Mapping the Yield gaps across Ukraine using time series of satellite derived biophysical variables

Tiago Soares Monge Pinho dos Santos
May 2015

Mapping the Yield gaps across Ukraine using time series of satellite derived biophysical variables

by

Tiago Soares Monge Pinho dos Santos

Thesis submitted to the University of Southampton, UK, in partial fulfilment of the requirements for the degree of Master of Science in Geo-information Science and Earth Observation.

Specialisation: Environmental Modelling and Management

DECLARATION OF AUTHORSHIP

I, **Tiago Soares Monge Pinho dos Santos**, declare that the thesis entitled "**Mapping the Yield gaps across Ukraine using time series of satellite derived biophysical variables**" and the work presented in the thesis are both my own and have been generated by me as the result of my own scholarship. I confirm that:

- This work was done wholly while in candidature for a Masters degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others accreditation has always been given.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.
- It parts of this work have been published, they are listed below.

1. I have read and understood the University's Academic Integrity Statement for Students, including the information on practice to avoid given in appendix 1 of the Statement and that in this thesis I have worked within the expectations of this Statement. http://www.calendar.soton.ac.uk/sectionIV/academic-integrity-statement.html

2. I am aware that failure to act in accordance with the Academic Integrity Statement for Students may lead to the imposition of penalties which, for the most serious cases, may include termination of programme.

3. I consent to the University copying and distributing my work and using third parties to verify whether my work contains plagiarised material. If a paper copy is also required for submission it must be identical to this electronic copy. Any discrepancies between these two copies may be considered as an act of cheating.

Signed ______________________ Date 26-05-2015

Disclaimer

This document describes work undertaken as part of a programme of study at the University of Southampton. All views and opinions expressed therein remain the sole responsibility of the author, and do not necessarily represent those of the University.

Abstract

The increase in global population and in demand for food and energy are expected to strongly lead to a raise in demand for agricultural outputs, which may have negative environmental impacts. It is therefore important to sustainably intensify agricultural production while reducing negative environmental impacts, by closing yield gaps on underperforming croplands, rather than converting new areas to agriculture. The collapse of the Soviet Union in 1991 led to widespread abandonment of agricultural lands, but the extent and spatial patterns of abandonment are unclear. Ukraine, one of the main agricultural producers in this region, has a significant unrealized grain production potential. A global analysis of yield gaps highlighted the croplands of most of Ukraine as performing at about 50% of their climatic potential. However, it was undertaken using data from around the year 2000; its spatial resolution was too coarse for efficient location of targeting resources; and it didn’t account for inter-annual variability in crop production.

Thus, the overall aim of this thesis is to develop a procedure to map Yield Gaps across Ukraine using time series of satellite derived biophysical variables, in order to identify underperforming croplands.

Wheat, barley, maize and sugar beet maps, were produced through the disaggregation of published agricultural statistics by using time series EVI derived from MODIS imagery. Actual crop yield was settled through correlation with maximum EVI, obtained with a phenology extraction algorithm. Potential crop yield was assessed through the 90^{th} percentile of yield in each previously determined homogeneous edapho-climatic zone. Yield Gaps across the croplands of Ukraine resulted from the difference between the potential yields and the actual yields.

Maximum EVI cluster areas were in general significantly correlated to official crop areas, whereas the maximum EVI itself didn’t greatly correlate to official crop yield or production, except for the moderately significant barley results. Most of the grain yield gaps were concentrated in the southern east steppe zone. Future work require the inclusion of higher spatial resolution imagery for individual crop fields mapping; improve extraction and mapping of the phenology parameter that best fits the production; use of updated and relevant ground control points to allow thematic validation.

Keywords: Time series analysis, Vegetation indices, Crop mapping, Crop yield gaps, Ukraine

Acknowledgements

I'd like to thank my supervisors, Dr. Jadu Dash and Dr. John Duncan, for accepting me as their pupil, their guidance, advice, sharing literature and scripts, and support at all stages of this work.

I would also like to thank to Dr. Kees de Bie, who generously offered help and advice, including introducing me to the field of the application of remote sensing to agriculture and food security.

Thank you to all my ITC and GEM friends for the great moments we shared in Enschede and Southampton over the last 22 months.

I would like to thank my family and friends who have always supported me to accomplish my goals and for their encouragement.

Table of Contents

List of figures

List of tables

Chapter 1

This chapter introduces the background of the study, concepts on yield gaps, monitoring, and literature review on previous studies.

1 Introduction

1.1 Global background – food security and climate change

According to the 2012 revision of the official United Nations population estimates and projections, the world population of 7.2 billion in mid-2013 is projected to reach 9.6 billion in 2050 (United Nations, 2013). This population growth and increasing consumption of calories and meat-intensive diets are expected to roughly double human food demand by 2050 (FAO, 2009)

This food demand will require an equivalent increase on food production, which may have important negative impacts, namely (Foley, 2010):

- Land use expansion/competition: 40% of Earth's land is used for agriculture, which competes with urbanization, other industries, forestry, non-food crops, as well as use of land for bio energy (Smith et al., 2010). Land conversion or inappropriate management may lead to loss of biodiversity (Godfray et al., 2010).
- Water shortage: about 70% of the planet's accessible freshwater is consumed by agricultural sector, which leads to water shortages.
- Pollution: 30% of global Green House Emissions are due to agriculture. There is also soil and water pollution due to nutrient run-off.

How, therefore, can the world double the availability of food while simultaneously cutting the environmental harm caused by agriculture? The Commission on Sustainable Agriculture and Climate Change, with representatives of all major regions of the world and a wide range of scientific backgrounds, made a summary for policy makers to achieve food security in an environmentally sustainable way and in the face of climate change (Beddington et al., 2011). Among other actions, they suggest to sustainably intensify agricultural production while reducing greenhouse gas emissions and other negative environmental impacts of agriculture. This measure places emphasis on closing yield gaps on underperforming lands, rather than converting new areas to agriculture.

It is thus of critical importance to know where and how best to increase crop yield on existing cropland area (Wart et al., 2013) by closing the yield gaps. Also, yield trends and variations among various regions should be analyzed to understand the sources of these variations (Makowski et al., 2013).

1.2 Local background – Ukraine opportunity

The collapse of the Soviet Union triggered widespread farmland abandonment, leading to a sharp decline of grain production during the past two decades (Alcantara et al., 2013) . In the context of the current economic and food-price crisis, Russia, Ukraine, and Kazakhstan might be presented with a window of opportunity to reemerge on the global agricultural market, if they succeed in increasing their productivity (Lioubimtseva & Henebry, 2012), by closing the yield gaps. More specifically, Ukraine, known as the breadbasket of Europe (Das, 2014), has huge agricultural potential due to its rich natural resources (soil, climate, and water) and a key geographical position, with access to the Black Sea and the key markets in the European Union (EU), Commonwealth of Independent States (CIS), the Middle East and North Africa (Leeuwen & et al., 2012). So, given this agricultural potential of Ukraine, it is important to quantify its crop yield and production capacity.

1.3 Concepts

During this chapter, except for the rest of the thesis, it is assumed that whenever yield is referred, production is included in the concept. It is noteworthy that crop yield refers to the harvested production per unit of harvested area for crop products, expressed in tonnes/Hectare. Whereas crop production refers to the actual harvested production from the field expressed in tonnes (FAO, 2015b).

Crop yield capacity can be evaluated by estimating yield potential and water-limited yield potential levels as benchmarks for crop production under, respectively, irrigated and rain fed conditions (Ittersum et al., 2013; FAO, 2015). The differences between these potential yield levels and actual or average farmers' yields over some specified spatial and temporal scale of interest define the yield gaps (Lobell et al., 2009; FAO, 2015). M. Van Ittersum et al., 2013 and Lobell et al., 2009 has extensive reviews about the concepts related with yield gaps.

Average yield or actual yield (Ya) represent variation in time and space in a defined geographical region achieved by farmers in the region under the most widely used management practices (sowing date, cultivar maturity, and plant density, nutrient management and crop protection)(Ittersum et al., 2013). Data on average yield are typically based on crop statistics (Hengsdijk & Langeveld, 2009), or by sampling farmers' fields, either directly or using remote sensing (Lobell, 2013; Hall et al., 2013). Crop data are generally summarized and aggregated at various levels of administrative districts (Schulthess et al., 2013; De Bie & Skidmore, 2010). The number of years utilized for estimating average yield must be a compromise between variability in yields and the necessity to avoid confounding effects of temporal yield trends due to technological or climate change (Ittersum et al., 2013).

Yield potential (Yp) is defined as the maximum attainable yield per unit land area that can be achieved by a particular crop cultivar in an irrigated system environment to which it is adapted when pests and diseases are effectively controlled and nutrients are non-limiting (Evans & Fischer, 1999). Potential yield is location specific because of climate factors such as solar radiation, temperature, carbon dioxide concentration, and genetic characteristics.

For rain fed crops, water-limited yield potential (Yw) is defined similarly to Yp, except crop growth is also limited by water supply, and hence influenced by soil type and field topography (Wart et al., 2013).

Yield potential estimation is based on crop models (Tuan, 2008; Brisson et al., 2010; Boogaard et al., 2013), field experiments and yield contests (Lobell & Burke, 2010) or maximum farmer yields within homogeneous zones (Lobell et al., 2009). Fischer et al. (2014) has a comprehensive list of crop specific global and local studies which includes the respective method of potential yield determination for yield gap calculation.

Yield gap is significantly affected by a number of environmental constraints - pests, diseases and management - and is decomposed into three parts, Yield gap 1, 2 and 3 (see Figure 1) (M. K. Van Ittersum & Rabbinge, 1997; Lobell et al., 2009b).

Yield gap 1 is the difference between observed best farmer yield and actual yield (Ya) under average farmer's practices, and can be narrowed with best practices. It has as reducing factors biological constraints (plant density, weeds, pests and diseases, problem soils, etc.) and outdated technology.

Yield gap 2 is the difference between on-farm experiment's maximum yield (attainable yield - Yt) and best farmer yield, and can be also narrowed with optimal existing technology. The attainable yield varies from season to season and year to year depending on climate (Pasuquin & Witt, 2007). It has as limiting factors water availability and nutrients.

Yield gap 3 is the difference between maximum yield estimated using crop growth models or experimentally through maximum yield trials (Yield potential - Yp), and an on-farm experiment's maximum yield. This yield gap arises from differences in environment (available rainfall, crop characteristics, temperature, soil and ground-water and/or by macro-nutrients) and some component technologies only available at research stations, which cannot be managed in the farmer's field (Dixon et al., 2001).

A management objective of farmers should be to minimize the difference between attainable and actual yield (Yt-Ya). To narrow this yield gap (1+2), farmers need to evaluate promising new technologies (e.g., planting density, nutrient management) that offer improvements in yield and/or productivity against current practices (Pasuquin & Witt, 2007).

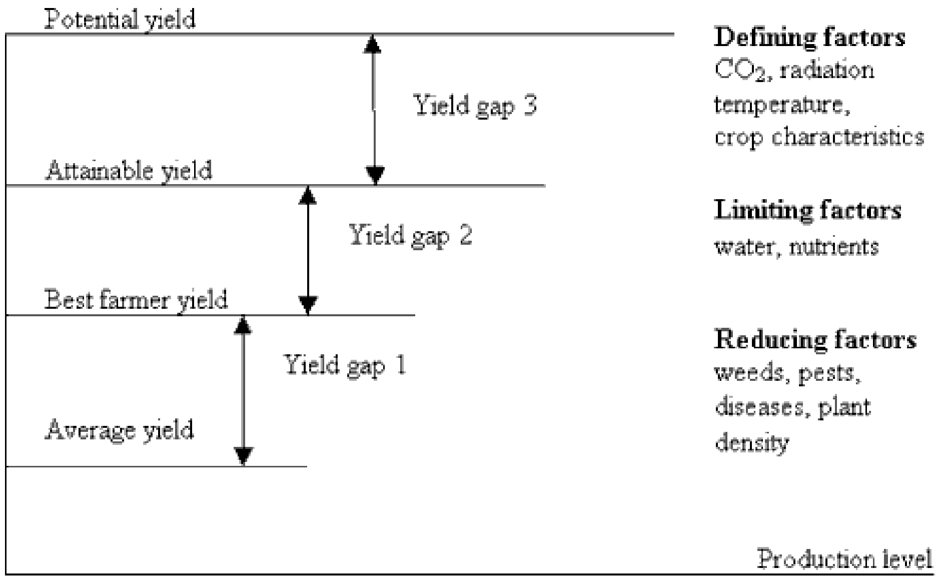


Figure 1: Types of yield gaps and related concepts (adapted from http://www.aglearn.net/isfmMod3.html)

1.4 Monitoring

To achieve these objectives, farmers and managers need to previously know where, when and what are the actual and potential yields through local or global agricultural monitoring programs, depending on the level of decision.

Remote sensing can provide data that help identify and monitor crops (Atzberger, 2013). When these data are organized in a Geographical Information System along with other types of data, they become an important tool that helps in making decisions about crops and agricultural strategies. Because of the particular manner vegetation reflects the electromagnetic radiation, we can assess the crop status by using remote sensing data (Tucker, 1979; Kalaitzidis & Manakos, 2015).

The vegetation spectral signature typically absorbs in the red and blue wavelengths, reflects in the green wavelength, strongly reflects in the near infrared (NIR) wavelength, and displays strong absorption features in wavelengths where atmospheric water is present. Different plant materials, water content, pigment, carbon content, nitrogen content, and other properties cause further variation across the spectrum (Silleos et al. , 2006; Elowitz, 2015). Measuring these variations and studying their relationship to one another can provide meaningful information about plant health (Zhaoqin et al., 2014; Martinelli et al., 2014), water content (Yebra et al., 2013), environmental stress (Qiu et al., 2009), biomass (Tucker, 1980; Silleos et al., 2006), and other important characteristics. Adding the temporal dimension to the vegetation spectral signature is useful for distinguishing land-cover types and for mapping land-use change, which makes phenological metrics useful within agricultural monitoring systems (Almond, 2009).

These measurements and relationships had resulted in a multitude of vegetation index (VI) equations that includes band ratios, normalized differences, linear band combinations, and optimized band combinations (Tucker, 1979; EXELIS, 2013). A wide review of VI can be found in Silleos et al. (2006), Basso et al. (2004), and textbooks like Jensen (2007).

Jackson & Huete (1991), classify VI into slope-based and distance-based vegetation indices. The slope-based VIs are simple arithmetic combinations that focus on the contrast between the spectral response patterns of vegetation in the red and near-infrared portions of the electromagnetic spectrum (Silleos et al., 2006). Whereas the distance-based group measures the degree of vegetation present by measuring the difference of any pixel's reflectance from the reflectance of bare soil.

Slope-based VIs indicates both the state and abundance of green vegetation cover and biomass. Hence, slope-based VIs are widely used in crop yield estimation (Taylor et al., 1997; Báez-González et al., 2002; Baez-Gonzalez et al., 2005; Funk & Budde, 2009). The slope-based VIs include the RVI (Ratio Vegetation Index), NDVI (Normalized Difference Vegetation Index) and EVI (Enhanced Vegetation Index), among many others referred in Silleos et al. (2006). One of the first index developed is the RVI (Jordan, 1969) which is the ratio between NIR and Red. The most commonly used index is the NDVI (Rouse, 1978), which is the ratio of the difference of the near-infrared and red reflectance, over the sum of those, and it ranges from -1 (no vegetation) to +1 (abundant vegetation). An advantage of the NDVI, as a ratio, is its ability to produce stable values by normalizing many extraneous sources of noise. The disadvantages with NDVI in landscape studies are related to the nonlinear behavior of ratios, sensitivity to soil background, and saturation at moderate to high vegetation densities. The Enhanced Vegetation Index (EVI), which also ranges from -1 to +1, was developed by the MODIS Land Discipline Group for use with MODIS data. It is a modified NDVI with a soil adjustment factor L and two coefficients C1 and C2, which describe the use of the blue band in correction of the red band for atmospheric aerosol scattering. This VI has improved sensitivity to high biomass regions and reduced atmospheric influence (Huete et al., 1999). Barroso & Monteiro (2010) further describes applications, concepts and advantages/disadvantages of NVDI and EVI indices. Slope-based VIs such as NDVI and EVI are commonly used as surrogate for crop yield (Bolton & Friedl, 2013; Son et al., 2014), and in crop yield gaps analysis (Kolotii et al., 2015).

Yield gaps have been estimated in previous studies with either a global or local focus (M. Van Ittersum, et al., 2013; Justin Van Wart et al., 2013; Oliver & Robertson, 2013) . Local scale studies include field experiments, growers' yield contests (Hochman et al., 2013), crop model simulations (Brisson et al. 2010) or farmers' maximum yields (Lobell et al. 2009). Global focus studies are based on remote sensing data and estimation from global crop datasets including yield values and climatic variables (Licker et al., 2010). Whereas global methods are generally coarse and provide worldwide coverage using a consistent method, local studies are based on location-specific environmental conditions and management, which give local relevance but are hard to compare across locations and studies because of inconsistent terminology, concepts and methods (M. Van Ittersum et al., 2013).

There are several complicating factors involved in characterizing croplands at the global/country level (Pittman et al., 2010). First, the spatial extent of croplands is highly variable within a large nation like Ukraine. Depending on the historical, political, social and technological context of agricultural development and natural factors such as landscape pattern, cropland characteristics such as field size can be highly variable, even for the same crop type. Second, each crop type has a specific growth phenology and structure, with significant seasonal variation between and even within individual crop types. Third, cropland is a land use and can be confused with natural vegetation cover types, such as cereal grains versus tall-grass prairie.

To overcome these limitations, high-temporal earth observation coverage at fine spatial scales is desired, but usually the available data has a compromise. For instance, to assess crop growth through extraction of crop phenology, high temporal resolution remote sensing images like MODIS (Pittman et al., 2010; Sakamoto et al., 2013), MERIS (Dash et al, 2010), NOAA-AVHRR (Huete et al., 2006) or SPOT-Vegetation (Kowalik et al., 2014) may be used in the time series analysis. Rembold et al.(2013), gives a further overview on using high temporal resolution remote sensing for crop monitoring and yield forecasting.

However, the spatial resolution needed to distinguish individual field may be low, depending on whether there are predominantly complex patterns with small and heterogeneous land covers or intensive agricultural large fields, i.e., each pixel may contain different land covers. Some processing techniques that combine different resolutions, which are referred latter in this section, can be used to overcome this loss of spatial resolution.

Another challenge of high temporal resolution imagery is to extract out of noisy random externalities related with cloud cover, poor atmospheric conditions, and unfavorable sun-sensor-surface viewing geometry, the crop phenological (vegetation green-up and senescence) data as a surrogate for yield estimation (Geng et al., 2014). Furthermore, croplands present a more complex phenology than natural land cover, due to their many peaks resulting from multiple crops planted sequentially within a growing season. Consequently, several studies have identified land cover based on specific properties of the observed green leaf phenology, such as start and end of season, moment of maximum vegetation index and amplitude of maximums (Dash et al., 2010; Atzberger, 2013). According to Beck et al. (2006), the different methods for phenology extraction can be grouped in two categories:

- methods estimating the timing of single phenological events (Reed et al., 1994; White & Thomton, 1997; Badeck et al., 2004);
- methods modeling the entire time series using a mathematical function (Jonsson & Eklundh, 2002; Stockli & Vidale, 2004).

Approaches belonging to the first group include the use of specific VI thresholds (Lloyd, 1990; White & Thomton, 1997); the detection of the largest VI increase between two consecutive observations (Kaduk & Heimann, 1996; Araya et al., 2013); backward-looking moving averages (Reed et al., 1994; Brown, 2015); and rate of change in the curvature of a locally fitted logistic model to identify phenological transition dates (Zhang et al., 2003). Methods for analyzing entire time series include principle component analysis (Hirosawa et al., 1996; Hall-Beyer, 2003); Fourier analysis (Azzali & Menenti, 2000; Atkinson et al., 2012); harmonic analysis (Rouse, 1978; Jakubauskas et al., 2001); wavelet decomposition (Li et al., 1996; Sakamoto et al., 2005); and curve fitting (X. Zhang et al., 2003; Jönsson &

Eklundh, 2004; Beck et al., 2006; Lu et al., 2013). A further overview on phenology extraction methods and its advantages/disadvantages is presented in de Beurs & Henebry (2010).

On the other hand, field studies (Hall et al., 2013), drones (C. Zhang & Kovacs, 2012), medium and high spatial resolution remote sensing platforms (Husak et al., 2008) bring forth data with the spatial resolution of individual crop fields or more, but compromising time resolution needed to access crop phenology for yield estimation at a field level. In this context, techniques that combine different spatial, temporal and spectral resolution allow combining information from multiple sensors or sources such as ground data (Pervez & Brown, 2010) to obtain image products with improved overall characteristics (Amorós-López et al., 2013). Thenkabail & Wu (2012) refers to several croplands mapping across resolutions methods, such as decision trees, neural network methods, etc. De Bie & Skidmore, (2010) suggest a crop mapping technique based on an unsupervised classification of the time series profiles of a vegetation index, followed by a stepwise linear regression between the vegetation index cluster areas and published agricultural statistics areas that aims to disaggregate those statistics (by crop, year, and administrative area), with the help of the vegetation index clusters product, in order to generate a crop-specific cropping intensity map (percentage crop per pixel).

The crop productivity on regional and global scales can be estimated with statistical models based on remote-sensing VI data with spatially and temporally continuous distributions (A. Li et al., 2007; Liang et al., 2012). Statistical models can be classified into two categories: direct establishment of the correlation between a vegetation index and vegetation productivity, which enables regional estimation (Prasad et al., 2006; Santin-Janin et al., 2009), and the establishment of a regression parameter vector for regional applications, which is realized through the utilization of vegetation indices and other environmental factors in regression trees (Lobell et al., 2005), neural networks (A. Li et al., 2007), or other complex statistical methods.

Thus, high temporal resolution remote sensing imagery used along with medium to high spatial resolution remote sensing imagery and/or data allows assessing crop yield temporal variation on crop extent areas, and further determining potential yield gaps at fine spatial resolution.

1.5 Literature review

Agricultural monitoring is important for all countries involved in crop production, especially those experiencing rapid changes in the extent of agricultural lands. In Ukraine the agricultural sector is undergoing rapid transformation with associated changes in agricultural land ownership and practices. National agricultural monitoring is particularly important for planning, where there is increasing competition for water and land resources (Justice & Becker-Reshef, 2007).

Many programs have been established by agricultural agencies to regularly provide agricultural statistics at different spatial and temporal scales (Group on Earth Observations, 2010), for example: global monitoring such as the "Monitoring Agriculture with Remote Sensing" (MARS) from the Institute for Environment and Sustainability, European Commission (IES, 2015) or the Foreign Agricultural Service (FAS) of the U.S. Department of Agriculture (USDA) from USA (USDA FSA, 2015); early-warning systems such as the "Global Information and Early Warning System" (GIEWS) from FAO (FAO, 2015a); or national monitoring such as the "National Centre of Space Research, Technologies" (NCRST) from Kazakhstan (NCSRT, 2015) or the "Space Research Institute of Russian Academy of Sciences" (ИКИ) from Russia (ИКИ РАН, 2015). The GEO-GLAM (global agriculture monitoring) project is working to harmonize remote sensing-based crop monitoring systems through the Joint Experiment of Crop Assessment and Monitoring (JECAM) project (Group on Earth Observations, 2013). Framed on this initiative, Ukraine has three test sites, in Kiev, Lviv and Pshenychne regions. This Ukraine's project aims to identify crops and estimate Crop Area; assess crop condition/stress; and perform yield prediction and forecasting (Kussul et al., 2014).

However, the USDA FAS with its GLAM system is currently the only provider of regular, timely, objective crop production forecasts at a global scale (Atzberger, 2013). This global monitoring system uses NASA's Moderate Resolution Imaging Spectroradiometer Rapid Response (MODIS RR) to monitor agricultural production (USDA, 2014).

Besides these monitoring systems, there are global scale studies that approached: the mapping of global cropland using multiple satellite sensor and ancillary data (P. Thenkabail et al., 2008; Pittman et al., 2010);and the global scale analysis of cropping intensity, crop duration and fallow land extent computed by using the global dataset on monthly irrigated and rainfed crop areas MIRCA2000 (Siebert et al., 2010). Whereas some local studies in Ukraine, addressed to: map and analyze changes of land management regimes (Kuemmerle et al., 2008;Baumann et al., 2011;Hostert et al., 2011;Alcantara et al., 2013;Stefanski et al., 2014); to assess efficiency of using satellite data for crop area estimation (Kravchenko & Moloshnii, 2012;Kussul et al., 2012; Kussul et al., 2014);and to forecast crops (Becker-Reshef et al., 2010;Kogan et al., 2013;Kussul et al., 2014;Kolotii et al., 2015;Franch et al., 2015).

Despite this information about crop production, and global and local studies, there are few reliable data on yield potential or water-limited yield potential for most major crop-producing countries, including data-rich regions such as the USA and Europe (M. Van Ittersum, Cassman, et al., 2013). Hence, the Global Yield Gap Atlas (GYGA) was created to provide best available estimates of the exploitable yield gap (M. Van Ittersum et al., 2013). But the European countries involved in the Global Yield Gap Atlas don’t include all Eastern Europe countries, except for Poland (M. Van Ittersum, 2013).

A global analysis of yield gaps across the world highlighted the wheat croplands of Eastern Europe as performing at about 50% of their climatic potential (Licker et al., 2010). However, this analysis was undertaken using data from the year 2000; its spatial resolution was too coarse for efficient location specific targeting of resources to boost yields; and it did not properly account for inter-annual variability in crop production (just averaged area harvested and yield data for the years 1997-2003).

All this suggests that the yield gaps across Ukraine are still poorly studied.

Chapter 2

This chapter includes general and specific objectives of the research and research questions.

2 Objectives and research questions

2.1 Objectives

2.1.1 General

The overall aim of this research is to develop a procedure to map Yield Gaps across Ukraine using time series of satellite derived biophysical variables, in order to identify underperforming croplands.

2.1.2 Specific

More specifically, this thesis aims to:

1. Produce Ukraine’s crop maps for wheat, barley, maize and sugar beet, through the disaggregation of published agricultural statistics by using time series of satellite MODIS derived biophysical variables;
2. Determine Actual crop yield or production through correlation with surrogate satellite derived biophysical variables, which are obtained with a phenology extraction algorithm;
3. Determine the Potential crop yield or production in each homogeneous edapho-climatic zone through the 90^{th} yield or production percentile;
4. Quantify the yield or production Gaps across the croplands of Ukraine through the difference between the Potential yields or production and the Actual yields or production.

2.2 Research questions

In order to achieve these enumerated specific objectives, the following research questions are stated:

Objective 1

- What phenology parameter best represent crop productivity?

Research questions

- What are and how many surrogate biophysical variable clusters represent crop distribution?

- What is the thematic accuracy of the crop maps?

Objective 2

- What is the level of adjustment of the predicted surrogate biophysical variable to the official yield or production statistics?

Objective 3

- What is the thematic resolution for soils and climate needed to suit the minimum crop environmental requirements?

Objective 4

- Is it possible to identify the Yield gaps across the croplands of Ukraine?

Chapter 3

This chapter describes the study area; the input data in terms of imagery, spatial thematic data and field data; the software tools used; and the methods for crop mapping, actual and potential yield calculation, yield gaps calculation and accuracy assessment.

3 Materials and methods

3.1 Study area

Ukraine, which is located between latitudes 44° and 53° N, and longitudes 22° and 41° E, is composed of 24 oblasts (Ukraine administrative division that can be translated into province or region) and the Autonomous Republic of Crimea, with the area of oblasts ranging from 8097 to 33,310 km2 (average area is approximately 24,000 km2). In general, Ukraine can be divided into the following agro-climatic zones (Figure 2): Plane-Polissya in the north (mixed forest zone, 26% of the entire Ukrainian territory), Forest-Steppe in the centre (34%) and Steppe in the south (the most intensive cultivated area, 40%) (Kogan et al., 2013).

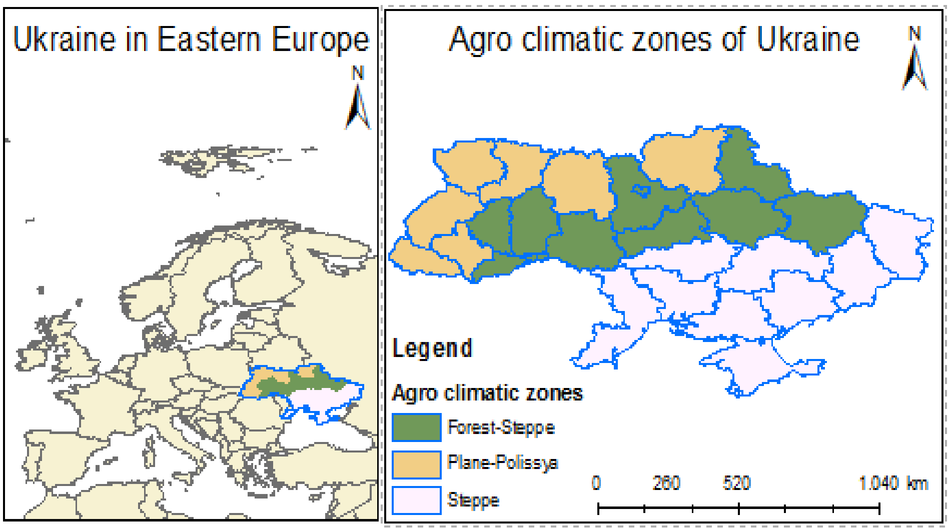


Figure 2: On left, relative location of the study area, Ukraine; on right, Agro-climatic zones of Ukraine.

From northwest to southeast the soils of Ukraine may be divided into three major aggregations: a zone of sandy podzolized soils; a central belt consisting of the black, extremely fertile Ukrainian chernozems; and a zone of chestnut and salinized soils (Kryzhanivsky, 2015).

Ukraine lies in a temperate climatic zone influenced by moderately warm, humid air from the Black sea. Winters in the west are considerably milder than those in the east. In summer, on the other hand, the east often experiences higher temperatures than the west (Kryzhanivsky, 2015).

Average annual precipitation in Ukraine is approximately 600 millimetres, including roughly 350 millimetres during the growing season (April through October). Amounts are typically higher in western and central Ukraine and lower in the south and east (WDC Ukraine, 2015).

3.2 Materials

3.2.1 Data

Table 1 summarises the data used in this research:

Type	**Material**	**Temporal frame**	**Spatial coverage /tiles**	**Spatial resolution/ scale**	**Source**
Images	MODIS MOD09A1 8day composites	2005-2014	h19v03, h19v04, h20v03, h20v04	500m x 500m	NASA
Spatial thematic	GlobCover	2009	Global	300m x 300m	ESA
	Administrative regions	2015	Ukraine	Oblast	www.diva-gis.org
	Climate data-maximum and minimum temperature and precipitation	1950-2000	16, 17	1km x 1km	www.worldclim.org
	SOVEUR Soil data	1988	Central and Eastern Europe	1:2.500.000	ISRIC World Soil Information
Field data	Crop official statistics-area, production & yield	2005-2013	Ukraine	Oblasts	State Statistical Committee of Ukraine
	Ground control points	Crowd source updated	Ukraine	73km x 111km spacing	www.geo-wiki.org

Table 1: Overview of the data used in this thesis.

Remote sensing imagery was MODIS MOD09A1 8-day surface reflectance composites, for a period ranging from March 2005 to December 2014, with a spatial resolution 500 meters, and provided by http://reverb.echo.nasa.gov/. The study area covers four MODIS tiles: h19v03, h19v04, h20v03 and h20v04. Each MOD09A1 pixel contains the best possible observation during an 8-day period as selected by high-observation coverage, low-view angle, the absence of clouds or cloud shadow, and aerosol loading. Data sets include reflectance values for Bands 1–7, quality assessment, and the day of the year for the pixel along with solar, view, and zenith angles (NASA, 2014).

All MODIS tiles were mosaiced to create an overview image of Ukraine, reprojected, masked with high quality data based upon MODIS QA to eliminate the obvious error noised data, Enhanced Vegetation Index (EVI) was calculated, and stacked by years. EVI is defined by:

$$\mathrm{EVI} = \mathrm{G} * \frac{(\mathrm{NIR}-\mathrm{Red})}{(\mathrm{NIR}+\mathrm{C1}*\mathrm{Red}-\mathrm{C2}*\mathrm{Blue}+\mathrm{L})} \qquad \text{Eq (1)}$$

Where Near-Infrared (NIR), Red, and Blue are atmospherically corrected (or partially atmospherically corrected) surface reflectance, and C1, C2, and L are coefficients to correct for atmospheric condition (i.e., aerosol resistance). For the standard MODIS EVI product, L=1, C1=6, C2=7.5 and G(gain factor)=2.5. EVI is less sensitive to soil and atmospheric effects than NDVI and simultaneously remain sensitive to increases in canopy density beyond where NDVI becomes saturated because it includes in the equation blue spectral wavelengths (Huete et al., 2002).

Crop mask was built with GlobCover 2009 global land cover map (see Figure 3), which is based on observations from the 300 meters spatial resolution MERIS sensor on board of the ENVISAT satellite mission, with an overall thematic accuracy of 70% (Defourny et al., 2011), and was provided by http://dup.esrin.esa.it/page_globcover.php. This land cover map was resampled to the same spatial resolution of MODIS imagery (500 meters), reclassified according to "Irrigate crop"="Rainfed crop"="Mosaic crop"=1 and "all the rest"=0, and clipped with the country Ukraine shapefile.

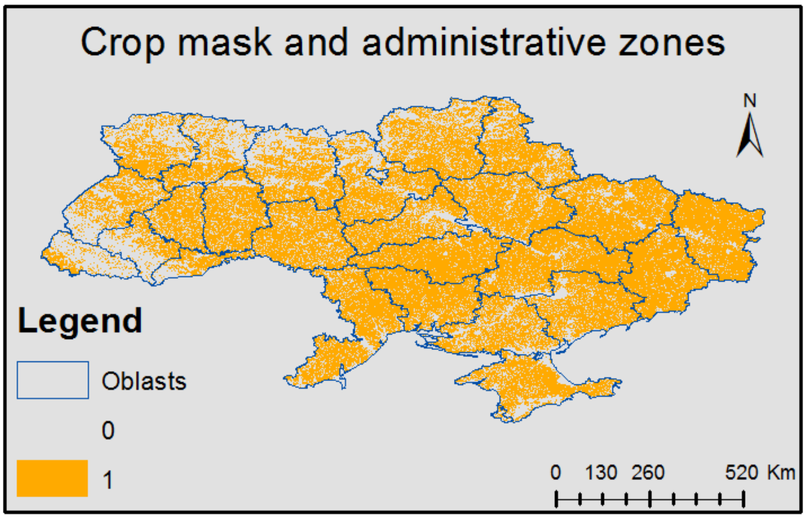


Figure 3: Ukraine crop mask based on reclassified and resampled GlobCover 2009 global land cover map. "Irrigate crop"="Rainfed crop"="Mosaic crop"=1 and "all the rest"=0.

Ground data with oblasts spatial resolution, yearly time resolution, a period ranging from 2005 to 2013 and provided by the State Statistical Committee of Ukraine (http://ukrstat.org/druk/publicat/kat_r/publ7_r.htm) were crop area (Ha), production (tonnes) and crop yield (tonnes/Ha). The crops used in this study were wheat, barley, maize and sugar beet. These official statistics were based on farm surveys collected from all the agricultural enterprises (large-scale farms that produce commodities exclusively for sale) which account for over 75% of Ukraine's grain production, and from a sample of household farms (small farms and household plots that produce crops both sale and for personal consumption) which account for the remainder of the grain production (Becker-Reshef et al., 2010).

Administrative regions of Ukraine, corresponding to boundaries of country and oblasts was a polygon shapefile provided by http://www.diva-gis.org/gdata (2015).

Climate data was raster imagery with current climate conditions (representative of 1950-2000) for maximum temperature, minimum temperature and precipitation with one square kilometre spatial resolution, and provided by http://www.worldclim.org/ (2015).

Soil data was a polygon shapefile dataset, scale 1:2.5 million, derived from the Soil and Terrain Database for Central and Eastern Europe (version 1.1)(SOVEUR), and provided by ISRIC World Soil Information (2015).

Ground control points (see Figure 4) were shapefile validation data for three global land cover datasets, which contains over 58000 features with a good spatial coverage of Ukraine, and provided by http://www.geo-wiki.org/ (2015).

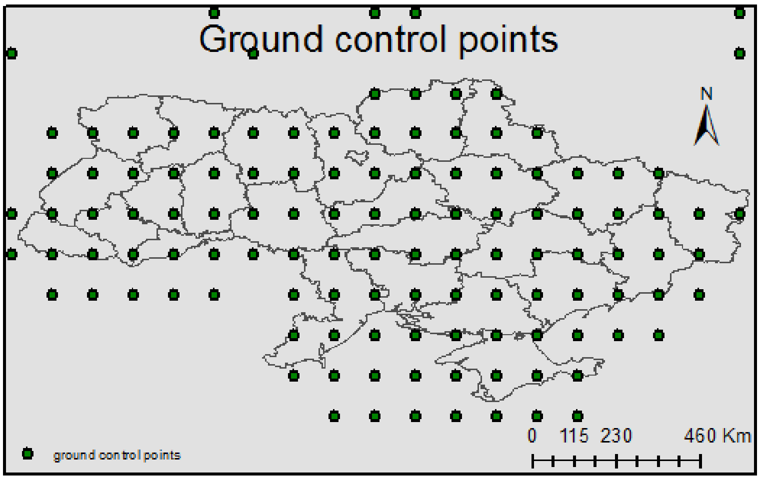


Figure 4: Location in Ukraine of the ground control points

3.2.2 Software used

The following software programs were applied within this research:

Software	**Usage**
ERDAS Imagine 2014	Unsupervised classification and image processing
ArcGIS 10.3	Data preparation, analysis and map
Matlab R2013	Image pre-processing and phenology extraction
SPSS Statistica 22	Stepwise linear regression analysis
MS Excel 2007	Data preparation and statistical analysis

Table 2: List of software used

3.3 Methods

A general overview of the followed workflow is illustrated in Figure 5. Detailed description of all the followed steps in the methodology is provided in the following sections (3.3.1 - 3.3.5).

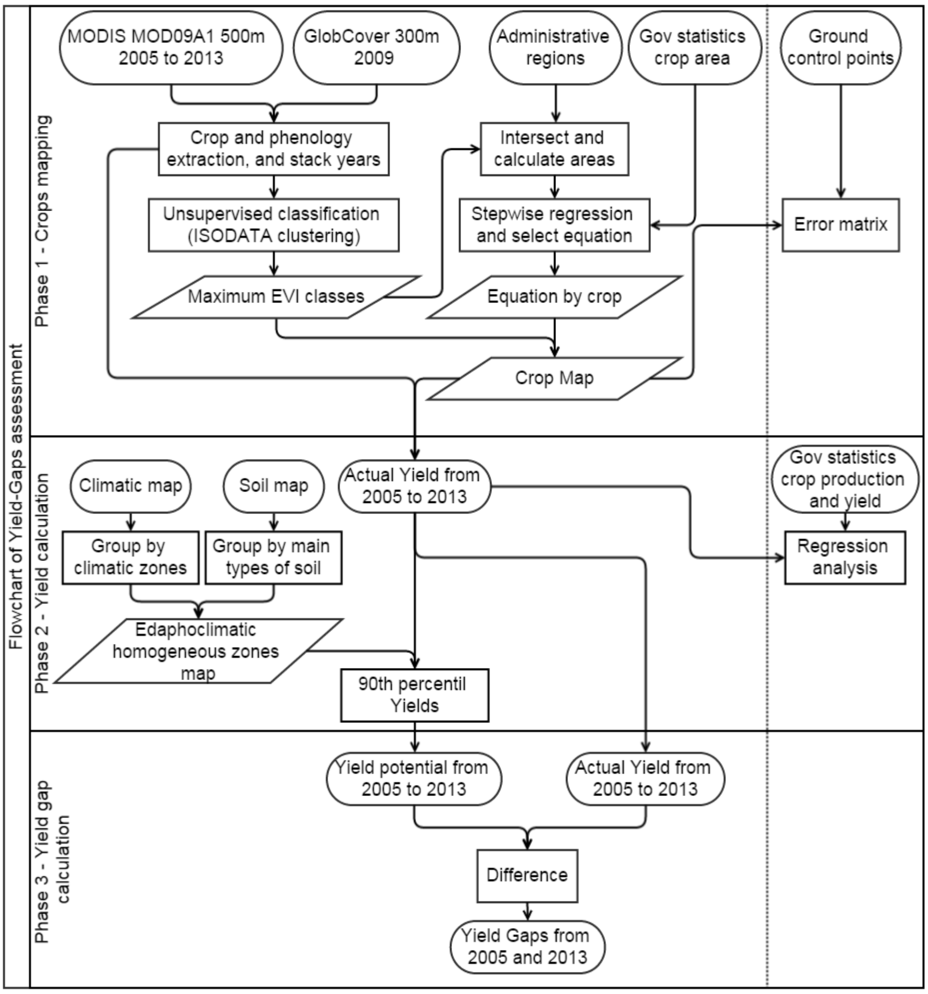


Figure 5: Flowchart of methods. The column on the right illustrates the accuracy assessment.

3.3.1 Phenology extraction

Cropping season was detected within each pixel using an adapted version of Dash et al, 2010: first a cleaning algorithm was used to remove missing data values from the original data and create a flag depending upon the quality of the temporal information available in each pixel; after, a phenology extraction algorithm was applied to smoothen the profile with a Fourier smoothing algorithm, search iteratively the phenology profile both for a peak of EVI, as well as for the start and end of season using first derivative, and finally perform cumulative sum or integrated growing season EVI. Thus, for each pixel, the algorithm output include start, end and maximum EVI dates in growing season as well as surrogate measures of ecosystem productivity such as maximum EVI and integrated EVI (Duncan et al., 2014).

To assess which phenology parameter best adjusts to crop productivity, a regression analysis was done between the sum of the phenological surrogate yield parameters (I_EVI and maximum EVI) for each oblast and year, and the official statistical crop data (yields and production), per corresponding oblast and year.

Crop mask was multiplied with the previously choosed output of phenology extraction parameter in order to select just crop areas.

3.3.2 Crop mapping

The methodology for crops mapping was adapted from (Khan et al., 2010).

The surrogate biophysical variable values were reported as digital number (DN) values, ranging between 0 and 255, using the following equation:

$$\mathrm{DN} = \frac{(\mathrm{EVI}+0.1)}{0.004} \qquad \text{Eq (2)}$$

EVI images from 2005 to 2008, 2010 and 2012 were gathered into one stack. The other years of the time range 2005-2013 were not included due to their abnormal loss of information (pixels) in relation to the crop mask.

The stack was then processed in ERDAS Imagine (Intergraph, 2014) using the Iterative Self-Organizing Data Analysis Technique (ISODATA) clustering algorithm in order to reduce the amount of data. Many unsupervised classification runs were carried out to generate maps with between 8 to 100 clusters. The maximum number of iterations of each unsupervised classification was 50 and the divergence threshold was set to 1, which were proved useful for optimal classification results in studies like for example's Khan et al., 2010; De Bie et al., 2012; Ali et al., 2013. The ISODATA algorithm tries to minimize the Euclidian distance to form clusters. Basically, this clustering method uses spectral distance and iteratively classifies the pixels into cluster mean vectors, redefines the criteria for each cluster or class, and classifies again, until the "change" between the iteration is small. This algorithm further performs splitting and merging of clusters. Clusters are merged if either the number of members (pixel) in a cluster is less than a certain threshold or if the centers of two clusters are closer than a certain threshold (Intergraph, 1997). In the end, the ISODATA algorithm provides by cluster or class an EVI-profile that contains information on past performance and cover changes.

The results of the different unsupervised classification runs are compared using the divergence separability which is a statistical measure of distance between the mean cluster vectors (Landgrebe, 2003); the 'best' number of clusters is the one corresponding to the run having the highest minimum and/or average divergence (Bie et al., 2010). A graphical presentation of these separability statistics was used to select which map produced, having 'what' number of pre-defined classes, is the map of choice.

The output of this unsupervised classification was an EVI map and intermediate legend that consists only of clusters that represent EVI profiles showing changes in vegetation greenness over time which is assumed to relate to the types of land cover and land use present. Once the number of clusters is known, the EVI profile clusters map was established.

The clusters map, was converted into a polygon shapefile. Using GIS spatial analysis functions from ArcGIS (ESRI, 2009), the oblasts and the maximum EVI profile clusters map were intersected to determine the respective areas (Hectares - Ha) of each EVI profile cluster per oblast. Weighted average for crop areas (Ha) from agricultural official statistics was calculated with more weight for 2013 data and decreasing weight in the direction of the beginning of the study time range, in order to adjust for trends in areas over the years.

The cluster areas were further used as explanatory or independent variables, in the stepwise linear regression (Neter et al, 1996), with the cropped areas (Ha) from agricultural statistics by season, crop, and oblast as dependent variable:

$$\mathrm{CA} = \sum_{i=1}^{n} \mathrm{C_i} * \mathrm{EVICluster_i} \qquad \text{Eq (3)}$$

With CA representing official cropped area (Ha) by oblast and $EVICluster_i$ representing the area (Ha) of the i^{th} EVI profile cluster.

Stepwise linear regression essentially does multiple regressions a number of times, each time removing the weakest correlated variable. At the end it is selected the variables that best explain the distribution (IBM Corp., 2013).

No constant was considered in the regression and the coefficients C_i were constrained to the 0 – 1 range in order to determine the estimated fraction or percentage of total area of a given EVI profile cluster where a given crop was grown at a specific oblast and season. Once the regression coefficients were estimated, the above equation 3 was used to generate maps showing cropped fractions by map units. Statistical computations were done using the SPSS Statistics 22 software (IBM Corp., 2013).

3.3.3 Yield calculation

3.3.3.1 Actual yield

Crop maps generated in the previous chapter were intersected with the maximum EVI maps to assess Actual yield for each pixel with crop.

Regression analysis for each type of crop between maximum EVI and official production (tonnes) and yield (tonnes/Ha) statistics, as well as respective level of adjustment R square were computed to assess which official parameter is best represented by the predicted surrogate biophysical variable maximum EVI.

3.3.3.2 Potential yield

First, the study area was divided into homogeneous climate zones. The stacked climate data images (maximum and minimum temperature, and precipitation) were classified with an unsupervised classification method, the ISODATA clustering algorithm. A series of classification runs corresponding to different number of clusters (5 to 15) were used. The maximum number of iterations was 50 and the divergence threshold was 1, which were proved useful for optimal classification results in studies like for example's Khan et al., 2010; De Bie et al., 2012; Ali et al., 2013. A graphical presentation of the separability statistics was used to select which map produced, having 'what' number of pre-defined classes of climatic zones, was the map of choice. Then, the chosen map was converted into a polygon shapefile.

The symbology of the soil map was set according to the Revised Legend of the Soil Map of the World (FAO, 1988) in the SOVEUR project. Symbology was reclassified by grouping soils with the same major class. After, the map was intersected with the climate zone polygon shapefile, giving homogeneous edapho-climatic zones.

Finally, it was determined for each homogeneous edapho-climatic zone the 90^{th} percentile value within the range of actual yield, which corresponds to the Potential Yield (M. Van Ittersum et al., 2013). Values above the 90^{th} percentile were not used to define the maximum in order to avoid erroneous or over-estimated values that may have been included in the yield datasets (Licker et al., 2010).

3.3.4 Yield gap calculation

The differences between the Potential yield levels and actual farmers' yield define the yield gaps (M. Van Ittersum et al., 2013). Crops yield gaps fraction was also calculated with the equation 4 (Licker et al., 2010).

$$\text{Yield Gap Fraction} = 1 - \frac{\text{Actual Yield}}{\text{Potential Yield}} \qquad \text{Eq (4)}$$

The yield gap fraction (a value from 0 to 1) tells us how close to the edapho-climatic potential any given location may be. Those places with a low yield gap (close to zero) have yields at or near their climatic potential.

Crop heat maps were produced for visualization of crops pattern distribution.

3.3.5 Accuracy assessment

Predicted sum of surrogate biophysical variable/crop type/oblast was compared to governmental production statistics/crop type/oblast through a regression analysis and the respective output level of adjustment R square.

For overall thematic accuracy assessment, a confusion matrix (or error matrix) between the predicted cropland areas derived from remote sensing data and ground control points was done, and it was calculated the overall thematic accuracy, i.e., the ratio between the correctly classified land cover area and the total classified land cover area. The level of thematic detail of the ground control points doesn't inform about the type of crop (wheat, barley, etc). Thus, the land covers related with crops (which were summed up for the calculation of overall thematic accuracy) from the ground control points, were:

- Cultivated and managed areas
- Mosaic Cropland/grass or shrub or forest
- Rainfed croplands
- Mosaic: Cropland/Shrub or Grass Cover
- Mosaic: Cropland/Tree Cover/natural vegetation
- Cropland/natural vegetation
- Mosaic grass or shrub or forest/Cropland
- Croplands

In consequence, it was just assessed whether unspecific croplands were well located in the map.

Chapter 4

This chapter describes in graphics, tables and maps the outputs for each step of the methods.

4 Results

4.1 Phenology extraction

The phenology parameters extracted from remote sensing MODIS imagery were start, end and maximum EVI dates in growing season as well as the surrogate measures of ecosystem productivity maximum EVI and integrated EVI. Figure 6 and Figure 7 illustrate the linear regression between surrogate measures of crop productivity (maximum EVI and integrated EVI) and grain production.

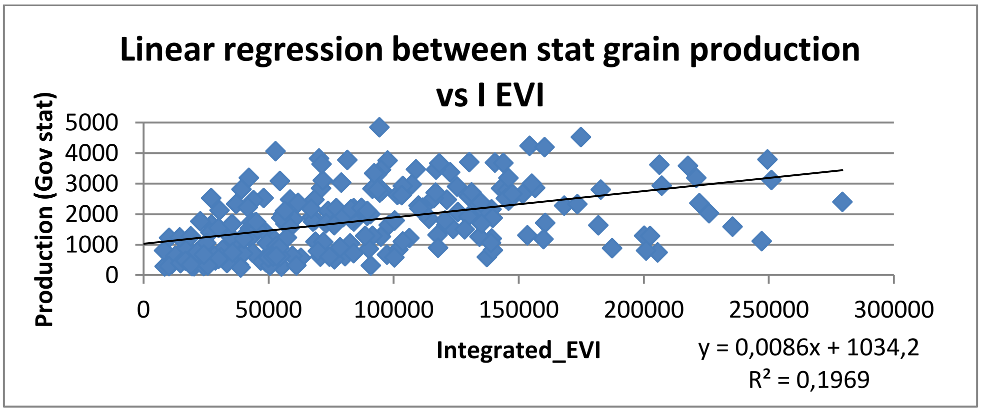


Figure 6: Regression analysis between sum of the I-EVI values for all oblast and years and the statistical data for the grain yields (centners (tons) per 1 ha of the harvested area). Grains were wheat, barley and maize.

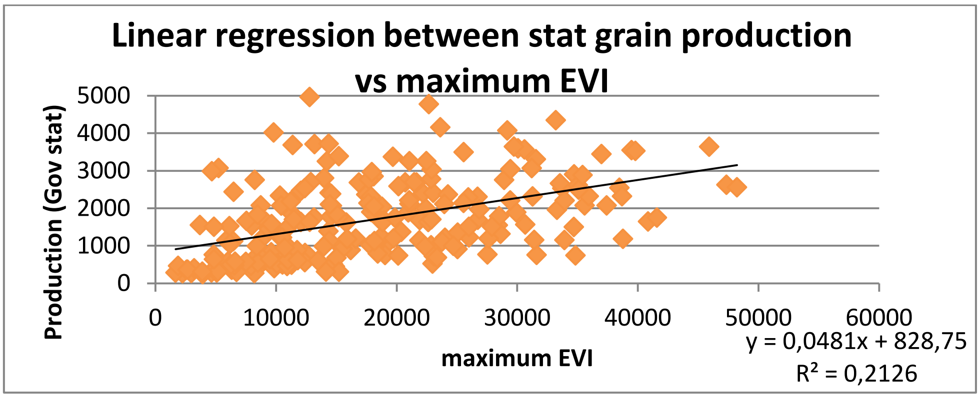


Figure 7: Regression analysis between sum of the maximum EVI values for all oblast and years and the statistical data for the grain yields (centners (tons) per 1 ha of the harvested area). Grains were wheat, barley and maize.

Maximum EVI represented slightly better grain production, with an adjustment R square of 0.2126. Furthermore, maximum EVI was the only phenological parameter that allowed to output regression equations for all study crops, which results are shown in the next section. Thus, maximum EVI was the phenology parameter used for the following step, crop mapping.

4.2 Crop mapping

4.2.1 Intermediate legend and cluster map

Figure 8 illustrates the separability analysis performed for each one of the maps containing between 8 and 100 clusters or classes, obtained in the ISODATA unsupervised classification. The choice of the number of clusters or classes presents almost always a no-win solution between: (i) keep the number of classes low to gain maximum data-reduction, and (ii) optimize separability between classes without information loss. Thus, the cluster map with 19 classes seemed to have good separability between classes because it had the highest average distance between clusters and the minimum separability is greater than 24, which is choice criteria as suggested by Erdas (1997) and De Bie et al.(2012).

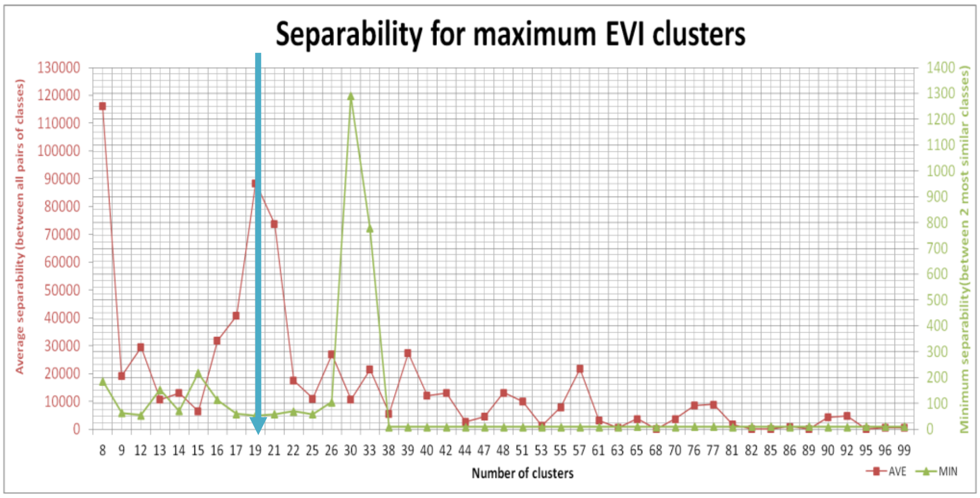


Figure 8: Graph of average and minimum cluster distance between 8 and 100 clusters. Those separability statistics indicate how different clusters or classes are between each other.

The maximum EVI map (Figure 9) and intermediate legend consists only of maximum EVI-profiles that are indicative of the crops location.

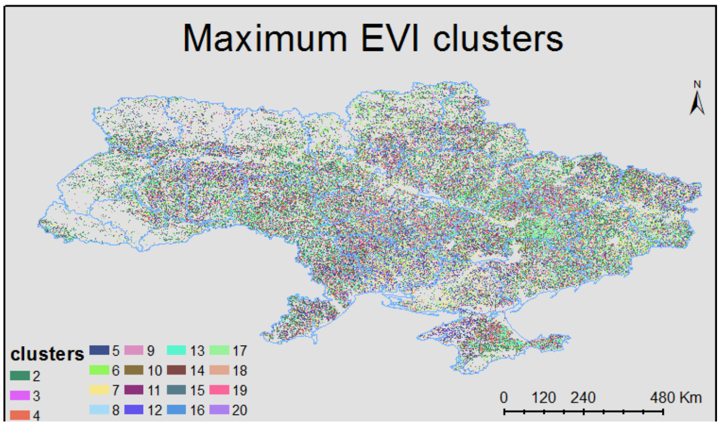


Figure 9: Map with 19 clusters or classes.

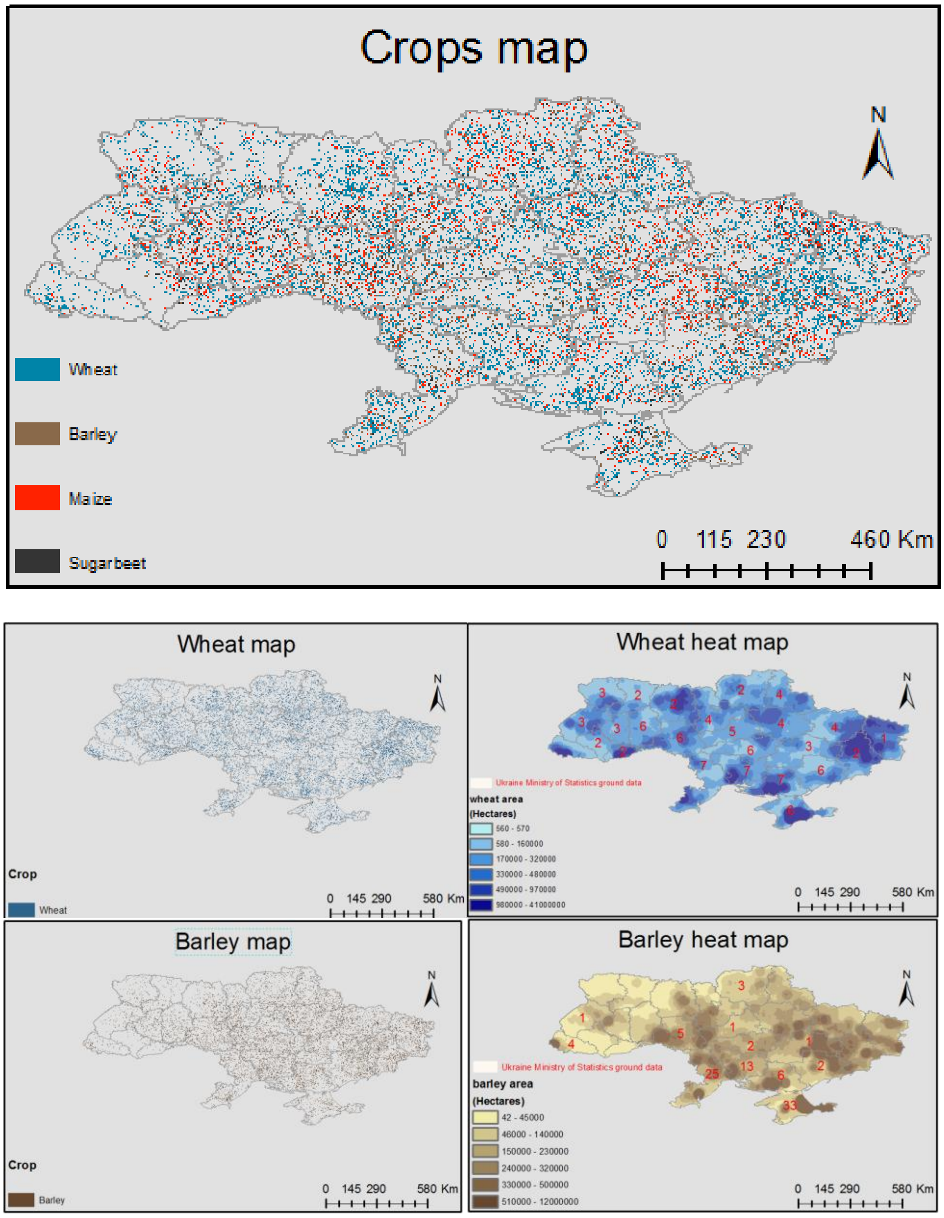
4.2.2 Crop maps

Table 3 shows the chosen clusters or classes; their coefficients; level of adjustment to the response variable; and crop regression equations resulted from the stepwise linear regressions between the explanatory variables area of pixels with a specific maximum EVI profile (clusters) and the response variables area of crop from governmental statistics. All the regression parameters for all the crops are significative (p≤ 0.05).

Crop	Maximum EVI classes (cluster predictors)	Coefficients	Adjusted R square	Crop linear regression equations
Wheat	Clt 2*; Clt 7**	0,689; 0,758	0,871	Wheat_area (ha)=clt2*0,689+clt7*0,758
Barley	Clt 7****; Clt 13***	0,364; 0,949	0,901	Barley_area (ha)=clt7*0,364+clt13*0,949
Maize	Clt 6***	0,874	0,722	Maize_ area(ha)=clt6*0,874
Sugarbeet	Clt 3***	0,462	0,474	Subarbeet_area (ha)=clt3*0,462

Table 3: Summary results of the stepwise linear regression for the different crops, and respective equations; each Maximum EVI class has the corresponding coefficient in the same writing order; * = Signifcant at P= 0.001; ** = Signifcant at P= 0.005; *** = Signifcant at P= 0.000;**** = Signifcant at P= 0.015; Regression parameters were considered to be significant at the p ≤ 0.05 level of significance. Independent variables that were significant at the p ≤ 0.05 level of significance were retained in the model. The adjusted R square compares the explanatory power of regression models that contain different numbers of predictors (Frost, 2013). All linear Regressions are through the Origin. Dependent variables are crop_area (ha).

The application of the crop linear regression equations resulted in the following maps.



Crops map
N
Wheat
Barley
Maize
Sugarbeet
0 115 230 460 Km
Wheat map
N
Crop
Wheat
0 145 290 580 Km
Wheat heat map
N
Ukraine Ministry of Statistics ground data
wheat area
(Hectares)
560 - 570
580 - 160000
170000 - 320000
330000 - 480000
490000 - 970000
980000 - 41000000
0 145 290 580 Km
Barley map
N
Crop
Barley
0 145 290 580 Km
Barley heat map
N
Ukraine Ministry of Statistics ground data
barley area
(Hectares)
42 - 45000
46000 - 140000
150000 - 230000
240000 - 320000
330000 - 500000
510000 - 12000000
0 145 290 580 Km

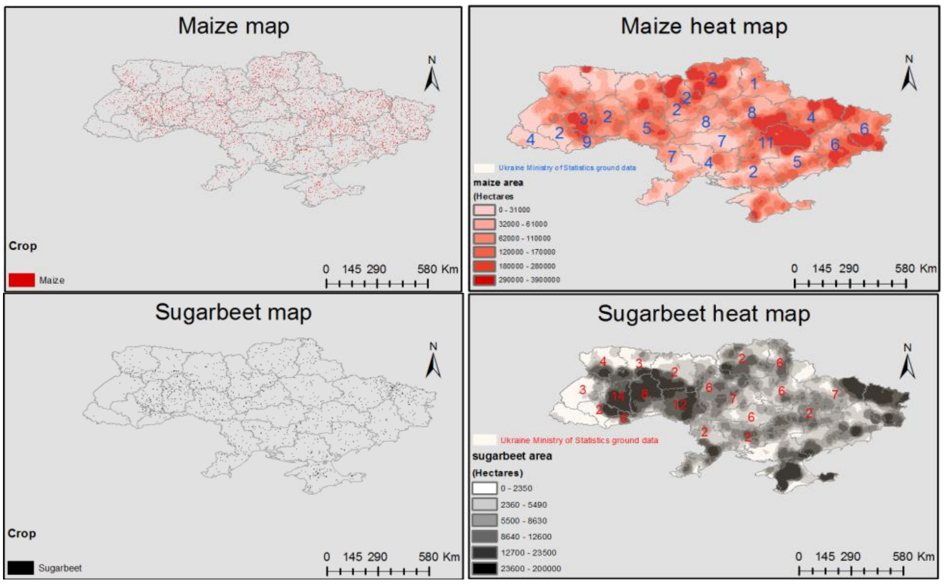


Figure 10: Combined crop map at top, specific crop maps and respective crop heat maps. The heat maps were produced for the sake of easier crop density visualisation. The red numbers in the heat maps indicate percentage of crop national total each oblast contributes to national area. Oblasts not numbered contribute less than 1% to the national total. These numbers are based upon averaged oblast-level data from the year 2000, obtained from the Ukraine Ministry of Statistics.

To assess the thematic accuracy of the crop maps, which is linked to the accuracy of crop type location, it would be necessary to have updated and statistically representative ground truth data of the studied crops to allow comparison in a matrix error, which unfortunately we didn't have available. Anyway, the result of the thematic accuracy for the undifferentiated crop map (that in the end corresponds to the crop mask), was 74.1%.

4.3 Yield calculation

Figure 11 illustrates the separability analysis performed for each one of the climate maps containing between 6 and 13 clusters or classes, obtained in the ISODATA unsupervised classification.

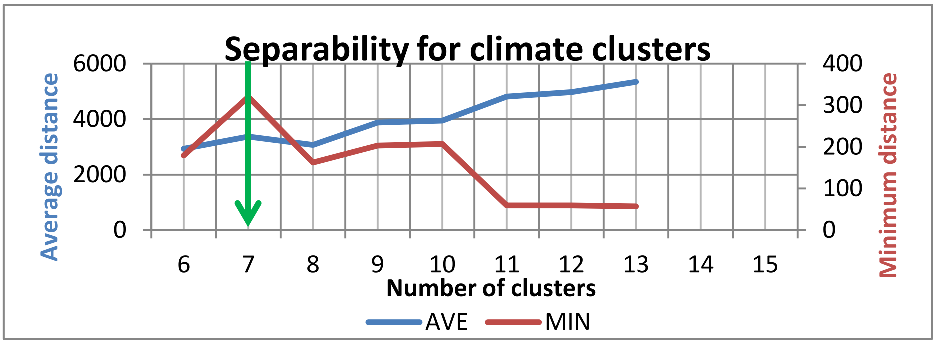


Figure 11: Graph of average and minimum cluster distance between 6 and 15 clusters. Those separability statistics indicate how different clusters or classes are between each other.

The cluster climate map with 7 classes seemed to have good separability between classes (highest minimum distance between clusters, and above 24 (requirement as suggested by Erdas (1997) and De Bie et al.(2012)), and represented a reasonable amount of climate information needed for the calculation of potential yield.
Figure 12 shows the edapho-climatic map resulted from the intersection of the cluster climate map with the soil map.

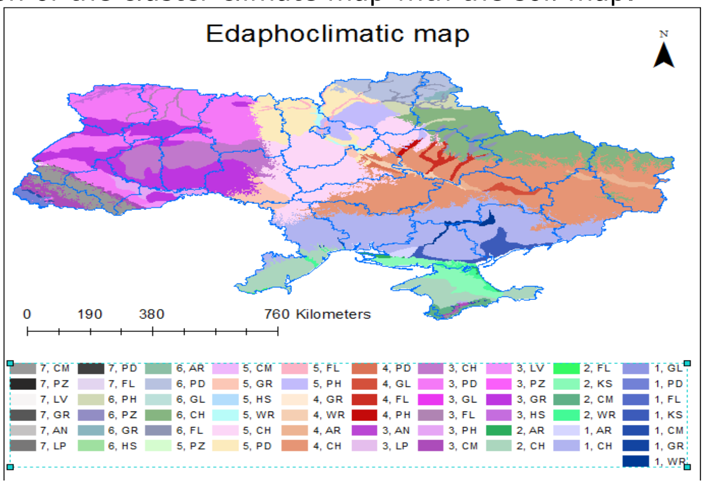


Figure 12: Edapho-climatic map of Ukraine

Table 4 shows that in all the studied years, except for barley in 2007, the governmental statistics that were best represented by the surrogate biophysical variable maximum EVI were the production. Overall, Barley has the best prediction regression models for production, with levels of adjustment greater than 0.42 (except for the year 2013). Whereas in the other crops, the prediction variable maximum EVI had small correlation with production in all the study years, with figures bellow 0.35 (except for the Wheat weighted average that was 0.4383).

R square	Wheat		Barley		Maize		Sugarbeet		Prod.> Yield?
	Prod.	**Yield**	**Prod.**	**Yield**	**Prod.**	**Yield**	**Prod.**	**Yield**	
2005	0,2846	0,1029	0,688	0,1791	0,2455	0,0014	0,2114	0,0024	Yes
2006	0,1763	0,0357	0,6046	0,0259	0,2344	0,0391	0,1928	0,0382	Yes
2007	0,3319	0,0034	0,4919	0,5753	0,2777	0,0213	0,2334	0,0166	No
2008	0,3033	0,2947	0,6159	0,0854	0,287	0,0888	0,1164	0,0003	Yes
2009	0,3077	0,0051	0,4701	0,2106	0,2035	0,046	0,0936	0,0005	Yes
2010	0,1342	0,0003	0,4235	0,198	0,2195	0,1558	0,186	0,00004	Yes
2011	0,187	0,1016	0,4238	0,3361	0,2598	0,002	0,1702	0,0042	Yes
2012	0,295	0,0103	0,5169	0,2979	0,2527	0,039	0,1633	0,0051	Yes
2013	0,2667	0,0347	0,3757	0,1531	0,2691	0,0243	0,1763	0,0538	Yes
Weighted average	0,4383	0,2777	0,5222	0,02	0,267	0,0437	0,2276	0,1768	Yes

Table 4: Correlation parameters R square resulted from the regression analysis between the governmental statistics crop production or crop yield, and the crop prediction value of Maximum EVI, for each year of the study time range and for the weighted average of the R square values (with more weight for 2013 data and decreasing weight in the direction of the beginning of the study time range, in order to adjust for trends in areas over the years). Prod. = Production.

4.4 Yield gap maps

Figure 13 illustrates the yield gaps fraction map for all the studied crops – wheat, barley, maize and sugar beet.

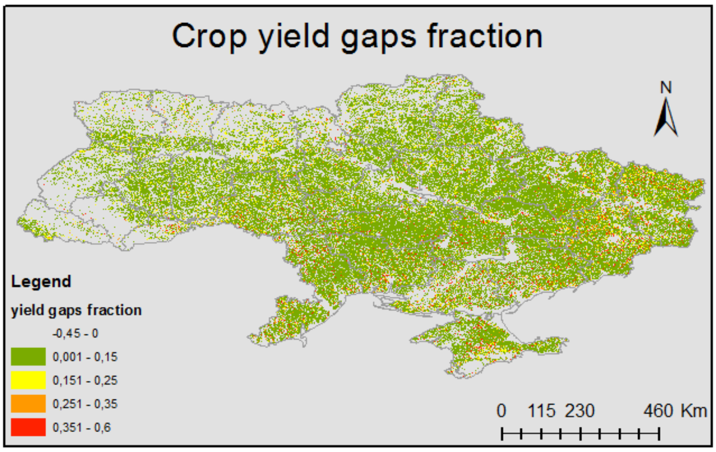
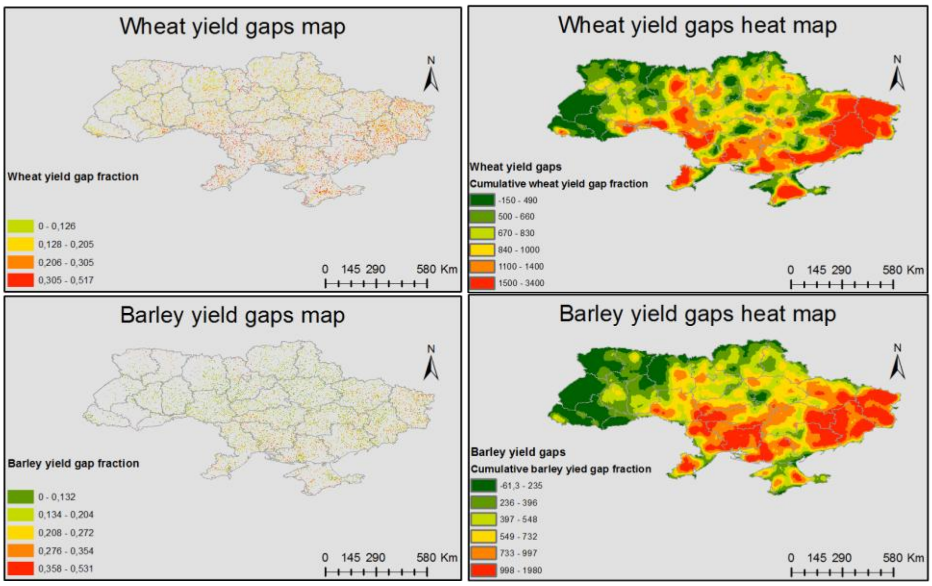


Figure 13: General crop yield gaps fraction map

The yield gap maps and heat maps with yield gap intensity are shown below. The color intensity of the heat maps is the cumulative (sum of pixels in each patch) yield gap fraction. The yield gaps for all the crops tends to be higher for the grains in the steppe zone, mainly in the eastern part of Ukraine, except for sugar beet where the yield gaps are higher in the western part of the forest-steppe.



Yield gap maps

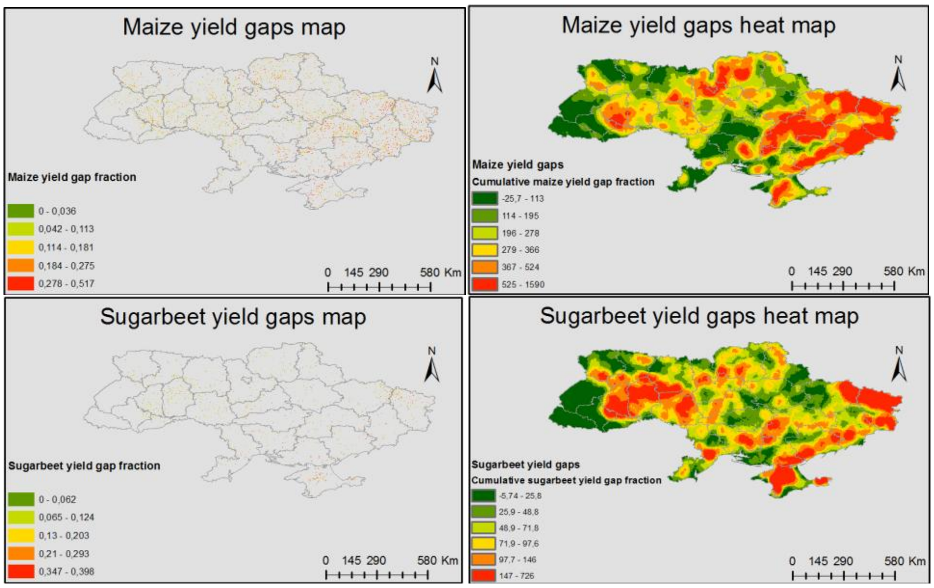


Figure 14: Specific crop yield gaps maps and respective heat maps with spatial pattern distribution of crops. The yield gaps heat maps are for the sake of easier visualisation. The more red it is the pixel, the greater is the yield gap.

Chapter 5

5 Discussion

In this thesis, four different crops were mapped in Ukraine– wheat, barley, maize and sugar beet - ,estimated their actual crop yield and the climatic potential crop yield. The difference between actual yield and the potential yield resulting in ‘yield gap’ was calculated for all the crops.

To get this ‘yield gap’ on a large study area like Ukraine, it was necessary to acquire large amounts of data (big data), such as remote sensing time series imagery and ground statistical data, and follow a time and resource consuming data processing and analysis procedure which included organizing data in mosaic and stacks, cleaning, phenological information extraction from the remote sensing data, iterative unsupervised classification, and disaggregating and mapping ground statistical data using the referred remote sensing time series. Furthermore, throughout this whole processing flow, there was the accumulation of error that may turn difficult the decision making process (Figure 15).

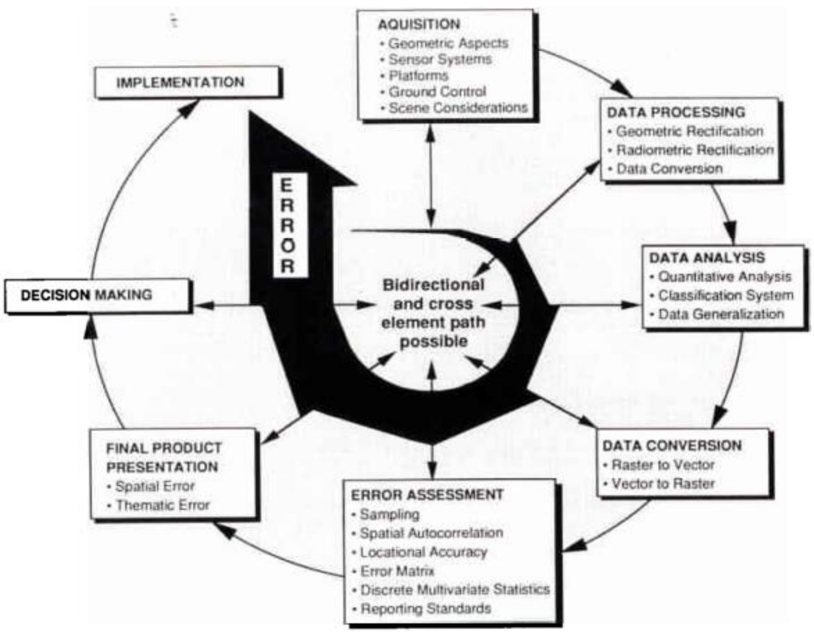


Figure 15: The accumulation of error in a “typical” remote sensing information processing flow (adapted from Lunetta et al., 1991).

Yield gap maps

The initial idea in this study (see Figure 16, flowchart) was to also use higher spatial resolution (30 meters) Landsat imagery, in order to be able to locate individual crop fields. This landsat imagery would be used in the mega file data cube (MFDC), along with time series MODIS imagery and ancillary data such as precipitation, temperature and terrain elevation. This MFDC would be loaded in the Automated Cropland Classification Algorithm (ACCA), which is an iterative decision tree based algorithm for crop mapping (Thenkabail & Wu, 2012). The threshold of ACCA rules (e.g., MODIS August NDVI ≥200; Figure 16, decision tree b) was written based on all available knowledge (or through trial and error), such as growing season, to capture as much cropland area and as many characteristics as possible, until 90% of the ACCA-derived cropland matched pixel-by-pixel with the truth cropland data layer. As the algorithm is further developed, greater complexity in rules/codes and larger number of datasets are involved in further delineating pure cropland areas from non-croplands. According to these authors, the ACCA algorithm computes total cropland areas as well as irrigated cropland areas consistently, rapidly (less than one hour of computer processing) and accurately, year after year. Despite these advantages of the ACCA algorithm, the development of rules through trial and error may be an overwhelming process that requires lots of expertise and ground information.

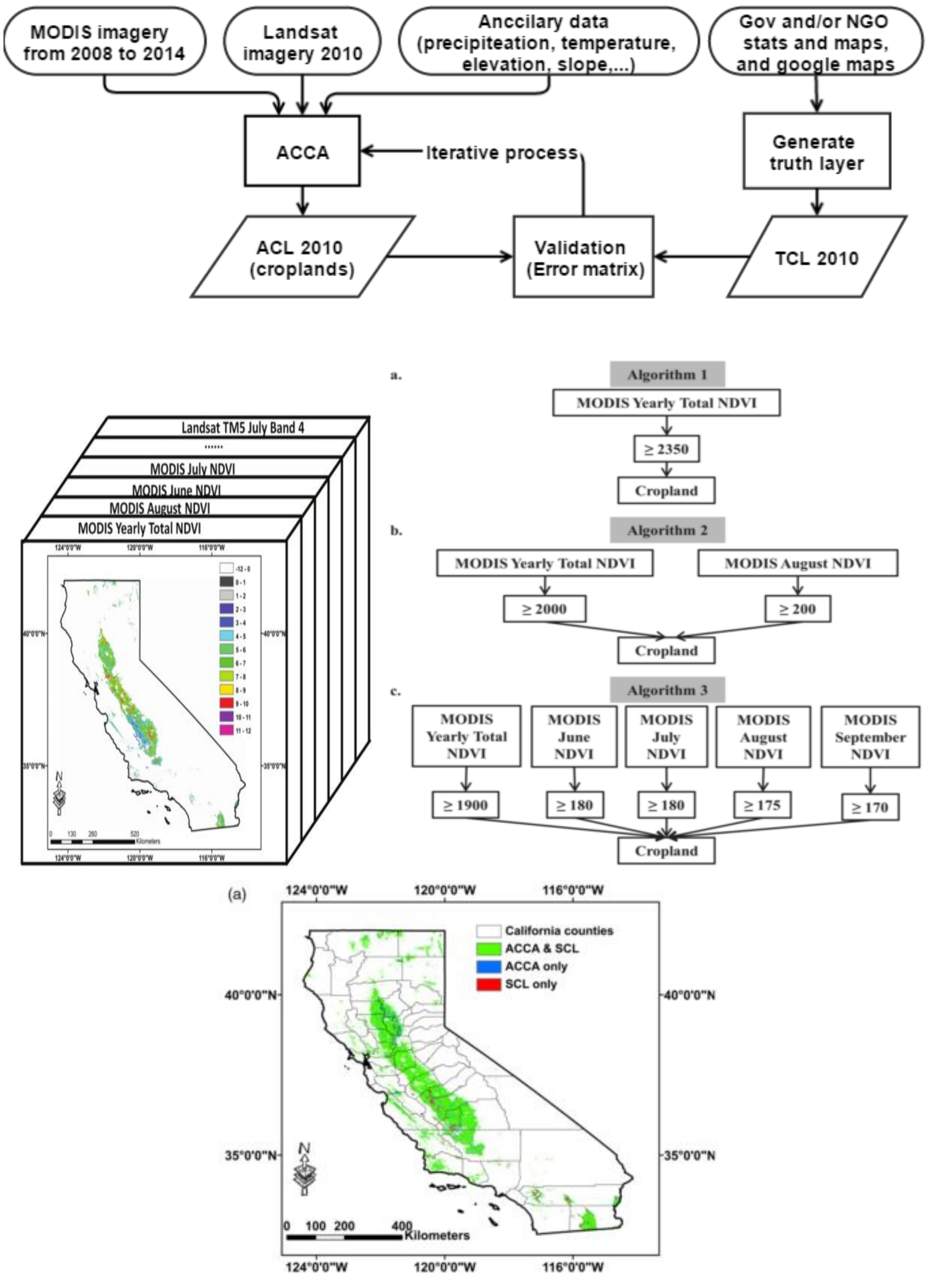


Figure 16: Crop mapping using the Automated Cropland Classification Algorithm. Upper diagram shows resumed flowchart for this procedure; below (adapted from Thenkabail & Wu (2012)) and starting from top left, an example that illustrates the mega file data cube (MFDC), the decision trees algorithm applied to the MFDC, and the resulted cropland map.

The pre processing of Landsat imagery required atmospheric and radiance correction, as well as mosaicing of large amount of images, in order to have consistent land cover characteristics throughout the study area Ukraine, which would allow crop field classification across Ukraine. Techniques for mosaicing include global colour balancing (so all scenes appear to be in the same colour range), feathering, cutlines along linear features etc. - all designed to create the illusion of one big seamless image. However the result of this pre processing, would normally be treated as a visual product and probably not suitable for analysis work such as the following procedure for image classification. Furthermore, it was resources, time and expertise consuming to perform such Landsat pre-processment for the short time scope of this study. So instead, we gave up from the ACCA algorithm that required as input the Landsat imagery and opted, despite the spatial resolution of 300 meters, to use a land cover classification map, the Globcover 2009, to be able to locate with an accuracy of 70% (Defourny et al., 2011) the crop fields. The higher spatial resolution Corine Land cover map doesn´t include Ukraine (Commission of the European Communities, 1995). Again, in this data acquisition phase, there was some intrinsic error due to the spatial resolution of the imagery in relation to the crop fields of Ukraine.

5.1 Phenology extraction

After locating the crop fields, the next general step was to know what kind of crop each field had. To achieve this general step, it was extracted from MODIS remote sensing imagery the crop phenology parameters per pixel and chosen the phenology parameter that best represent crop yield or production.

The first version of phenology extraction algorithm (Dash et al, 2010) used in this study had a cleaning algorithm for atmospheric correction based on a temporal moving average window function, which produced a lower than expected phenology extraction retrieval in relation to the crop mask. Thus, instead, it was used an improved version with the cleaning process based on a Fourier smoothing algorithm, which smoothens data using a sum of weighted sine and cosine terms of increasing frequency. However, according to the De Bie et al. (2012), the use of a Fourier algorithm, assumes that behavior between years remains stable, averaging out changes in cropping patterns or heavy variability in weather patterns. Nevertheless, a difference analysis between the two extraction algorithm versions outputs showed that there was an increased retrieval with the Fourier algorithm.

Cumulative sum, or integrated vegetation index (VI) values, and maximum VI values are used commonly as surrogate measures of vegetation productivity and crop yield (Pettorelli et al., 2005; Funk & Budde, 2009; Vrieling et al., 2011; Rembold et al., 2013 as cited in Duncan et al., 2014). Integrated EVI would be the best phenology parameter choice, since vegetation index values post-peak growing season often provide more accurate predictions of crop yield as they correspond to the reproductive and grain-filling development stages of cereal crops (Funk & Budde, 2009; Rojas et al., 2011 cited in Duncan et al., 2014). Duncan et al. (2014) found that integrated-EVI was significantly correlated with district-wise wheat crop yield and production during their study time, with an R square value for the integrated-EVI crop yield model of 0.6. However, in this study, the regression analysis between the phenological parameters maximum EVI and integrated EVI, and the ground statistical crop data resulted in a slightly better adjustment of the maximum EVI to the ground statistical data.

5.2 Crop mapping

After extracting the phenology and choosing the surrogate VI maximum EVI for crop yield, the data was stacked and the pixels VI profiles were iteratively classified using an ISODATA clustering algorithm. This clustering procedure, despite being time consuming (about more than one day of computing) allowed reducing the enormous amount of data in the stack (big data), and further identifying and mapping land cover gradients based on hyper temporal VI profiles similarities. Thereby, according to Ali et al. (2014), hyper-temporal imagery is also found effective at mapping the spatial patterns in vegetation cover that represent gradual changes in the form of gradients, which are originated due to the local vegetation seasonal trends.

The ISODATA clustering iteration was set to produce cluster maps with between 8 and 100 clusters. The output analysis of separability didn’t have all the required results because of an undetermined problem in the ERDAS software. After the 38 clusters map, the minimum separability consistently was below 24, so we can assume that the cluster map can’t have more than 38 clusters (Erdas, 1997; De Bie et al., 2012). Before that figure, there were missing 15 cluster maps due to the referred software error, but with the large average separability peak in 19 clusters map we can assume a reasonable separability for this number of clusters. This uncertainty brings more error factor to the whole process.

The analysis of separability with the remaining results for Maximum EVI clusters typically fits a heterogeneous landscape, i.e. where very clear spatial partitions exist in cover and use, corresponding to clear peaks in the average separability graph, similarly to the results of De Bie et al., 2012. This is because some zones of Ukraine, such as the Plane-Polissya zone, has a more complex landscape with forests, grasslands and abandoned fields, and less winter wheat crop area comparing to other zones. Furthermore, there is temporal variability within each pixel because farms in Ukraine employ a variety of crop-rotation schemes, some including four or more crops, some only two (Rogovska, 2009).

To improve the process of selection of clusters, decision trees with crop calendar criteria for crop phenology characteristics such as peak of EVI, and start and end of growing season, could have been tested in this thesis after the ISODATA unsupervised clustering classification and separability analysis, and in case we had more study time available. Thenkabail & Wu (2012) used with success decision tree algorithms in the ACCA process (see page 37 of this thesis) to resolve mixed classes derived from unsupervised classification.

5.3 Yield calculation

As to the separability analysis of the climate data, despite the number of iterations was less, the graph shows through the low peaks that climate data is typically characterized for smooth transitions and gradients.

The regression analysis between the variable maximum EVI and the official crop production and yield showed that maximum EVI is generally better adjusted to the crop production values. These results, along with the good level of fit of the choosen cluster areas to the official crop area, suggest that the prediction variable may be well estimated with the adjusted R square values greater then 0.70 (except for sugar beet), but their location based on the maximum EVI values are probably wrong. The only crop that may be closer to a correct location and area estimation is barley, which had an adjusted R square of 0.901 for the cluster regression, and levels of adjustment of the prediction maximum EVI to the explanatory variable production generally greater than 0.42. In other words, the barley biomass (production) is well represented by the surrogate biophysical variable maximum EVI (with a specific location) and the barley crop areas are also well explained by the respective clusters areas.

The R square results for wheat yield prediction were in overall lower than 0.3, which is lower than Kogan et al. (2013) results on winter wheat yield forecasting study in Ukraine. This study was based on a regression model that uses as predictor 16-day NDVI (vegetation index with same range and similar behavior as EVI) composites derived from MODIS with 250 m resolution, and ESA Global Land Cover map (GlobCover) with 300 m resolution for 2009 as crop mask. These authors predicted the winter wheat yield distribution with a level of adjustment R square to the observed governmental yield statistics of 0.69 in a regression model from 2000 to 2010. They used minimum root mean square error (RMSE) value as predictor in the regression model in order to find the day of the year for which to select the NDVI value that best predicts winter wheat yield.

Another study in Ukraine that adopted a regression model derived from Kansas USA data to assess winter wheat production, with a time range from the year 2000 until 2008, and using MODIS seasonal maximum NDVI data as explanatory variable and governmental wheat yield and production statistics as response variable (Becker-Reshef et al., 2010), had as results for the production a regression coefficient R of 0.88 (R square of 0.7744), and for the yield a R of 0.94 (R square of 0.8836). These comparatively better than this thesis figures resulted from a method based on the assumption that the yield is positively and linearly correlated to the seasonal maximum NDVI (adjusted for background noise) at the administrative unit (oblast) level and to the purity of the wheat signal. The purity of the wheat signal was accomplished by deriving a set of relationships between yield and maximum NDVI and then generalizing the slopes of these regressions as a function of wheat percent in order to un-mix the maximum NDVI for wheat signal from the maximum NDVI for other land covers signal.

Franch et al. (2015) made a study on Ukraine's winter wheat, with a time range from the year 2000 until 2011, and with a method based on the previously referred Becker-Reshef et al. (2010) study, but enhanced by including the growing degree day information to get an earlier forecast of the winter wheat production at the national scale. These study results had comparatively to Becker-Reshef et al., (2010) a lower error in the yield and a slightly higher error in the production. However, it can be considered that both results are equivalent reasserting the good performance of the method (Franch et al., 2015).

These previously referenced studies on Ukraine's winter wheat used the same type of remote sensing imagery and official statistical data as this thesis, except the time range and the methodology to extract and map crops were different, and they achieved better results. Therefore, maybe it would be a better and more reliable alternative mapping method to apply in this thesis for wheat and maybe other crops.

Thus, these low R square results for the regression between the production and maximum EVI may be due to some problem in the methodology for the location of the correct surrogate biophysical variable, such as the phenology extraction algorithm didn’t detect the correct surrogate biophysical phenology signal due to the spatial complexity inherent to the coarse resolution MODIS imagery; to the temporal variability of crops related with crop rotation; to the structure of the vegetation (for instance, sugar beet which has a fleshy root with a crown of leaves lying near the soil, is a plant that may be difficult to assess through remote sensing); to the lack of responsiveness of the phenology extraction algorithm; or the crop mapping, mainly the analysis of separability to assess the number of cluster, was not accurate due to the aforementioned problems with the ERDAS software.

Comparing visually the crop heat maps with the year 2000 ground data statistics from the Ukraine Ministry of Statistics, it can be noticed that the highest crop areas overlap with the oblasts with the highest percentage contribution to national area from the governmental statistics, except for Eastern Ukraine where for wheat, barley and sugar beet, the predicted crop area seems to be overestimated in relation to the official ground statistics. However, in a report from Ukrainian Agribusiness Club (Strohm et al., 2010), it shows in a map (see Figure 17) that this eastern steppe zone of Ukraine has one of the largest national share of winter wheat in the arable land for the year 2009.

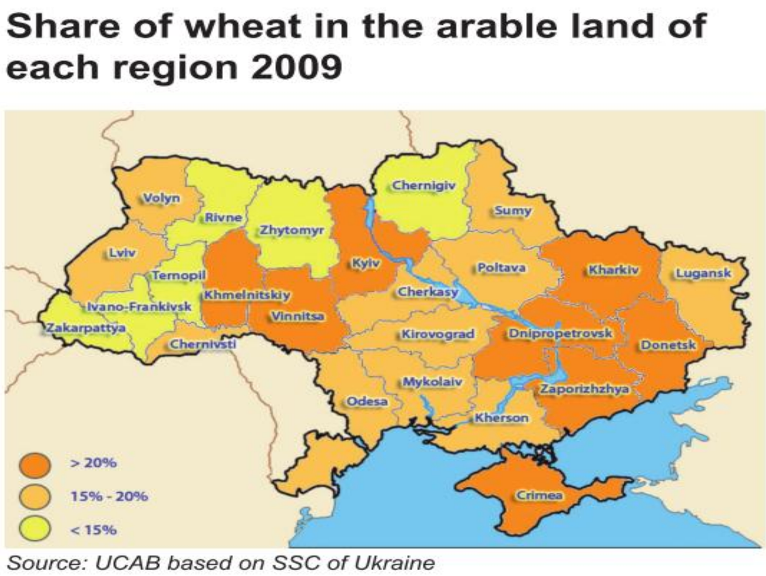


Figure 17: Wheat map adapted from Strohm et al., (2010)

This is a very coarse validation of the crop maps that doesn't allow to test if "what crop" and "where cropfield" is well estimated, but gives a broad idea of the crops location and area that can maybe be used on a national decision level.

5.4 Yield gap maps

The Licker et al. (2010) global study on maize, wheat, barley and other crop's yield gaps used as input 10km spatial resolution datasets based on global sensus data, along with detailed remote sensing data of global land cover, representing conditions around the year 2000 (to account for inter-annaul variability, averaged area harvested and yield data for the years 1997-2003), in order to: map crop yields into climate zones based on crop growing degree days and crop soil moisture index; determine maximum potential yield within each climate zone (90th yield percentile) and the actual yield; and finally determine the yield gap through the difference between the potential yield and actual yield, and yield gap fraction using equation 4. The maize, wheat and barley output yield gap results for Eastern Europe were generally high. A more specific analysis (see Figure 18) of a zoomed map of Europe and Ukraine, shows that the wheat yield gap fraction in Ukraine is widespread and more than 0.6, whereas the maize yield gap fraction in Ukraine are also widespread and very high in the western part (more than 0.8) and more than 0.6 in the rest of the Ukranian territory.

Yield gap maps

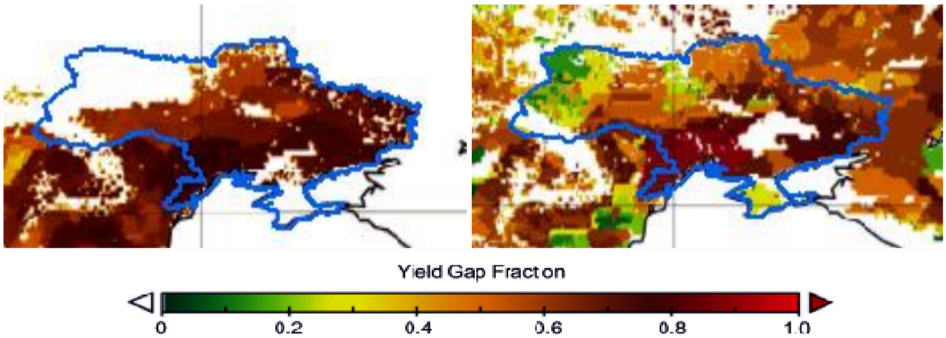


Figure 18: Yield gap fraction detail on a 5’ grid with an equirectangular projection for wheat (left) and maize (right) in eastern Europe (results from Licker et al. (2010)). Ukraine boarder is the blue line.

This thesis had yield gap results with higher spatial resolution, when compared to Licker et al. (2010) study. The majority of the grain yield gaps are concentrated in the southern east steppe zone (which is the most intensive cultivated area of Ukraine), with some minor yield gap patches in the forest-steppe zone; whereas in most of the western part, that includes the plane-polissya climatic zone, the yield gaps are lower, but also the area of cropland in this climate zone is also lower. The higher spatial resolution and the time series analysis of this thesis results allows to have a more spatially and thematically detailed view on the yield gaps of Ukraine.

However it would be further needed updated and representative ground control data to assess the thematic accuracy of the crop maps and yield maps.

The project Geo-wiki (Fritz et al., 2012) aims to provide through crowd sourcing a map validation service for land cover maps, such as cropland maps, currently being developed by the scientific community. It would be further useful to, whenever possible, add the crop type as attribute for each control point.

Furthermore, the JECAM project (Kussul et al., 2014), which has as goal to reach a convergence of approaches, develop monitoring and reporting protocols and best practices for global agriculture systems, may be a promising platform to systematize methodology and to fill this knowledge gap, which is of great importance for decision making on agriculture at global and local level.

With concerted initiatives such as the Geo-wiki crowd sourcing validation and the JECAM harmonization for crop monitoring procedures, we may be heading for a more supported, accurate and efficient global and local crop yield forecast and management system.

Chapter 6

6 Conclusion and recommendations

A procedure to map crop yield gaps across Ukraine using time series of satellite derived biophysical variables was successfully developed. On the other hand, as a result of some processing problems that are next referred, the capacity to identify underperforming croplands was just partially achieved.

The best separability achieved between clusters, after the unsupervised classification of the phenological parameter maximum EVI, was for the intermediate map with 19 cluster. However, due to inconsistent running of the ERDAS software this was based on somehow faulty incomplete results.

The stepwise linear regression analysis between these 19 maximum EVI clusters area and official crop statistics area resulted in a significative fit between both variables, with an adjusted R square greater than 0.7. Whereas the cluster areas for sugar beet had a low level of fit to their official ground statistical area, with an adjusted R square of 0.474.

The thematic accuracy of the crop maps was not successfully resolved, because the attributes of the ground control points didn't specify the type of crop.

Among all the crops, barley had the best prediction regression models for production, with levels of adjustment greater than 0.42 (except for the year 2013). All the other prediction models for crop yield and production had low level of fit, which suggest that crop areas were well estimated, but crop location based on surrogate yield or production data, i.e, on maximum EVI profiles, were not accurate.

Finally, and supported by these weak foundations, it was calculated the potential yield in each homogeneous edapho-climatic zone, in order to calculate the final yield gap maps output. Despite southern east steppe zone being the most intensive cultivated area of Ukraine, it was observed that the majority of the grain yield gaps were concentrated in this zone.

This was one of the first attempts to map crop yield gaps with a spatial resolution almost field size in Ukraine. However, despite of the importance of this topic, the final output is far from satisfactory, because the design of this resource and time consuming methodology that had as input big amounts of remote sensing data and extensive official statistics, had large challenges, more specifically, i)phenology extraction, ii) classification and location of the remote sensing derived crop yield classes or clusters that best represents the crop, iii) validation of yield or production model with updated and relevant ground control data.

In future Works, firstly, higher spatial resolution imagery such as landsat should also be used besides the time series, to allow identification of individual crop fields; secondly, use of decision trees with crop calendars criteria, and as input data the phenological parameters extracted previously, in order to improve the selection of clusters for the making of the intermediate map; finally create or find an updated and relevant ground control database that would allow thematic validation.

7 References

Alcantara, C., Kuemmerle, T., Baumann, M., Bragina, E. V, Griffiths, P., Hostert, P., … Radeloff, V. (2013). Mapping the extent of abandoned farmland in Central and Eastern Europe using MODIS time series satellite data. *Environmental Research Letters*, *8*(3), 035035. http://doi.org/10.1088/1748-9326/8/3/035035

Alfredo Huete, Chris Justice, Wim van Leeuwen, A. (1999). Modis Vegetation Index Algorithm Theoretical Basis. *Environmental Sciences*, (Mod 13), 129.

Ali, A., Bie, K. De, Skidmore, a. K. K., Scarrott, R. G. G., & Lymberakis, P. (2014). Mapping the heterogeneity of natural and semi-natural landscapes. *International Journal of Applied Earth Observation and Geoinformation*, *26*, 176–183. http://doi.org/10.1016/j.jag.2013.06.007

Ali, A., Bie, K. De, Skidmore, A. K., Scarrott, R. G., Hamad, A., Venus, V., & Lymberakis, P. (2013). Mapping land cover gradients through analysis of hyper-temporal NDVI imagery. *International Journal of Applied Earth Observation and Geoinformation*, *23*, 301–312. http://doi.org/10.1016/j.jag.2012.10.001

Almond, S. F. (2009). Validation and application of the MERIS Terrestrial Chlorophyll Index. Retrieved from http://eprints.bournemouth.ac.uk/16208/4/licence.txt

Amorós-López, J., Gómez-Chova, L., Alonso, L., Guanter, L., Zurita-Milla, R., Moreno, J., & Camps-Valls, G. (2013). Multitemporal fusion of Landsat/TM and ENVISAT/MERIS for crop monitoring. *International Journal of Applied Earth Observation and Geoinformation*, *23*, 132–141. http://doi.org/10.1016/j.jag.2012.12.004

Araya, S., Ostendorf, B., Lyle, G., & Lewis, M. (2013). Crop phenology based on MODIS satellite imagery as an indicator of plant available water content, (December), 1–6.

Atkinson, P. M., Jeganathan, C., Dash, J., & Atzberger, C. (2012). Inter-comparison of four models for smoothing satellite sensor time-series data to estimate vegetation phenology. *Remote*

Sensing of Environment, *123*, 400–417. http://doi.org/10.1016/j.rse.2012.04.001

Atzberger, C. (2013). Advances in Remote Sensing of Agriculture: Context Description, Existing Operational Monitoring Systems and Major Information Needs. *Remote Sensing*, *5*(2), 949–981. http://doi.org/10.3390/rs5020949

Azzali, S., & Menenti, M. (2000). Mapping vegetation-soil-climate complexes in southern Africa using temporal Fourier analysis of NOAA-AVHRR NDVI data. *International Journal of Remote Sensing*, *21*(5), 973–996. http://doi.org/10.1080/014311600210380

Badeck, F. W., Bondeau, A., Böttcher, K., Doktor, D., Lucht, W., Schaber, J., & Sitch, S. (2004). Responses of spring phenology to climate change. *New Phytologist*, *162*(2), 295–309. http://doi.org/10.1111/j.1469-8137.2004.01059.x

Baez-Gonzalez, A. D., Chen, P., Tiscareno, M. L., & Srinivasan, R. (2002). Using Satellite and Field Data with Crop Growth Modeling to Monitor and Estimate Corn Yield in Mexico. *Crop Science*, *42*(6), 1943. http://doi.org/10.2135/cropsci2002.1943

Baez-Gonzalez, A. D., Kiniry, J. R., Maas, S. J., Tiscareno, M. L., Macias C., J., Mendoza, J. L., … Manjarrez, J. R. (2005). Large-Area Maize Yield Forecasting Using Leaf Area Index Based Yield Model. *Agronomy Journal*, *97*(2), 418. http://doi.org/10.2134/agronj2005.0418

Barroso, C., & Monteiro, I. (2010). Monitoring vegetation from space. Retrieved from http://www.eumetrain.org/data/3/36/navmenu.php?page=4.2.0

Basso, B., Cammarano, D., & Vita, P. De. (2004). Remotely Sensed Vegetation Indices : Theory and applications for crop management, *53*(1), 36–49.

Baumann, M., Kuemmerle, T., Elbakidze, M., Ozdogan, M., Radeloff, V. C., Keuler, N. S., … Hostert, P. (2011). Patterns and drivers of post-socialist farmland abandonment in Western Ukraine. *Land Use Policy*, *28*(3), 552–562. http://doi.org/10.1016/j.landusepol.2010.11.003

Beck, P. S. a, Atzberger, C., Høgda, K. A., Johansen, B., & Skidmore, A. K. (2006). Improved monitoring of vegetation dynamics at very high latitudes: A new method using MODIS NDVI. *Remote Sensing of Environment*, *100*(3), 321–334. http://doi.org/10.1016/j.rse.2005.10.021

Becker-Reshef, I., Vermote, E., Lindeman, M., & Justice, C. (2010). A generalized regression-based model for forecasting winter wheat yields in Kansas and Ukraine using MODIS data. *Remote Sensing of Environment*, *114*(6), 1312–1323. http://doi.org/10.1016/j.rse.2010.01.010

Beddington, J., Asaduzzaman, M., Clark, M., Fernández, A., Guillou, M., Jahn, M., … Wakhungu, J. (2011). *Achieving food security in the face of climate change: Summary for policy makers from the Commission on Sustainable Agriculture and Climate Change*. Copenhagen,. Retrieved from www.ccafs.cgiar.org/commission

Bie, K. De, Nguyen, T. T. H., Ali, A., Scarrott, R., & Skidmore, A. K. (2012). LaHMa: a landscape heterogeneity mapping method using hyper-temporal datasets. *International Journal of Geographical Information Science*, *26*(11), 2177–2192. http://doi.org/10.1080/13658816.2012.712126

Bie, K. De, Venus, V., & Skidmore, A. K. (2010). Improved mapping and monitoring with hyper-temporal imagery, (Wfp).

Bolton, D. K., & Friedl, M. a. (2013). Forecasting crop yield using remotely sensed vegetation indices and crop phenology metrics. *Agricultural and Forest Meteorology*, *173*, 74–84. http://doi.org/10.1016/j.agrformet.2013.01.007

Boogaard, H., Wolf, J., Supit, I., Niemeyer, S., & van Ittersum, M. (2013). A regional implementation of WOFOST for calculating yield gaps of autumn-sown wheat across the European Union. *Field Crops Research*, *143*(SI), 130–142. http://doi.org/10.1016/j.fcr.2012.11.005

Brisson, N., Gate, P., Gouache, D., Charmet, G., Oury, F. X., & Huard, F. (2010). Why are wheat yields stagnating in Europe? A comprehensive data analysis for France. *Field Crops Research*, *119*, 201–212. http://doi.org/10.1016/j.fcr.2010.07.012

Brown, J. F. (2015). Methods for Deriving Metrics. Retrieved from http://phenology.cr.usgs.gov/methods_deriving.php

Commission of the European Communities. (1995). Corine Land Cover. Retrieved from http://www.eea.europa.eu/publications/COR0-landcover

Das, K. B. (2014). Ukraine : Stand-of Threatens Europe Breadbasket, *3*(3), 375–381.

Dash, J., Jeganathan, C., & Atkinson, P. M. (2010). The use of MERIS Terrestrial Chlorophyll Index to study spatio-temporal variation in vegetation phenology over India. *Remote Sensing of Environment*, *114*(7), 1388–1402. http://doi.org/10.1016/j.rse.2010.01.021

De Beurs, K. M., & Henebry, G. M. (2010). *Spatio-temporal statistical methods for modelling land surface phenology*. *Media*. http://doi.org/10.1007/978-90-481-3335-2

Defourny, P., Bogaert, E. Van, Kalogirou, V., & Perez, J. R. (2011). GLOBCOVER 2009 Products Description and Validation Report, 53. http://doi.org/10013/epic.39884.d016

Dixon, J., Gulliver, A., & Gibbon, D. (2001). *Farming Systems and Poverty*. *FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS*.

Duncan, J. M. a, Dash, J., & Atkinson, P. M. (2014a). Elucidating the impact of temperature variability and extremes on cereal croplands through remote sensing, 1–11. http://doi.org/10.1111/gcb.12660

Duncan, J. M. a, Dash, J., & Atkinson, P. M. (2014b). ScienceDirect Spatio-temporal dynamics in the phenology of croplands across the Indo-Gangetic Plains. *Advances in Space Research*, *54*(4), 710–725. http://doi.org/10.1016/j.asr.2014.05.003

Elowitz, M. (2015). What is Imaging Spectroscopy (Hyperspectral Imaging)? Retrieved from http://www.markelowitz.com/Hyperspectral.html

Evans, L. T., & Fischer, R. A. (1999). Yield Potential: Its Definition, Measurement, and Significance. *Crop Science*, *39*, 1544–1551. http://doi.org/10.2135/cropsci1999.3961544x

EXELIS. (2013). Vegetation Analysis: Using Vegetation Indices in ENVI. Retrieved from https://www.exelisvis.com/Learn/WhitepapersDetail/TabId/802/ArtMID/2627/ArticleID/13742/Vegetation-Analysis-Using-Vegetation-Indices-in-ENVI.aspx

FAO. (1988). *FAO/UNESCO Soil Map of the World, Revised Legend, with corrections and updates*. Rome.

FAO. (2009). *Global agriculture towards 2050*. Rome.

FAO. (2015a). Global Information and Early Warning System (GIEWS). Retrieved from http://www.fao.org/giews/english/index.htm

FAO. (2015b). Methods & Standards - Glossary list. Retrieved from http://faostat3.fao.org/mes/glossary/E

FAO. (2015c). Yield and Production Gaps. Retrieved from http://www.fao.org/nr/gaez/about-data-portal/yield-and-production-gaps/en/#

Foley, J. a. (2010). The other inconvenient truth. Retrieved from http://www.ted.com/talks/jonathan_foley_the_other_inconvenient_truth#t-548012

Franch, B., Vermote, E. F., Becker-Reshef, I., Claverie, M., Huang, J., Zhang, J., … Sobrino, J. a. (2015). Improving the timeliness of winter wheat production forecast in the United States of America, Ukraine and China using MODIS data and NCAR Growing Degree Day information. *Remote Sensing of Environment*, *161*, 131–148. http://doi.org/10.1016/j.rse.2015.02.014

Fritz, S., Mccallum, I., See, L., Perger, C., Schill, C., Schepaschenko, D., & Albrecht, F. (2012). Geo-Wiki: A tool for crowdsourcing land cover. *GLP News*, 1–3.

Frost, J. (2013). Multiple Regression Analysis: Use Adjusted R-Squared and Predicted R-Squared to Include the Correct Number of Variables. Retrieved from http://blog.minitab.com/blog/adventures-in-statistics/multiple-regession-analysis-use-adjusted-r-squared-and-predicted-r-squared-to-include-the-correct-number-of-variables

Funk, C., & Budde, M. E. (2009). Phenologically-tuned MODIS NDVI-based production anomaly estimates for Zimbabwe. *Remote Sensing of Environment*, *113*(1), 115–125. http://doi.org/10.1016/j.rse.2008.08.015

Geng, L., Ma, M., Wang, X., Yu, W., Jia, S., & Wang, H. (2014). Comparison of Eight Techniques for Reconstructing Multi-Satellite Sensor Time-Series NDVI Data Sets in the Heihe River Basin, China. *Remote Sensing*, *6*(3), 2024–2049. http://doi.org/10.3390/rs6032024

Godfray, H. C. J., Beddington, J. R., Crute, I. R., Haddad, L., Lawrence, D., Muir, J. F., … Toulmin, C. (2010). Food security: the challenge of feeding 9 billion people. *Science (New York, N.Y.)*, *327*(5967), 812–8. http://doi.org/10.1126/science.1185383

Group on Earth Observations. (2010). Global Agriculture Monitoring. Retrieved June 1, 2014, from http://biblio.central.ucv.ro/bib_web/bib_pdf/EU_books/0143.pdf

Group on Earth Observations. (2013). *Progress on GEOGLAM Implementation First steps towards Implementation 2013-2014 Phase I and II* (Vol. 10). Retrieved from http://www.earthobservations.org/documents/geoglam/GEOGLAM_Implementation_Plan.pdf

Hall, a. J., Feoli, C., Ingaramo, J., & Balzarini, M. (2013). Gaps between farmer and attainable yields across rainfed sunflower growing regions of Argentina. *Field Crops Research*, *143*, 119–129. http://doi.org/10.1016/j.fcr.2012.05.003

Hall-Beyer, M. (2003). Comparison of Single-Year and Multiyear NDVI Time Series Principal Components in Cold Temperate Biomes, *41*(11), 1–8.

Hengsdijk, H., & Langeveld, J. (2009). *Yield trends and yield gap analysis of major crops in the world*. Wageningen.

Hirosawa, Y., Marsh, S. E., & Kliman, D. H. (1996). Application of standardized principal component analysis of land-cover characterization using multitemporal AVHRR data. *Remote Sensing of Environment*, *58*(3), 267–281. http://doi.org/10.1016/S0034-4257(96)00068-5

Hochman, Z., Gobbett, D., Holzworth, D., McClelland, T., van Rees, H., Marinoni, O., … Horan, H. (2013). Reprint of "Quantifying yield gaps in rainfed cropping systems: A case study of wheat in Australia." *Field Crops Research*, *143*, 65–75. http://doi.org/10.1016/j.fcr.2013.02.001

Hostert, P., Kuemmerle, T., Prishchepov, A., Sieber, A., Lambin, E. F., & 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. http://doi.org/10.1088/1748-9326/6/4/045201

Huete, A., Didan, K., Miura, T., Rodriguez, E. P., Gao, X., & Ferreira, L. G. (2002). Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment*, *83*(1-2), 195–213. http://doi.org/10.1016/S0034-4257(02)00096-2

Huete, A., Didan, K., Shimabukuro, Y. E., Ratana, P., Saleska, S. R., Hutyra, L. R., … Myneni, R. (2006). Amazon rainforests green-up with sunlight in dry season. *Geophysical Research Letters*, *33*(6), 2–5. http://doi.org/10.1029/2005GL025583

Husak, G. J., Marshall, M. T., Michaelsen, J., Pedreros, D., Funk, C., & Galu, G. (2008). Crop area estimation using high and medium resolution satellite imagery in areas with complex topography. *Journal of Geophysical Research: Atmospheres*, *113*(October 2007). http://doi.org/10.1029/2007JD009175

IBM Corp. (2013). IBM SPSS Statistics for Windows, Version 22.0. Armonk, NY: IBM Corp.

IES. (2015). Monitoring Agricultural ResourceS (MARS). Retrieved from http://mars.jrc.ec.europa.eu/

Intergraph. (1997). ERDAS Field Guide. *Imagine*, 645.

Intergraph. (2014). ERDAS Imagine.

ISRIC World Soil Information. (2015). SOIL AND TERRAIN DATABASE FOR CENTRAL AND EASTERN EUROPE (VERSION 1.1) - SCALE 1:2.5 MILLION (SOVEUR). Retrieved from http://meta2.isric.org/geonetwork/srv/en/main.home

Jackson, R., & Huete, A. (1991). Interpreting vegetation indices. *Preventive Veterinary Medicine*.

Jakubauskas, M. E., Legates, D. R., & Kastens, J. H. (2001). Harmonic analysis of time - series AVHRR NDVI data. *Photogrammetric Engineering and Remote Sensinghotogrammetric Engineering and Remote Sensing*, *67*(4), 461 – 470.

Jensen, J. R. (2007). *Remote sensing of the environment - An earth resource perspective* (2nd ed.). Pearson Education, Inc.

Jonsson, P., & Eklundh, L. (2002). Seasonality extraction by function fitting to time-series of satellite sensor data. *IEEE Transactions on Geoscience and Remote Sensing*, *40*(8), 1824–1832. http://doi.org/10.1109/TGRS.2002.802519

Jonsson, P., & Eklundh, L. (2004). TIMESAT - A program for analyzing time-series of satellite sensor data. *Computers and Geosciences*, *30*(8), 833–845. http://doi.org/10.1016/j.cageo.2004.05.006

Jordan, C. F. (1969). Derivation of Leaf-Area Index from Quality of Light on the Forest Floor. *Ecology*, *50*(4), 663. http://doi.org/10.2307/1936256

Justice, C. O., & Becker-Reshef, I. (2007). *Report from the Workshop on Developing a Strategy for Global Agricultural Monitoring in the framework of Group on Earth Observations (GEO), UN FAO*. Rome.

Kaduk, J., & Heimann, M. (1996). A prognostic phenology scheme for global terrestrial carbon cycle models. *Climate Research*, *6*, 1–19. http://doi.org/10.3354/cr006001

Kalaitzidis, C., & Manakos, I. (2015). Crop Yield Estimation. Retrieved from http://www.seos-project.eu/modules/agriculture/agriculture-c02-p01.html

Khan, M. R., Bie, K. De, Keulen, H. Van, Smaling, E. M. A., & Real, R. (2010). Disaggregating and mapping crop statistics using hypertemporal remote sensing. *International Journal of Applied Earth Observation and Geoinformation*, *12*, 36–46. http://doi.org/10.1016/j.jag.2009.09.010

Kogan, F., Kussul, N., Adamenko, T., Skakun, S., Kravchenko, O., Kryvobok, O., … Lavrenyuk, A. (2013). Winter wheat yield forecasting in Ukraine based on Earth observation, meteorological data and biophysical models. *International Journal of Applied Earth Observation and Geoinformation*, *23*, 192–203. http://doi.org/10.1016/j.jag.2013.01.002

Kolotii, A., Kussul, N., Shelestov, A., Skakun, S., Yailymov, B., Basarab, R., … Service, S. (2015). COMPARISON OF BIOPHYSICAL AND SATELLITE PREDICTORS FOR WHEAT YIELD FORECASTING IN UKRAINE, (May), 11–15.

Kowalik, W., Dabrowska-Zielinska, K., Meroni, M., Raczka, T. U., & de Wit, A. (2014). Yield estimation using SPOT-VEGETATION products: A case study of wheat in European countries. *International Journal of Applied Earth Observation and Geoinformation*, *32*(1), 228–239. http://doi.org/10.1016/j.jag.2014.03.011

Kravchenko, O., & Moloshnii, B. (2012). CROP STATE AND AREA ESTIMATION IN UKRAINE BASED ON REMOTE AND IN- SITU OBSERVATIONS Nataliia Kussul , Andrii Shelestov , Sergii Skakun , A Network of JECAM Test Sites in Ukraine, *1*, 251–259.

Kryzhanivsky, S. (2015). Ukraine. Retrieved April 7, 2015, from http://www.britannica.com/EBchecked/topic/612921/Ukraine/30095/Soils

Kuemmerle, T., Hostert, P., Radeloff, V. C., Linden, S., Perzanowski, K., & Kruhlov, I. (2008). Cross-border Comparison of Post-socialist Farmland Abandonment in the Carpathians. *Ecosystems*, *11*(4), 614–628. http://doi.org/10.1007/s10021-008-9146-z

Kussul, N., Kolotii, A., & Skakun, S. (2014). Estimating relative efficiency of use multisource satellite data for winter wheat yield forecasting in Ukraine, *4*(i), 62–64.

Kussul, N., Shelestov, A., & Skakun, S. (2014). JECAM Site Ukraine. Ottawa, Canada: Group on Earth Observation.

Kussul, N., Skakun, S., Shelestov, A., Kravchenko, O., Gallego, J. F., & Kussul, O. (2012). CROP AREA ESTIMATION IN UKRAINE USING SATELLITE DATA WITHIN THE MARS PROJECT. *IEEE INTERNATIONAL GEOSCIENCE AND REMOTE SENSING SYMPOSIUM (IGARSS)*, 3756–3759. Retrieved from http://apps.webofknowledge.com/full_record.do?product=WOS&search_mode=GeneralSearch&qid=3&SID=U28S287d7UKbUvcHhdb&page=1&doc=20

Kussul, N., Skakun, S., Shelestov, A., Lavreniuk, M., Yailymov, B., & Kussul, O. (2014). REGIONAL SCALE CROP MAPPING USING MULTI-TEMPORAL SATELLITE IMAGERY.

Landgrebe, D. (2003). *Signal theory methods in multispectral remote sensing*. New Jersey: Wiley. Retrieved from https://books.google.co.uk/books?id=4Kcfg1od3twC&dq=Signal+theory+methods+in+multispectral+remote+sensing&lr=&hl=pt-PT&source=gbs_navlinks_s

Leeuwen, M. Van, & Et Al. (2012). *The Agri-food Sector in Ukraine: Current Situation and Market Outlook until 2025*. *JRC Scientific and Policy Reports*. http://doi.org/Report EUR 25444 EN

Li, A., Liang, S., Wang, A., & Qin, J. (2007). Estimating Crop Yield from Multi-temporal Satellite Data Using Multivariate Regression and Neural Network Techniques. *Photogrammetric Engineering & Remote Sensing*, *73*(10), 1149–1157. http://doi.org/10.14358/PERS.73.10.1149

Li, Z., & Kafatos, M. (1996). In terann ual V ariabilit y of V egetation in the United States and its Relation to El Niño/Southern Oscillation. *International Journal of Remote Sensing*, *17*(13), 2533–2548.

Li, Z., Xu, D., & Guo, X. (2014). Remote Sensing of Ecosystem Health: Opportunities, Challenges, and Future Perspectives, (1), 21117–21139. http://doi.org/10.3390/s141121117

Liang, S., Wang, J., & Li, X. (2012). Vegetation Production in Terrestrial Ecosystems (pp. 501–531). http://doi.org/10.1016/B978-0-12-385954-9.00016-2

Licker, R., Johnston, M., Foley, J. a., Barford, C., Kucharik, C. J., Monfreda, C., & Ramankutty, N. (2010). Mind the gap: how do climate and agricultural management explain the "yield gap" of croplands around the world? *Global Ecology and Biogeography*, *19*(6), 769–782. http://doi.org/10.1111/j.1466-8238.2010.00563.x

Lioubimtseva, E., & Henebry, G. M. (2012). Grain production trends in Russia, Ukraine and Kazakhstan: New opportunities in an increasingly unstable world? *Frontiers of Earth Science*, *6*(2), 157–166. http://doi.org/10.1007/s11707-012-0318-y

Lloyd, D. (1990). A phenological classification of terrestrial vegetation cover using shortwave vegetation index imagery. *International Journal of Remote Sensing*, *11*(12), 2269–2279. http://doi.org/10.1080/01431169008955174

Lobell, D. B. (2013). The use of satellite data for crop yield gap analysis. *Field Crops Research*, *143*, 56–64. http://doi.org/10.1016/j.fcr.2012.08.008

Lobell, D. B., & Burke, M. B. (2010). *Climate Change and Food Security - Adapting Agriculture to a Warmer World*. Springer Netherlands. http://doi.org/10.1007/978-90-481-2953-9

Lobell, D. B., Cassman, K. G., & Field, C. B. (2009). Crop Yield Gaps: Their Importance, Magnitudes, and Causes. *Annual Review of Environment and Resources*, *34*(1), 179–204. http://doi.org/10.1146/annurev.environ.041008.093740

Lobell, D. B., Ortiz-Monasterio, J. I., Asner, G. P., Naylor, R. L., & Falcon, W. P. (2005). Combining field surveys, remote sensing, and regression trees to understand yield variations in an irrigated wheat landscape. *Agronomy Journal*, *97*, 241–249. http://doi.org/10.2134/agronj2005.0241

Lu, L., Wang, C., Guo, H., & Li, Q. (2013). Detecting winter wheat phenology with SPOT-VEGETATION data in the North China Plain. *Geocarto International*, (April 2013), 1–12. http://doi.org/10.1080/10106049.2012.760004

Makowski, D., Nesme, T., Papy, F., & Doré, T. (2013). Global agronomy, a new field of research. A review. *Agronomy for*

Sustainable Development, *34*(2), 293–307. http://doi.org/10.1007/s13593-013-0179-0

Martinelli, F., Scalenghe, R., Davino, S., Panno, S., Scuderi, G., Ruisi, P., … Boschetti, M. (2014). Advanced methods of plant disease detection. A review. *Agron. Sustain. Dev.*, 1–25. http://doi.org/10.1007/s13593-014-0246-1

NASA. (2014). Surface Reflectance 8-Day L3 Global 500m MOD09A1. Retrieved from https://lpdaac.usgs.gov/products/modis_products_table

NCSRT. (2015). National Centre of Space Research and Technologies (NCSRT) of the National Space Agency of Kazakhstan. Retrieved from http://izk.gzi.kz/

Neter, J., Kutner, M., Nachtsheim, C., & Wasserman, W. (1996). *Applied linear statistical models* (4th ed.). Boston, Massachusetts: McGraw-Hill.

Oliver, Y. M., & Robertson, M. J. (2013). Quantifying the spatial pattern of the yield gap within a farm in a low rainfall Mediterranean climate. *Field Crops Research*, *150*, 29–41. http://doi.org/10.1016/j.fcr.2013.06.008

Pasuquin, J. M., & Witt, C. (2007). Yield potential and yield gaps of maize in Southeast Asia. Retrieved from http://www.ipipotash.org/pt/eifc/2007/14/4/english

Pervez, M. S., & Brown, J. F. (2010). Mapping Irrigated Lands at 250-m Scale by Merging MODIS Data and National Agricultural Statistics. *Remote Sensing*, *2*(10), 2388–2412. http://doi.org/10.3390/rs2102388

Pittman, K., Hansen, M. C., Becker-Reshef, I., Potapov, P. V., & Justice, C. O. (2010). Estimating Global Cropland Extent with Multi-year MODIS Data. *Remote Sensing*, *2*(7), 1844–1863. http://doi.org/10.3390/rs2071844

Prasad, A. K., Chai, L., Singh, R. P., & Kafatos, M. (2006). Crop yield estimation model for Iowa using remote sensing and surface parameters. *International Journal of Applied Earth Observation and Geoinformation*, *8*(1), 26–33. http://doi.org/10.1016/j.jag.2005.06.002

Qiu, G. Y., Omasa, K., & Sase, S. (2009). An infrared-based coefficient to screen plant environmental stress: concept, test and applications. *Functional Plant Biology*, *36*(11), 990. http://doi.org/10.1071/FP09132

Reed, B. C., Brown, J. F., VanderZee, D., Loveland, T. R., Merchant, J. W., & Ohlen, D. O. (1994). Measuring phenological variability from satellite imagery. *Journal of Vegetation Science*, *5*(5), 703–714. http://doi.org/10.2307/3235884

Rembold, F., Atzberger, C., Savin, I., & Rojas, O. (2013). Using low resolution satellite imagery for yield prediction and yield anomaly detection. *Remote Sensing*, *5*(4), 1704–1733. http://doi.org/10.3390/rs5041704

Rogovska, N. (2009). Crop Production in Ukraine. Retrieved from http://www.public.iastate.edu/~natashar/Crop production in Ukraine.pdf

Rouse, J. W. (1978). *Monitoring the vernal advancement and retrogradation (green wave effect) of natural vegetation*. Greenbelt, Maryland.

Sakamoto, T., Gitelson, A. a., & Arkebauer, T. J. (2013). MODIS-based corn grain yield estimation model incorporating crop phenology information. *Remote Sensing of Environment*, *131*, 215–231. http://doi.org/10.1016/j.rse.2012.12.017

Sakamoto, T., Yokozawa, M., Toritani, H., Shibayama, M., Ishitsuka, N., & Ohno, H. (2005). A crop phenology detection method using time-series MODIS data. *Remote Sensing of Environment*, *96*(3-4), 366–374. http://doi.org/10.1016/j.rse.2005.03.008

Santin-Janin, H., Garel, M., Chapuis, J. L., & Pontier, D. (2009). Assessing the performance of NDVI as a proxy for plant biomass using non-linear models: A case study on the kerguelen archipelago. *Polar Biology*, *32*, 861–871. http://doi.org/10.1007/s00300-009-0586-5

Schulthess, U., Timsina, J., Herrera, J. M., & McDonald, A. (2013). Mapping field-scale yield gaps for maize: An example from Bangladesh. *Field Crops Research*, *143*, 151–156. http://doi.org/10.1016/j.fcr.2012.11.004

Siebert, S., Portmann, F. T., & Döll, P. (2010). Global Patterns of Cropland Use Intensity. *Remote Sensing*, *2*(7), 1625–1643. http://doi.org/10.3390/rs2071625

Silleos, N. G., Alexandridis, T. K., Gitas, I. Z., & Perakis, K. (2006). Vegetation Indices: Advances Made in Biomass Estimation and Vegetation Monitoring in the Last 30 Years. *Geocarto International*, *21*(4), 21–28. http://doi.org/10.1080/10106040608542399

Smith, P., Gregory, P. J., van Vuuren, D., Obersteiner, M., Havlík, P., Rounsevell, M., … Bellarby, J. (2010). Competition for land. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *365*(1554), 2941–57. http://doi.org/10.1098/rstb.2010.0127

Son, N. T., Chen, C. F., Chen, C. R., Minh, V. Q., & Trung, N. H. (2014). A comparative analysis of multitemporal MODIS EVI and NDVI data for large-scale rice yield estimation. *Agricultural and Forest Meteorology*, *197*, 52–64. http://doi.org/10.1016/j.agrformet.2014.06.007

Stefanski, J., Chaskovskyy, O., & Waske, B. (2014). Mapping and monitoring of land use changes in post-Soviet western Ukraine using remote sensing data. *Applied Geography*, *55*, 155–164. http://doi.org/10.1016/j.apgeog.2014.08.003

Stockli, R., & Vidale, P. L. (2004). European plant phenology and climate as seen in a 20-year AVHRR land-surface parameter dataset, (May 2015), 37–41. http://doi.org/10.1080/01431160310001618149

Strohm, K., Tovstopyat, A., & Walther, S. (2010). *Arable farming in Ukraine*.

Taylor, J. C., Wood, G. A., & Thomas, G. (1997). Mapping yield potential with remote sensing. *Precision Agriculture '97 : Papers Presented at the First European Conference on Precision Agriculture, Warwick University Conference Centre, UK, 7-10 September 1997*. Retrieved from http://agris.fao.org/agris-search/search.do?recordID=US201302883268

Thenkabail, P., Biradar, C. M., Noojipady, P., Dheeravath, V., Li, Y. J., Velpuri, M., … Dutta, R. (2008). *A Global Irrigated Area Map*

(GIAM) Using Remote Sensing at the End of the Last Millennium. *Water Management*.

Thenkabail, P., & Wu, Z. (2012). An Automated Cropland Classification Algorithm (ACCA) for Tajikistan by Combining Landsat, MODIS, and Secondary Data. *Remote Sensing*, *4*(12), 2890–2918. http://doi.org/10.3390/rs4102890

Tuan, T. Van. (2008). *Quantifying spatial and temporal patterns of the water related yield gap , using synoptic data and a dynamic Crop Growth Model (PSn) A Case Study of Sunflower in Andalucia , Spain Quantifying spatial and temporal patterns of the water related yield gap*.

Tucker, C. J. (1979). Red and photographic infrared linear combinations for monitoring vegetation, *150*, 127–150. http://doi.org/10.1016/0034-4257(79)90013-0

Tucker, C. J. (1980). A critical review of remote sensing and other methods for non-destructive estimation of standing crop biomass. *Grass and Forage Science*, *35*(3), 177–182. http://doi.org/10.1111/j.1365-2494.1980.tb01509.x

United Nations. (2013). *World population projected to reach 9 . 6 billion by 2050 with most growth in developing regions , especially Africa – says UN*.

USDA. (2014). GLAM—Global Agricultural Monitoring. Retrieved June 2, 2014, from http://www.pecad.fas.usda.gov/glam.htm#background

USDA FSA. (2015). Foreign Agricultural Service (FSA). Retrieved from http://www.fas.usda.gov/

Van Ittersum, M., Bussel, L., Wolf, J., Cassman, K., Wart, J., Grassini, P., & Yang, H. (2013). Global Yield Gap Atlas. Retrieved May 23, 2014, from http://www.yieldgap.org/

Van Ittersum, M., Cassman, K. G., Grassini, P., Wart, J., Bussel, L., Wolf, J., … Hochman, Z. (2013). Yield gap analysis with local to global relevance—A review. *Field Crops Research*, *143*, 4–17. http://doi.org/10.1016/j.fcr.2012.09.009

Van Ittersum, M. K., & Rabbinge, R. (1997). Concepts in production ecology for analysis and quantification of agricultural input-output combinations. *Field Crops Research*, *52*, 197–208. http://doi.org/10.1016/S0378-4290(97)00037-3

Van Wart, J., Kersebaum, K. C., Peng, S., Milner, M., Cassman, K. G., Wart, J., … Cassman, K. G. (2013). Estimating crop yield potential at regional to national scales. *Field Crops Research*, *143*, 34–43. http://doi.org/10.1016/j.fcr.2012.11.018

Wart, J., Kersebaum, K. C., Peng, S., Milner, M., & Cassman, K. G. (2013). Estimating crop yield potential at regional to national scales. *Field Crops Research*, *143*, 34–43. http://doi.org/10.1016/j.fcr.2012.11.018

Wart, J., van Bussel, L. G. J., Wolf, J., Licker, R., Grassini, P., Nelson, A., … van Wart, J. (2013). Use of agro-climatic zones to upscale simulated crop yield potential. *Field Crops Research*, *143*, 44–55. http://doi.org/10.1016/j.fcr.2012.11.023

WDC Ukraine. (2015). Ukraine: Agricultural Overview. Retrieved from http://wdc.org.ua/en/node/29

White, a, & Thomton, P. E. (1997). A continental responses phenology model climatic for monitoring variability vegetation to interannual. *Global Biogeochemical Cycles*, *11*(2), 217–234. http://doi.org/10.1029/97GB00330

Yebra, M., Dennison, P. E., Chuvieco, E., Riaño, D., Zylstra, P., Hunt, E. R., … Jurdao, S. (2013). A global review of remote sensing of live fuel moisture content for fire danger assessment: Moving towards operational products. *Remote Sensing of Environment*, *136*, 455–468. http://doi.org/10.1016/j.rse.2013.05.029

Zhang, C., & Kovacs, J. M. (2012). The application of small unmanned aerial systems for precision agriculture: a review. *Precision Agriculture*, *13*(6), 693–712. Retrieved from http://static-content.springer.com/lookinside/art%3A10.1007%2Fs11119-012-9274-5/000.png

Zhang, X., Friedl, M. a., Schaaf, C. B., Strahler, A. H., Hodges, J. C. F., Gao, F., … Huete, A. (2003). Monitoring vegetation phenology using MODIS. *Remote Sensing of Environment*, *84*(3), 471–475. http://doi.org/10.1016/S0034-4257(02)00135-9

ИКИ РАН. (2015). Space Research Institute of Russian Academy of Sciences. Retrieved from http://iki.rssi.ru/eng/monitor.html