha, averaged over the years 1986 – 1990 (a), and 1997 – 2000 (b)

However, not only crops were affected by the reduction of productivity - when regarding the milk yields in kilogram per cow per year of the different rayons, a similar, despite less dramatic, picture becomes visible. The mean has been reduced from 2746 kg/cow to 2216 kg/cow, as visible in the Box-Whisker Plot in Figure 53, and are strongly statistically different according to a two-tailed dependent t-test.

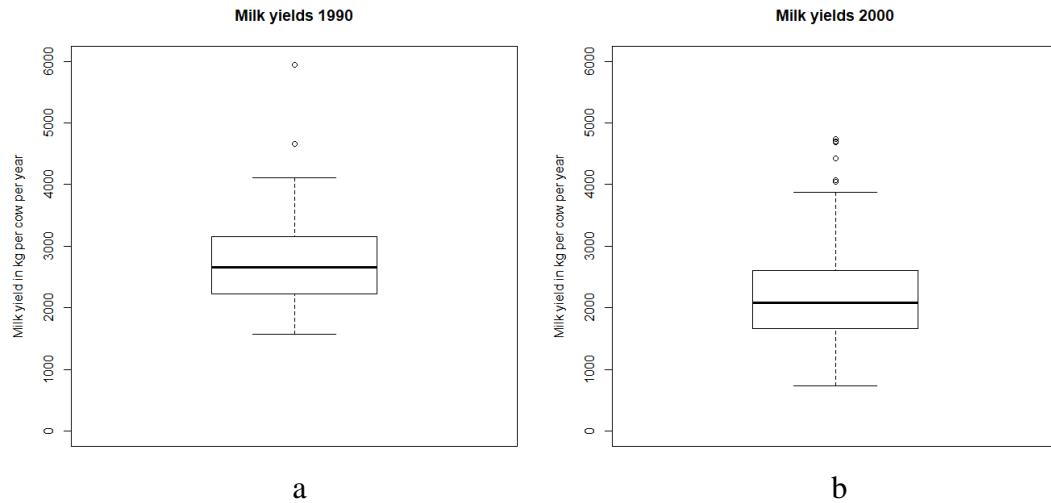


Figure 53: Box-Whisker-Plots of the milk yield in kg per cow and year, for 1990 (a) and 2000 (b)

Again, an analysis of the four sub-basins was also carried out. This time, there was data available in 24 rayons within the Kuban river basin, 93 in the Upper Don basin, 70 in the Lower Don basin, and 28 in the remaining basin. Again, milk yields decreased for all subbasins. However, for this variable, the Kuban mean decreased only by 21%, the Lower Don by 28%, Upper Don by 26%, and the remaining parts only by 16%. The Box-Whisker plots can be seen in Figure 54.

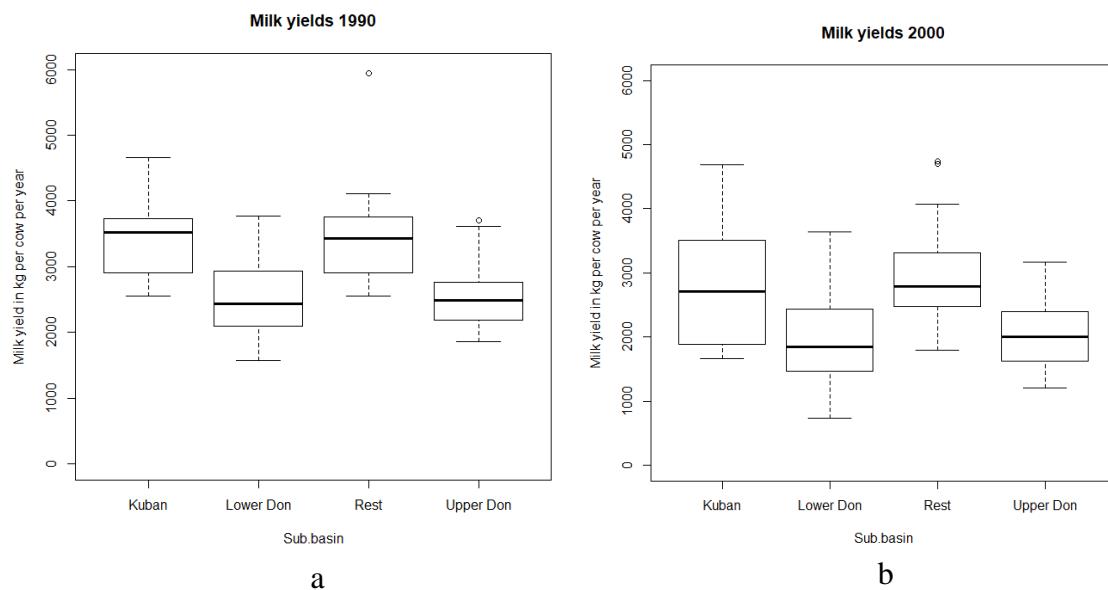


Figure 54: Box-Whisker Plots of the average milk yield in kg per cow and year for 1990 (a) and 2000 (b)

The decreasing productivity over several agricultural products could be a determining driving force in the socio-ecological system of the Azov Sea Basin. This could explain the decreasing

agricultural area of the total basin, especially between 1995 and 2000, as more and more farmers left the region and agricultural areas were taken over by natural vegetation and forests. A return later on could be explained by the agricultural importance of the basin – while all of Russia and Ukraine were in an economic crisis at the time, the fact that this area has the most fertile soil - chernozem - in both these countries, and is one of the only regions in Russia that has a climate suitable for agriculture, could have motivated people to return to this region first, which is why the overall trend of the basin is not strongly negative, but merely fluctuating over time. Nevertheless, even nowadays, studies show that the agricultural productivity and yield rates of Russia and Ukraine could still be expanded vastly and hold possibilities for the further food security of the region (Deppermann *et al.* 2018).

7.1.2 Abandoned Land

In the light of the previous chapter, the exact analysis of abandoned agricultural land is worth discussing. For this study, the created land use maps explained above have been used for reference. As the consideration of all seven maps would have increased the graph's options to 128 possibilities, only the land use maps of 1985, 2000, and 2015 have been used, creating only 8 possible scenarios and therefore a simplified version of reality, but a more concise one for this study's purposes. Therefore, the categories have been reduced to “Agriculture” or “Not Agriculture”, without splitting them up in all five different classes as above. Similarly to the code in the previous chapters, *.expression()* was used in combination with comparison and ternary operators to determine those areas where agriculture has always been there (1 – 1 – 1), never been there (0 – 0 – 0), and all other combinations as explained below. The full code can be found in Appendix D. One of the eight options was so small that it can be neglected, which is areas that have not been agriculturally used in 1985 or 2015, but that have been agriculture

in 2000 (0 – 1 – 0). The results for all other 7 combinations can be viewed in Figure 55.

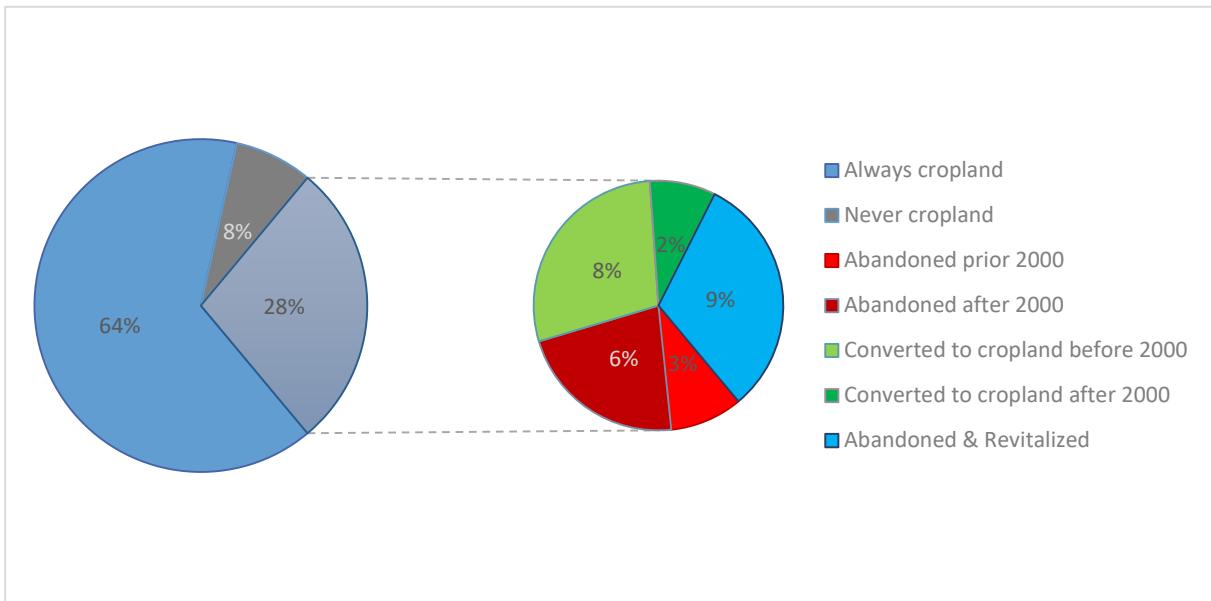


Figure 55: Development of agricultural areas in the Azov sea basin, in percent of the total area

Between 1985 and 2015, the case study area has been covered permanently to 64% with agricultural area; on the other hand, there are only 8% that have never been agriculture at all, so that have either permanently been forest, water, or urban/industrial areas. The remaining 28% have changed between agricultural and non-agricultural land use types over time, and can pretty equally be split into three groups. According to these maps, around 9% of the land in agricultural areas have been abandoned over time and have not been converted back to agriculture. However, around 10% of all areas have also been converted to cropland after being other land use types before, and the other 9% were agricultural land before and after 2000, but not in the year 2000 itself.

Therefore, a total of 18% of land has been abandoned at least at some point between 1985 and 2015, even though as explained above, this has not been permanent in half the cases, probably due to the critical importance of this area to Russia's food supply. Conversely, almost as much agricultural area has been added between 1985 and 2015, as has been abandoned during the same timeframe. This means that despite land abandonment being a temporal problem during

the fall of the Soviet Union, the area is mostly not returning to natural habitats and afforestation is not taking over agriculturally used lands.

Figure 56 furthermore shows where what kind of land transition took place. All land that was either permanently agriculture or never agriculture is visualized in grey. Permanently abandoned land is pictured in red, land that was converted from other land use types to agriculture is yellow, and blue is the colour for land that used to be agriculture before 1990, and was only temporarily abandoned thereafter. The red strip in the north-western corner of the map is due to a cloudy area in the map of 1985 and is, therefore, not relevant. Other than that, it is interesting to see that land was abandoned and added all around the basin more or less equally; however, the majority of only temporarily abandoned land (blue) is clearly located in the area west and north of the Tsimlyansk reservoir. This area is located in the Upper and Lower Don catchment area, and therefore, the area of the least primary grain yield, but also of the least productivity reduction as explained above. It is, therefore, plausible, that the agricultural area was first reduced there as yields have not been high, but also, that it has been re-developed quite soon as the fall in productivity was less dramatic than in other regions.

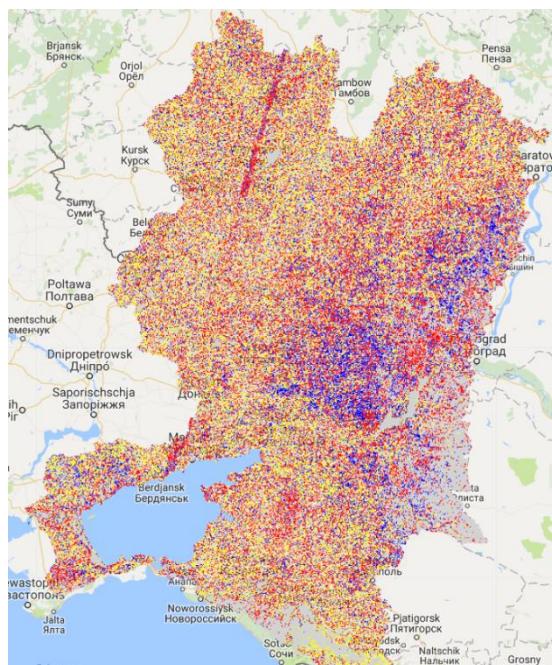


Figure 56: Visualization of Abandoned agricultural land (red), added agricultural land (yellow), and land that was temporarily abandoned and re-developed (blue).

In respect to literature, it makes sense that in the Azov Sea Basin (one of the most agriculturally productive areas of the wider region), land abandonment was relatively low in comparison to the rest of the former Soviet Union. This is plausible as a market-oriented economy demands that only the most efficient farms can succeed when state support and subsidies fall away (Ioffe *et al.* 2014; Prishchepov *et al.* 2013; Deppermann *et al.* 2018). Studies investigating land abandonment in the former Soviet Union reached the conclusion that the chernozem zone was the least damaged of the economic transitions, due to its soil fertility, relatively mild climate, and proximity to regional centres and the ports of the Black and Azov Sea (Ioffe and Nefedova 2000; Deppermann *et al.* 2018).

A detailed study only regarding the Stavropol oblast (Ioffe *et al.* 2014), which is located in the Lower Don as well as in the Kuban river basin, also concluded that it has been much more successful in keeping up a relatively high living standard and, hand in hand with that, a lot of actively farmed lands. Even private family farms have largely not been leased to bigger corporations, but have stayed in the hands of the original owners. The loss in productivity during the 1990s has by now been restored in those areas (Ioffe *et al.* 2014). Overall agricultural productivity and yields per area are likely to further increase in the future, as less productive areas are being abandoned and agricultural areas in the Azov Sea basin and other productive areas are not (Ioffe *et al.* 2004).

7.2 Drivers: Population Change and Urbanization

Most commonly, an increase in built-up land either stems from changes in the amount/location of the population, or from the development of the local industry and transportation infrastructure (Meyer and Turner 1992). To rule out some of the potential reasons and draw conclusions for the case study, the population development in the Azov Sea basin was examined.

For this purpose, the spatially explicit population datasets found at the NASA Socioeconomic Data and Applications Center (SEDAC) and created by the Center for International Earth Science Information Network (CIESIN) at the Columbia University were used. All used data was available in a 30 arc second resolution (approximately one square kilometre per pixel), has been downloaded in the GeoTIFF format and has subsequently been uploaded as several assets to the Google Earth Engine servers.

Unfortunately, no single dataset spanning the whole study time frame was available. Therefore, the used data came from two different sets:

The first set, spanning a time frame from 1970 to 2000 and having data available in ten-year steps, was used for the years 1980 to 2000. The set is called “Global Population Count Grid Time Series Estimates, v1 (1970 – 2000)”, and is based on the official UN World Population Prospects (UNWPP) of the years 2000 and 2008, to then project the data backwards. The data sets are believed to be good estimates; nonetheless, they do not depict reality absolutely correctly, which should be taken into account when reading the next few pages (CIESIN 2011; CIESIN 2017a).

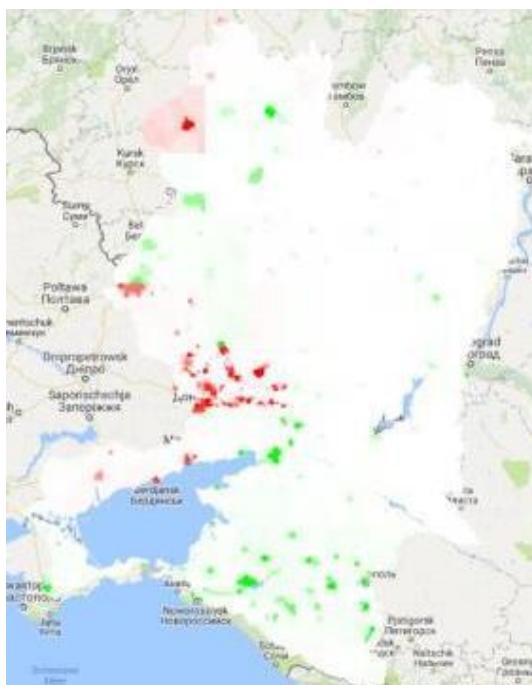
The second data set used was available in 5-year-time steps from 2000 until 2020 and is based on the UNWPP of 2015 and was then projected back- and forwards. The dataset is called “UN WPP-Adjusted Population Count, v4.10”. It is available in several resolutions; for an optimal comparability, the 30 arc second resolution data sets have been chosen (CIESIN 2017b).

According to the official guidelines of the data sets, it is not recommended to use them in combination with one another, which is why they will be treated separately. The year 2000 will be used for both parts of the analysis, once with the first dataset and once with the second to avoid confusion, as they are based on slightly different assumptions. However, the methods of analysis were the same for both datasets:

First, the data sets were imported and clipped to the basin; then, they were all put into a single image as bands with the different years in order to do calculations. Subsequently, the bands were subtracted from one another, added to the layer and visualized as green for population increase, red for a decrease, and white for no change. Furthermore, the mean population per pixel of the subbasins have been calculated.

7.2.1 Population Development 1980 – 2000

Figure 57 shows the development of the region as a single image from 1980 to 2000 as according to the CIESIN 2017a data set explained above. Red, in this case, means a decrease in population, white means no change, and green indicates a population increase. The maximum decrease observed in those twenty years was located in the urban Donetsk region, where the population density decreased by around 67 people per 1 km^2 grid cells. The maximum increase, on the other hand, was around 75 people per 1 km^2 grid cell in the Upper Don as well as in the Kuban region in the South of the basin.



*Figure 57: Population development per grid cell (30 arc seconds resolution) from 1980 to 2000.
Data Source: CIESIN 2017a*

Figure 58 shows the development of the mean population per grid cell in the basin. While all areas showed a small increase in mean population from 1980 to 1990, only the Kuban area

continued growing beyond that up to 2000, while the Don and the Upper Don decreased slightly, pushing the overall mean population down too. This graph also makes clear that the Kuban has the highest average population density as compared to other regions, despite its high forest cover and relatively low agricultural land use cover.

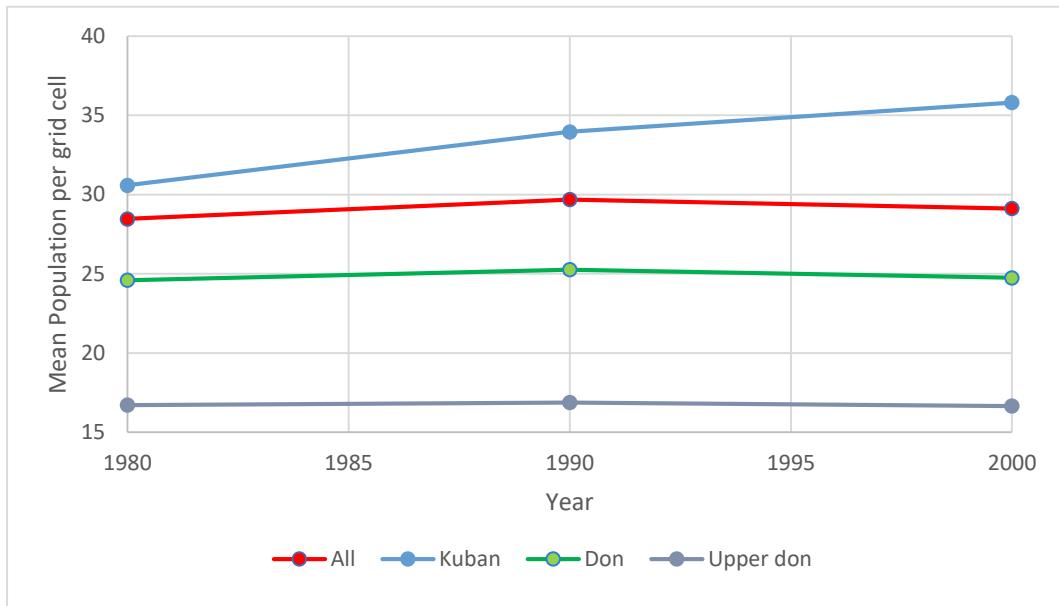


Figure 58: Population development from 1980 to 2000 in the different regions of the Azov Sea Basin, in No. of people per 30 arc seconds grid cell. Data Source: CIESIN 2017a

The development of the total population of the regions is in line with that, as visible in Figure 59 (not cumulative). This figure also shows that while the Kubas has the highest average population density, it has in total the lowest population of all.

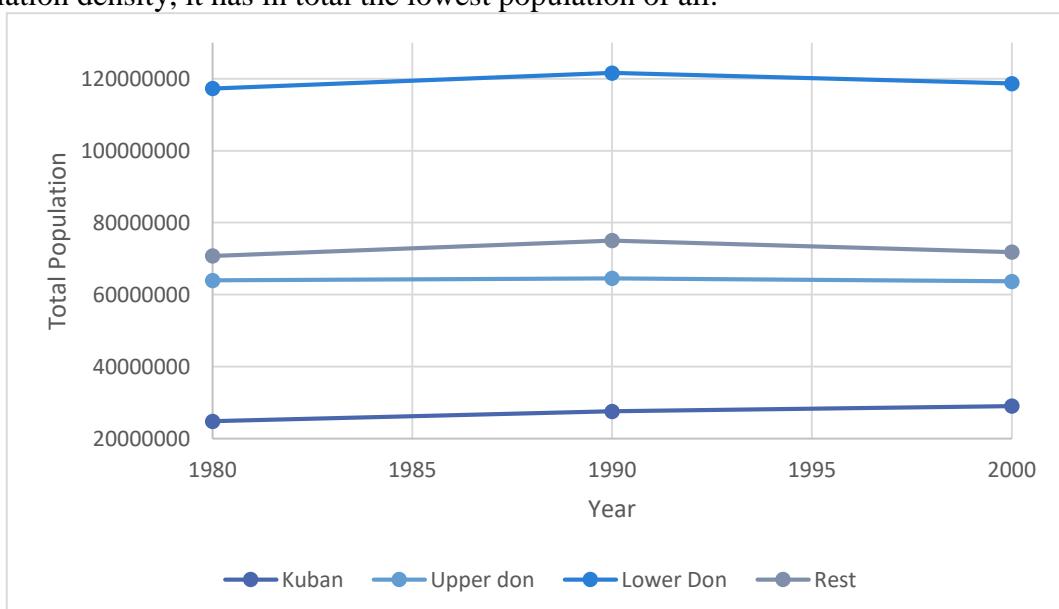


Figure 59: Population development from 1980 to 2000 of the different sub-basins within the Azov Sea Basin. Data Source: CIESIN 2017a

7.2.2 Population Development 2000 – 2015

From 2000 to 2015, the overall population has decreased significantly; the maximum decrease in one grid cell was 123 people, whereas the maximum increase was 60 people. Especially in the Donbas region, the change has been dramatic; not least because of the political changes described in Chapter 7.1. Figure 60 is showing where which changes occurred – again, especially the Kuban basin has increased in those fifteen years, together with regions in the North-West of the basin around the city Belgorod. The maps picturing the change in five-year steps can be found in Appendix E.

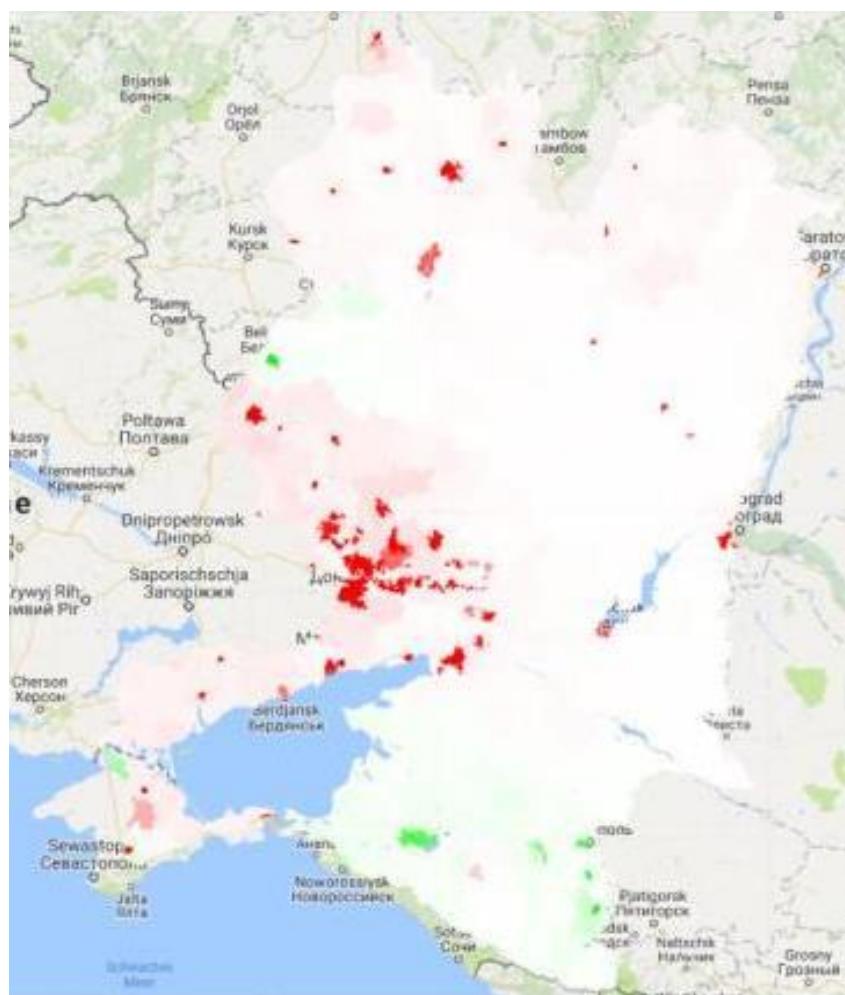


Figure 60: Population change from 2000 to 2015. Areas with a decrease in population are pictured in red, areas with an increase are pictured in green. The darker the colour, the stronger the change has been. Data Source: CIESIN 2017b

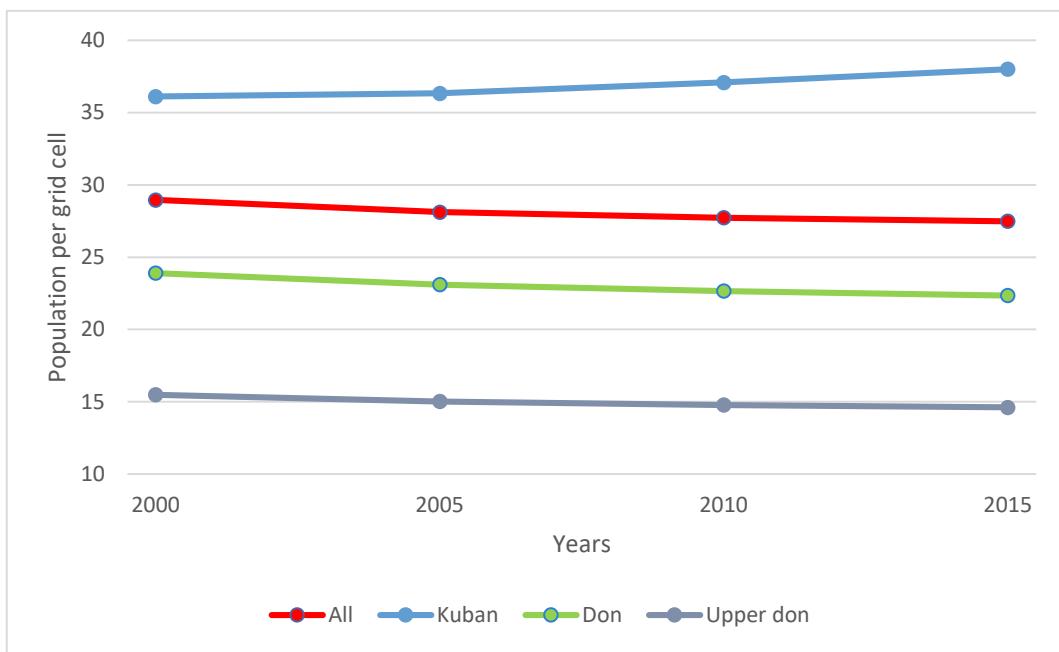


Figure 62: Population development from 2000 to 2015 in the different regions of the Azov Sea Basin, in No. of people per 30 arc seconds grid cell. Data Source: CIESIN 2017b

Again, the Kuban area, being the sub-basin with the highest average population density, has a steadily increasing mean population density from 2000 to 2015 (Figure 62). In absolute (non-cumulative) numbers (Figure 61), the situation is similar – in the total area, as well as in the Don and Upper Don area, the overall population is decreasing. Only in the Kuban area, the overall population is slightly increasing.

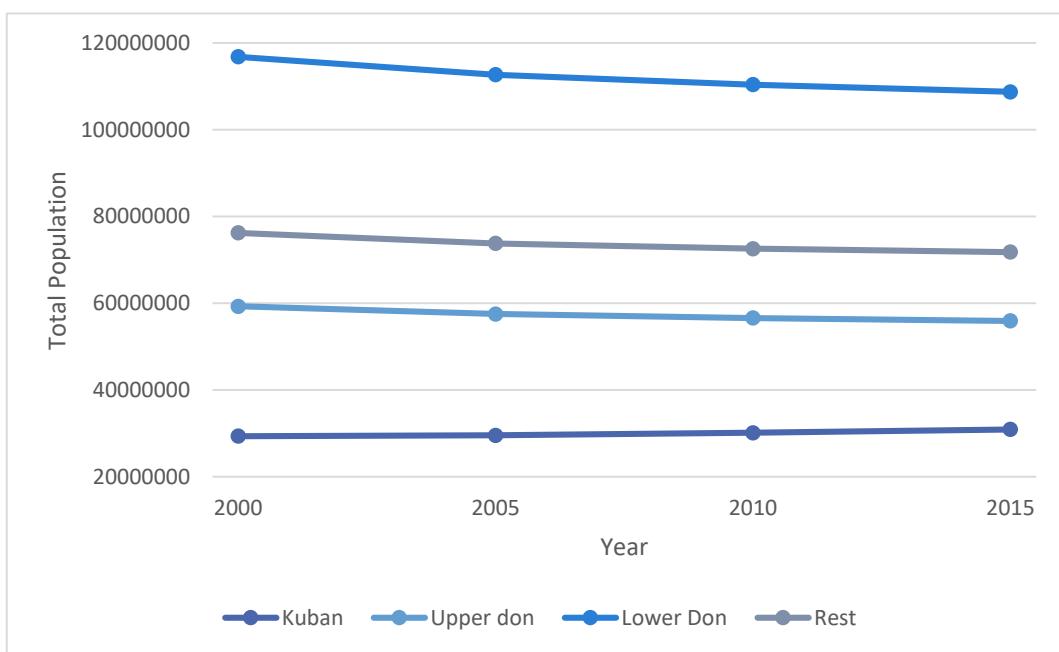


Figure 61: Population development from 2000 to 2015 of the different sub-basins within the Azov Sea Basin Data Source: CIESIN 2017b

7.2.3 Correlation Population Development and Urbanization

Similar to the analysis of change in agricultural land, the correlation of urbanized land and population development was analysed with the data sets described above.

Again, the land use maps of 1985, 2000 and 2015 were imported as assets to GEE and divided into urban land (1) and non-urban land (2). Following that, an image was created that has three land use bands of the three years, as well as 2 population datasets – the first one being the population change between 1980 to 2000 (CIESIN 2011; CIESIN 2017a), and the second one being the change of 2000 to 2015 (CIESIN 2017b). Then, with the help of *.expression()*, different categories were created and the area of each of the categories was calculated. The four main categories of interest were as follows:

- Land that was urbanized in 1985, however, not in 2000 or 2015, therefore leading to the conclusion that it was permanently abandoned before 2000. This category was named “Abandoned Before 2000”, and was split into land where population either decreased (“Decrease”) from 1980 to 2000, or not (“No Decrease”).
- Land that was urbanized in 1985 and 2000, but not in 2015; therefore labelled as “Abandoned after 2000”. This category was again split into land where the population decreased between 2000 and 2015 (“Decrease”), or not (“No Decrease”).
- Land that was not urbanized in 1985, but that was urbanized in 2000 and 2015; this category is labeled as “Added before 2000”. This category is split into an increasing population between 1980 and 2000 (“Increase”), and a stable or decreasing part (“No Increase”).
- Land that was not urbanized in 1985 and 2000, but that was urban in 2015 (“Added After 2000”). Again, it was split into an increase between 2000 and 2015 (“Increase”), and a stable/decreasing, but urbanized part (“No Increase”).

All other combinations are either not of interest, or occupying an area too small to be worth analysis.

Figure 63 shows the total added land as explained above. It is interesting to note that much more urbanized areas were added post-2000 than before 2000, which could most likely be driven by the transition to a market-based economy as explained above, which caused the livelihoods of people to decrease in the 1990s. However, the areas that were added before 2000 were predominantly added due to a population increase in the same area, whereas after 2000, urbanization mostly did not go hand in hand with a population increase. This could be related to the taking over of commercially owned farms and industrialization (Ioffe *et al.* 2004). The majority of the built-up area is, therefore, not housing, but either transportation infrastructure for trade or industrial buildings in areas of little population change.

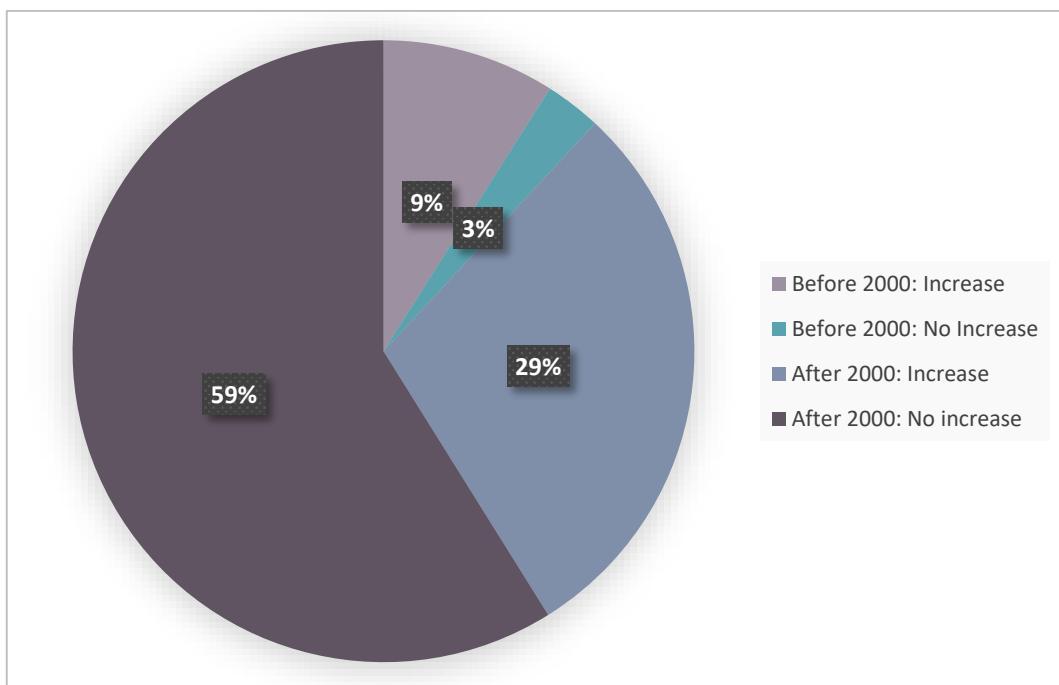


Figure 63: Urbanized land divided into the point of time of urbanization and relationship with population increase

The graph showing the total abandoned land (Figure 64) is showing quite the opposite of the added land, as is plausible. Most land was abandoned prior to 2000, which makes sense with the huge economic transitions the area had to go through between 1990 and 2000. However,

around two-thirds of the pre-2000 abandoned land does not show a decrease in population at the same time. This could be explained by the loss of infrastructure and industry in that time frame; while people were not necessarily moving away, they did focus more on subsistence agriculture instead of exporting goods (Baumann *et al.* 2011), which potentially decreased the amount of built-up area in the sea basin. After 2000, much less land was abandoned, some of which was related to a population decrease, while other land was just abandoned without a decrease in population, which could, again, be due to infrastructure changes.

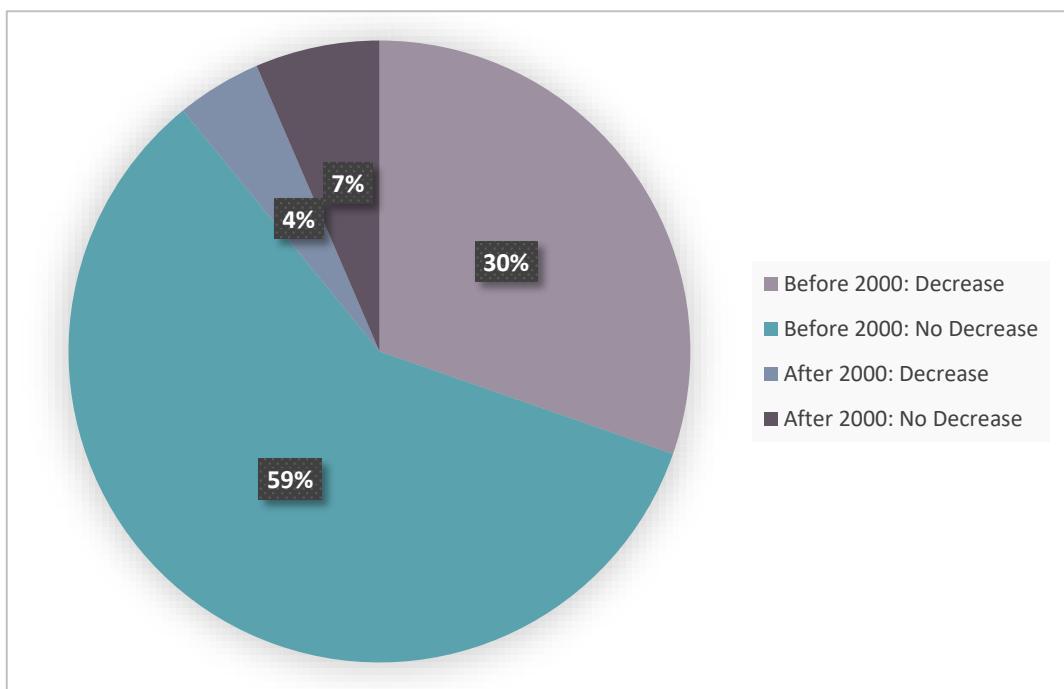


Figure 64: Land switching from urbanized to other land use types, split into point of time of the transition, and relationship with population decrease

7.3 Impact: Freshwater Ecosystem Services

Due to the complex relationships between land use and land cover changes and the hydrology of a watershed, data for the Upper Don area will be extracted from the overall results, and only changes in that part of the case study area will be discussed. Since the Tsimlyansk reservoir is offering those ecosystem services that have the most direct and important effect on the well-being of the whole case study region, such as energy production, drinking water provision, irrigation, et cetera, it makes sense to focus on everything upstream of the reservoir; also, this

way, the actual water level of the reservoir can be analysed, as reliable water level data is only available for this reservoir in the region.

First, the literature covering land use/land cover changes and their effects on the water yield will be discussed, to subsequently set them into context with data acquired by NASA's Global Land Data Assimilation System (Rodell *et al.* 2004). This dataset is also embedded in GEE and ready for spatially explicit analysis there; however, as the datasets there are only available in 3-hour-values as opposed to monthly values, long-term analysis is impossible as it will always lead to a computation time out, or to an exceedance of the user memory limit. An analysis over 26 years as targeted here (from 1984 to 2010) in 3-hour-time steps would lead to 75 920 data points in the set, and cannot easily be averaged over each month, as looping is not recommended in Earth Engine and can lead to errors. As, however, the same dataset is available for analysis in the Giovanni online data system, developed and maintained by the NASA GES DISC, the monthly-averaged datasets were downloaded from there and analysed offline (Acker and Leptoukh 2007).

Unfortunately, no reliable data discussing a change of water quality in the Tsimlyansk reservoir or the Don River could be found; therefore, only changes of water yield will be discussed.

To understand the effects of land use changes on stream flow and water yield, the different ways water can reach the flow must be taken into account. The water yield of a river depends on different processes above and below ground; most importantly surface runoff (consisting of snowmelt runoff and overland flow), evapotranspiration (often split into transpiration and evaporation), infiltration, base flow, lateral groundwater inflow, discharge, and effective precipitation (Kresic and Mikszewski 2013; Brutsaert 2005).

The water cycle, visualising the relationship between these parameters as a simplified version of reality, is shown in Figure 65 (NASA 2018).

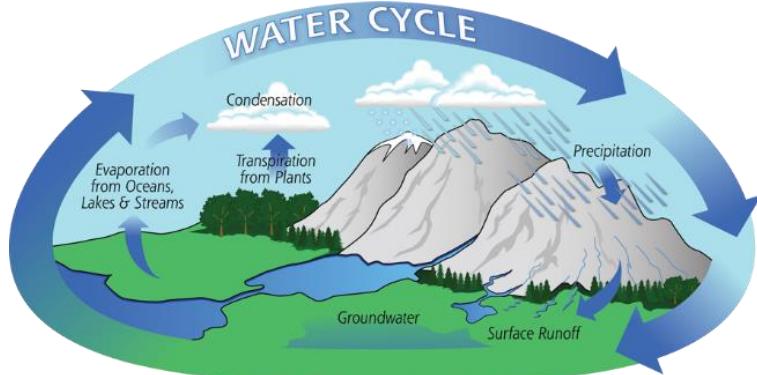


Figure 65: The water cycle and its most important components
Source: NASA 2018

Simplified, all these parameters can be summarized into 4 major variables – runoff (R), precipitation (P), evapotranspiration (ET), and the change of water stored in rivers, ponds, lakes, or underground (dS), which is necessarily the most important factor for water ecosystem services. The relationship between those four variables can be calculated as (Jenkins *et al.* 1994):

$$R = P - ET + dS$$

The streamflow and water yield of a river is, therefore, positively influenced by increased precipitation and runoff, and can be negatively influenced by increased evapotranspiration. The causes of a change in these parameters are complex, with soil type and moisture, climatic variables, and topography being in interaction with land use and land cover changes, making the differentiation of the importance of all of these causes difficult to determine (Wei and Zhang 2010). Especially changes in temperature and precipitation can have immense influences, as they individually influence all of the above-mentioned flows. Subtracting out the climatic variables, however, is a difficult task. It is useful to e.g. consider the run-off coefficient, which is commonly defined as the ratio of annual rainfall to annual runoff, therefore excluding the effect increased rainfall has; however, measuring and determining the annual runoff is an extremely difficult task for a larger case study area (Wei and Zhang 2010).

7.3.1 Literature Review of LULCC in a Watershed

For the sake of staying within the scope of this research, only effects of changes between the land use types forests, cropland, pastures, and urban/built-up land will be discussed, as these were also the land use categories investigated in this study.

The effects of deforestation and afforestation on the water cycle have been studied well in the past years, especially on a smaller, mostly sub-basin, scale, but with partly contradictory results. In general, forests are known to retain more water in their soil and increase evapotranspiration, decrease run-off and soil erosion, and decrease the local temperature (Sun *et al.* 2016). This can especially be helpful when managing smaller floods, as the peak run-off is often decreased (Hundecha and Bárdossy 2004). Deforestation, on the other hand, almost independently of whether the forest is transformed to agricultural land or urban land, usually leads to an increased water yield all year around. This is due to increased run-off and decreased evapotranspiration, which leads to higher peak water flow accompanied by a general higher total discharge amount and a higher groundwater flow with decreasing forest area in a catchment basin (Anaba *et al.* 2017; Dwarakish and Ganassi 2015; Sun *et al.* 2016; Fohrer *et al.* 2001). Additionally, the evapotranspiration has been shown to decrease more when a forest is transformed to agricultural cropland than to pasture (Fohrer *et al.* 2001).

Contrarily, however, other recent studies found woodland and forests to actually increase run-off, explaining this e.g. with steeper slopes in the forested area, and a too small hummus layer of the observed forests (Liu *et al.* 2017; Wang *et al.* 2014). Furthermore, while forests buffer smaller floods, during larger flood events, forest area could actually be positively correlated to the amount of run-off, as – once saturated – the soil can give off more water at once. The resilience of forests is, thus, limited. While they do balance the stream flow during smaller flood events, this can actually back-fire in extreme cases (Sriwongsitanon *et al.* 2011).

Moreover, there are also indirect effects of land use changes on runoff and discharge. By transforming forests to agricultural land, the increased albedo can change the micro-climate and thus lead to more evaporation (Vitousek *et al.* 1997). At the same time, built-up areas can store heat and further increase the local temperature, leading to a heat-island effect (Imhoff *et al.* 2010). Human settlements and irrigated agricultural lands can change a river by their increased usage of the water and therefore water extraction (Vitousek *et al.* 1997).

One study (Liu *et al.* 2017) found that agricultural land decreases run-off also due to a usually ploughed and therefore loose top-soil, which increases filtration of the water into the soil and prevents run-off; another one (Fohrer *et al.* 2001) found the opposite – directly after harvesting and sowing, the soil is the most vulnerable due to a lack of vegetation, which leads to more run-off together with soil erosion.

Urbanization on the expense of agricultural land, on the other hand, has been shown to increase low peak run-off in the summer months dramatically after a storm; however, the effect on high peak run-off in winter was negligible. This was explained by a higher evapotranspiration in summer in agricultural land, and drier soils in summer that can take up more water, therefore buffering the flood event (Hundecha and Bárdossy 2004).

When talking about increased surface run-off, the severe effects of erosion and sediment yield cannot be ignored that are directly positively correlated to run-off (Jain and Kothyari 2000; Wang *et al.* 2009; Sabri *et al.* 2017). This can have effects on the depth and health of the soil, and therefore potentially reduce the areas suitable for growing crops and vegetation, but also has an impact on the river flow. The risk of erosion is usually largest in least vegetated areas, and in areas with the biggest slope (Jain and Kothyari 2000; Wang *et al.* 2009; Sabri *et al.* 2017).

7.3.2 Calculating the Precipitation Deficit

When looking at causal relationships, only land use changes in the Upper Don area can have a significant influence on the hydrology of the Tsimlyansk reservoir, as all other areas are downstream of the reservoir or not connected to it at all. The Upper Don catchment basin has gone, similarly to the rest of the region, through continuous urbanization from 1990 to 2010, and showed a fluctuating forest area with a significant deforestation trend at the same time. Agricultural areas have fluctuated with a decrease of bare/harvested land in August, and an increase of cropped land. As several different parameters could influence the water level of the Tsimlyansk reservoir, it is important to compare different hydrological parameters.

The following data has all been acquired from the Global Land Data Assimilation Systems provided by NASA (Rodell *et al.* 2004) that is embedded in Google Earth Engine and can be found under the Image ID “NASA/GLDAS/V20/NOAH/G025/T3H”. While the data was visualized and verified there, the results of the following pages finally come from an offline-calculation after the monthly-averaged data was downloaded from NASA GES DISC and the Giovanni online system (Acker and Leptoukh 2007) to avoid an exceedance of the user memory limit and the computation time. However, an example visualization of the average evapotranspiration of the year 1984 as available in Google Earth Engine can be seen in Figure 66.

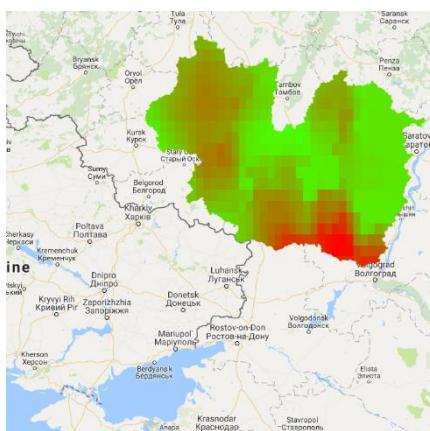


Figure 66: Visualization of the GLDAS Evapotranspiration dataset in GEE. Red = high ET, green = low ET

As GLDAS 2.0, with a time frame from 1948 to 2010, and GLDAS 2.1, from 2010 until today, can hardly be compared due to a different methodology in the processing of the data, only GLDAS-2.0 will be used, as it covers a bigger time frame of what has been addressed in this study. The data is available in monthly time steps and over a 0.25 arc-degrees spatial resolution. The data has been averaged over the whole Upper Don river basin for each monthly time step from January 1984 to December 2010.

The data sets provided by GLDAS, unfortunately, do not calculate overall run-off; run-off data is only available as base flow groundwater runoff, and storm surface runoff. Both of these datasets peak at irregular times over the whole area and are close to zero the rest of the time; analysis in respect to long-term trends is therefore difficult, as the low number of storm events can hardly be analysed and can have multiple factors. Figure 67 shows the storm surface runoff as an example; it is clear that no further trends of run-off can be analysed.

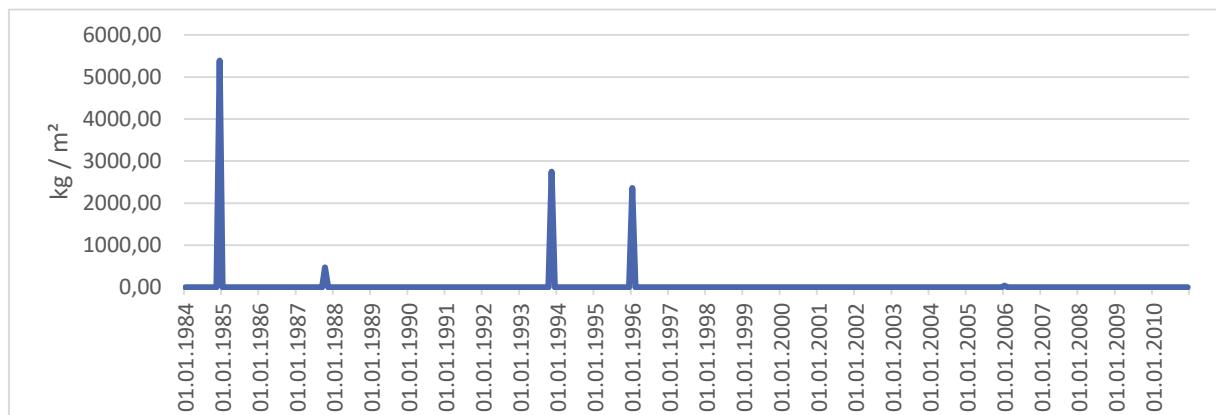


Figure 67: Area-averaged storm surface runoff of the Upper Don river basin. Values are monthly with a 1deg spatial resolution. Data Source: GLDAS, Rondell et al. 2004

More interesting is GLDAS data concerning evapotranspiration and precipitation, as well as the so-called precipitation deficit as used by Harmsen *et al.* (2009) ($PD = P - ET$); hereby, a negative PD indicates an arid environment, while a positive PD actually indicates a precipitation excess that can increase the water stored in rivers and lakes, or that can contribute to the groundwater flow (Harmsen *et al.* 2009). Figure 68 is visualizing the average monthly evapotranspiration and precipitation in the Upper Don river basin from January 1984 to

December 2010. It is clearly visible that the evapotranspiration has always succeeded precipitation in the summer months, and has usually been below it in the winter months.

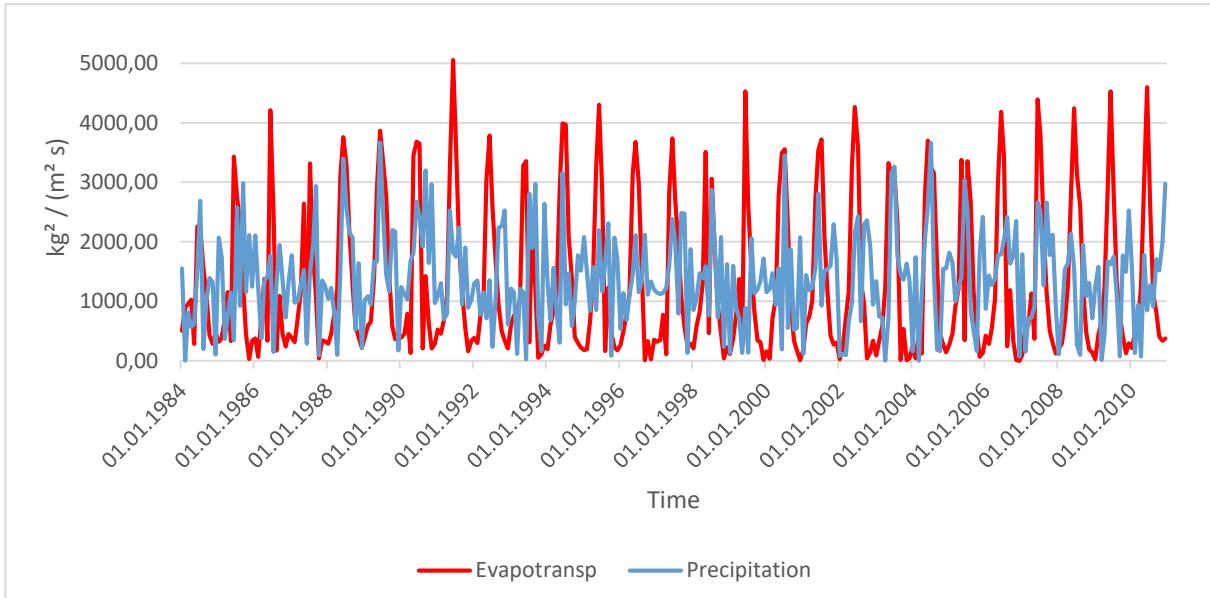


Figure 68: Change of ET and P in the Upper Don river basin over time. Data Source: GLDAS, Rondell et al. 2004

To determine trends over time more clearly, the monthly values were added up in order to visualize the annual precipitation and evapotranspiration rates. Then, the precipitation deficit of the whole area was calculated and visualized in Figure 69. A small downward trend in the PD is visible, driving it towards values below zero. This means that the water that stays in the soil or runs off to lakes and rivers must be decreasing, as the difference between the amounts of water precipitating versus the amount that is evaporating is decreasing.

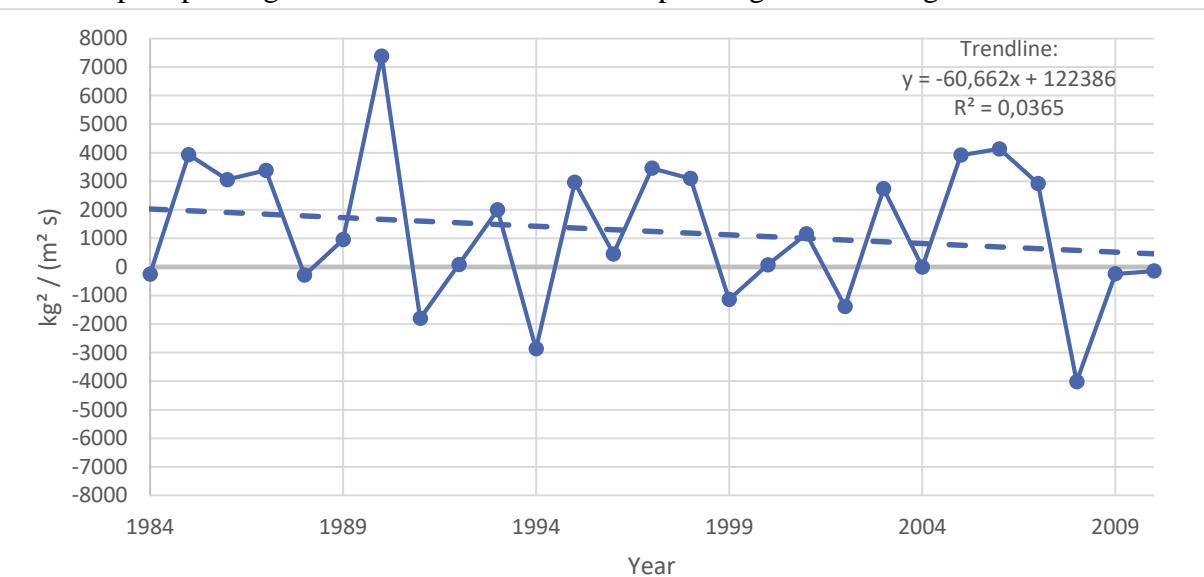


Figure 69: PD of the Upper Don river basin in $\text{kg}^2 \text{ m}^{-2} \text{ s}^{-1}$, from 1984 to 2010, summed up monthly values per year. Data Source: GLDAS; Rondell et al. 2004

In the Upper Don catchment basin, forested land has been clearly and quite constantly decreased until 2010, while urbanized areas have increased. Nevertheless, and contrarily to the majority of the literature, evapotranspiration is increasing – this most likely means that the local climatic changes enhanced by these land use changes override other effects, as the average annual air temperature over the Upper Don basin shows an increasing trend (Figure 70). The heat island effect and a potentially higher albedo could cause the ET to rise despite the lower vegetated area. Additionally, the higher cropped/vegetated agricultural area, as opposed to harvested/bare agricultural area, could also increase evapotranspiration and use up more water under certain circumstances.

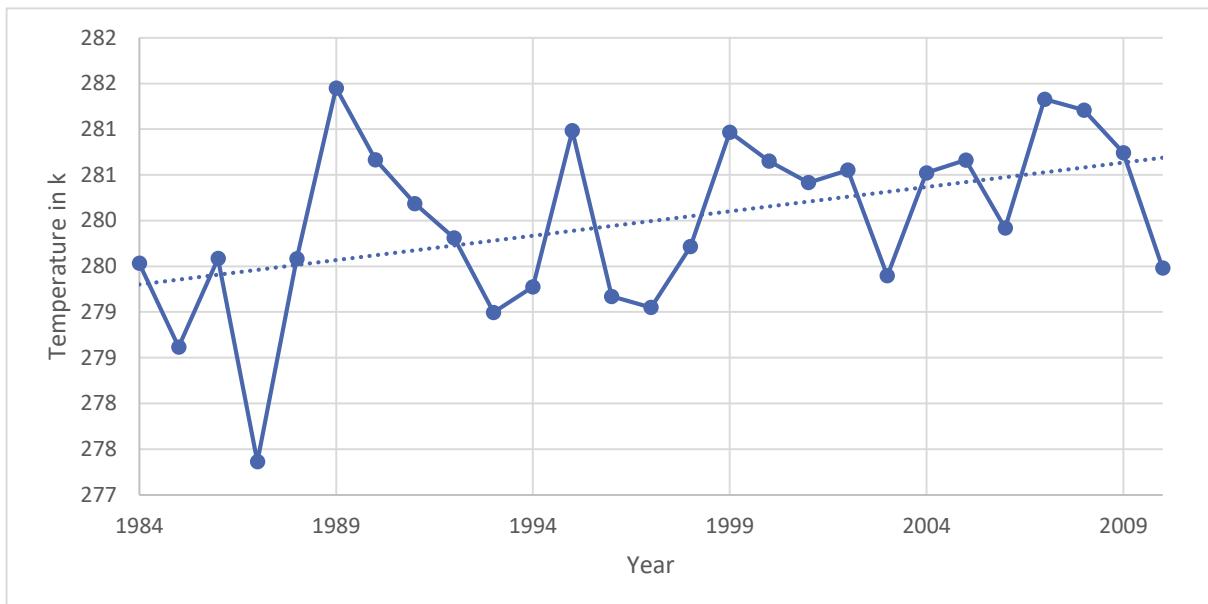


Figure 70: Mean annual air temperature in Kelvin over the Upper Don river basin. Data Source: GLDAS, Rondell et al. 2004

7.3.3 Effects of the Precipitation Deficit on the Tsimlyansk Reservoir

The potential effects of the precipitation, evapotranspiration, and temperature on the Tsimlyansk reservoir are shown in Figure 71 (Theia Land 2018). The graph shows the change in the water level of the Tsimlyansk reservoir, the most important reservoir of the case study region, from 1992 to 2018; unfortunately, no pre-1992 data is available for this region. This data was calculated with remote sensing techniques, and due to cloudiness, the possible error of certain data points, as pictured in red, is quite high at times.

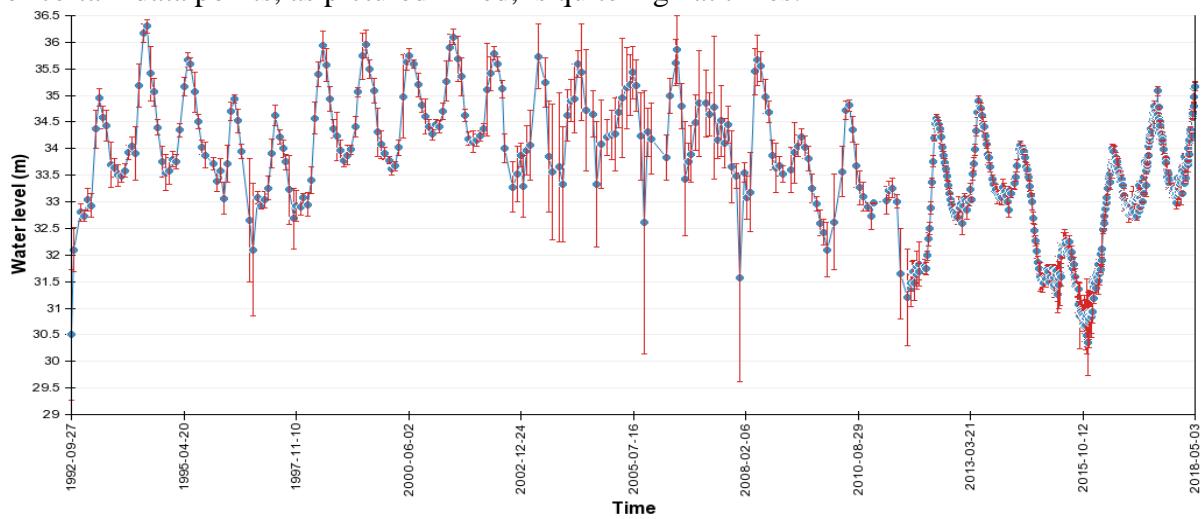


Figure 71: Water level in meters of the Tsimlyansk reservoir from 1992 to 2018. Source: Hydroweb, Theia Land 2018

After calculating the average water level for each calendar year, a general downward trend is visible; a linear trend line has an R^2 value of 0.36 as visible in Figure 72. The years 1992 and 2018 have been kept out of this graph, as the data is not available for the whole calendar year and would be biased due to the seasonal variety of the data sets. The water level change could result from the precipitation decrease (not influenced by the land use changes), as well as from the evapotranspiration and temperature increase. With this visualization, the especially low values of 2011 and 2015 are notable, which could influence and explain the different water extent measured in the land use classification scheme partly.

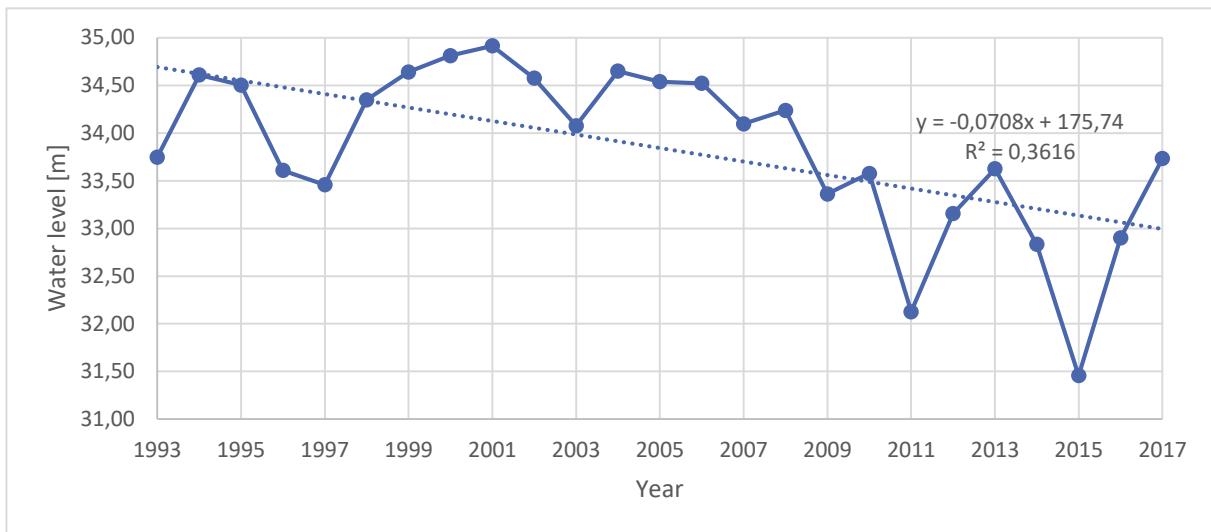


Figure 72: Water level of the Tsimlyansk reservoir, averaged per calendar year. Ten data points have been averaged for each year, and are evenly distributed from January to December. Data Source: Theia Land 2018. This is also in line with a report by the Federal Agency of Water Resources (FAWR) that claims

that the last two decades have been arider than previous ones, which made the Tsimlyansk reservoir shallower with an actual water level of 32 meters from 2010 until 2012 (31 meters is usually considered the absolute minimum) (FAWR 2012). Furthermore, a higher pollution level was recognised (FAWR 2012). Also other papers claim that the increasing temperature in recent years in the region is not met with an equally increasing precipitation, therefore lowering the PD stronger than a change in land use could (Dronin and Kirilenko 2012).

However, it has to be noted here that this data does not necessarily explain the fall in water extent in the whole case study area; the swampy areas directly surrounding the Azov Sea, and the extent of the sea itself, could potentially make a much bigger difference. As explained in previous chapters, potential misclassifications due to eutrophication over time of the reservoir and other water areas in the basin could also add to the loss in perceived water extent; this does not necessarily mean a lower water content generally. Furthermore, fluctuating water extraction from the reservoir, as well as temperature differences over time, could also have an immense influence that can further alter the water level.

Nevertheless, the conclusion can be made that the lower forest area and the increased urbanization of the Upper Don catchment basin did not have the expected effect of a higher

runoff and therefore, a higher general water level. More likely, the climatic variables offset any land use change effects, and the higher albedo and the heat-island effect of urbanized land furthermore increased the temperature and evapotranspiration rates. All of this most likely had an effect on the water level of the Tsimlyansk reservoir, which tends to dry out. If this trend continues in the same manner, the reservoir will most likely not be able to support all the ecosystem services it provides today. Currently, over 450 thousand inhabitants are using the freshwater provided by the reservoir, and more than 150 thousand hectares of agricultural land are irrigated with water from the reservoir (Sharvak *et al.* 2012). Beyond that, it is supporting the energy production of the region, and is used to cool the Rostov nuclear power plant (Sharvak *et al.* 2012). All of these goods and services are not yet significantly reduced, however, if trends continue as they did in the past thirty years, a decrease in the provision of EGS cannot be ruled out.

8 Conclusion

In respect of the first research question, it was found that GEE is a useful tool that should be applied more by researchers of all levels and all institutions that are willing to invest the time to learn the specific server-sided version of JavaScript. Despite some issues of frequent computation time outs and a sometimes insufficient maximum memory limit, Earth Engine has the potential to improve remote sensing analysis and make it more efficient and accessible for low-resource research projects. It is an excellent tool that takes away many constraints that junior researchers otherwise face when it comes to computational power, memory space, and time, and is in that sense a true game changer. It has been shown that the results obtained by this tool are of satisfactory accuracy and are in line with real directly observed events, such as the Donbas conflict. By incorporating many different data sets by default, and by giving researchers the opportunity to save results as assets for new scripts or upload their own data sets as assets, results for many different fields can be achieved and combined. This has the potential to make spatially explicit research generally easier to carry out.

With this tool, the second research question was answered, and the changes the Azov Sea basin has gone through in the past 30 years have successfully been detected. A consistent change has been the urbanization and industrialization of the region – the urban percentage has more or less steadily increased from 1985 until 2015 at the expense of natural vegetation. While agriculture has fluctuated a lot and has especially decreased up until the year 2000, there is no clear trend recognisable. In particular, the fall of the Soviet Union has decreased the average agricultural yield per hectare, which led to the partial abandonment of agricultural land; however, just as much land has been converted back to agriculture at some point of the time frame, or has been converted to agricultural land from other land use types. The Azov Sea basin is most likely less affected by land abandonment than other parts of the former Soviet Union, as it has the highest yields, most fertile soils, and a relatively mild climate as opposed to other

areas. As the transition to a market-based economy caused only the most efficient farms to survive, the Azov Sea basin was less affected than other areas.

The ongoing Ukrainian crisis has and is still influencing the area of conflict in the Donbas region immensely. Thus, urban areas, as well as agricultural areas, have been decreasing dramatically since 2014, while the forest cover/natural vegetation increased due to the abandonment of the area by large parts of the population.

The further urbanization of the Upper Don area could have added to a change in the local climate; the increasing temperature is not met with an increase in precipitation, which leads to a potential future precipitation deficit. The Tsimlyansk reservoir has been observed to get drier and more polluted in recent years, which could have a negative effect on the ecosystem services the reservoir provides. As such, the LULCC in the Upper Don area have been shown to have a measurable effect on the provision of ecosystem services and, therefore, human well-being.

The geospatial databases developed during the making of this study are all available at the CEU Environmental Systems Laboratory² for further investigation.

² <https://syslab.ceu.edu/projects/GEE-land-use-azov-basin>

Bibliography

- Acker, J.G., and Leptoukh, G. 2007. Online Analysis Enhances Use of NASA Earth Science Data. *Eos, Transactions of the American Geophysical Union* Vol. 88 (No. 2 (9 January 2007)): 14–17.
- Aitkenhead, M.J., and Aalders, I.H. 2011. Automating Land Cover Mapping of Scotland Using Expert System and Knowledge Integration Methods. *Remote Sensing of Environment* 115 (5): 1285–95. doi:10.1016/j.rse.2011.01.012.
- Allan, J.D. 2004. Landscapes and Riverscapes: The Influence of Land Use on Stream Ecosystems. *Annu. Rev. Ecol. Evol. Syst.* 35: 257–284.
- Anaba, L.A., Banadda, N., Kiggundu, N., Wanyama, J., Engel, B., and Moriasi, D. 2017. Application of SWAT to Assess the Effects of Land Use Change in the Murchison Bay Catchment in Uganda. *Computational Water, Energy, and Environmental Engineering* 06 (01): 24–40. doi:10.4236/cweee.2017.61003.
- Anderson, J.R., Hardy, E.E., Roach, J.T., and Witmer, R.E. 1983. *A Land Use and Land Cover Classification System for Use with Remote Sensor Data*. Geological Survey Professional Paper 964. Washington DC. <https://pubs.usgs.gov/pp/0964/report.pdf>.
- Aylward, B., Bandyopadhyay, J., Belausteguigotia, J.-C., Borkey, P., Cassar, A.Z., Meadors, L., Saade, L., Siebentritt, M., Stein, R., and Tognetti, S. 2005. Freshwater Ecosystem Services. *Ecosystems and Human Well-Being: Policy Responses* 3: 213–256.
- Barbier, E.B., Burgess, J.C., and Grainger, A. 2010. The Forest Transition: Towards a More Comprehensive Theoretical Framework. *Land Use Policy*, Forest transitions, 27 (2): 98–107. doi:10.1016/j.landusepol.2009.02.001.
- Barsi, J.A., Lee, K., Kvaran, G., Markham, B.L., and Pedelty, J.A. 2014. The Spectral Response of the Landsat-8 Operational Land Imager. *Remote Sensing* 6: 10232–51. doi:doi:10.3390/rs61010232.
- Baumann, M., Kuemmerle, T., Elbakidze, M., Ozdogan, M., Radloff, V.C., Keuler, N.S., Prishchepov, A.V., Kruhlov, I., and Hostert, P. 2011. Patterns and Drivers of Post-Socialist Farmland Abandonment in Western Ukraine. *Land Use Policy* 28 (3): 552–62. doi:10.1016/j.landusepol.2010.11.003.
- Bekturganov, N.S., and Bolaev, A.V. 2017. The Eurasia Canal as a Factor of Economic Prosperity for the Caspian Region. *GEOGRAPHY, ENVIRONMENT, SUSTAINABILITY* 10 (1): 34–43. doi:10.24057/2071-9388-2017-10-1-34-43.
- Benini, L., Bandini, V., Marazza, D., and Contin, A. 2010. Assessment of Land Use Changes through an Indicator-Based Approach: A Case Study from the Lamone River Basin in Northern Italy. *Ecological Indicators, Landscape Assessment for Sustainable Planning*, 10 (1): 4–14. doi:10.1016/j.ecolind.2009.03.016.
- Bobylev, S., and Libert, B. 1994. Prospects For Agricultural and Environmental Policy Integration in Russia. In *Agriculture and the Environment in the Transition to a Market Economy*, edited by OECD. Vol. 4. OECD Documents. Paris.
- Brutsaert, W. 2005. *Hydrology: An Introduction*. 3rd ed. New York: Cambridge University Press.
- Buckley, C., Clem, R., Fox, J., and Herron, E. 2018. Analysis | The War in Ukraine Is More Devastating than You Know. *Washington Post*, April 9, sec. <https://www.washingtonpost.com/news/monkey-cage/wp/2018/04/09/the-war-in-ukraine-is-more-devastating-than-you-know/>. Accessed May 10, 2018
- Butler, K. 2013. Band Combinations for Landsat 8. *Esri ArcGIS Blog*. <https://www.esri.com/arcgis-blog/products/product/imagery/band-combinations-for-landsat-8/>. Accessed May 26, 2018

- Chen, J., Chen, J., Liao, A., Cao, X., Chen, L., Chen, X., He, C., Han, G., Peng, S., Lu, M., Zhang, W., Tong, X., and Mills, J. 2015. Global Land Cover Mapping at 30m Resolution: A POK-Based Operational Approach. *ISPRS Journal of Photogrammetry and Remote Sensing* 103 (May): 7–27. doi:10.1016/j.isprsjprs.2014.09.002.
- Center for International Earth Science Information Network (CIESIN). 2011. *Foresight Project on Migration and Global Environmental Change, Report MR4: Estimating Net Migration by Ecosystem and by Decade, 1970-2010*. Center for International Earth Science Information Network - CIESIN - Columbia University. London: UK Government Foresight.
- . 2017a. Global Population Count Grid Time Series Estimates. NASA Socioeconomic Data and Applications Center (SEDAC) & Center for International Earth Science Information Network - CIESIN - Columbia University. <https://doi.org/10.7927/H4CC0XNV>.
- . 2017b. Gridded Population of the World, Version 4 (GPWv4): Population Count Adjusted to Match 2015 Revision of UN WPP Country Totals, Revision 10. NASA Socioeconomic Data and Applications Center (SEDAC) & Center for International Earth Science Information Network (CIESIN) Columbia University. <https://doi.org/10.7927/H4JQ0XZW>.
- Coman, J. 2017. On the Frontline of Europe's Forgotten War in Ukraine. *The Guardian*, November 12, sec. World news. <http://www.theguardian.com/world/2017/nov/12/ukraine-on-the-front-line-of-europes-forgotten-war>. Accessed May 8, 2018
- Cosgrove, W.J., and Rijsberman, F.R. 2000. Chapter Two: The Use of Water Today. In *World Water Vision: Making Water Everybody's Business*, edited by World Water Council. London: Earthscan Publications Ltd. <http://www.worldwatercouncil.org/fileadmin/wwc/Library/WWVision/Chapter2.pdf>.
- Dai, L., Vorselen, D., Korolev, K.S., and Gore, J. 2012. Generic Indicators for Loss of Resilience Before a Tipping Point Leading to Population Collapse. *Science* 336 (6085): 1175–77. doi:10.1126/science.1219805.
- Daily, G.C., and Matson, P.A. 2008. Ecosystem Services: From Theory to Implementation. *Proceedings of the National Academy of Sciences* 105 (28): 2.
- DeFries, R., and Eshleman, K.N. 2004. Land-Use Change and Hydrologic Processes: A Major Focus for the Future. *Hydrological Processes* 18 (11): 2183–86. doi:10.1002/hyp.5584.
- Delanoe, I. 2014. After the Crimean Crisis: Towards a Greater Russian Maritime Power in the Black Sea. *Southeast European and Black Sea Studies* 14 (3): 367–82. doi:10.1080/14683857.2014.944386.
- Deppermann, A., Balkovič, J., Bundle, S.-C., Di Fulvio, F., Havlik, P., Leclère, D., Lesiv, M., Prishchepov, A.V., and Schepaschenko, D. 2018. Increasing Crop Production in Russia and Ukraine—regional and Global Impacts from Intensification and Recultivation. *Environmental Research Letters* 13 (2): 025008. doi:10.1088/1748-9326/aaa4a4.
- Dronin, N., and Kirilenko, A. 2012. Climate Change, Water and Agriculture in the Azov Sea Basin. In *Environmental Security in Watersheds: The Sea of Azov*, edited by Viktor Lagutov. NATO Science for Peace and Security Series C: Environmental Security. Dordrecht: Springer.
- Dubinina, V.G., and Kozlitina, S.V. 2000. Water Resources Management of the Southern Rivers of Russia with Reference to Fisheries Requirements. *Fisheries Management and Ecology* 7 (1-2): 157–65. doi:10.1046/j.1365-2400.2000.00190.x.
- Dupliak, V. 1994. Rational Water Use for Irrigation Purposes in Ukraine. In *Agriculture and the Environment in the Transition to a Market Economy*, edited by OECD. OECD Documents.

- Dwarakish, G.S., and Ganasri, B.P. 2015. Impact of Land Use Change on Hydrological Systems: A Review of Current Modeling Approaches. Edited by Lucia De Stefano. *Cogent Geoscience* 1 (1). doi:10.1080/23312041.2015.1115691.
- EnviroGRIDS. 2012. <http://www.envirogrids.net/index.html>. Accessed May 20, 2018.
- European Environment Agency (EEA). 2016. Figure 1: The DPSIR Framework. Page. European Environment Agency. <https://www.eea.europa.eu/publications/92-9167-059-6-sum/page002.html>. Accessed May 16, 2018.
- Food and Agriculture Organization (FAO). 2010. FAO - News Article: The World's next Breadbasket. *Food and Agriculture Organization of the United Nations*. <http://www.fao.org/news/story/en/item/46401/icode/>. Accessed May 7, 2018.
- Federal Agency of Water Resources (FAWR). 2012. Правила Использования Водных Ресурсов Цимлянского Водохранилища [Terms of Use of Water Resources of Tsimlyansk Reservoir]. Federal Agency of Water Resources, Russia.
- Fohrer, N., Haverkamp, S., Eckhardt, K., and Frede, H.-G. 2001. Hydrologic Response to Land Use Changes on the Catchment Scale. *Physics and Chemistry of the Earth, Part B: Hydrology, Oceans and Atmosphere* 26 (7–8): 577–82. doi:10.1016/S1464-1909(01)00052-1.
- Gibbons-Neff, T. 2016. As Focus Remains on Syria, Ukraine Sees Heaviest Fighting in Months. *Washington Post*, February 16, sec. Checkpoint. <https://www.washingtonpost.com/news/checkpoint/wp/2016/02/16/as-focus-remains-on-syria-ukraine-sees-heaviest-fighting-in-months/>. Accessed May 24, 2018.
- Goldman, E., Harris, N., and Maschler, T. 2015. *Predicting Future Forest Loss in the Democratic Republic of the Congo's CARPE Landscapes*. Washington DC: World Resources Institute.
- Google. 2018a. Google Earth Engine Code Editor. <https://code.earthengine.google.com/>. Accessed March 7, 2018.
- Google. 2018b. Introduction | Google Earth Engine API. *Google Developers*. <https://developers.google.com/earth-engine/>. Accessed April 5, 2018
- Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., and Moore, R. 2017. Google Earth Engine: Planetary-Scale Geospatial Analysis for Everyone. *Remote Sensing of Environment* 202 (December): 18–27. doi:10.1016/j.rse.2017.06.031.
- Gray, R. 2017. Russian Federation Grain and Feed Annual Annual 2017. *USDA Foreign Agricultural Service - Global Agricultural Information Network*, no. RS 1726 (April): 34.
- Guo, Z., Xiao, X., and Li, D. 2000. An Assessment of Ecosystem Services: Water Flow Regulation and Hydroelectric Power Production. *Ecological Applications* 10 (3): 925. doi:10.2307/2641057.
- Gutman, G., Janetos, A.C., Justice, C.O., Moran, E.F., Mustard, J.F., Rindfuss, R.R., Skole, D., II, B.L.T., and Cochrane, M.A. 2004. *Land Change Science: Observing, Monitoring and Understanding Trajectories of Change on the Earth's Surface*. Springer Science & Business Media.
- Hägerstrand, T. 2001. A Look at the Political Geography of Environmental Management. In *Sustainable Landscapes and Lifeways: Scale and Appropriateness*, edited by Anne Buttmer. Cork, Ireland: Cork University Press.
- Hansen, M.C., Potapov, P.V., Moore, R., Hancher, M., Turubanova, S.A., Tyukavina, A., Thau, D., Stehman, S.V., Goetz, S.J., Loveland, T.R., Kommareddy, A., Egorov, A., Chini, L., Justice, C.O., and Townshend, J.R. 2013. High-Resolution Global Maps of 21st-Century Forest Cover Change. *Science* 342 (6160): 850–53. doi:10.1126/science.1244693.

- Harmsen, E.W., Miller, N.L., Schlegel, N.J., and Gonzalez, J.E. 2009. Seasonal Climate Change Impacts on Evapotranspiration, Precipitation Deficit and Crop Yield in Puerto Rico. *Agricultural Water Management* 96 (7): 1085–95. doi:10.1016/j.agwat.2009.02.006.
- Havlik, P., and Astrov, V. 2014. Economic Consequences of the Ukraine Conflict. The Vienna Institute for International Economic Studies (wiiw).
- Hostert, P., Kuemmerle, T., Prishchepov, A., Sieber, A., Lambin, E.F., and Radeloff, V.C. 2011. Rapid Land Use Change after Socio-Economic Disturbances: The Collapse of the Soviet Union versus Chernobyl. *Environmental Research Letters* 6 (4): 045201. doi:10.1088/1748-9326/6/4/045201.
- Huang, K., Yi, C., Wu, D., Zhou, T., Zhao, X., Blanford, W.J., Wei, S., Wu, H., Du Ling, and Li, Z. 2015. Tipping Point of a Conifer Forest Ecosystem under Severe Drought. *Environmental Research Letters* 10 (2): 024011. doi:10.1088/1748-9326/10/2/024011.
- Hundecha, Y., and Bárdossy, A. 2004. Modeling of the Effect of Land Use Changes on the Runoff Generation of a River Basin through Parameter Regionalization of a Watershed Model. *Journal of Hydrology* 292 (1–4): 281–95. doi:10.1016/j.jhydrol.2004.01.002.
- Imhoff, M.L., Zhang, P., Wolfe, R.E., and Bounoua, L. 2010. Remote Sensing of the Urban Heat Island Effect across Biomes in the Continental USA. *Remote Sensing of Environment* 114 (3): 504–13. doi:10.1016/j.rse.2009.10.008.
- Ioffe, G. 2005. The Downsizing of Russian Agriculture. *Europe-Asia Studies* 57 (2): 179–208. doi:10.1080/09668130500051627.
- Ioffe, G., and Nefedova, T. 2000. Areas of Crisis in Russian Agriculture: A Geographic Perspective, 19.
- Ioffe, G., Nefedova, T., and de Beurs, K. 2014. Agrarian Transformation in the Russian Breadbasket: Contemporary Trends as Manifest in Stavropol'. *Post-Soviet Affairs* 30 (6): 441–63. doi:10.1080/1060586X.2013.858509.
- Ioffe, G., Nefedova, T., and Zaslavsky, I. 2004. From Spatial Continuity to Fragmentation: The Case of Russian Farming. *Annals of the Association of American Geographers* 94 (4): 913–43. doi:10.1111/j.1467-8306.2004.00441.x.
- Jain, M.K., and Kothyari, U.C. 2000. Estimation of Soil Erosion and Sediment Yield Using GIS. *Hydrological Sciences Journal* 45 (5): 771–86. doi:10.1080/02626660009492376.
- Jenkins, A., Peters, N.E., and Rodhe, A. 1994. Hydrology. In *Biogeochemistry of Small Catchments: A Tool for Environmental Research*, edited by Bedrich Moldan and Jiri Cerny. Vol. 51. SCOPE Report - Scientific Committee on Problems of the Environment. West Sussex: Wiley.
- Jetz, W., Thau, D., and Guralnick, R. 2015. Map of Life: A Preview of How to Evaluate Species Conservation with Google Earth Engine. *Google AI Blog*. <http://ai.googleblog.com/2015/01/map-of-life-preview-of-how-to-evaluate.html>. Accessed April 26, 2018
- Katchanovski, I. 2016. The Separatist War in Donbas: A Violent Break-up of Ukraine? *European Politics and Society* 17 (4): 473–89. doi:10.1080/23745118.2016.1154131.
- Kates, R.W., Clark, W.C., Corell, R., Hall, J.M., Jaeger, C.C., Lowe, I., McCarthy, J.J., Schellnhuber, H.J., Bolin, B., Dickson, N.M., Faucheur, S., Gallopin, G.C., Grübler, A., Huntley, B., Jäger, J., Jodha, N.S., Kasperson, R.E., Mabogunje, A., Matson, P., Mooney, H., Moore, B., O'Riordan, T., and Svedin, U. 2001. Sustainability Science. *Science* 292 (5517): 641–42. doi:10.1126/science.1059386.
- Keeler, B.L., Polasky, S., Brauman, K.A., Johnson, K.A., Finlay, J.C., O'Neill, A., Kovacs, K., and Dalzell, B. 2012. Linking Water Quality and Well-Being for Improved Assessment and Valuation of Ecosystem Services. *Proceedings of the National Academy of Sciences* 109 (45): 18619–24. doi:10.1073/pnas.1215991109.

- Kramer, A.E. 2017. Truce Unravels as Fighting Picks Up in Ukraine. *The New York Times*, December 21, sec. World.
<https://www.nytimes.com/2016/02/22/world/europe/fighting-in-ukraine-picks-up-sharply.html>. Accessed April 23, 2018.
- Kresic, N., and Mikszewski, A. 2013. *Hydrogeological Conceptual Site Models - Data Analysis and Visualization*. Boca Raton: CRC Press - Taylor & Francis Group.
- Kurakov, Y.I., Samofalov, V.S., Malikov, I.N., and Kolomiets, V.A. 2010. Coal Mining in the Russian Donetsk Basin. *Coke and Chemistry* 53 (4): 121–23. doi:10.3103/S1068364X10040010.
- Kurtzman, L. 2014. UCSF, Google Earth Engine Making Maps to Predict Malaria. *UC San Francisco News Center*. <https://www.ucsf.edu/news/2014/09/116906/ucsf-google-earth-engine-making-maps-predict-malaria>. Accessed May 4, 2018.
- Kuzio, T. 2015. Competing Nationalisms, Euromaidan, and the Russian-Ukrainian Conflict. *Studies in Ethnicity and Nationalism* 15 (1): 157–69. doi:10.1111/sena.12137.
- Lagutov, V., and Lagutov, V. 2012. The Azov Ecosystem: Resources and Threats. In *Environmental Security in Watersheds: The Sea of Azov*, edited by Viktor Lagutov. NATO Science for Peace and Security Series C: Environmental Security. Dordrecht: Springer.
- Lambin, E.F., and Geist, H.J. 2008. *Land-Use and Land-Cover Change: Local Processes and Global Impacts*. Springer Science & Business Media.
- Lambin, E.F., and Meyfroidt, P. 2010. Land Use Transitions: Socio-Ecological Feedback versus Socio-Economic Change. *Land Use Policy*, Forest transitions, 27 (2): 108–18. doi:10.1016/j.landusepol.2009.09.003.
- . 2011. Global Land Use Change, Economic Globalization, and the Looming Land Scarcity. *Proceedings of the National Academy of Sciences* 108 (9): 3465–72. doi:10.1073/pnas.1100480108.
- Lambin, E.F., Turner, B.L., Geist, H.J., Agbola, S.B., Angelsen, A., Bruce, J.W., Coomes, O.T., Dirzo, R., Fischer, G., Folke, C., George, P.S., Homewood, K., Imbernon, J., Leemans, R., Li, X., Moran, E. F., Mortimore, M., Ramakrishnan, P.S., Richards, J.F., Skanes, H., Steffen, W., Stone, G.D., Svedin, U., Veldkamp, T.A., Vogel, C., and Xu, J. 2001. The Causes of Land-Use and Land-Cover Change: Moving beyond the Myths. *Global Environmental Change* 11 (4): 261–69. doi:10.1016/S0959-3780(01)00007-3.
- Lange, B. 2016. *Brian Lange | It's Not Magic: Explaining Classification Algorithms*. <https://www.youtube.com/watch?v=y8J6ggsLSfw>. Accessed April 6, 2018.
- Lary, D.J., Alavi, A.H., Gandomi, A.H., and Walker, A.L. 2016. Machine Learning in Geosciences and Remote Sensing. *Geoscience Frontiers* 7 (1): 3–10. doi:10.1016/j.gsf.2015.07.003.
- Liu, J., Zhang, C., Kou, L., and Zhou, Q. 2017. Effects of Climate and Land Use Changes on Water Resources in the Taoer River. Research article. *Advances in Meteorology*. doi:10.1155/2017/1031854.
- Lu, D., and Weng, Q. 2007. A Survey of Image Classification Methods and Techniques for Improving Classification Performance. *International Journal of Remote Sensing* 28 (5): 823–70. doi:10.1080/01431160600746456.
- Magliocca, N.R., Rudel, T.K., Verburg, P.H., McConnell, W.J., Mertz, O., Gerstner, K., Heinemann, A., and Ellis, E.C. 2015. Synthesis in Land Change Science: Methodological Patterns, Challenges, and Guidelines. *Regional Environmental Change* 15 (2): 211–26. doi:10.1007/s10113-014-0626-8.
- Mather, A.S. 1990. *Global forest resources*. London: Belhaven.
- Meyer, W.B., and Turner, B.L. 1992. Human Population Growth and Global Land-Use/Cover Change. *Annual Review of Ecology and Systematics* 23: 39–61.

- Moore, R.T., and Hansen, M.C. 2011. Google Earth Engine: A New Cloud-Computing Platform for Global-Scale Earth Observation Data and Analysis. *AGU Fall Meeting Abstracts* 43 (December): IN43C-02.
- Munroe, D.K., McSweeney, K., Olson, J.L., and Mansfield, B. 2014. Using Economic Geography to Reinvigorate Land-Change Science. *Geoforum* 52 (March): 12–21. doi:10.1016/j.geoforum.2013.12.005.
- National Aeronautics and Space Administration (NASA). 2018. The Water Cycle | Precipitation Education. <https://pmm.nasa.gov/education/water-cycle>. Accessed May 24.
- Navarro, J.A. 2017. First Experiences with Google Earth Engine: In , 250–55. SCITEPRESS - Science and Technology Publications. doi:10.5220/0006352702500255.
- Ness, B., Anderberg, S., and Olsson, L. 2010. Structuring Problems in Sustainability Science: The Multi-Level DPSIR Framework. *Geoforum* 41 (3): 479–88. doi:10.1016/j.geoforum.2009.12.005.
- Nielsen, M.A. 2015. Neural Networks and Deep Learning. <http://neuralnetworksanddeeplearning.com>. Accessed April 4, 2018.
- Office for the Coordination of Humanitarian Affairs (OCHA) Ukraine. 2017. Ukraine Administrative Level 0, 1, 2, 3 and 4 Boundaries. *Humanitarian Data Exchange*. <https://data.humdata.org/dataset/ukraine-administrative-boundaries-as-of-q2-2017>.
- Organisation for Economic Co-Operation and Development (OECD). 2003. *OECD Environmental Indicators. Development, Meausrement and Use*. <https://www.oecd.org/env/indicators-modelling-outlooks/24993546.pdf>.
- Pallot, J., and Nefedova, T. 2003. Geographical Differentiation in Household Plot Production in Rural Russia. *Eurasian Geography and Economics* 44 (1): 40–64. doi:10.2747/1538-7216.44.1.40.
- Pekel, J.-F., Cottam, A., Gorelick, N., and Belward, A.S. 2016. High-Resolution Mapping of Global Surface Water and Its Long-Term Changes. *Nature* 540 (7633): 418–22. doi:10.1038/nature20584.
- Pettorelli, N., Vik, J.O., Mysterud, A., Gaillard, J.-M., Tucker, C.J., and Stenseth, N.C. 2005. Using the Satellite-Derived NDVI to Assess Ecological Responses to Environmental Change. *Trends in Ecology & Evolution* 20 (9): 503–10. doi:10.1016/j.tree.2005.05.011.
- Pielke, R.A. 2005. Land Use and Climate Change. *Science* 310 (5754): 1625–26. doi:10.1126/science.1120529.
- Potschin, M. 2009. Land Use and the State of the Natural Environment. *Land Use Policy, Land Use Futures*, 26 (December): S170–77. doi:10.1016/j.landusepol.2009.08.008.
- Prishchepov, A.V., Müller, D., Dubinin, M., Baumann, M., and Radeloff, V.C. 2013. Determinants of Agricultural Land Abandonment in Post-Soviet European Russia. *Land Use Policy* 30 (1): 873–84. doi:10.1016/j.landusepol.2012.06.011.
- Privalov, V.A., Sachsenhofer, R.F., Panova, E.A., and Antsiferov, V.A. 2004. Coal Geology of the Donets Basin (Ukraine/Russia): An Overview. *BHM*, no. 149. Jg. (2004), Heft 6: 212–22.
- Rashleigh, B., Lagutov, V., and Salathe, T. 2012. Ecosystem Services of Rivers: The Don River (Russian Federation) and the Roanoke River (USA). In *Environmental Security in Watersheds: The Sea of Azov*, ed. Viktor Lagutov. NATO Science for Peace and Security Series C: Environmental Security. Dordrecht: Springer.
- Richards, J.A., and Jia, X. 2013. *Remote Sensing Digital Image Analysis*. Fifth Edition. Berlin, Heidelberg: Springer Berlin Heidelberg. doi:10.1007/978-3-662-03978-6.
- Rindfuss, R.R., Walsh, S.J., Turner, B.L., Fox, J., and Mishra, V. 2004. Developing a Science of Land Change: Challenges and Methodological Issues. *Proceedings of the National Academy of Sciences* 101 (39): 13976–81. doi:10.1073/pnas.0401545101.

- Rodell, M., Houser, P.R., Jambor, U., Gottschalck, J., Mitchell, K., Meng, C.-J., Arsenault, K., Cosgrove, B., Radakovich, J., Bosilovich, M., Entin, J. K., Walker, J.P., Lohmann, D., and Toll, D. 2004. The Global Land Data Assimilation System. *Bulletin of the American Meteorological Society* 85 (3): 381–94. doi:10.1175/BAMS-85-3-381.
- Sabri, E.M., Boukdir, A., Meslouhi, R.E., Mabrouki, M., Mahboul, A.E., Mbaki, V.R.E., Zitouni, A., Baite, W., and Echakraoui, Z. 2017. Predicting Soil Erosion and Sediment Yield in Oued El Abid Watershed, Morocco 4 (3): 18.
- Sala, O.E., Chapin, F.S., Iii, Armesto, J.J., Berlow, E., Bloomfield, J., Dirzo, R., Huber-Sanwald, E., Huenneke, L.F., Jackson, R.B., Kinzig, A., Leemans, R., Lodge, D.M., Mooney, H.A., Oesterheld, M., Poff, N.L., Sykes, M.T., Walker, B.H., Walker, M., and Wall, D.H. 2000. Global Biodiversity Scenarios for the Year 2100. *Science* 287 (5459): 1770–74. doi:10.1126/science.287.5459.1770.
- Sedik, D., Trueblood, M., and Arnade, C. 1999. Corporate Farm Performance in Russia, 1991–1995: An Efficiency Analysis. *Journal of Comparative Economics* 27 (3): 514–33. doi:10.1006/jcec.1999.1599.
- Sharvak, E., Fesenko, L., Generalenko, I., and Ilyashenko, N. 2012. Multipurpose Water Scheme Assessment of the Tsimlyansk Reservoir. In *Environmental Security in Watersheds: The Sea of Azov*, edited by Viktor Lagutov. NATO Science for Peace and Security Series C: Environmental Security. Dordrecht: Springer.
- Smeets, E., and Weterings, R. 1999. Environmental Indicators: Typology and Overview, 19. Technical Report No. 25. Copenhagen: European Environment Agency.
- Sriwongsitanon, N., and Taesombat, W. 2011. Effects of Land Cover on Runoff Coefficient. *Journal of Hydrology* 410 (3–4): 226–38. doi:10.1016/j.jhydrol.2011.09.021.
- Steffen, W., Jäger, J., Carson, D.J., and Bradshaw, C. 2002. *Challenges of a Changing Earth: Proceedings of the Global Change Open Science Conference, Amsterdam, The Netherlands, 10–13 July 2001*. Springer Science & Business Media.
- Sun, G., Amatya, D.M., and McNulty, S.G. 2016. Forest Hydrology. In *Handbook of Applied Hydrology*, edited by V.P. Singh, 2nd Edition:8. McGraw-Hill Education.
- Svarstad, H., Petersen, L.K., Rothman, D., Siepel, H., and Wätzold, F. 2008. Discursive Biases of the Environmental Research Framework DPSIR. *Land Use Policy* 25 (1): 116–25. doi:10.1016/j.landusepol.2007.03.005.
- The Economist. 2016. Russia Is Negotiating with Germany and France over Ukraine. *The Economist*, October 22. <https://www.economist.com/news/europe/21709066-while-foreign-powers-argue-about-its-future-ukrainians-struggle-take-their-destiny>. Accessed May 10, 2018.
- Theia Land. 2018. Hydroweb. Database. *Time Series of Water Levels in the Rivers and Lakes around the World*. <http://hydroweb.theia-land.fr/hydroweb/view/6c6792a3-6a85-50c1-8296-c01a820a966b?lang=en&basin=DON>. Accessed May 8, 2018.
- Trofimov, I.A., Trofimova, L.S., and Yakovleva, E.P. 2014. Agrolandscapes of the Central Chernozem Region. *Biology Bulletin* 41 (10): 901–6. doi:10.1134/S1062359014100100.
- Trueblood, M.A., and Arnade, C. 2001. Crop Yield Convergence: How Russia's Yield Performance Has Compared to Global Yield Leaders. *Comparative Economic Studies (Association for Comparative Economic Studies)* 43 (2): 59.
- Tscherning, K., Helming, K., Krippner, B., Sieber, S., and Paloma, S.G. y. 2012. Does Research Applying the DPSIR Framework Support Decision Making? *Land Use Policy* 29 (1): 102–10. doi:10.1016/j.landusepol.2011.05.009.
- Turner, B.L. 2002. Toward Integrated Land-Change Science: Advances in 1.5 Decades of Sustained International Research on Land-Use and Land-Cover Change. In *Challenges*

- of a Changing Earth*, 21–26. Global Change — The IGBP Series. Springer, Berlin, Heidelberg. doi:10.1007/978-3-642-19016-2_3.
- Turner, B.L., Lambin, E.F., and Reenberg, A. 2007. The Emergence of Land Change Science for Global Environmental Change and Sustainability. *Proceedings of the National Academy of Sciences* 104 (52): 20666–71. doi:10.1073/pnas.0704119104.
- Turner, B.L., and Robbins, P. 2008. Land-Change Science and Political Ecology: Similarities, Differences, and Implications for Sustainability Science. *Annual Review of Environment and Resources* 33 (1): 295–316. doi:10.1146/annurev.energy.33.022207.104943.
- United States Geological Survey (USGS). 2018. Landsat 7 | Landsat Missions. <https://landsat.usgs.gov/landsat-7>. Accessed May 15, 2018.
- Veraart, A.J., Faassen, E.J., Dakos, V., van Nes, E.H., Lürling, M., and Scheffer, M. 2012. Recovery Rates Reflect Distance to a Tipping Point in a Living System. *Nature* 481 (7381): 357–59. doi:10.1038/nature10723.
- Vitousek, P.M., Mooney, H.A., Lubchenco, J., and Melillo, J.M. 1997. Human Domination of Earth's Ecosystems. *Science* 277 (5325): 494–99.
- Wang, G., Hapuarachchi, P., Ishidaira, H., Kiem, A.S., and Takeuchi, K. 2009. Estimation of Soil Erosion and Sediment Yield During Individual Rainstorms at Catchment Scale. *Water Resources Management* 23 (8): 1447–65. doi:10.1007/s11269-008-9335-8.
- Wang, G., Yang, H., Wang, L., Xu, Z., and Xue, B. 2014. Using the SWAT Model to Assess Impacts of Land Use Changes on Runoff Generation in Headwaters. *Hydrological Processes* 28 (3): 1032–42. doi:10.1002/hyp.9645.
- Wei, X., and Zhang, M. 2010. Quantifying Streamflow Change Caused by Forest Disturbance at a Large Spatial Scale: A Single Watershed Study. *Water Resources Research* 46 (12). doi:10.1029/2010WR009250.
- Weisse, M., Petersen, R., Sargent, S., and Gibbes, S. 2017. Places to Watch: Identifying High-Priority Forest Disturbance from Near-Real Time Satellite Data. *World Resources Institute*, 12.
- Xiong, J., Thenkabail, P.S., Gumma, M.K., Teluguntla, P., Poehnelt, J., Congalton, R.G., Yadav, K., and Thau, D. 2017. Automated Cropland Mapping of Continental Africa Using Google Earth Engine Cloud Computing. *ISPRS Journal of Photogrammetry and Remote Sensing* 126 (April): 225–44. doi:10.1016/j.isprsjprs.2017.01.019.
- Xue, H., Li, S., and Chang, J. 2015. Combining Ecosystem Service Relationships and DPSIR Framework to Manage Multiple Ecosystem Services. *Environmental Monitoring and Assessment* 187 (3). doi:10.1007/s10661-015-4303-2.
- Zhulidov, A.V., Robarts, R.D., Holmes, R.M., Peterson, B.J., Kämäri, J., Meriläinen, J.J., and Headley, J.V. 2003. Water Quality Monitoring in the Former Soviet Union and the Russian Federation: Assessment of Analytical Methods, 51.
- Zinets, N., and Prentice, A. 2016. Ukraine Military Reports Highest Daily Death Toll since November. *Reuters*, February 16. <https://www.reuters.com/article/us-ukraine-crisis-casualties/ukraine-military-reports-highest-daily-death-toll-since-november-idUSKCN0VP1D1>. Accessed May 25, 2018.

Appendices

Appendix A: Main code for land use classification

```
2015 Get Link Save Run Reset
  - Imports (9 entries) ↗
    > var shape_basin: Table users/saraprückner/Azov_watershed_borders
    > var water: FeatureCollection (46 elements) ↗ ↘
    > var croplandVEG: FeatureCollection (122 elements) ↗ ↘
    > var urban: FeatureCollection (278 elements) ↗ ↘
    > var croplandBARE: FeatureCollection (97 elements) ↗ ↘
    > var naturalVEG: FeatureCollection (74 elements) ↗ ↘
    > var don_basin: Table users/saraprückner/don_basin
    > var upperdon: Table users/saraprückner/upperdon
    > var kuban: Table users/saraprückner/watershed_kuban

1 // Load Landsat 8 surface reflectance data
2 // Look for adequate image filtered by date and cloud cover and region
3 var l8sr = ee.ImageCollection('LANDSAT/LC08/C01/T1_SR')
4   .filterBounds(shape_basin)
5     .filter(ee.Filter.calendarRange(2014, 2016, 'year'))
6       .filter(ee.Filter.calendarRange(8, 8, 'month')).sort('CLOUD_COVER', false);
7
8 // Function to cloud mask from the Fmask band of Landsat 8 SR data.
9 function maskL8sr(image) {
10   // Bits 3 and 5 are cloud shadow and cloud, respectively.
11   var cloudShadowBitMask = ee.Number(2).pow(3).int();
12   var cloudsBitMask = ee.Number(2).pow(5).int();
13
14   // Get the pixel QA band.
15   var qa = image.select('pixel_qa');
16
17   // Both flags should be set to zero, indicating clear conditions.
18   var mask = qa.bitwiseAnd(cloudShadowBitMask).eq(0)
19     .and(qa.bitwiseAnd(cloudsBitMask).eq(0));
20
21   // Return the masked image, scaled to [0, 1].
22   return image.updateMask(mask).divide(10000);
23 }
24
25 // Map the function over one year of data and take the median.
26 var composite = l8sr.map(maskL8sr)
27   .median();
28
29 // NDVI band added to collection
30 var addNDVI = function(composite) {
31   var ndvi = composite.normalizedDifference(['B5', 'B4']).rename('NDVI');
32   return composite.addBands(ndvi);
33 };
34
35 // Test the addNDVI function on a single image.
36 var withndvi = addNDVI(composite);
37 var onlyndvi = addNDVI(composite).select('NDVI').clip(shape_basin);
38
39 // Clip it to the exact region
40 var clipped = withndvi.clip(shape_basin);
41
42 // Display the results.
43 Map.centerObject(shape_basin, 6);
44 Map.addLayer(clipped, {bands: ['B6', 'B5', 'B4'], min: 0, max: 0.4}, 'Landsat 8 image');
45 Map.addLayer(onlyndvi, {min:0,max:0.8, palette: ['FFFFFF', 'FF0000']}, 'NDVI');
46
47 // Use these bands for prediction.
48 var bands = ['B2', 'B3', 'B4', 'B5', 'B6', 'B7', 'NDVI'];
49
50 // Calculating the amount of pixels I have in each polygon
51 var mask = clipped.select(0).mask().rename('mask');
52 var area = ee.Image.pixelArea().multiply(mask).rename('area');
53
54 var sumDictionarypolygon1 = mask.addBands(area).reduceRegion({
55   reducer: ee.Reducer.sum(),
56   geometry: water.geometry(),
57   scale: 30,
58   maxPixels: 1e9
59 });
60 print('sum for water polygons, scale=30', sumDictionarypolygon1);
61
62 var sumDictionarypolygon2 = mask.addBands(area).reduceRegion({
63   reducer: ee.Reducer.sum(),
64   geometry: croplandVEG.geometry(),
65   scale: 30,
66   maxPixels: 1e9
67 });
68 print('sum for cropland polygons, scale=30', sumDictionarypolygon2);
69
70 var sumDictionarypolygon3 = mask.addBands(area).reduceRegion({
71   reducer: ee.Reducer.sum(),
72   geometry: croplandBARE.geometry(),
73   scale: 30,
74   maxPixels: 1e9
75 });
76 print('sum for bareland polygons, scale=30', sumDictionarypolygon3);
77
78 var sumDictionarypolygon4 = mask.addBands(area).reduceRegion({
79   reducer: ee.Reducer.sum(),
80   geometry: urban.geometry(),
81   scale: 30,
82   maxPixels: 1e9
83 });

```

```

86  var sumDictionarypolygons = mask.addBands(area).reduceRegion({
87    reducer: ee.Reducer.sum(),
88    geometry: naturalVEG.geometry(),
89    scale: 30,
90    maxPixels: 1e9
91  });
92  print('sum for forest polygons, scale=30', sumDictionarypolygons);
93
94 //merging feature collections
95
96 var newfc = water.merge(croplandVEG).merge(croplandBARE).merge(urban).merge(naturalVEG);
97 //print(newfc, 'newfc');
98
99
100 // Get the values for all pixels in each polygon in the training.
101 var training = clipped.sampleRegions({
102   // Get the sample from the polygons FeatureCollection.
103   collection: newfc,
104   // Keep this list of properties from the polygons.
105   properties: ['class'],
106   // Set the scale to get Landsat pixels in the polygons.
107   scale: 300
108 });
109
110 // column of random uniforms to the training dataset.
111 var withRandom = training.randomColumn('random');
112
113 // We want to reserve some of the data for testing, to avoid overfitting the model.
114 var split = 0.8; // Roughly 80% training, 20% testing.
115 var trainingPartition = withRandom.filter(ee.Filter.lt('random', split));
116 var testingPartition = withRandom.filter(ee.Filter.gte('random', split));
117
118 // Create a random forest classifier with custom parameters.
119 var classifier = ee.Classifier.randomForest().train({
120   features: trainingPartition,
121   classProperty: 'class',
122   inputProperties: bands
123 });
124
125 // Train the classifier.
126 var trained = classifier.train(trainingPartition, 'class', bands);
127
128 // Classify the image.
129 var classified = clipped.classify(trained);
130
131 // Create a palette to display the classes.
132 var palette =['0000FF', '00FF00','FF0000','FFFF00','006400'];
133
140 // display the classification result and the input image.
141 Map.addLayer(classified, {min: 1, max: 5, palette: palette}, 'Landuse Type 2015');
142
143 // Classify the test FeatureCollection to assess the accuracy!
144 var test = testingPartition.classify(trained);
145
146 // Print the confusion matrix.
147 var confusionMatrix = test.errorMatrix('class', 'classification');
148 print('Confusion Matrix', confusionMatrix);
149 print('Validation overall accuracy: ', confusionMatrix.accuracy());
150 print('Kappa value: ', confusionMatrix.kappa());
151
152 //Exporting confusion matrix to avoid time out
153 var confusionMatrixExport = ee.Feature(null, {matrix: test.errorMatrix('class','classification').array()});
154
155 // Export the FeatureCollection.
156 Export.table.toDrive({
157   collection: ee.FeatureCollection(confusionMatrixExport),
158   description: 'exportAccuracy2015',
159   fileFormat: 'CSV'
160 });
161
162 // Export the image, specifying scale and region.
163 Export.image.toDrive({
164   image: classified,
165   description: 'landusemap2015',
166   scale: 300,
167 });
168
169 // Export the image to an asset in order to analyse the abandoned land
170 Export.image.toAsset({
171   image: classified,
172   description: 'landusasset2015',
173   assetId: 'landuse2015',
174   scale: 300,
175   region: shape_basin.geometry(),
176 });
177
178 //getting statistics
179 // class band added to image
180 var addclass = function(clipped) {
181   var classes = classified.select('classification').rename('classband');
182   return clipped.addBands(classes);
183 };
184
185 var withclass = addclass(classified);
186
187 //clip it to the exact region and select only the bands of interest
188 var resultWithpixqa = withclass.clip(shape_basin);

```

```

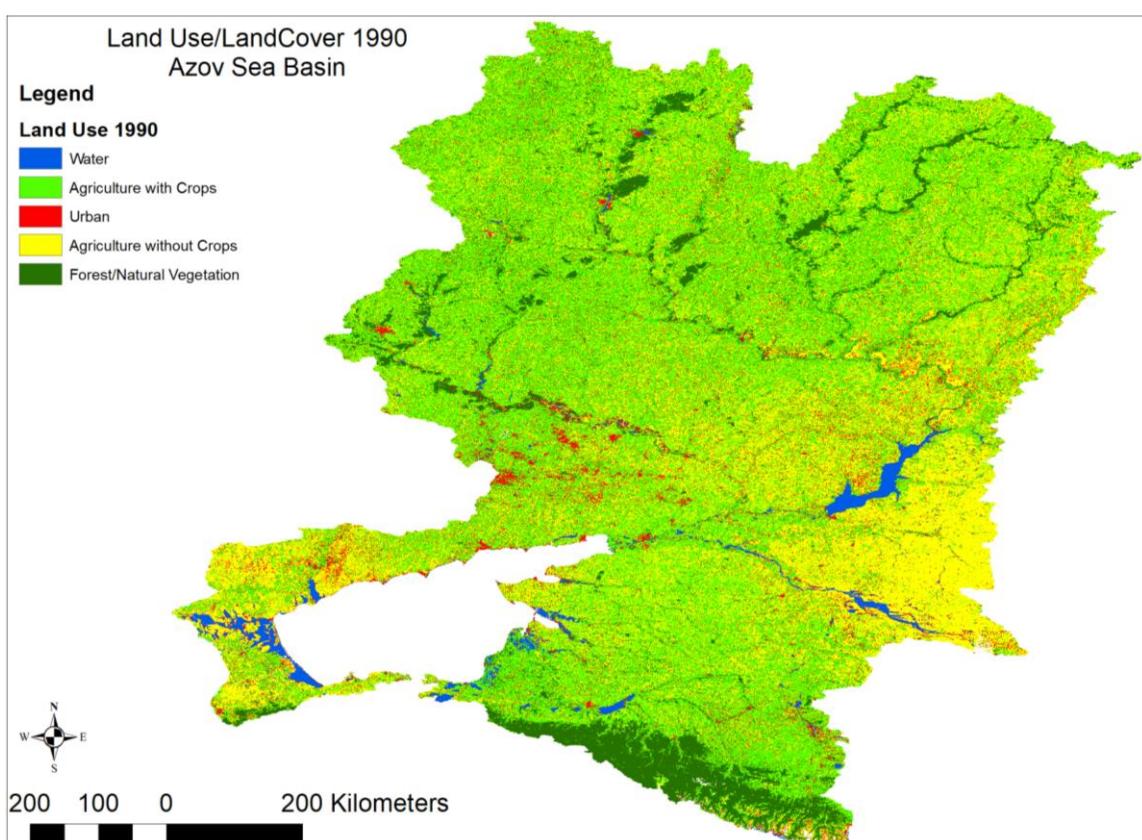
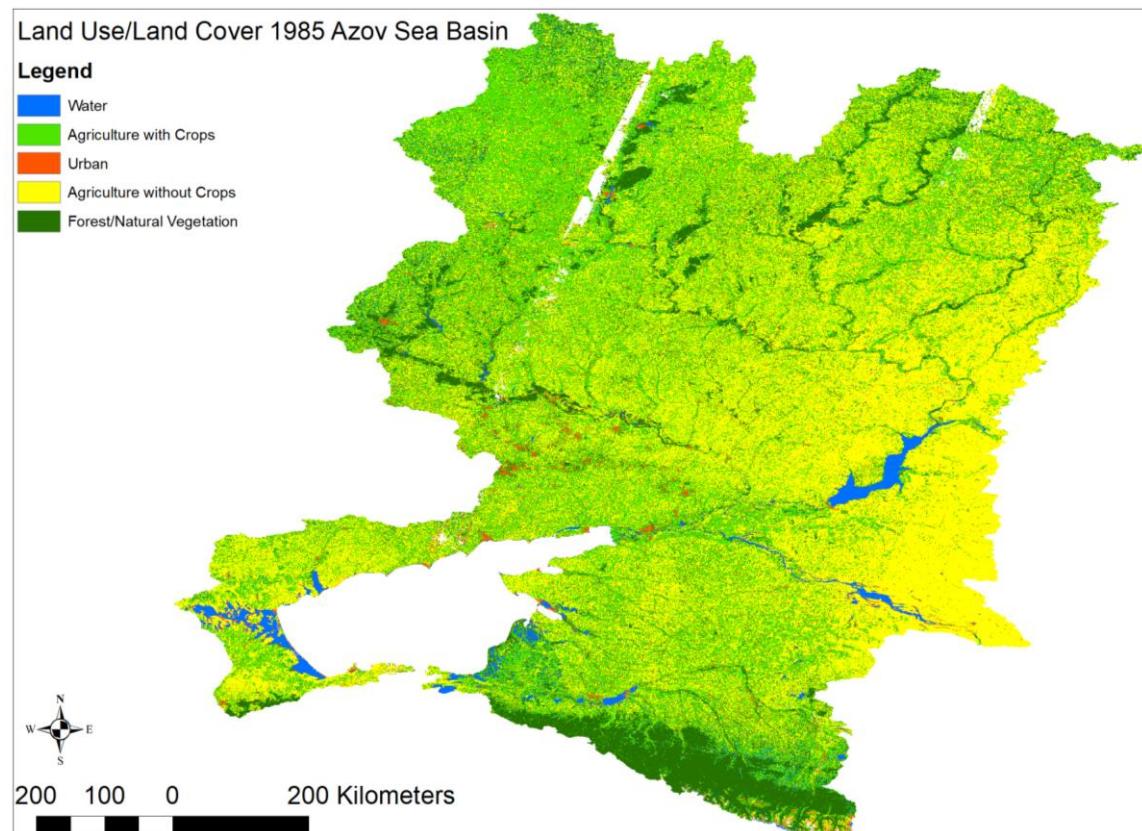
189 var result = resultWithpixqa.select('B2','B3','B4','B5','B6','B10','B7','classband');
190 // Define a broad list of land cover categories.
191 var classNames = ee.List(['0','Water', 'Agriculture', 'Urban', 'Bare', 'Forest']);
192
193 // Define chart customization options.
194 var options = {
195   lineWidth: 1,
196   pointSize: 2,
197   hAxis: {title: 'Band'},
198   vAxis: {title: 'Reflectance'},
199   title: '2015 Mean Spectra Azov Sea Basin'
200 };
201
202 // Make the chart, set the options.
203 var chart = ui.Chart.image.byClass({
204   image: result,
205   classBand: 'classband',
206   region:shape_basin,
207   reducer: ee.Reducer.mean(),
208   scale:2000,
209   xLabels: ['B1','B2','B3','B4','B5','B6','B7'],
210   classLabels: classNames)
211 .setOptions(options).setChartType('Histogram');
212
213 //this is just because it wont work otherwise if I put "setChartType('LineChart')"- this is a bug in GEE!
214 var linechart = chart.setChartType('LineChart');
215 print(linechart);
216
217 // calculating the areas by class
218 var classes = result.select('classband');
219 var area = ee.Image.pixelArea().divide(1000000);
220
221 //water
222 var waterMask = result.updateMask(classes.eq(1));
223 var waterarea = waterMask.multiply(area).select('classband').rename('water');
224 //crops
225 var cropMask = result.updateMask(classes.eq(2));
226 var croparea = cropMask.multiply(area).select('classband').rename('cropland');
227 //urban
228 var urbanMask = result.updateMask(classes.eq(3));
229 var urbanarea = urbanMask.multiply(area).select('classband').rename('urban');
230 //bare
231 var bareMask = result.updateMask(classes.eq(4));
232 var barearea = bareMask.multiply(area).select('classband').rename('bareland');
233 //forest
234 var forestMask = result.updateMask(classes.eq(5));
235 var forestarea = forestMask.multiply(area).select('classband').rename('forest');

252 //Doing the same three more times now - for the Don, Upper Don and Kuban, and then that is it.
253 var donresult = result.clip(don_basin);
254 var donclasses = classes.clip(don_basin);
255 //Map.addLayer(donclasses, {min: 1, max: 5, palette: palette}, 'Landuse Type 1985');
256 //water
257 var donwaterMask = donresult.updateMask(donclasses.eq(1));
258 var donwaterarea = donwaterMask.multiply(area).select('classband').rename('donwater');
259 //crops
260 var doncropMask = donresult.updateMask(donclasses.eq(2));
261 var doncroparea = doncropMask.multiply(area).select('classband').rename('doncropland');
262 //urban
263 var donurbanMask = donresult.updateMask(donclasses.eq(3));
264 var donurbanarea = donurbanMask.multiply(area).select('classband').rename('donurban');
265 //bare
266 var donbareMask = donresult.updateMask(donclasses.eq(4));
267 var donbarearea = donbareMask.multiply(area).select('classband').rename('donbareland');
268 //forest
269 var donforestMask = donresult.updateMask(donclasses.eq(5));
270 var donforestarea = donforestMask.multiply(area).select('classband').rename('donforest');

272 //calculating the area of the waterclasses
273 var donarea_image = donwaterarea.addBands(doncroparea)
274   .addBands(donurbanarea)
275   .addBands(donbarearea)
276   .addBands(donforestarea);
277
278 var donareas = donarea_image.reduceRegion({
279   reducer:ee.Reducer.sum(),
280   geometry: don_basin,
281   scale: 300,
282   maxPixels:1e15
283 });
284 print('Areas of the don catchment', donareas);
285
286 //Same for upperdon only
287 var upperdonresult = result.clip(upperdon);
288 var upperdonclasses = classes.clip(upperdon);
289 //water
290 var upperdonwaterMask = upperdonresult.updateMask(upperdonclasses.eq(1));
291 var upperdonwaterarea = upperdonwaterMask.multiply(area).select('classband').rename('upperdonwater');
292 //crops
293 var upperdoncropMask = upperdonresult.updateMask(upperdonclasses.eq(2));
294 var upperdoncroparea = upperdoncropMask.multiply(area).select('classband').rename('upperdoncropland');
295 //urban
296 var upperdonurbanMask = upperdonresult.updateMask(upperdonclasses.eq(3));
297 var upperdonurbanarea = upperdonurbanMask.multiply(area).select('classband').rename('upperdonurban');
298 //bare
299 var upperdonbareMask = upperdonresult.updateMask(upperdonclasses.eq(4));
300 var upperdonbarearea = upperdonbareMask.multiply(area).select('classband').rename('upperdonbareland');

```

Appendix B: Land Use Maps



Land Use/LandCover 1995 Azov Sea Basin

Legend

Land use 1995

- [Blue square] Water
- [Green square] Agriculture with Crops
- [Red square] Urban
- [Yellow square] Agriculture without Crops
- [Dark Green square] Forest/Natural Vegetation



200 100 0 200 Kilometers

Land Use/LandCover 2000
Azov Sea Basin

Legend

Land use 2000

- [Blue square] Water
- [Green square] Agriculture with Crops
- [Red square] Urban
- [Yellow square] Agriculture without Crops
- [Dark Green square] Forest/Natural Vegetation



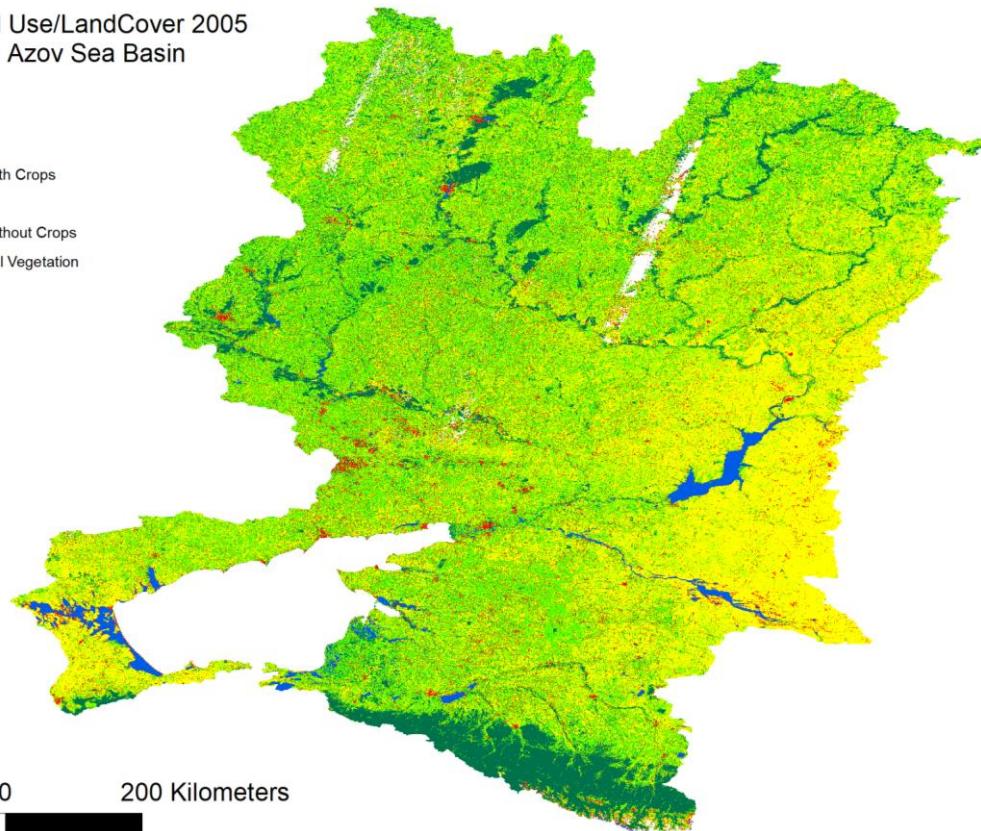
200 100 0 200 Kilometers

Land Use/LandCover 2005
Azov Sea Basin

Legend

Land Use 2005

- [Blue square] Water
- [Green square] Agriculture with Crops
- [Red square] Urban
- [Yellow square] Agriculture without Crops
- [Dark Green square] Forest/Natural Vegetation

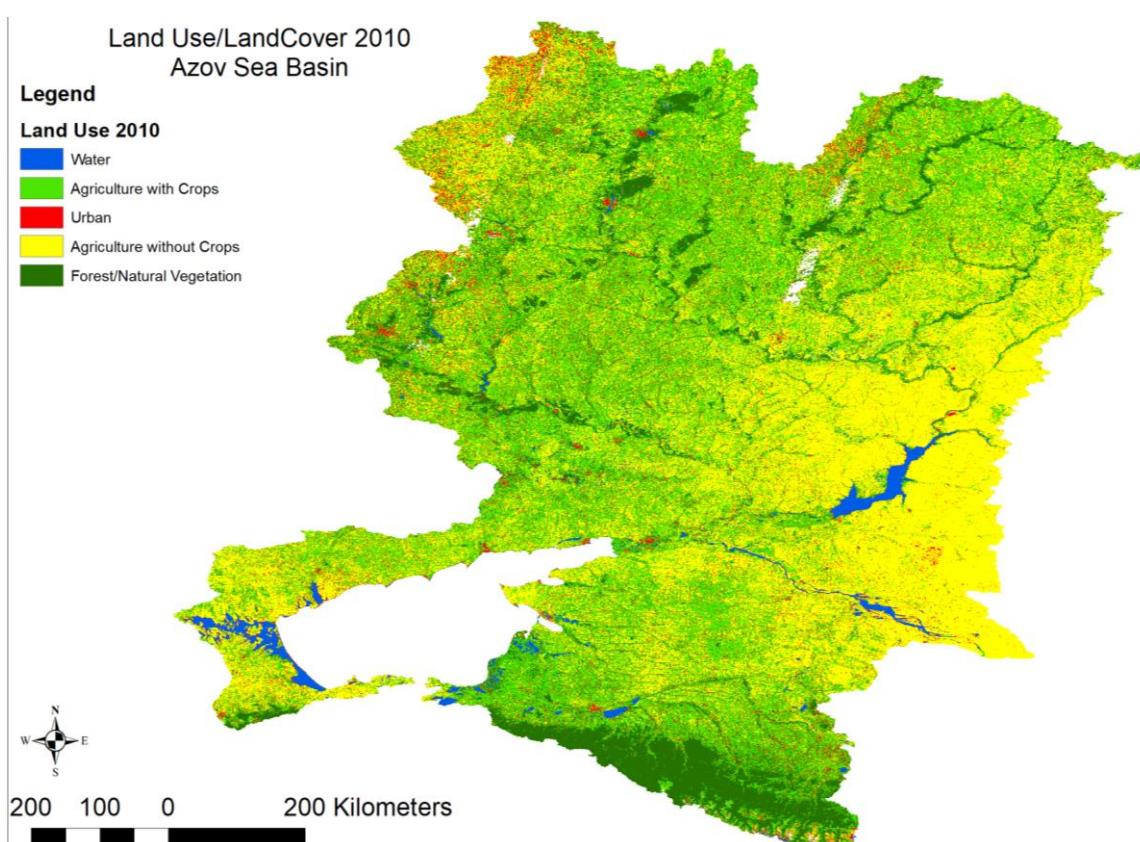


Land Use/LandCover 2010
Azov Sea Basin

Legend

Land Use 2010

- [Blue square] Water
- [Green square] Agriculture with Crops
- [Red square] Urban
- [Yellow square] Agriculture without Crops
- [Dark Green square] Forest/Natural Vegetation



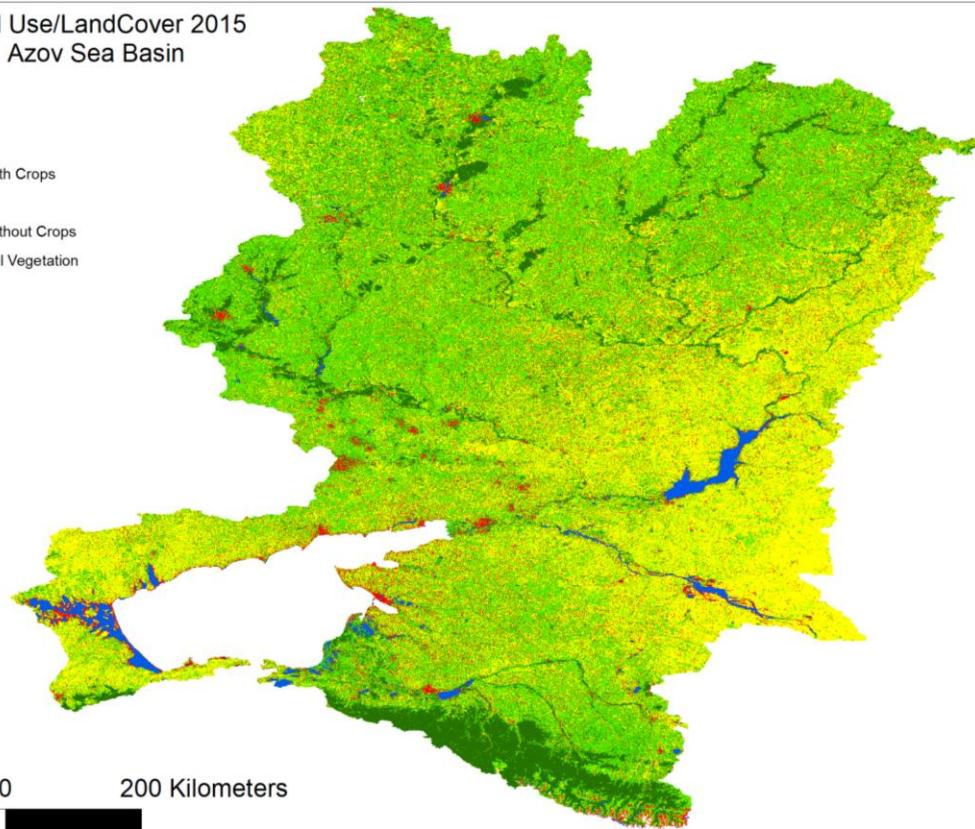
Land Use/LandCover 2015

Azov Sea Basin

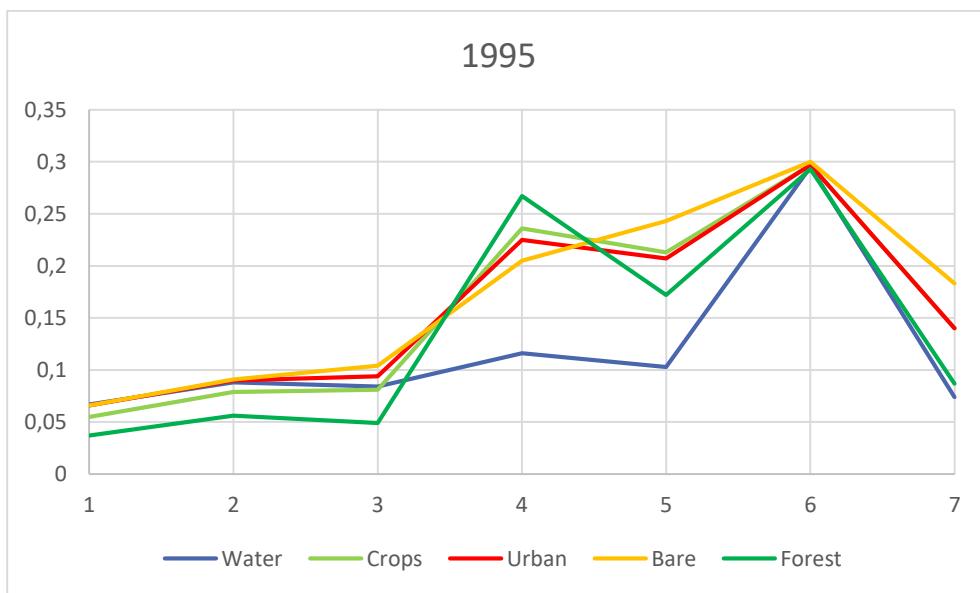
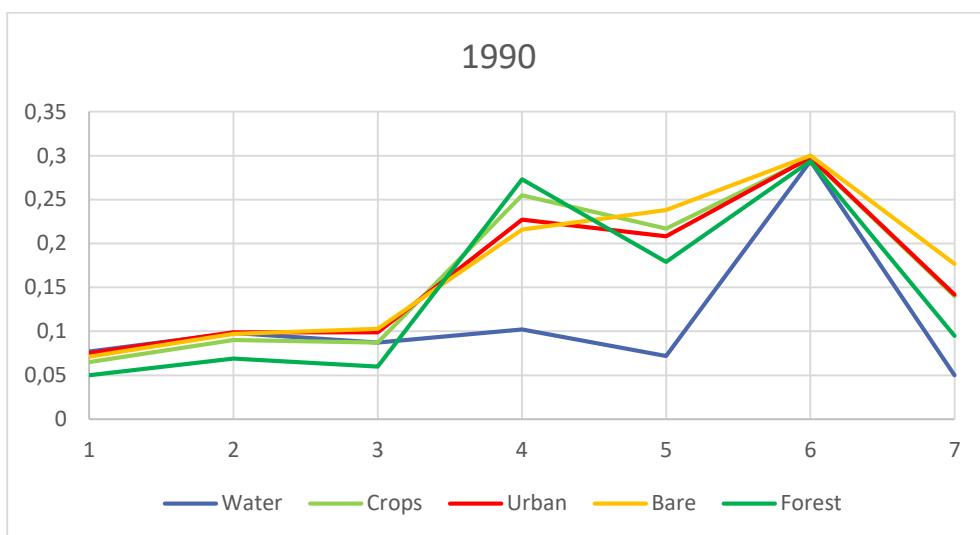
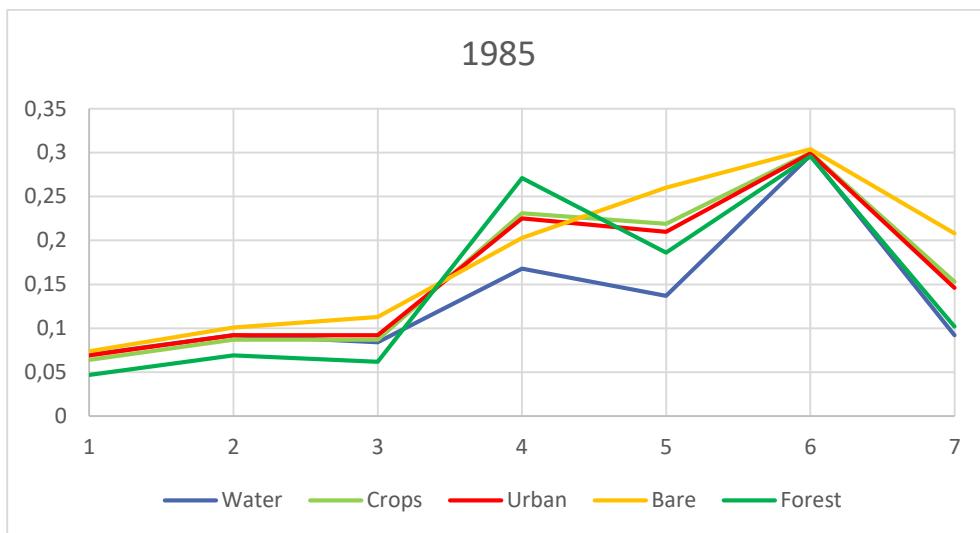
Legend

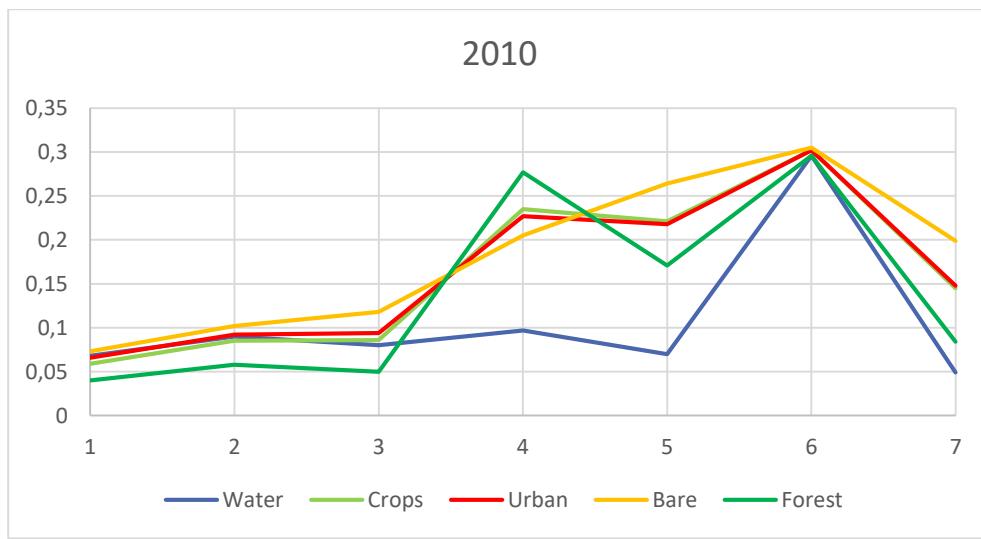
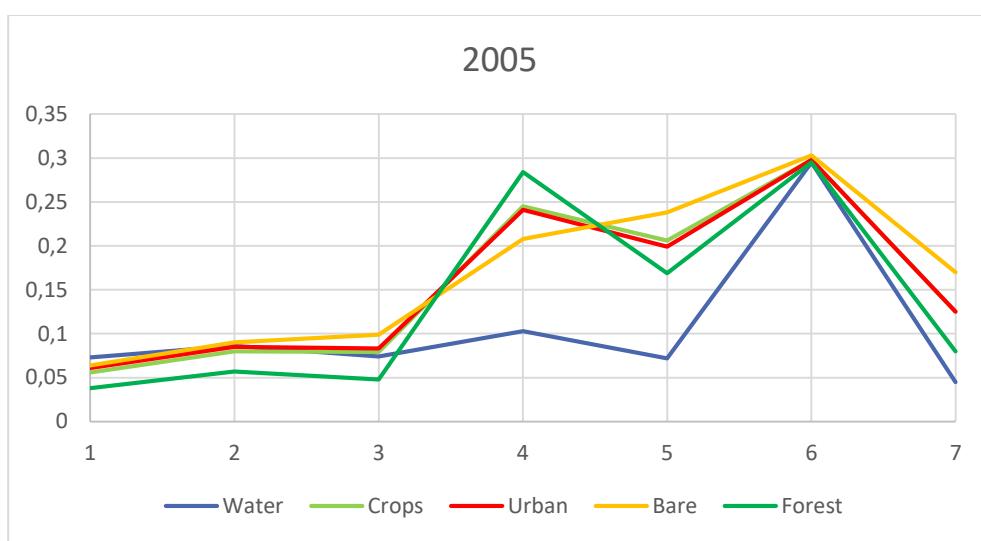
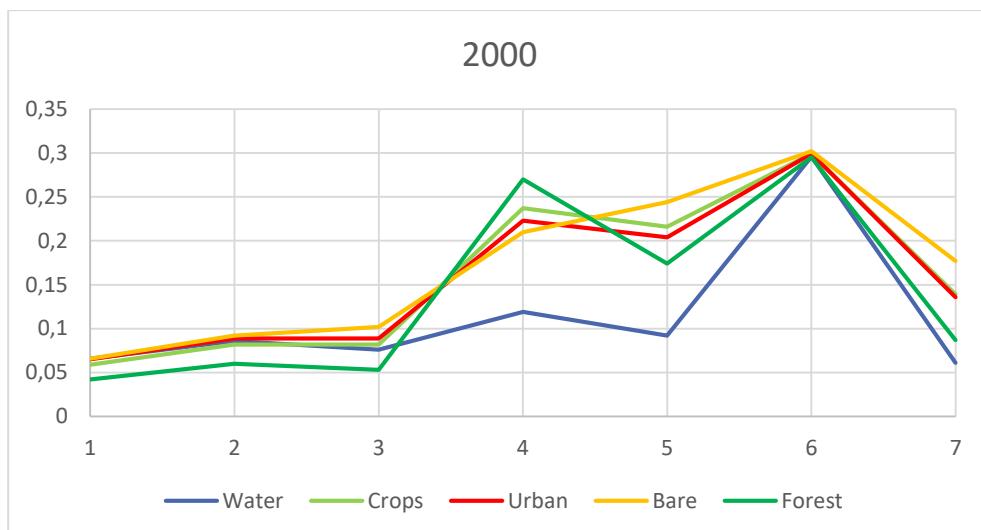
Land Use 2015

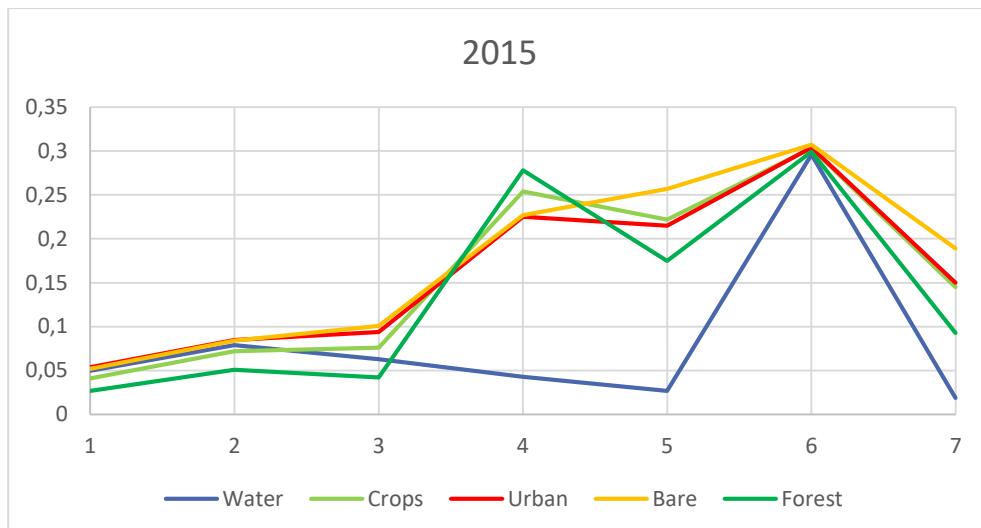
- █ Water
- █ Agriculture with Crops
- █ Urban
- █ Agriculture without Crops
- █ Forest/Natural Vegetation



Appendix C: Mean spectra of all 7 years of all classes







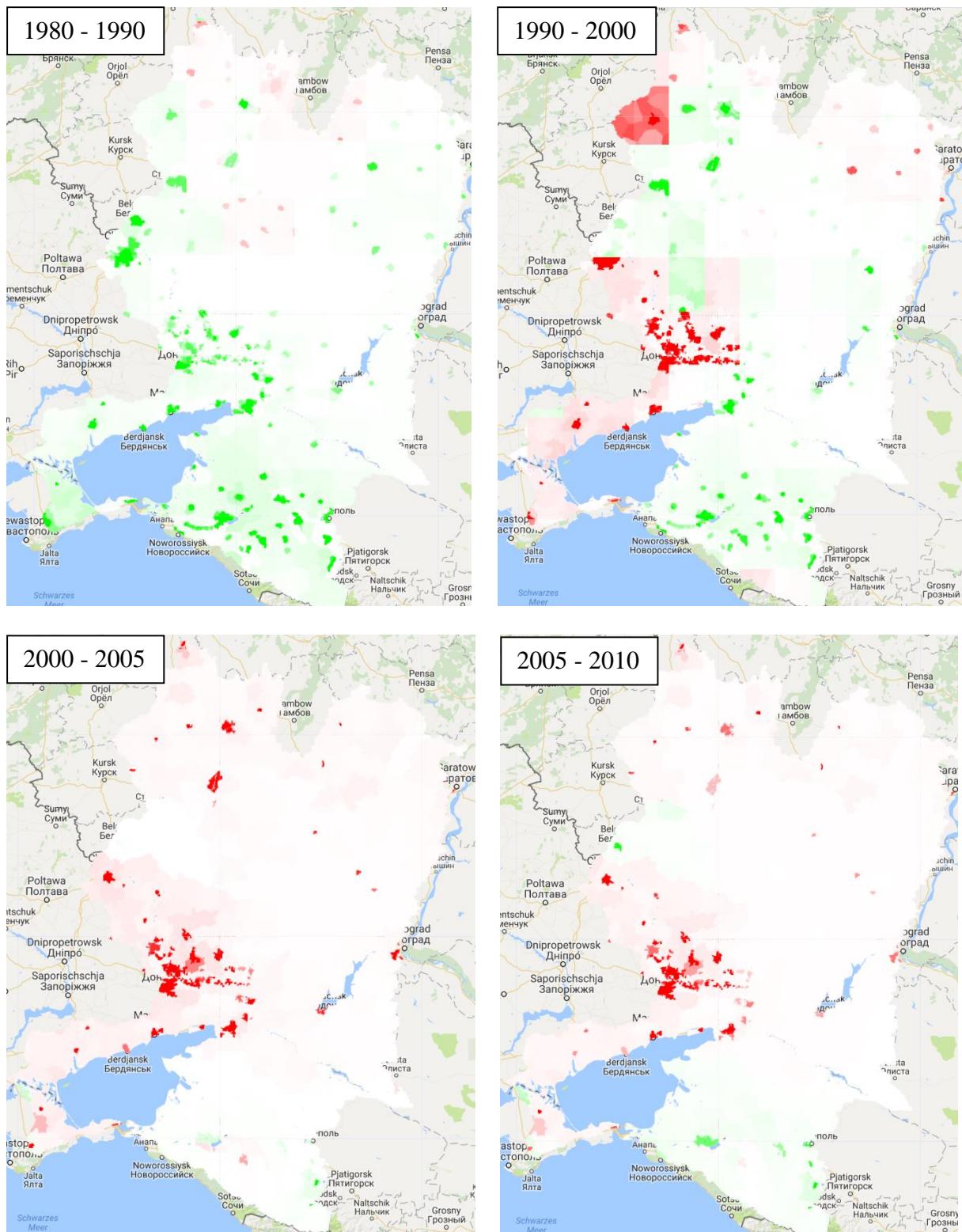
Appendix D: Code for Abandoned land

```

Abandoned_Land *
 1  var shape_Basin: Table users/sarapruckner/Azov_watershed_borders
 2  var landuse1985: Image users/sarapruckner/landuse1985_new (1 band) ☐ ☐
 3  var landuse2000: Image users/sarapruckner/landuse2000 (1 band) ☐ ☐
 4  var landuse2015: Image users/sarapruckner/landuse2015_ (1 band) ☐ ☐
 5
 6 //Map.centerObj(landuse1985,s);
 7 //put all three years into one image with 3 different bands that are named like the years.
 8 var landuse = landuse1985.addBands(landuse2000).addBands(landuse2015);
 9 var landuse = landuse.rename('1985','2000','2015');
10
11 //Everything in class 1 (water), 3 (urban), or 5 (natural vegetation) will be put as 1 = not agriculture.
12 // the rest - 2 (crops) and 4 (bare agricultural land) will be classified as 2 = agriculture
13
14 var not1985 = landuse.expression(
15   "(b('1985') == 1) || (b('1985') == 3) || (b('1985') == 5) ? 1" +
16   ": 2"
17 );
18 var not2000 = landuse.expression(
19   "(b('2000') == 1) || (b('2000') == 3) || (b('2000') == 5) ? 1" +
20   ": 2"
21 );
22 var not2015 = landuse.expression(
23   "(b('2015') == 1) || (b('2015') == 3) || (b('2015') == 5) ? 1" +
24   ": 2"
25 );
26
27 // those 3 images will again be combined into a single image with the three bands.
28 var agriculture = not1985.addBands(not2000).addBands(not2015);
29 var agriculture = agriculture.rename('1985','2000','2015').clip(shape_Basin);
30 print(agriculture);
31 //Map.addLayer(agriculture, {min: 1, max: 2}, 'Agriculture, yes no');
32
33 // "never" is identified when in all 3 years, there was no agriculture present. grey
34 var never = agriculture.expression(
35   "(b('1985') == 1) & (b('2000') == 1) & (b('2015') == 1) ? 1 : 0"
36 );
37 var mask = never.eq(1);
38 var never = never.updateMask(mask).rename('never');
39 Map.addLayer(never, {palette: ['#D0D0D0']}, 'Never');
40
41 // "always" is when in all 3 years, there was agriculture at the spot. Grey
42 var always = agriculture.expression(
43   "(b('1985') == 2) & (b('2000') == 2) & (b('2015') == 2) ? 1 : 0"
44 );
45 var mask = always.eq(1);
46 var always = always.updateMask(mask).rename('always');
47 Map.addLayer(always, {palette: ['#D0D0D0']}, 'Always agriculture');
48
49 // "abandoned1" is when it was agriculture in 1985, but NOT in 2000 or 2015 .red
50 var abandoned1 = agriculture.expression(
51   "(b('1985') == 2) & (b('2000') == 1) & (b('2015') == 1) ? 1 : 0"
52 );
53 var mask = abandoned1.eq(1);
54 var abandoned1 = abandoned1.updateMask(mask).rename('abandoned1');
55 Map.addLayer(abandoned1, {palette: ['#ff0000']}, 'Abandoned between 1985 and 2000, not revitalized');
56
57 // "abandoned2" is when it was agriculture in 1985 and 2000, but NOT in 2015. red
58 var abandoned2 = agriculture.expression(
59   "(b('1985') == 2) & (b('2000') == 2) & (b('2015') == 1) ? 1 : 0"
60 );
61 var mask = abandoned2.eq(1);
62 var abandoned2 = abandoned2.updateMask(mask).rename('abandoned2');
63 Map.addLayer(abandoned2, {palette: ['#ff0000']}, 'Abandoned between 2000 and 2015 only');
64
65 //revitalized is when it was agriculture in 1985 and 2015, but not in 2000. Blue
66 var revita = agriculture.expression(
67   "(b('1985') == 2) & (b('2000') == 1) & (b('2015') == 2) ? 1 : 0"
68 );
69 var mask = revita.eq(1);
70 var revita = revita.updateMask(mask).rename('revita');
71 Map.addLayer(revita, {palette: ['#0000ff']}, 'Abandoned between 1985 and 2000, later revitalized');
72
73 //added1 is when it was agriculture in 2015 and 2000. Yellow
74 var added1 = agriculture.expression(
75   "(b('1985') == 1) & (b('2000') == 2) & (b('2015') == 2) ? 1 : 0"
76 );
77 var mask = added1.eq(1);
78 var added1 = added1.updateMask(mask).rename('added1');
79 Map.addLayer(added1, {palette: ['#FFFF00']}, 'Not agriculture in 1985, but added in 2015');
80
81 //added is when it was agriculture only in 2015. Yellow
82 var added2 = agriculture.expression(
83   "(b('1985') == 1) & (b('2000') == 1) & (b('2015') == 2) ? 1 : 0"
84 );
85 var mask = added2.eq(1);
86 var added2 = added2.updateMask(mask).rename('added2');
87 Map.addLayer(added2, {palette: ['#FFFF00']}, 'Not agriculture in 1985, but added in 2000');
88
89 //other is the last opportunity - no agriculture in 1985 and 2015, but yes in 2000. Grey
90 var other = agriculture.expression(
91   "(b('1985') == 1) & (b('2000') == 2) & (b('2015') == 1) ? 1 : 0"
92 );
93 var mask = other.eq(1);
94 var other = other.updateMask(mask).rename('other');
95 Map.addLayer(other, {palette: ['#D0D0D0']}, 'Only Ag in 2000');
96
97 //calculating the area of the waterclasses
98 var area_image = never.addBands(always)
99   .addBands(abandoned1)
100  .addBands(abandoned2)
101  .addBands(revita)
102  .addBands(added1)
103  .addBands(added2)
104  .addBands(other);
105
106 var areas = area_image.reduceRegion(
107   reducer ee.Reducer.sum(),
108   geometry: shape_Basin,
109   scale: 300,
110   maxPixels: 1e13
111 );
112 print(areas);

```

Appendix E: Population Change Maps in ten- and five-year-steps, respectively



2010 - 2015

