CROPS FROM SPACE

IMPROVED EARTH OBSERVATION CAPACITY TO MAP CROP AREAS AND TO QUANTIFY PRODUCTION

MOBUSHIR RIAZ KHAN

Examination Committee

Prof. Dr. Ing. W. Verhoef, University of Twente
Prof. Dr. Ir. V.G. Jetten, University of Twente
Prof. Dr. Ir. A.K. Bregt, Wageningen University and Research Centre
Dr. Ir. B.G.J.S. Sonneveld, Free University of Amsterdam

UNIVERSITY OF TWENTE.
ITC
FACULTY OF GEO-INFORMATION SCIENCE AND EARTH OBSERVATION

ITC dissertation number 180
ITC, P.O. Box 217, 7500 AE Enschede, The Netherlands

ISBN 978-90-6164-301-2
Cover designed by Benno Masselink
Cover illustration: Vincent van Gogh, 1888, Sower at Sunset
Printed by ITC Printing Department

CROPS FROM SPACE

IMPROVED EARTH OBSERVATION CAPACITY TO MAP CROP AREAS AND TO QUANTIFY PRODUCTION

DISSERTATION

to obtain
the degree of doctor at the University of Twente,
on the authority of the Rector Magnificus,
prof. dr. H. Brinksma,
on account of the decision of the graduation committee,
to be publicly defended
on Wednesday 23 February 2011 at 14:45 hrs

by

Mobushir Riaz Khan

born on 27 November 1974

in Faisalabad, Pakistan

This thesis is approved by

Prof. Dr. Ir. E.M.A. Smaling, promotor
Faculty of Geo-information Science and Earth Observation (ITC),
University of Twente, the Netherlands

Prof. Dr. Ir. H. Van Keulen, promotor
PRI Agrosystems Research, Wageningen University and Research
Centre, the Netherlands

Dr. Ir. C.A.J.M. De Bie, assistant promotor
Faculty of Geo-information Science and Earth Observation (ITC),
University of Twente, the Netherlands

Table of Contents

List of figures

List of tables

Acknowledgments

I am heartily thankful to my promotor, Professor Eric Smaling, whose encouragement, guidance and support made me to accomplish this uphill task. It has been an honour to work with him. I am also thankful to my second promoter Prof Herman Van Keulen. Eric and Herman I really appreciate your critical and thorough insights while you were reviewing my work. I am highly indebted to Dr. Kees de Bie, my co-promotor, who enabled me from the initial to the final level to develop an understanding of the subject and successfully finish it. I could not have wished for a better coach. Kees, your contributions, detailed comments and insight have been of great value to me. I hope our collaboration continues for long. My wishes and prayers are always there for you and your family.

I would especially like to thank Valentijn Venus for his valuable suggestions and great support to me while I was working on crop modelling.
My special thanks to Higher Education Commission of Pakistan for providing me funding without which it might not be possible for me to carry on my project.

Thanks also to the ladies and gentlemen of the supporting departments. My special thanks are to Loes for helping me a lot right from the beginning till the end of my doctoral study.

Research without data? No way! Thanks Alfredo with your kind help in data collection and in the arrangements of meetings with the officials of the Ministry of Agriculture and Fisheries, Andalucía. Your kind support was always there whenever I needed, Thank you for that.

I am and will forever be deeply indebted to my parents, especially my Late Father, for what they have enabled me to achieve today through their endless love and prayers. All I can say is it would take another thesis to express my deep love for them. Dear Father, I miss you sooo much and I pray that your soul rests in peace.

Many thanks to my loving siblings and their families: Afshan, Mansoor Bhai, Kashif & Ayesha, Ashi & Adil and Bushra & Nadeem. Api you always gave me the confidence and courage to stay put and keep faith in my abilities – thanks for showing such a tremendous trust in me! You have played a ‘key’ role in my life right from the beginning of education and till now – Api and Ashi I don’t see myself anywhere without such a loving sisters like you in every walk of life! Above all I am really thankful to the great kids Haris, Zainab, Rubab, Hanan, Mannan, Ali, Eisha and Zoha for their all-time love for me.

Heartiest gratitude to my loving life partner Dr. Amna for helping me to concentrate on this dissertation and supporting mentally during the course of this work. I am also thankful to God for blessing me with 'the best wife' of this whole world. Love you sweetheart. I am also thankful to my extended family, Uncle Yusuf, Ali, Abdul Baseer and Fatima, for supporting me and Amna very warmly and for being enthusiastic and wishful for my studies. Uncle, you took care of Noor while we were studying abroad…Thanks a lot.

Noor-ul-Ain -the light of my Eyes – my daughter your innocent patience and love have upheld me, particularly in those many a times when I spent more moments with my computer than with you.

Many thanks to Dr. Tariq for his all-time support, friendship and guidance in my professional career.

Thanks to my friends for sharing your experiences with me and listening to mine – great support no doubt. Thank you Farrukh Abdullah, Ikram Ali, Ha, Amjad, Abel, Jazi, Mrs and Yasin, Khurram, Shafique, Saleem, Azmat…..

I really appreciate my colleagues of Ph.D. and Masters who travelled with me to Andalucía for field works. Ha, Amjad, Tuan, Mumun, PK, Yanti and Amit, you guys have been a great company.

Finally, I offer my regards and blessings to all of you who supported me in any respect during the completion of the project. I express my apology that I could not mention personally all names one by one.

1

General Introduction

1.1 Food security problems

Agriculture, as the basis of the supply of food for the human population, is one of the most important human activities, its origin dating back to the Neolithic Revolution, more than 10 000 years ago (Janick, 1974). Since then, agriculture has developed substantially, while in the last centuries the field of agricultural research has continuously been advancing in response to the demands of human society. Progress, however, has not been as smooth as desired: the world has witnessed many food crises and man has continuously been battling against hunger in efforts to achieve food security. Food security exists "when all people, at all times, have physical and economic access to sufficient, safe and nutritious food to meet their dietary needs and food preferences for an active and healthy life" (FAO, 1983; World Food Summit, 1996).

Every generation has suffered hunger, famines are of all times, and hence feelings of food insecurity have always been part of human history. The endeavours to achieve food security to avoid hunger and famine are as old as civilization. In this section, a brief account is provided of various food crises since the third decade of the 20th century, followed by an overview of the main causes underlying such food crises. Finally, some global efforts to achieve food security are discussed.

Food availability, at times, became precariously low for the rapidly increasing population. In 1928-29, famine in northern China caused about 2 million deaths. Drought in 1932 and 1933 resulted in severe shortages of food which affected more than 4 million people in many parts of the former Union of Soviet Socialist Republics (USSR). China was affected again in 1936 by a massive famine with an estimated 5 million fatalities. During World War II, many parts of Europe were affected and experienced severe shortages of food, such as Poland (1940-43) and Greece (1941-44). The Netherlands experienced the 'hongerwinter' in the transition from 1944 to 1945. Countries in Africa and Asia such as India, Vietnam, Rwanda and Malawi also were adversely affected by the ravages of World War II (Shaw, 2007).

In the 1960s, southern Asia experienced monsoon failures from 1965-67, so that massive food aid was required to prevent large-scale starvation, and fears of impending world famine were widespread. In the early 1970s, a feeling of

despair prevailed, as famines ravaged hundreds of millions of the poorest citizens of the world. Ethiopia and the West African Sahel experienced large-scale famine because of a persistent drought (Morgan and Solarz, 1994). In 1972, world cereal production fell by 3% compared to 1971, instead of increasing by 2%, as required to keep pace with the growth of the population (Schnittker, 1973). Cereal stocks in the major exporting countries dropped by 50% during 1972-73 crop year leaving almost no food reserves (Sarris and Taylor, 1976). The economic situation was further aggravated by the increasing oil prices in the same year. Consequently, world cereal prices increased fourfold. The combination of these events created serious financial problems for the food-deficit developing countries, a situation that was worsened by a simultaneous cutback in food aid supplies. The result was a real threat of worldwide food shortages, and even famine (FAO, 1974; UN, 1974).

There was a widespread famine in Ethiopia in 1984-85, affecting the inhabitants of today's Eritrea and Ethiopia. Over three million people were affected by the Somalian famine of 1991-93. More than 4 million people died, mostly from starvation and disease during the period of 1998-2004 in Congo and Ethiopia. About seven million inhabitants of Zimbabwe faced starvation because of its food crisis during 2000–2009 caused by the post-2000 land reforms. The severe food security crisis in Malawi affected more than five million people. In 2006, an acute shortage of food affected Somalia, Djibouti and Ethiopia and the northeastern part of Kenya. In 2007-2008, the world witnessed a food crisis, associated with steeply rising food prices that aggravated the situation for many countries already in need of emergency interventions and food aid, because of natural disasters and local conflicts. In some countries, this resulted in food riots and political instability. Recently, in 2010, Sub-Saharan countries such as Niger, Mali and Chad – already among the poorest countries in the world – are in the grip of a food crisis.

Major causes underlying food crises

Following are some of the main causes that have always played a role in food crises, in combination with poverty.

***a) Population growth and urbanization*:** The world population continues to increase and is projected to reach 10.0 billion in the year 2050 (Lutz *et al*.,

1997). Land is being overexploited to meet the increasing demand for food, fibre and shelter. Expansion of agriculture into marginal lands and clearance of natural vegetation, forests and wetlands, caused land degradation and resource depletion (Luck, 2007; Hedal Kløverpris, 2009; Bagan *et al*., 2010).

Population growth, combined with industrialization, has triggered urbanization. It is expected that most population growth between now and 2030 will be urban (Dixon *et al*., 2001). To meet the demands of the urban consumers, that form the backbone of the electorate practically everywhere, governments want to keep food prices low. This negatively affects food production, because farmers' decisions on growing crops strongly depend on the market situation. Moreover, urbanization causes changes in dietary habits and results for instance in a sharp rise in the consumption of animal products (CAST, 1999; Ma *et al*., 2004). Between 1993 and 2000, meat consumption in China almost doubled and is expected to grow in many other countries as well. Meat production is largely based on consumption of feed grains, an inefficient process, as 2 to 20 kg of feed dry matter is needed to produce 1 kg of meat. Currently, about one-third of the total world grain production is used for animal consumption (Keyzer *et al*., 2005), which contributes to high food prices.

b) ***Natural disasters*****:** Natural disasters such as adverse weather, pest and disease attacks, floods and earthquakes have always affected world food production, e.g., droughts in major wheat-producing countries in 2005-06 were among the causes of the recent food crisis of 2007-08.

c) ***Increasing oil prices*****:** High oil prices put pressure on agriculture by increasing the prices of inputs such as fertilizers and pesticides. Increasing oil prices played an important role in the food crises of the early 1970s and 2007-08.

d) ***Biofuels*****:** The production of fuels from plant material is increasing, to reduce the dependence on fossil fuels and to cut carbon dioxide emissions, requiring ever more land. In 2008, one third of the US maize production was used for biofuel production. The industry is encouraged by high subsidies in the United States and Europe. Land use for the production of biofuels, currently occupying about 15% in the EU and about 10% in North America, is projected to expand three- to four-fold at global level (IEA, 2007; Searchinger *et al.*,

2008). However, in Europe, emphasis is shifting towards second generation biofuels for reasons of sustainability, while addressing social aspects (such as preventing land grabs and environmental protection) and food security (EU Council, 2009).

Global efforts to achieve food security

Many initiatives have been taken to improve food security at both the global and individual country levels. For a very long time, concerted efforts have been made by international organizations and global governance groups such as the Food and Agriculture Organization (FAO), the World Food Programme (WFP), the International Fund for Agricultural Development (IFAD), the World Bank (WB), the World Health Organization (WHO) and the Consultative Group on International Agricultural Research (CGIAR) to strengthen global food security through research and aid programmes (Von Braun, 2010). Increasing agricultural production and reducing population growth are vital to achieve food security. High yielding varieties of crops are being developed through mutation breeding and biotechnology (Borlaug, 2000; Ingram *et al*., 2008). Increased crop production in many countries, as a result of the “Green Revolution”, has strongly contributed to their self-sufficiency in food. Breeding and improved agronomic practices, especially the increasing use of eternal inputs, such as fertilizers and irrigation, were successfully combined to improve the standard of living of billions of people through spectacular increases in crop yields. Since the 1960s, average yields of staples such as rice, wheat and maize have more than doubled in both developing and developed countries (Alexandratos, 1995; Dyson, 1996; Conway, 1997; Hafner, 2003).

The countries that have been successful in achieving food security are developed countries, characterised by rapid economic growth, especially in the agricultural sector and relatively low population growth. However, many developing countries, exposed to unfavourable agro-ecological and/or economic conditions, are still facing the problem of food insecurity. In 2001, UN member states recognized the need to more strongly focus on eradicating poverty to fight against hunger, and adopted the UN Millennium Development Goals (MDG’s), of which the first one (MDG1) is "to eradicate extreme hunger and poverty" (UN, 2000). MDG1 calls for halving hunger and poverty by 2015 in comparison to 1990.

Summarizing, the most important strategies that have been adopted to address food security issues are:

- Promote agricultural research: increase food production;
- Improve food aid: monitor the situation around the world and assist the needy through food aid;
- Stabilize food prices: prevent high prices, among others through control of the unpredictable and erratic behaviour of the food market;
- Regulate food trade: increase the quickly diminishing food reserves;
- Protect natural resources: safeguard natural resources such as land, water and biodiversity and increase water use efficiency;
- Adapt to and mitigate climate change: minimize adverse effects of climate change.

1.2 Need for enhanced agricultural monitoring

The threat of imbalances in food supply and demand, globally and/or locally, will continue to rise, as crop production faces the challenges of climate change, limited and dwindling resources, increasing energy needs and energy prices and population growth. The focus of many national and international organizations is on how to minimize this threat, while guaranteeing sustainable use of natural resources (Becker-Platen, 1976; IGOL, 2006; Justice and Becker-Reshef, 2007).

Sustainable use of land, one of the main natural resources, critically depends on continuous assessment and monitoring of the status of the land resources. To reduce the threat of local and/or temporary imbalances between food supply and demand, timely and accurate information on areas of crops and estimates of their production is needed. Policy makers, responsible for food security and land use planning, crop insurance companies and agricultural scientists need accurate and timely information on crop production at regional level. Such information is not only required to formulate policies aiming at food security and land use planning, but also to assess the success of policies implemented in the past. For trade organizations this is also relevant information, as a basis for making decisions on trade volumes, trade flows and price control and management of agricultural markets.

In this context, knowledge on the areas devoted to the various food crops is necessary baseline information for regional, national and/or international food production assessments. Such information should be regularly updated to account for the seasonality in agricultural production and for temporary or permanent agricultural land use changes. The value and relevance of such information substantially increases when it is geo-referenced. Following establishment of agricultural land use areas, regional crop production estimates are necessary. Currently, such information is lacking, not readily available and/or not compatible with other sources of information such as soil maps, interpolated weather data, thematic maps, etc.

A literature search has been performed to prepare an inventory of currently available information on land use at various spatial levels and the methods used in its acquisition. The results are briefly summarized here.

***a) Farmers' interviews*:** One of the simplest methods of acquiring land use and crop yield information is farmers' interviews, in which a representative sample of farmers is asked to estimate the land area occupied by different commodities and their expected production. However, this method has limitations. Farmers can be suspicious of enumerators, especially when they feel that the information might be used for tax purposes. Moreover, farmers in developing countries often lack resources and skills for quantitative estimation at the required level of accuracy (De Groote and Traoré, 2005).

***b) Land use surveys*:** Land use surveys are carried out for inventorying the specific land use practiced on a known unit of land that is considered homogenous in land resources, as a basis for the preparation of land use maps. Conventional methods of surveying are labour-intensive and time-consuming. The information on agricultural land use from such surveys soon becomes outdated, particularly in rapidly developing areas. The frequency of preparation of land use maps is generally very low, and moreover their spatial resolution is insufficient. In Europe for instance, land cover/land use maps, comprising only generalized classes (mix of land cover/land use, e.g. all field crops are grouped as one class "arable land"), are prepared at 10-year intervals (Feranec *et al.*, 2007). Useful information from land use surveys should include at least reliable estimates, at various spatial scales, of crop areas and crop yields. Aerial photographs and other remotely sensed sources of information are being used

extensively in various types of land use surveys. Aerial photographs are used to prepare the inventories of current land use, for instance, digitized boundaries of fields on aerial photographs are used in statistical sampling methods to collect data on crop areas and yield, livestock densities, forests, etc. (Carfagna and Gallego, 2005; Claggett *et al*., 2010), Aerial photographs and satellite data are also used for mapping of soil units, water resources and topographical features.

***c) Crop statistical data*:** Currently, many governments annually compile statistical information on cropped areas and crop yields in tabular form. Internationally comparable series of annual crop production data are available at national scale from the FAO, the European Commission's Directorate-General of Eurostat and the United States Department of Agriculture (USDA). While rich in commodity coverage, these data give no insight in the geographical distribution of crop areas and yields within countries. Several (sub-)regional efforts have been made by centres of the Consultative Group on International Agricultural Research (CGIAR) and by FAO (Carter *et al*., 1992; , FAO, 1994; IFPRI, 2001; IGOL, 2006) to collect and use crop statistical data. However, even when available, such data give no clue as to where "exactly" various crops are grown. This lack of spatial explicitness prevents the use of such data for monitoring of crop conditions and estimation of crop production, but they constitute an important source of information that can be used in planning and decision making on land use. FAO stresses the importance of such statistical data and highlights their importance in recognizing their collection and storage as one of the three primary roles of the organization (FAO, 2000).

Crop statistical data, though rich in information, lack detailed (mapped) information about specific land uses which makes their usefulness limited (Figures 1- 1 and 1- 2). The wheat area map of Spain (Figure 1- 1), based on the agro-maps initiative of FAO (**http://www.fao.org/ag/agl/default.stm)**, shows the land area under wheat in various municipalities of Spain, without indicating exactly where those areas are located. The rainfed wheat area map at municipal scale (% of the agricultural area as defined by the CORINE land cover map (Figure 1- 2), does not show "exactly where" in the various municipalities the crop is grown. Such maps are not suitable for monitoring and early warning systems on crop production (Verburg *et al*., 2002; Aalders and Aitkenhead, 2006).

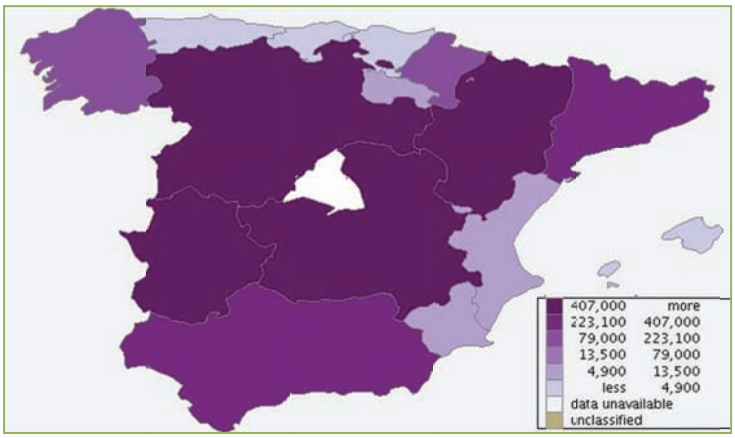


Figure 1- 1: Wheat area map of Spain, downloaded from FAO's agro-maps (Source: http://www.fao.org/ag/agl/default.stm)

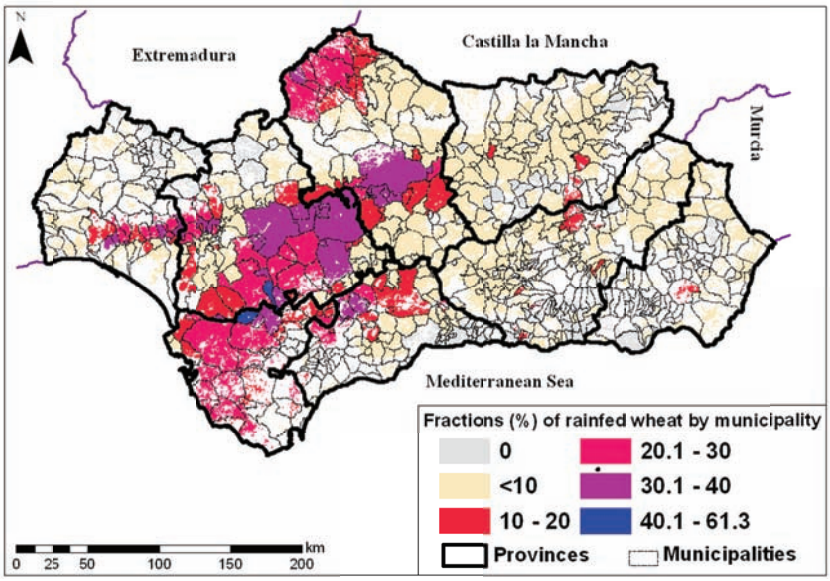


Figure 1- 2: Rainfed wheat areas of Andalucía, Spain based on crop statistical data (2001-05) (Source: Ministry of Agriculture and Fisheries, Andalucía, Spain)

***d) Crop yield estimation methods*:** A wide range of crop yield estimation methods are being used, such as expert estimates, empirical models and process-based biophysical models. In the 1950s, regression models were

developed to estimate crop yield, using weather conditions as the predictors such as regression models of relationships between monthly rainfall, monthly temperatures, and crop yields (e.g., Runge and Odell, 1958; Thompson, 1962). In the late 1960s, development of dynamic simulation models started. The rapid development and application of computers facilitated research on crop simulation models. As a result, in the past forty years, a wide variety of crop models has been developed all over the world to serve many different purposes (Matthews and Stephens, 2002). DSSAT (Decision Support System for Agro-technology Transfer) was developed in the USA by IBSNAT (International Benchmark Sites Network for Agrotechnology Transfer) (Jones *et al.*, 2003). The APSIM (Agricultural Production system SIMulator) modeling framework was developed by APSRU (Agricultural Production Systems Research Unit) in Australia (Keating *et al.*, 2003; Thorburn *et al.*, 2010). In the Netherlands, the late C. T. de Wit started work on crop growth modeling at the Department of Theoretical Production Ecology of Wageningen Agricultural University (Van Ittersum *et al*., 2003; Van Keulen *et al*., 2008). One of the crop growth models developed in the 'Wageningen School' (Bouman *et al*., 1996), in the framework of the Centre for WOrld FOod STudies, is WOFOST, which is the core of the Monitoring Agriculture with Remote Sensing (MARS) program of the European Union implemented at the Joint Research Center (JRC) (Van Diepen *et al*., 1989; Reidsma *et al*., 2009). Other models developed in Wageningen are SUCROS (Simple Universal CROp growth Simulator; Van Laar *et al*., 1997) and ORYZA (a crop growth model for rice; Bouman *et al*., 2001).

However, only a limited number of systems for quantitative yield assessment are operational.

1.3 Remote sensing data for acquiring agricultural land use information

Remote sensing (RS) started in 1858 when Gaspard-Felix Tournachon first took aerial photographs of Paris from a hot air balloon. The initial planned uses of remote sensing were performed during the U.S. Civil War, when messenger pigeons, kites, and unmanned balloons were flown over enemy territory with cameras attached. The term 'remote sensing' was coined in the early 1960's by the staff of the Office of Naval Research and Geography (Pruitt, 1979). Though the initial uses of remote sensing were military, the work on vegetation also

started in the middle of the 20th century, with the use of Colour Infra-Red (CIR) photographs for vegetation studies, such as classification of vegetation types, detection of diseased and damaged vegetation and detection of severe crop stress symptoms (Ustin and Gamon, 2010). Concurrently, the radar technology was being developed and side-looking airborne radar (SLAR) and Synthetic Aperture Radar (SAR) were developed to improve the angular resolution for better image interpretation (Jongschaap, 2006).

The use of satellites started with the space race between the former USSR and the United States of America (USA) in the late 1950s, while the use of satellite-based crop imagery began in 1971 through the launch of Landsat 1 by the National Aeronautics and Space Administration of the USA (Williams *et al*., 2006). Data are acquired by sensors on board of such satellites. Since then, many satellites with sensors have been launched to acquire data. Of special importance are the National Oceanic and Atmospheric Administration (NOAA) satellite with the Advanced Very High Resolution Radiometer (AVHRR) sensor on board, the Terra satellite that carries the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) sensor and the Moderate Resolution Imaging Spectroradiometer (MODIS). Another system is the Système Probatorie de la Observation de la Terre (SPOT) satellite that also carries a vegetation sensor (VGT). Currently, highly advanced satellite systems such as Quickbird and IKONOS obtain data from space that are comparable to aerial photography in terms of spatial resolution.

Satellite remote sensing has enabled acquisition of land use/land cover and vegetation information at different spatial and temporal scales, including relevant information on agricultural land use. Remote sensing in combination with Geographical Information Systems (GIS) and Global Positioning Systems (GPS) can be used to assess the characteristics and growth of vegetation. Vegetation behaviour depends on its specific (physical, physiological and biochemical) characteristics and its interactions with the aerial and soil environments, characterized by weather conditions, such as solar radiation, temperature, humidity and rainfall, and the availability of plant nutrients and (soil) water. Vegetation indices (VIs) have been extensively used for monitoring vegetation and land cover changes (DeFries *et al*., 1995). Development of vegetation indices is based on differential absorption, transmittance, and reflectance of energy by the vegetation in the red and near-infrared regions of

the electromagnetic spectrum (Jensen, 1996). Many remote sensing studies of vegetation have focused on the use of spectral vegetation indices such as the Normalized Difference Vegetation Index (NDVI) and other simple ratios, calculated as combinations of near infrared (NIR) and red reflectance. These indices have been shown to correlate with plant variables related to primary production such as biomass, leaf area index (LAI) and absorbed photosynthetically active radiation (APAR). Dry matter production in plants is based on the process of photosynthesis, the rate of which is determined by the plant's biochemical characteristics and absorbed solar radiation. Remote sensing systems, i.e. meteorological and earth-observing satellites can provide inputs for crop production estimates (Ray and Dadhwal, 2001; Lobell *et al*., 2003; Prenzel, 2004; Launay and Guerif, 2005). The use of multitemporal images results in higher classification accuracy and leads to consistent accuracy in all classes being mapped. Multitemporal data are especially advantageous in areas where vegetation or land use changes rapidly. This offers many opportunities for more complete vegetation description than could be achieved with a single image. For example, the differences between evergreen and deciduous trees can be identified, as the former appear uniform throughout the year, whereas the latter strongly differ between leaf-on and leaf-off periods. The discriminative capabilities of multitemporal observations are based on their characterization of seasonal dynamics of vegetation growth (Agrawal *et al*., 2001).

Hypertemporal (long temporal sequences of regularly acquired data) remote sensing data serve as an important source of information on crops, because of their spatial and temporal resolution. Such imagery meets the requirements for land cover mapping at regional scale.

Applications of remote sensing in agriculture

Following is the brief account of various applications of remote sensing in agriculture:

***a) Land cover/land use mapping*:** Regional land cover mapping is performed to obtain an inventory of its vegetation to be used as a baseline map for monitoring of the dynamics of land use and land management. Satellite observations acquired by the hypertemporal remote sensor can be translated into

land cover information (Latifovic *et al*., 2004; Ran *et al*., 2010). Satellite remote sensing techniques have been shown to be effective in preparing accurate land use/land cover maps and in monitoring changes at regular intervals (Yang *et al*., 2007; Zhang and Zhang, 2007).

Remote sensing offers an efficient and reliable means of collecting the information required to map crop type and area. Hypertemporal imagery facilitates classification by taking into account changes in reflectance as a function of plant phenology (stage of growth) (De Bie *et al*., 2010).

In areas of persistent cloud cover or haze, radar is an excellent tool for observing and distinguishing crop types, due to its active sensing capabilities and long wavelengths (McNairn *et al*., 2009). High resolution satellite imagery also provides an efficient tool for mapping crop types and is being used for area stratification purposes (Gallego and Bamps, 2008; Knorn *et al*., 2009).

***b) Crop monitoring and damage assessment*:** Remote sensing can provide information about possible pest and disease infestations of the vegetation. The spectral reflection of a field varies with changes in phenology (growth stage), and possible stress conditions can be monitored by multispectral sensors. The wavelengths in the optical (VIR) range are highly sensitive to crop vigour, as affected by stress and crop damage. Recent advances in communication technology allow support in timely decision-making on crop management from images of agricultural fields. NDVI time series have also been used for monitoring anomalies, drought, phenology, land cover characteristics and crop yields (Goerner *et al*., 2009).

***c) Remote sensing and crop growth modelling*:** Information from remote sensing observations can effectively be integrated into crop modeling methodologies. Such data have been used in crop models for regional yield assessment (Roebeling *et al.* 2004; Doraiswamy *et al.* 2005; Jongschaap, 2006; De Wit and Van Diepen, 2008). The use of satellite-based inputs highly simplifies the process, considering the amount of time and labor that regional level data collection requires. The remote sensing images can also be used for aggregation of results of crop growth models to regional scales.

1.4 Problem statement

To increase food security for an increasing population, information is required on current land use, and the dynamics of land use should be monitored. Land use maps, including detailed temporal and spatial information, are needed for land use planning, crop monitoring and food security issues. At regional to global scales, land use data are either missing or, if available, the information is in most cases outdated and therefore not useful. Moreover, the quality of the available information is variable and often poor (Dalal-Clayton and Dent, 1993; Sombroek and Antoine, 1994; Fresco *et al*., 1997; De Bie, 2000; Nachtergaele, 2000). This lack of information on land use is one of the major obstacles hampering efficient policy making and research to achieve food security. With these problems in mind, I identified the following objectives for my Ph.D. research.

1.5 Objectives of the study

This study aims at contributing to development of an operational methodology for quantitative mapping and monitoring of agricultural land use, including the assessment of the possibilities for crop growth modelling based on remotely sensed input data. The objective is to develop remote sensing- and GIS-based methods that result in timely availability of accurate agricultural land use/land cover information, as required by agricultural land use planners, policy makers, donor agencies and crop insurance companies for a variety of purposes.

The specific objectives of the study are:

1a To develop a method for compiling spatial and temporal land use data sets by combining hypertemporal NDVI images with other available data sources.

1b To further develop and validate a method for generating spatially explicit crop area maps by combining hypertemporal NDIV images, crop statistical data and primary field data.

1c To test whether the information derived from NDVI images, used in generating spatially explicit crop area maps, is affected by soil type.

2 To compare crop yields estimated by a simple and new (under development) crop growth model with those generated through the Crop Growth Monitoring System (CGMS) of the European Union's Monitoring Agriculture with Remote Sensing (MARS) project (Van Diepen *et al*., 1989; Reidsma *et al*., 2009).

3 To test whether the intended stakeholders of the research have benefited from its results through top-down valorisation of the produced outputs.

1.6 Outline of this thesis

The present study makes use of satellite-based remotely sensed data, crop statistics, crop calendar information and *in situ* observations to allow reliable, unambiguous and quantitative interpretation of the data.

The innovative characteristic of this study is identification of a combination of remote sensing, GIS and crop modelling as a basis for development of an operational system to monitor and map agricultural land use (Figure 1- 3).

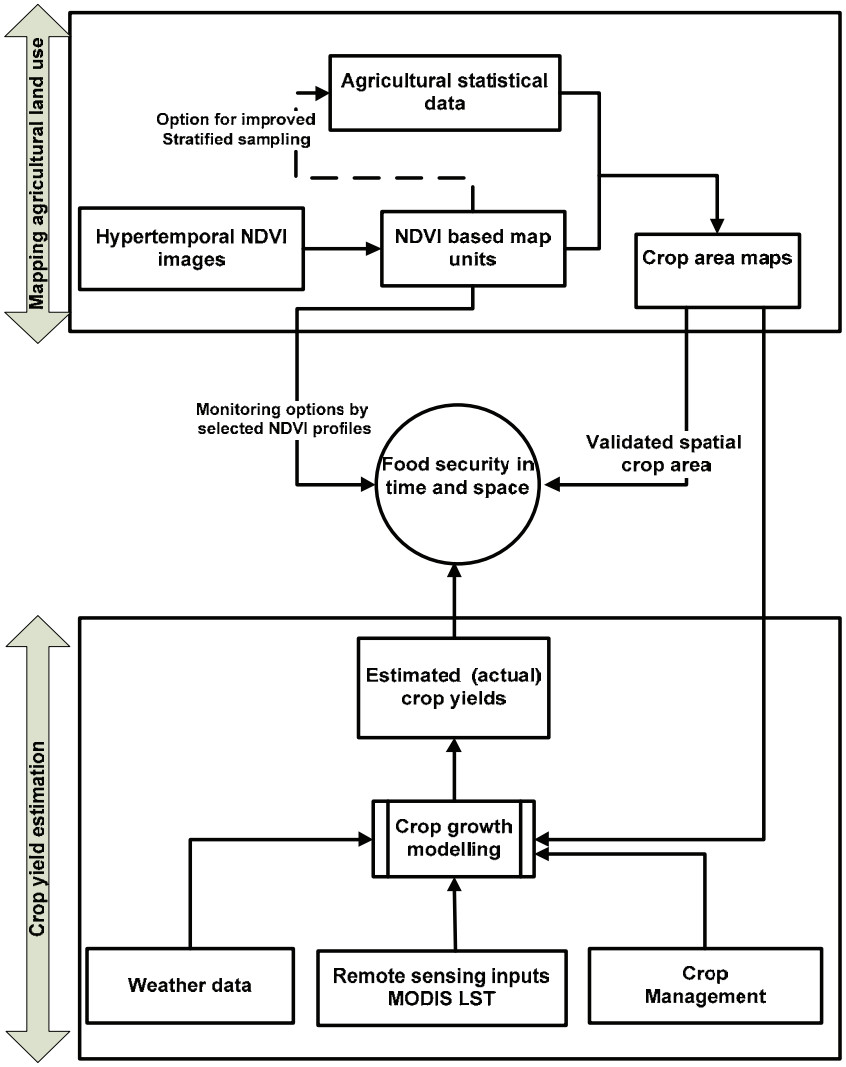


Figure 1- 3: Schematic diagram of the research framework

This thesis is organized in seven chapters (Figure 1- 4).

***Chapter 2*:** This chapter deals with mapping of agricultural land use on the basis of various data sources such as crop calendar information, agricultural statistics, maps of land cover and satellite-based hypertemporal SPOT NDVI images. The method illustrated in Chapter 2 has the potential to be incorporated in remote sensing- and GIS-based drought monitoring systems.

***Chapter 3*:** In this chapter, a method is presented to generate crop area maps by combining hypertemporal SPOT NDVI images, agricultural statistics and primary field data. It is assumed that the NDVI data express the combined influences of all environmental variables such as soil, terrain, weather and land use conditions.

***Chapter 4*:** In this chapter, some of the assumptions from Chapters 2 and 3 are tested by incorporating soil information derived from the soil type map and the soil geogenisis map. Moreover, the methods developed in Chapters 3 and 4 are evaluated by selecting the "best crop map" from the maps generated, based on their accuracy.

***Chapter 5*:** In this chapter, the output of a new crop growth model (C*f*-Water) is compared with that of an operational model, WOFOST, as incorporated in the Crop Growth Monitoring System, by using published crop statistical data and primary field data.

***Chapter 6*:** In this chapter, stakeholders, with a wide range of interests, have been involved in valorizing the products generated in this thesis (outputs of 3rd, 4th and 5th chapter) and in comparing and analyzing their opinions about currently available options.

***Chapter 7*:** In this chapter, a synthesis is presented, evaluating the strengths and weaknesses of the developed methods along with recommendations for further research to acquire and present timely, accurate and meaningful information on agricultural land use.

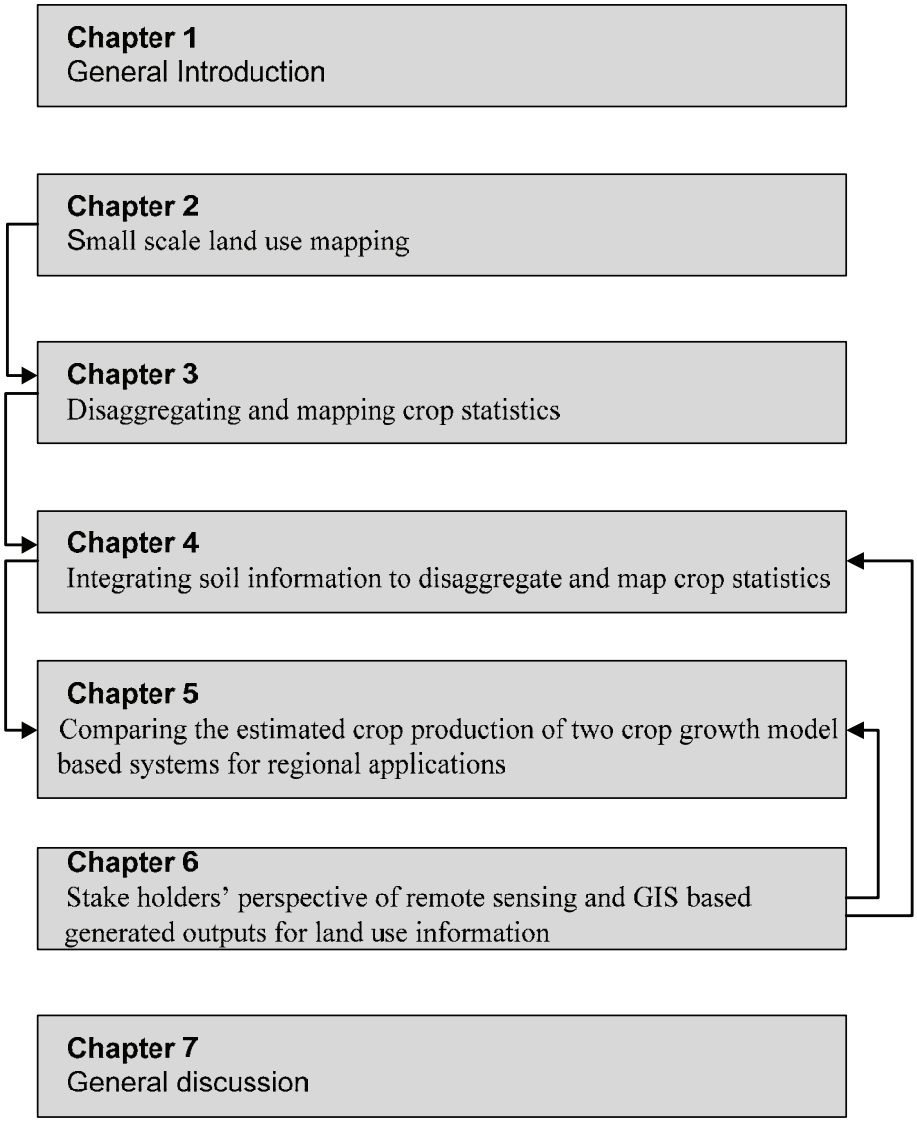


Figure 1- 4: Coherence of thesis chapters

2

*Small Scale Land use mapping

* This chapter is based on de Bie, C.A.J.M., Khan, M. R., Smakhtin, V.U., Venus, V., Weir, M.J.C., Smaling, E.M.A. (accepted 10-Jul-2010). Analysis of multi temporal NDVI SPOT images for small scale land use mapping In: International journal of remote sensing. DOI: 10.1080/01431161.2010.512939.

Abstract

Land use information is required for a number of purposes such as to address food security issues, to ensure the sustainable use of natural resources and to support decisions regarding food trade and crop insurance. Suitable land use maps often either do not exist or are not readily available. This chapter presents a novel method to compile spatial and temporal land use data sets using hypertemporal remote sensing in combination with existing data sources. SPOT-Vegetation 10-day composite NDVI images (1998-2002) at 1-km^2 resolution for a part of Nizamabad district, Andhra Pradesh state, India were linked with available crop calendars and information about cropping patterns. The NDVI images were used to stratify the study area into map units represented by 11 distinct NDVI classes. These were then related to an existing land-cover map compiled from high resolution IRS-images (Liss-III on IRS-1C), reported crop areas by sub-district and practiced crop calendar information. This resulted in an improved map containing baseline information on both land-cover and land use. It is concluded that each defined NDVI class represents a varying but distinct mix of land-cover classes and that the existing land-cover map consists of too many detailed 'year-specific' features. Four groups of the NDVI classes present in agricultural areas match well with four categories of practiced crop calendars. Differences within a group of NDVI classes reveal area specific variations in cropping intensities. The remaining groups of NDVI classes represent other land-cover complexes. The method illustrated in this chapter has the potential to be incorporated into RS/GIS based drought monitoring systems.

2.1 Introduction

Population growth drives the increasing demand for food, consequently putting pressure on land resources for higher food production. This over exploitation of land resources causes land degradation, harvest shortfalls and ultimately the food shortage. Timely and accurate assessment of food production systems is required to counter the threats posed by land degradation, food shortages, adverse effects of climate change and natural disasters. Agronomic baseline maps are required not only for accurate assessment of the agricultural land use systems but also for optimization of inputs in these systems (Hatfield, 2001; Tilman *et al*., 2002).

Timely and reliable information on crop areas and yields is required by the governments of all the countries and also by international organizations for planning and management of land resources. The international organizations such as the Food and Agriculture Organization (FAO) of the United Nations, the European Union (EU), the United Nations Environment Programme (UNEP) and the United States Geological Survey (USGS) stress on the importance of regional to global databases on crop areas and production (Cohen and Shoshany, 2002; George and Nachtergaele, 2002; Townshend *et al*., 2008). Spatial land use data is also needed for understanding the role of land use in natural resource management and environmental change studies. Many countries monitor land use change as a basis for policy guidelines and for implementing plans of action. Land use data has been considered as the second most fundamental set of statistics after the population data (Young, 1998). In many developing countries, however, there is a general paucity of reliable land use information or, even if such information exists, it is often difficult for all interested clients to access it. Preparation of conventional land use maps is both expensive and time consuming and the information contained in them quickly becomes outdated.

Annual estimates of areas of agricultural land use (e.g., crop areas) are usually available in the form of tabular statistical data collected at various administrative levels. The user of such data, however, remains uninformed about the exact geographic locations of specific land uses within the administrative units (Jansen and DiGregorio, 2003). Consequently, such information is poorly suited for monitoring of crop production and food security

studies because of this lack of spatial reference. Remote sensing and GIS techniques provide means for analysing and monitoring the spatial and temporal aspects of land use (Walker and Mallawaarachchi, 1998; Oetter *et al*., 2001).

The relation between remotely sensed data and crop characteristics is frequently described by means of vegetation indices. Vegetation indices are indictors of vegetation conditions and have been used to derive information about land use and land-cover (Maseli, 2004). The Normalized Difference Vegetation Index (NDVI) is calculated as the ratio of the difference between the red and infrared reflectance to their sum. NDVI expresses the photosynthetic activity of vegetation (Justice *et al*., 1985; Sellers, 1985; Drenge and Tucker, 1988; Ringrose *et al*. 1996; Unganai and Kogan, 1998; Maggi and Stroppiana, 2002; Archer, 2004; Weiss *et al*., 2004). Several studies have also discussed the suitability of temporal NDVI data for studying vegetation phenology, especially that of crops (Gorham, 1998; Murakami *et al*., 2001; Uchida, 2001; Groten and Octare, 2002; Hill and Donald, 2003; Huang and Siegert, 2006). Various authors have attempted to map land-cover phenology, land-cover dynamics and land degradation through the analysis of multi-temporal NDVI data (e.g. Cayrol *et al*., 2000; Ledwith, 2000; Eerens *et al*., 2001; Brand and Malthus, 2004; Budde *et al*., 2004; Fraser *et al*., 2009; Julien *et al*., 2009). Others have aimed at disaggregating agricultural statistics using NDVI data (Walker and Mallawaarachchi, 1998; Khan *et al*., 2010). The 'MARS Food Aid' project of the European Union's Joint Research Centre (JRC) has even integrated SPOT images into their methods for integrated agricultural monitoring and yield forecasting for Africa (Rembold, 2004).

In this chapter, we utilized 10-days temporal resolution SPOT-Vegetation NDVI data to develop a method for identification and describing areas with different vegetation cover types and agricultural areas that follow different crop calendars. Normally, crops are on the ground for 6-7 months and substantial information in terms of number of images is required. Therefore, the SPOT VGT satellite data with a spatial resolution of 1-km^2 and a temporal resolution of one day was selected for this research. Although land use and land-cover change is an important issue (Lepers *et al*. 2005), the focus of present research was on mapping land use and land-cover rather than on change detection. The aim of the research is to develop methods to compile spatial and temporal land

use data sets using both existing, non-spatial data sources and improved RS/GIS based analyses.

2.2 Study area

The study area is situated on the Deccan plateau in the western part of Nizamabad district of Andhra Pradesh state, India (Figure 2- 1). This area was selected because of its heterogeneity not only in terms of the environmental conditions but also with respect to resources per farming unit, levels of development, and practised land use (see also Figure 2- 3). The total arable land of the area comprises about 900 km^2. The soils in the study area belong to four major orders according to the United States Department of Agriculture (USDA) soil classification system: Inceptisols (67%), Alfisols (15%), Vertisols (10%) and Entisols (8%). The area is relatively flat, with nearly 69% of the area having a slope of less than 1% and a further 12% of the area having slopes between 1% and 3%. The climate of the area is tropical. The annual average rainfall of about 900 mm occurs over a period of about 60 days and 95% of this falls during the southwest monsoon.

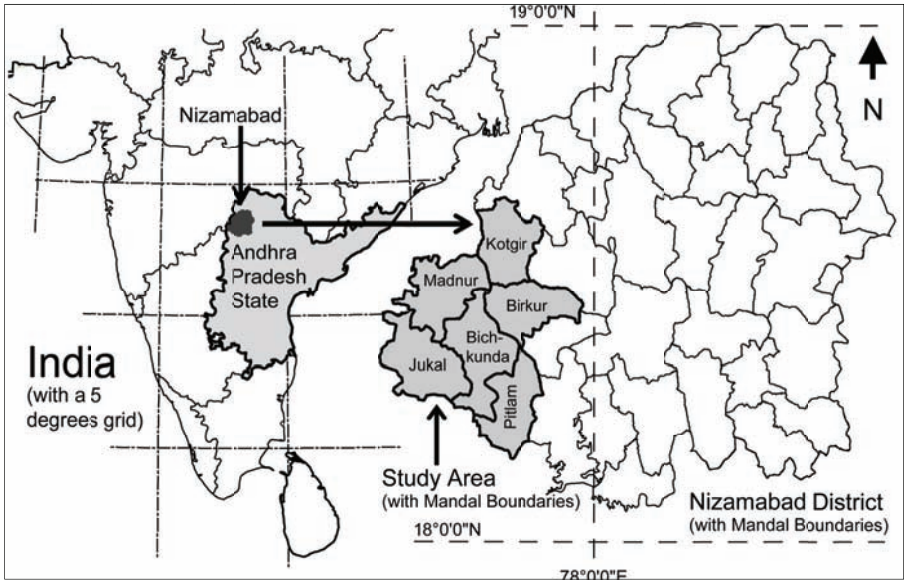


Figure 2- 1: Study area location

Administratively, the study area comprises six *Mandals* (sub-divisions of districts): Kotgir, Birkur, Bichkunda, Madnur, Jukal and Pitlam, with a total area of about 1300 km^2. These six *Mandals* contain 220 villages and have a total population of 294 000 (Census of India, 2001). Traditionally, the primary occupation of the local population is agriculture with about 80% of the inhabitants relying on agriculture for their livelihood.

2.3 Data used

In the research, six main data sets were used:

a) In total, 147 geo-referenced and de-clouded SPOT 4 Vegetation 10-days maximum value composite (MVC) NDVI images (S-10 product) at 1-km^2 resolution from April 1998 to April 2002 were used. The word 'de-clouded' means that only those pixels satisfying quality requirements were retained. These requirements are *i*) a 'good' radiometric quality for bands 2 (red; 0.61-0.68 µm) and 3 (near IR; 0.78-0.89 µm) and *ii*) the pixels were not labelled as 'shadow', 'cloud' or 'uncertain' but as 'clear' in the quality flags by pixels accompanying the NDVI data. The removed pixels were labelled as 'missing'. These images were obtained from www.VGT.vito.be (figure 2- 2).

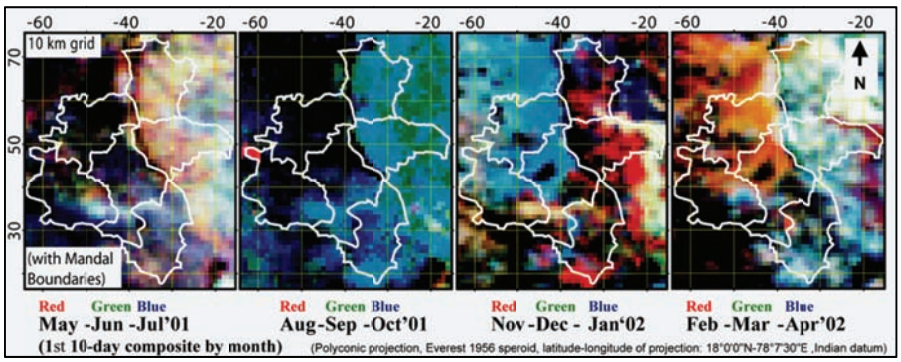


Figure 2- 2: Sequence of original SPOT 4 Vegetation 10-days composite NDVI images

10-days composite data is made by using the best quality daily images from the 10 day period. NDVI indicates chlorophyll activity and is calculated from the formula:

$$(\text{band } 3 - \text{band } 2) / (\text{band } 3 + \text{band } 2).$$

The index is converted to a digital number (DN-value) in the 0-255 data range using the standard formula:

$$DN = (NDVI + 0.1) / 0.004.$$

b) A land-cover map of the area at 1:50 000 scale, based on high resolution 1994-95 IRS-C images (Liss-III; 23m resolution) of two periods, *i.e.*, *Kharif* (monsoon) and *Rabi* (post monsoon) was used. This map was compiled as a part of the Integrated Mission for Sustainable Development (IMSD) project which is an operational land use planning programme in the area (NRSA, 1995). The map was based on visual interpretation techniques in conjunction with collateral data such as topographic maps, district records, and the field investigations. The original land-cover map has 18 legend entries (at 3 levels). In this study only seven renamed legend entries (at 2 levels) were used (Table 1; Figure 2- 3).

Table 2- 1: Land cover classes used in this study

Original Legend			km-sq	Used Legend		Assigned Colour
Level-1	**Level-2**	**Level-3**		**Level-1**	**Level-2**	
Agriculture	Crop land	Kharif	363.5	Fields	With green cover in Rabi only	orange
		Kharif+Rabi(double Cropped)	219.2		With green cover in Kharif only	l.green
		Rabi	294.2		With green cover in Rabi and Khari	d.green
	Fallow		1.5		Units are ignored (too small)	
	Plantations		0.8			
Built-up	Built-up		3.8	Built-up (towns)		red
	Towns/cities(Urban)		0.6			
	Villages(Rural)		5.7			
Forest	Deciduous (Moist/Dry)	Dense/Closed	77.4	Non-Fields	With trees (forests)	brown
		Forest Blanks	2.3			
		Open	58.6			
		Scrub Forest	2.3			
Wastelands	Barren Rocky/Stony Waste/Sheet		1.2		Almost bare	white
	Land With Scrub		94.5			
	Land Without scrub		121.6			
Water bodies	River	River	31.9		With water (water bodies)	blue
		Sandy area	12.2			
	Tanks		26.6			

(The 18 legend entries of the original land cover map as mapped at 1:50,000 by the IMSD project are re-assigned to seven legend entries)

c) Crop calendar data were obtained from published literature (APAU, 1989; Rao, 1995). The crop calendar data refer to major events (agronomic practices in the cropping system per year in a specific area) such as field preparation, sowing time and harvesting time.

d) Crop statistics were collected by a District Agricultural Officer and his staff, who, as part of their regular duties, compile the data on a seasonal basis at village level and then aggregate them to different administrative levels. These data are published as an official record in the District Gazetteer by the Chief Planning Officer of the District (CPO, 2001).

f) The meteorological data for this study were compiled from official monthly records. Rainfall data on a daily basis are available from the meteorological station located in the study area.

e) Soil maps, produced at 1:50 000 scale, were obtained from the Integrated Mission for Sustainable Development (IMSD) project in India. The procedures used to compile these maps are discussed in detail in the IMSD Technical Guidelines (NRSA, 1995).

2.4 Methods

The methods of this study comprise four distinct steps: image classification, linking hypertemporal RS images with the existing land-cover map, linking hypertemporal RS images with crop statistical data by sub-district and finally the legend construction and matching with known crop calendars.

2.4.1 Image classification

The iterative self organized unsupervised clustering algorithm (ISODATA) of ERDAS IMAGINE software was used to perform an unsupervised classification of all 147 NDVI images. The ISODATA clustering method uses the minimum spectral distance formula to form clusters. The algorithm repeats the clustering of the image until either a maximum number of iterations have been performed, or a maximum percentage of unchanged pixel assignments have been reached between two iterations.

A series of unsupervised classification runs was carried out to generate classified NDVI maps with a pre-defined number of classes (5 to 35) with an increase of 1 in every proceeding step. The maximum number of iterations was set to 50 and the convergence threshold was set to 1.0. Each iteration produced a classification result with the desired number of classes. The algorithm is called "self organizing" because the procedure is not influenced by additional data or by expert's knowledge. The ISODATA algorithm minimizes the Euclidian distances inherent in the data and locates clusters (ERDAS 2003). The divergence statistical measure of distance between generated cluster signatures by each run was used to select the optimal run (ERDAS, 2003; Swain and Davis, 1978). The optimal run with a distinguishable peak in divergence separability was selected for further study. From that optimal run, comparable NDVI classes (annual averaged classes) were explored visually and in this way a supervised grouping of the NDVI classes was performed.

2.4.2 Linking hypertemporal RS images with existing land cover map

After the supervised grouping of the NDVI classes, the resultant NDVI map was compared with the existing land-cover map of the area (Figure 2- 3). This comparison was made to derive quantified data on the mix of land-cover classes represented by each NDVI class. The land-cover map at 1:50 000 scale shows field patterns and specifically, reflects the cover status of those fields for the period of the imagery (1994-95) used to compile the map. Differences with the 18 January 2000 pan sharpened IRS-1D image (Liss-III and PAN; 5.8m resolution; Figure 2- 3) can be detected visually taking the seasonality aspect of the IRS-1D image into consideration. This indicates the degree of land use change in the area from 1994/95 to 2000.

2.4.3 Linking hypertemporal RS images with crop statistical data by sub-district

The area of all the NDVI classes in each *Mandal* was calculated through GIS analysis by combining the classified NDVI map and the administrative map of study area.

Table 2- 2 reports crop statistical data for six *Mandals* (averaged from 1998-2001) and the area of NDVI classes in each *Mandal*. These data were used to estimate the function:

Cultivated Area by a Crop = f {extent of NDVI classes} [both in ha/Mandal]

This function was estimated through stepwise forward multiple regression for each crop with no constant and coefficients constrained between 0.0 and 1.0. This constraint is introduced because the area of a crop in a *Mandal* can neither be negative nor more than the area of that particular *Mandal*. The model is thus as described by (Khan *et al.* 2010):

$$Y = \sum_{i=1}^{n} b_i x_i + \varepsilon_i \qquad \text{(Eq.1)}$$

where,

Y	= Average area of a crop (ha) per *Mandal* (1998-01)
b_i	= Regression coefficient for NDVI class i per *Mandal*
x_i	= Average area (ha) of NDVI class i per *Mandal* from
n	= Number of NDVI classes
ε_i	= Residual error

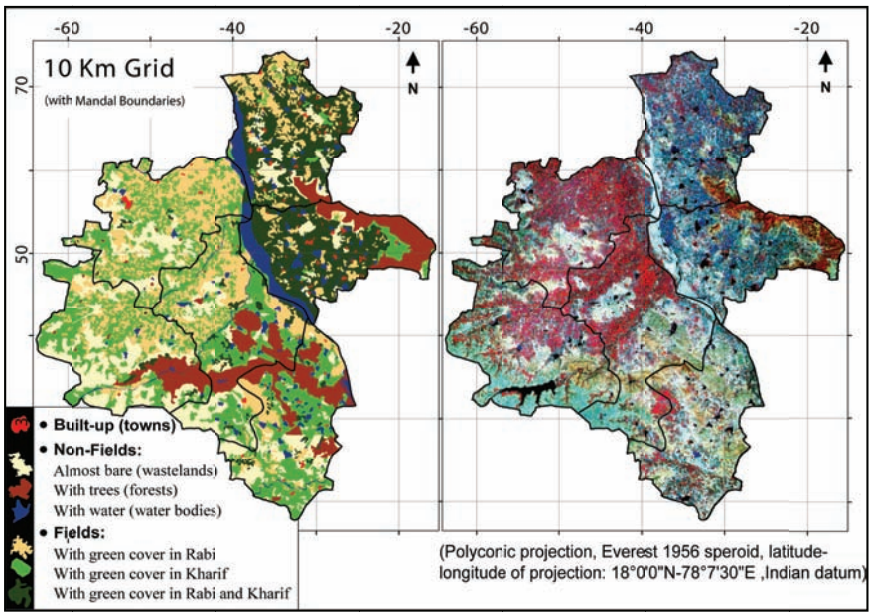


Figure 2- 3: Existing land cover map (Left) and a pan-sharpened IRS-1D image of 18 Jan. 2000 (Right)

Table 2- 2: Crop statistics for 6 Mandals

KHARIF	Bichkunda	Birkoor	Jukkal	Kotagir	Madnoor	Pitlam	Total(ha)
Irrig.Rice	3,255	6,919	793	7,325	1,282	3,200	22,774
Sorghum	718	4	2,305	1	371	564	3,963
Maize	12	27	52	7	0	384	482
Pulses	5,986	93	6,609	1,312	9,268	5,853	29,122
Sugarcane	230	372	4	951	64	774	2,395
Cotton	628	11	2,524	313	2,910	474	6,860
Other Crops	218	16	518	151	150	50	1,103
Total(ha)	11,047	7,442	12,805	10,060	14045	11,299	66,699
RABI							
Irrig.Rice	651	3,995	247	4,153	335	2,100	11,481
Sorghum	2,462	16	3,035	1,181	6,987	1,773	15,454
Millet	0	4	0	34	0	0	38
Maize	6	10	0	0	0	61	77
Pulses	885	4	368	243	1,001	323	2,823
Sugarcane	245	244	0	803	0	668	1,960
Groundnut	657	820	469	1,205	658	2,133	5,942
Other Crops	1,153	187	1,771	337	1,966	433	5,847
Total(ha)	6,059	5,280	5,890	7,956	10,947	7,490	43,622

NDVI-Class*	Bichkunda	Birkoor	Jukkal	Kotagir	Madnoor	Pitlam	Total (km^2)
1,2	0	1	0	1	0	0	2
11	16	18	2	16	28	2	81
3,4,5,6	33	0	91	0	10	3	136
19,20	4	34	9	3	0	0	50
16	3	8	6	3	0	15	34
10,12,14	71	2	17	3	1	140	234
7,8	112	0	105	13	173	4	407
9	8	1	5	6	10	0	29
13,15	7	15	1	30	0	38	91
17	0	109	0	102	0	1	212
18	0	12	0	28	0	0	40
Total (km^2)	254	200	237	204	222	202	1,318

** Product of the image classification step discussed under 'methods'.*

The regression coefficients in the equation provided an estimate of the fractions (percent per unit area) of each related NDVI classes representing a specific crop.

2.4.4 Legend construction and matching with known crop calendars

The final stage of the study – compiling a legend and matching it with known crop calendars - was limited only to those areas where crops are grown as indicated by the statistical data and where 'green' fields were dominant in the existing land-cover map. For these areas, a legend for the NDVI map was compiled from the data generated by the method explained in sections 2.4.3 and 2.4.2. Known crop calendars were assigned to the relevant legend entries and cross-checked by using the relevant NDVI classes.

2.5 Results

2.5.1 Image classification

As described in section 2.4.1, image classification involved two steps. Firstly, the extent to which the generated signatures are separable across various classification runs was compared by using the divergence statistical values. Figure 2- 4 shows by each run *i)* the average and *ii)* the minimum of all generated divergence values between the defined clusters. The 'average' separability values reflect a comparison of all generated clusters by run simultaneously and the 'minimum' separability values compare only those that are the most similar. The figure shows the best 'average' values of the runs for 22 and 28 classes and the best 'minimum' values of the runs for 18 and 23 classes. We have considered the 'minimum divergence' as more appropriate than the 'average divergence' because of clearer peaks in values. Thus, in an unsupervised classification of the study area, somewhere between 18 and 23 classes is optimum to classify the 1-km^2 NDVI images from 1998-2002. Within that range, variation in the divergence values can be considered as 'noise' in the satellite data.

Secondly, as input into the next step, the 20 classes map was selected and a supervised grouping of the 20 NDVI classes was carried out. The supervised (visual) grouping was undertaken to repair possible deficits because one single peak could not be identified in both the minimum and average divergence

separability values. Therefore, the next step after unsupervised classification was exploring visually the average annual NDVI class profiles of the selected 20 classes NDVI map. Based on the similarity of the average annual NDVI class profiles, several of them were grouped (on an area weighted basis) in such a way that the areas represented by them were always adjacent to one another. Figure 2- 5 shows that, on a visual basis alone, 15 NDVI classes could be grouped into six generalized NDVI classes without losing important details. The remaining five classes were retained as standalone classes.

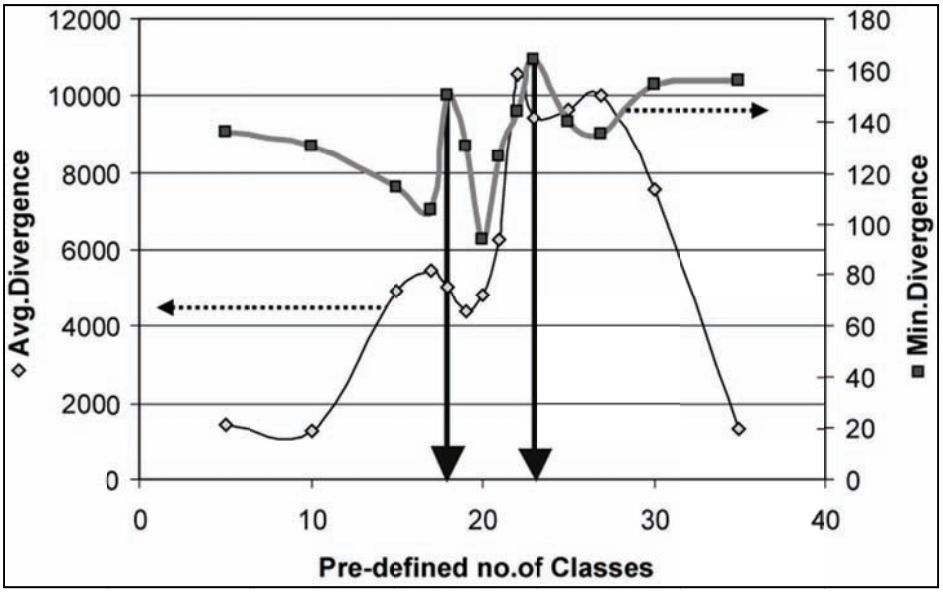


Figure 2- 4: Average and minimum divergence values produced through a number of unsupervised classification runs

The resulting 11 NDVI classes (6 grouped and 5 standalone) and the 1-km^2 resolution NDVI unit map are presented in figure 2- 6. By this process, annual variations in the original NDVI classes were eliminated that could have caused a problem for the ISODATA algorithm to differentiate the classes. The larger, relatively uniform, areas are indicated in figure 2- 6 as A, B and C.

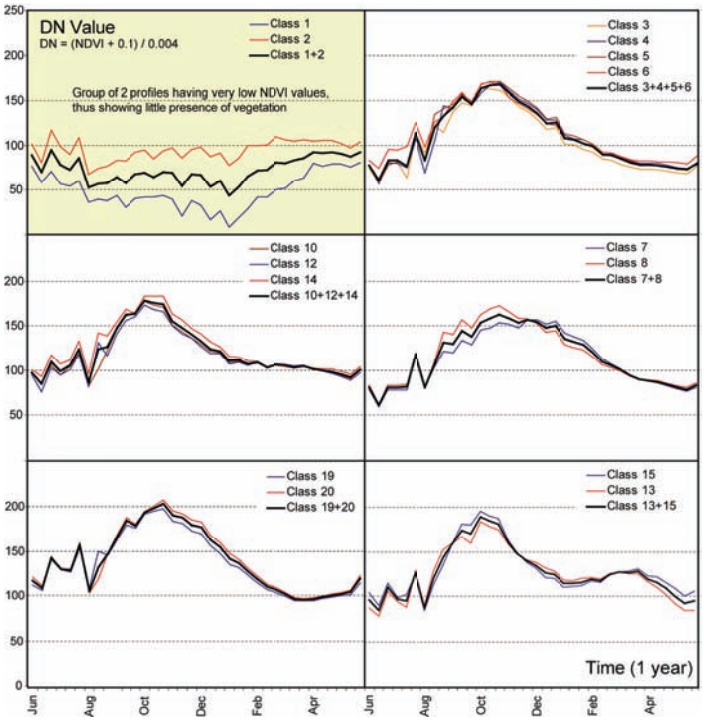


Figure 2- 5: Supervised grouping of the annual averages of the generated unsupervised NDVI classes

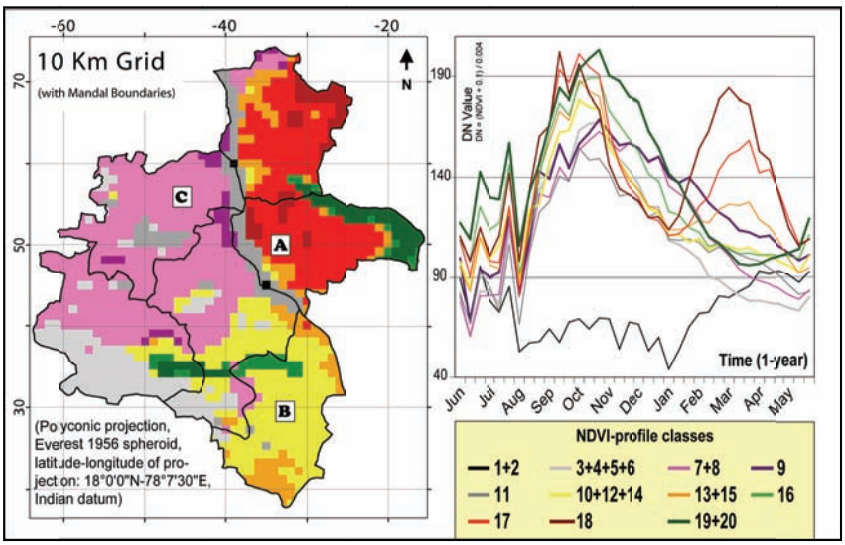


Figure 2- 6: 1 km^2 NDVI unit map with 11 NDVI classes

2.5.2 Linking hypertemporal RS images with existing land cover map

In the second stage of the methods, the NDVI unit map was compared with an existing map of land-cover. For this purpose, the area of each class of the existing land-cover map (Figure 2- 3**)** was compared with the area of each NDVI class. The results of this step are presented in table 2- 3. The difference in scales between the two maps was circumvented by using fractions only (in percentages). Using the area statistics for the six *Mandals*, an analysis of variance (ANOVA) was used to test the association between the NDVI classes (11x) and the seven underlying land-cover classes. This ANOVA revealed a significant relation (at the 1% confidence level) for both factors and for their interaction (385 degrees of freedom and 7 outliers).

The original land-cover map reported 'cropped' fields based on image interpretation. 'Cropped' is a land use term and in table 2- 1 it is replaced by the term 'green' (as observed in the images).

2.5.3 Linking hypertemporal RS images with crop statistical data by sub-district

Having linked the hypertemporal images with the existing land-cover map, the third stage of the work involved relating these images to crop statistical data. Stepwise multiple linear regression analysis was performed for each crop to estimate the function described in equation 1 using the data reported in table 2- 2. The limited number of *Mandals* studied (6 statistical records versus 11 NDVI classes as parameters) provided only a limited number of statistical degrees of freedom. This called for further grouping of the NDVI classes as indicated in table 2- 4 (using the area coding presented in figure 2- 3**)**.

The process of grouping consisted of merging classes [7, 8] and 9 (Area A) on the basis of their relatively high fractions of green fields that are cropped only for one season per year and merging of classes [13, 15], 17 and 18 (Area C) on the basis of the dominant occurrence of green fields that are cropped during both seasons (*Kharif* and *Rabi*). This process resulted in a matrix of six statistical records versus eight NDVI classes as parameters. Table 2- 4 reports the average proportion (in percent) that a specific crop is grown within each map unit represented by a specific group of NDVI classes. NDVI classes [1, 2], 11, [19, 20] and 16 seem to be related with non-crop areas.

Table 2- 3: The fraction of each land-cover class by NDVI-unit (A) and vice-versa (B), plus totals (%) and areas (km^2)

A	**Green Fields**			**Other**				**Total**	**Area**
NDVI-Unit	**Kharif only**	**Rabi only**	**Kharif +Rabi**	**Trees**	**Bare**	**Water**	**Built -up**	**(%)**	**(km^2)**
1,2	0.6	3.8	0.0	0.0	0.0	95.6	0.0	**100**	2
11	17.6	16.1	11.9	0.0	18.2	36.1	0.1	**100**	81
3,4,5,6	32.9	8.7	1.0	4.9	51.3	1.2	0.0	**100**	136
19,20	4.0	1.1	2.0	89.2	2.8	0.9	0.0	**100**	50
16	16.6	5.8	1.3	66.2	7.4	2.6	0.1	**100**	34
10,12,14	39.2	8.0	5.3	22.5	19.9	4.5	0.6	**100**	234
7,8	39.2	43.6	1.9	1.2	12.2	1.4	0.5	**100**	407
9	44.8	36.7	5.5	0.2	3.4	9.0	0.4	**100**	29
13,15	31.8	22.7	23.0	6.4	9.1	5.8	1.2	**100**	91
17	2.6	15.0	63.6	1.4	10.7	4.6	2.1	**100**	212
18	0.8	16.5	72.8	0.0	1.0	6.8	2.1	**100**	40

B	**Green Fields**			**Other**				
NDVI-Unit	**Kharif only**	**Rabi only**	**Kharif +Rabi**	**Trees**	**Bare**	**Water**	**Built -up**	
1,2	0.0	0.0	0.0	0.0	0.0	2.7	0.0	
11	3.9	4.5	4.4	0.0	6.8	41.1	0.6	
3,4,5,6	12.3	4.0	0.6	4.7	32.2	2.3	0.0	
19,20	0.5	0.2	0.5	31.8	0.6	0.6	0.0	
16	1.6	0.7	0.2	16.2	1.2	1.3	0.5	
10,12,14	25.1	6.4	5.6	37.5	21.4	14.8	14.0	
7,8	43.5	60.4	3.6	3.6	22.8	8.2	21.0	
9	3.6	3.7	0.7	0.0	0.5	3.8	1.2	
13,15	7.9	7.0	9.6	4.1	3.8	7.4	11.1	
17	1.5	10.8	61.5	2.1	10.5	13.9	43.3	
18	0.1	2.3	13.3	0.0	0.2	3.9	8.3	
Total (%)	**100**	**100**	**100**	**100**	**100**	**100**	**100**	
Area (km^2)	366	294	219	141	217	71	10	**1,318**

In figure 2- 7, two NDVI map units comprising a mix of rice and sugarcane land-cover are shown. These match closely with the results presented in section 5.3 (tables 2- 4, 2- 5 and 2- 6 and Figure 2- 6).

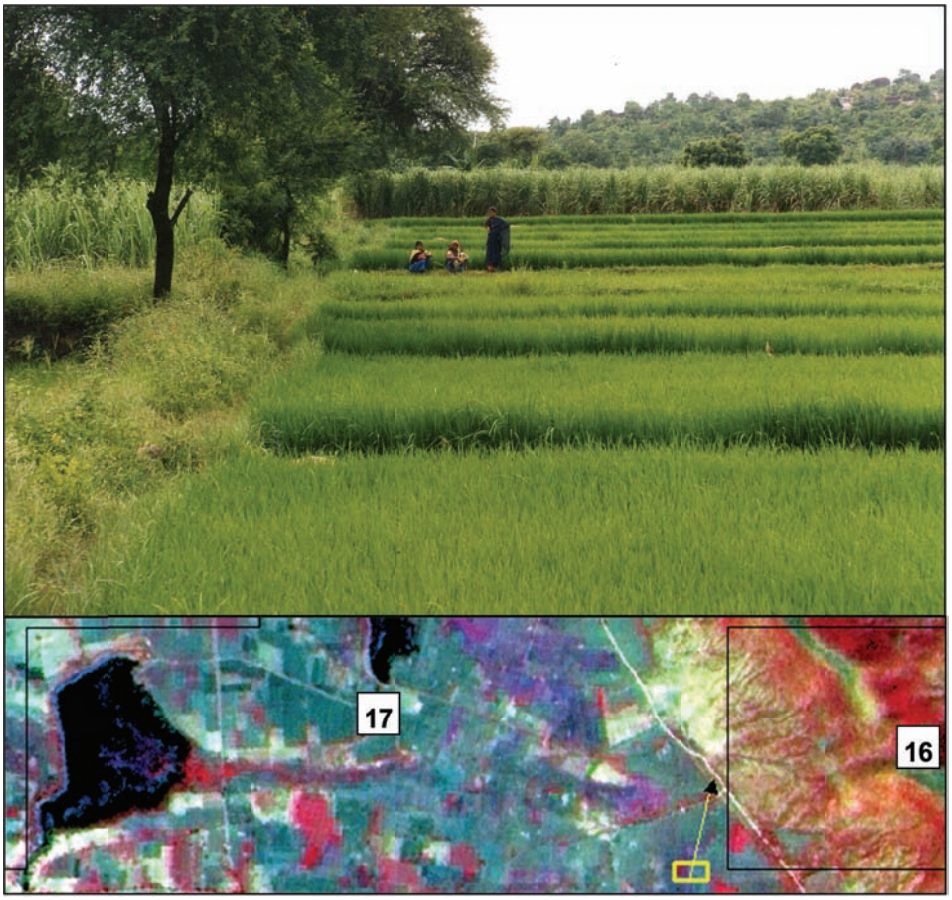


Figure 2- 7: Details about a mapping unit

The extent to which a crop is grown has been fully quantified by the regression results, e.g. during *Kharif*, irrigated rice was grown on 49.6% of the area of map units represented by the group of NDVI classes {[13, 15], 17 and 18} while during *Rabi* this figure was 28.4%. Sorghum was grown on different map units during different seasons: during *Kharif* on NDVI unit [3, 4, 5, 6] (25%), and during *Rabi* on the map unit of the NDVI class group {[7, 8] and 9} (32%).

Table 2- 4: Results of the multiple stepwise linear regression analysis

		NDVI-profile Class Groups				
			Area B:	Area C:	Area A:	
KHARIF	**Adj.R^2**	**3,4,5,6**	**10,12,14**	**'7,8',9**	**'13,15',17,18**	**Area (ha)**
Irrig.Rice	88%				49.6%	22,774
Sorghum	95%	25.2%				3,963
Maize	78%		2.2%			482
Pulses	95%		34.0%	46.6%		29,122
Sugarcane	89%		3.9%		4.6%	2,395
Cotton	82%			14.8%		6,860
Other crops	89%	5.9%				1,103
RABI	**Adj.R^2**	**3,4,5,6**	**10,12,14**	**'7,8',9**	**'13,15',17,18**	**Area (ha)**
Irrig.Rice	95%				28.4%	11,481
Sorghum	89%			32.2%		15,454
Pulses	87%			5.6%		2,823
Sugarcane	85%		3.6%		3.6%	1,960
Groundnut	88%		12.6%		6.7%	5,942
Other crops	92%			11.7%		5,847
	Area (km^2)	136	234	436	343	

***Note**: the Adjusted R^2, when regression through the origin is forced, cannot be compared to R^2s for models that include an intercept. Reported cropped areas are estimated by the extent of NDVI-profile classes. Coefficients are reported as percentages that are confined to the 0-100% range (0 to 1); each was significance at 5%.*

2.5.4 Legend construction and matching with known crop calendars

In the final stage, the data presented in tables 2- 2, 2- 3 and 2- 4 were compiled into two comprehensive legends (Table 2- 5and table 2- 6) organized by the NDVI classes. Each legend item is linked to different land-cover and land use complexes. NDVI class areas are taken from table 2- 2, land-cover data from table 2- 3 and agricultural land use data from table 2- 4. In those cases where the 11 NDVI classes were grouped (areas A, B, and C) to derive cropped area

estimates by crop and season, the regression results are reported by group. Tables 2- 5 and 2- 6 also report the cropping intensity (%) of the area under cultivation during *Rabi* for areas A and C based on land-cover statistics and the relative shape of the NDVI class profiles. Crop calendar information was included in table 2- 6 based on the aforementioned information regarding where a crop is grown. Following facts are present in tables 2- 5 and 2- 6.

- **Area A**: the bi-model shape (reflecting cropping in both *Kharif* and *Rabi* seasons) of the three underlying NDVI classes indicates sequential cropping of irrigated rice at three cropping intensities during the *Rabi* season. This cropping intensity relates to the farmers' decision whether or not to grow rice depending on the availability of irrigation water during the *Rabi* season.

- **Area B**: the NDVI class profiles of this area show high values during *Kharif* and low to very low values during the remaining part of the year. In this area, mainly pulses (grams) such as black, green, or red gram having a short growth cycle were cultivated (34% during *Kharif*). During *Rabi*, groundnuts were grown on about 13% of the area which was moderately reflected by the NDVI class profile [10, 12, 14].

- **Area C**: this area represents two NDVI classes that retain relatively high values in the *Rabi* season. This matched with either i) the growing of crops with a relatively long growing period such as cotton (15% of the area) or ii) late planting of sorghum (32%) and Bengal gram (chickpea; 6%). During *Rabi*, the NDVI class profile 9 remained significantly higher than the profile of [7, 8] due to relatively high cropping intensity in NDVI class 9. This indicates that a larger area was allocated to 'late planted' crops like sorghum, chickpea, *Rabi* sunflower, or safflower.

Table 2- 5: Land cover legend of the NDVI Unit Map

Map Unit				related land info	Land Cover Complexes					Zones
NDVI Unit			Area (ha)		Fields-Green		Other (%) *			
					Period	%	Trees	Bare	Water	
1,2			198	Soils with 70-90% sand	All year	4	0	0	96	Other Areas (little agriculture)
11			8,114		Kharif	18	0	18	36	
					Rabi	16				
					Kh+Rabi	12				
3,4,5,6			13,620		Kharif	33	5	51	1	
					Rabi	9				
					Kh+Rabi	1				
19,20			5,030	Undulating terrain	All year	7	89	3	1	
16			3,443		Kharif	17	66	7	3	
					Rabi	6				
					Kh+Rabi	1				
10,12, 14			23,414	sandy loams/ loamy sands	Kharif	39	22	20	4	Area B
					Rabi	8				
					Kh+Rabi	5				
7,8	Cropping Intensities during Rabi	Medium	40,741	Vertisols; topsoil with 35-50% clay	Kharif	39	1	12	1	Area C
					Rabi	44				
					Kh+Rabi	2				
9		High	2,946		Kharif	45	0	3	9	
					Rabi	37				
					Kh+Rabi	6				
13,15		Low	9,103	Areas with irrigation canals and tubewells; heavy textured clay-loam soils	Kharif	32	6	9	6	Area A
					Rabi	23				
					Kh+Rabi	23				
17		Medium	21,191		Kharif	3	1	11	5	
					Rabi	15				
					Kh+Rabi	64				
18		High	4,016		Kharif	1	0	1	7	
					Rabi	17				
					Kh+Rabi	73				

Source: Soil and Land Cover maps (NRSA, 1995).
* Built-up area is not reported here.

Table 2- 6: Crop calendars of agricultural zones of the NDVI Unit Map

Zones	Agricultural Land Use Complexes		
	Area Cropped		Crop calendars
	Crop	%-of Area	Ju Ju Au Se Oc No De Ja Fe Ma Ap Ma
Other	Sorghum	25	
	Other (Kharif)	6	Grams, Sunflower, Gr.Nuts, Sesame
Area B	Pulses	34	Bl.Gr.Red Gram
	Maize	2	
	Groundnut	13	Irrigated:
	Sugarcane	4	
Area C	Pulses	47	Bl.Gr.Red Gram
	Cotton	15	
	Sorghum	32	
	Pulses	6	Chickpea
	Other (Rabi)	12	Safflower, Sunflower
Area A	Irrig.Rice	21	
	Groundnut	7	
	2xIrrig.Rice	28	
	Sugarcane	5	

Crop calendar details of "Other" crops:	Kharif	Rabi
Kharif sunflower:		
Rabi sunflower:		
Black gram:		
Green gram:		
Red gram:		
Safflower:		
Kharif groundnut:		
Sesame:		

Source crop statistics: interpreted 1998/99-2001/02 data (CPO, 2001).
Source crop calendars: Apau (1989) and Rao (1995).

- **Other areas**: only the area with NDVI classes [3, 4, 5, 6] was under a significant cultivation of sorghum during *Kharif* (25% of its area). The NDVI class profile [3, 4, 5, 6] showed a noticeable, but relatively low, increase during the October-December period.

2.6 Discussion and conclusions

The results of series of unsupervised classifications of the SPOT NDVI images helped to draw the conclusion that the study area can optimally be stratified into i) between 22 and 28 classes on the basis of average divergence and ii) between 18 and 23 classes on the basis of minimum divergence. Therefore, 'visual supervised grouping' of the annual averaged classes produced by the unsupervised ISODATA algorithm was performed to repair possible deficits in classification that were due to the absence of a very clear single peak in the divergence values. This method shows promise for application where similar kinds of uncertainties can arise because of, for example, missing satellite data as a result of cloud cover. Supervised grouping after unsupervised classification helps to eliminate the annual variations within the original classes without any loss of important detail. The results shown in table 2- 4 are comparable even with those where a clear peak in the average divergence has been reported for a larger study area (Khan *et al*. 2010). The method developed in this study resulted in a map that represents the 1998-2002 periods without considering land use changes that occurred during that period.

In 2002, the first author carried out other research work in the Nizamabad area for several months. In the course of this work, area specific expert knowledge was obtained that closely matched with the results of this study (Figure 2- 7, which was taken during that period). The results of this research show promise for a variety of uses in agriculture. Our method enables the monitoring of the crop performances and/or cropping intensities over the years by comparing annual NDVI class profiles for selected map units. Furthermore, our approach provides a method to monitor changes in the extent of specific map units over time and thus to detect land use conversions. When required, more specific and detailed surveys on actual system management, user inputs and obtained outputs can be performed. This method of producing land use maps has the potential to support a broad range of applications in agriculture such as early warning about food security, yield gap analysis and regional to global assessment of agricultural productivity.

In our research, we found that annual changes in the NDVI class profiles most frequently reflect changes in cropped areas (for the 1-km^2 pixels). These changes could not be related directly to the monsoon rainfall amounts in the

region. Nevertheless, it would be incorrect to conclude that declining NDVI values over the years show crops being suffered from additional stress and that yields per hectare declined. On the other hand, it would be correct to conclude that the overall biomass by area (km^2) declined causing lower production of the studied crops in the region.

In other studies, NDVI classes have been used in combination with the rainfall data and crop characteristics to monitor drought (Ji and Peters, 2003; Bhuiyan *et al*., 2006; Ghulam *et al*., 2007). The IWMI's prototype Drought Monitoring System (DMS) provides monitoring reports either by administrative areas (districts and provinces) or by pixels (Thenkabail *et al*. 2004). In its present version, however, the DMS does not distinguish between land use classes. It has been stressed that more reliable thresholds of numerical values of remote sensing indices should be developed to quantify drought severity (Smakhtin *et al*. 2005). Keeping this requirement in mind, the results of our study were used as an example of monitoring. An example related to the monitoring of land use modifications over time is provided in figure 2- 8 which shows profiles of NDVI classes [13, 15], 17, and 18 present in the irrigated areas where rice is the dominant crop during the *Rabi* season. The figure 2- 8 shows a decline in NDVI (DN values) starting in 2001 for the NDVI group [13, 15] and expanding one year later (2002) in all three NDVI classes. These declines are either a result of a poor performance of the rice crop or by a substantial number of fields not being cultivated during *Rabi*. From interviews with the farmers concerned (undertaken in September 2002), it became clear that the latter was the case. The interviewed farmers vehemently complained about the unreliable availability of electricity to run their water pumps. This was due to the practice of load shedding (power supply cuts) which started in 2001 as a result of low water levels in reservoirs across the state. Consequently, many farmers decided to decrease the paddy cultivation so that their pumps could pump up sufficient water for the cropped area. The method illustrated in this chapter can be incorporated into RS/GIS based drought monitoring systems like Famine Early Warning System Network (FEWSNET), US Drought Monitor (www.drought.unl.edu) and Southern African Development Community's (SADC) Regional Remote Sensing Unit Drought Monitoring Center.

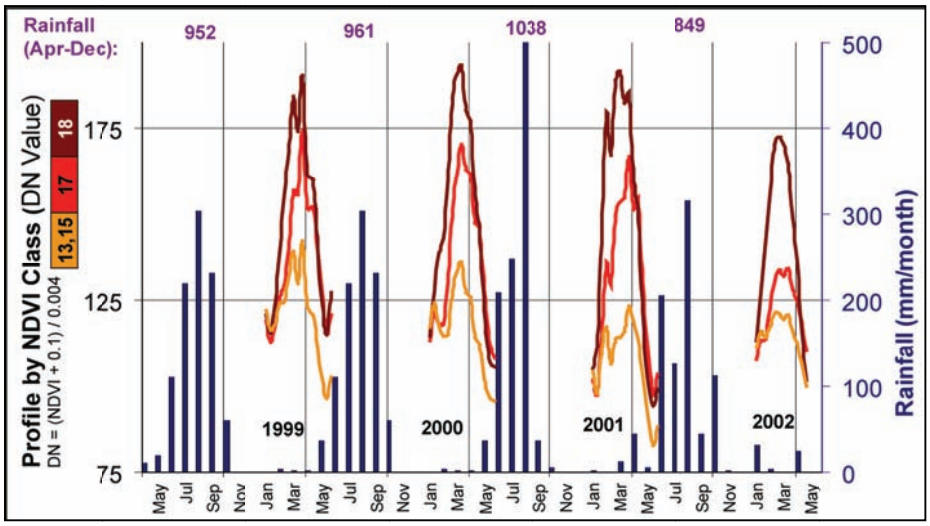


Figure 2- 8: Monitoring land use modifications

Once the NDVI map units have been delineated and linked with land use of a district or a larger geographical region, continuous monitoring of the land use systems over time can be achieved. Stratified monitoring - based on NDVI units - allows preparing more specific early warning bulletins. The anomalies in the NDVI classes in time and duration of their state below pre-defined thresholds can be monitored continuously. This can effectively lead to a continuous monitoring and prediction of various crop conditions at different scales (from pixel to region levels).

Acknowledgements

We thank Dr. R.R. Navalgund, former Director, National Remote Sensing Agency, Hyderabad, Dr. Nidumolu, Spatial and Simulation Scientist (Climate Risk), CISRO Sustainable Ecosystems and the Chief Planning Officer of Nizamabad District and his staff for facilitating and supporting this research. We also acknowledge the cooperation of the farmers in Nizamabad District.

3

Disaggregating and Mapping Crop Statistics using Hypertemporal Remote Sensing

* This chapter is based on Khan, M.R., de Bie, C.A.J.M., van Keulen, H., Smaling, E.M.A. and Real, R., 2010. Disaggregating and mapping crop statistics using hypertemporal remote sensing. International Journal of Applied Earth Observation and Geoinformation 12, 36-46.

Abstract

Governments compile their agricultural statistics in tabular form by administrative area, which gives no clue to the exact locations where specific crops are actually grown. Such data are poorly suited for early warning and assessment of crop production. 10-daily satellite image time series of Andalucia, Spain, acquired since 1998 by the SPOT Vegetation Instrument in combination with reported crop area statistics were used to produce the required crop maps. Firstly, the 10-daily (1998-2006) 1-km resolution SPOT-Vegetation NDVI-images were used to stratify the study area in 45 map units through an iterative unsupervised classification process. Each unit represents an NDVI profile showing changes in vegetation greenness over time which is assumed to relate to the types of land cover and land use present. Secondly, the areas of NDVI units and the reported cropped areas by municipality were used to disaggregate the crop statistics. Adjusted R-squares were 98.8% for rainfed wheat, 97.5% for rainfed sunflower, and 76.5% for barley. Relating statistical data on areas cropped by municipality with the NDVI-based unit map showed that the selected crops were significantly related to specific NDVI-based map units. Other NDVI-profiles did not relate to the studied crops and represented other types of land use or land cover. The results were validated by using primary field data. These data were collected by the Spanish government from 2001-2005 through grid sampling within agricultural areas; each grid (block) contains three 700x700m segments. The validation showed 68, 31 and 23 percent variability explained (Adjusted R-squares) between the three produced maps and the thousands of segment data. Mainly variability within the delineated NDVI-units caused relatively low values; the units are internally heterogeneous. Variability between units is properly captured. The maps must accordingly be considered “small scale maps”. These maps can be used to monitor crop performance of specific cropped areas because of using hyper-temporal images. Early warning thus becomes more location and crop specific because of using hyper-temporal remote sensing.

3.1 Introduction

Agricultural scientists, resource managers and policy makers require updated information on agricultural land use to address a broad range of issues. Increased production of food is required to meet the needs of global population estimated at over 10 billion by the year 2050 (Engelman *et al.*, 2000; United Nations, 2005). Land resources are finite and overexploitation is leading to land degradation, declining crop yields and the risk of food shortages. Adequate knowledge on crop producing areas allows decision makers to locate populations that are most vulnerable to food insecurity and poverty. Crop monitoring is equally important for the developed countries to effectively and sustainably use their resources (Dymond *et al*., 2001; George and Nachtergaele, 2002). The first component of improved monitoring of agricultural production is exact knowledge of where "what" is being grown (IAWG, 2000; Dixon *et al.*, 2001; FAO, 2003).

Currently, annual estimates of land use per administrative unit are being compiled by relevant (government) agencies, and are generally available in tabular form. Such data lack information on the spatial distribution of specific land uses (Jansen and Di Gregorio, 2003). The agricultural statistics data contain important information and can be transformed into desired spatially explicit land use maps. Agricultural statistics have been used to produce agricultural land use density maps. Statistical information for agricultural land use at level 2 of so called Nomenclature of Territorial Units for Statistics (NUTS 2) regions, which correspond to administrative areas of 160 km^2 to 440 km^2 was used in the economic model CAPRI (Common Agricultural Policy Regional Impact assessment) to distribute different crops to the individual HSMUs "Homogeneous Spatial Mapping Unit" (HSMU) i.e. soil, slope, land cover and administrative boundaries (Leip *et al*., 2008; Kempen *et al*., 2005).

Detailed land use maps are not readily available for many countries (Fresco *et al.*, 1994; Wood *et al.*, 2000). Conventional methods of land use mapping are labor–intensive and time-consuming and as a consequence expensive. Land use maps are therefore infrequently prepared with often insufficient detail. In Europe for instance, land cover/use maps are prepared at 10 year intervals that contain only very generalized classes of agricultural activities (Feranec *et al.*,

2007). Moreover, such maps soon become out-of-date, particularly in rapidly changing environments.

In recent years, satellite remote sensing techniques have been shown to be effective in preparing accurate land use/land cover maps and for monitoring changes at regular intervals through multi-temporal remote sensing data (e.g. Loveland *et al.*, 2000; Souza *et al.*, 2003; Brand and Malthus, 2004; Budde *et al.*, 2004; Wessels *et al.*, 2004). The dynamic nature of agriculture, like seasonality and its occurrence almost everywhere are the strongest incentives for scientists to monitor agriculture from space (De Bie, 2000; Oetter *et al.*, 2001; Mayaux *et al.*, 2004).

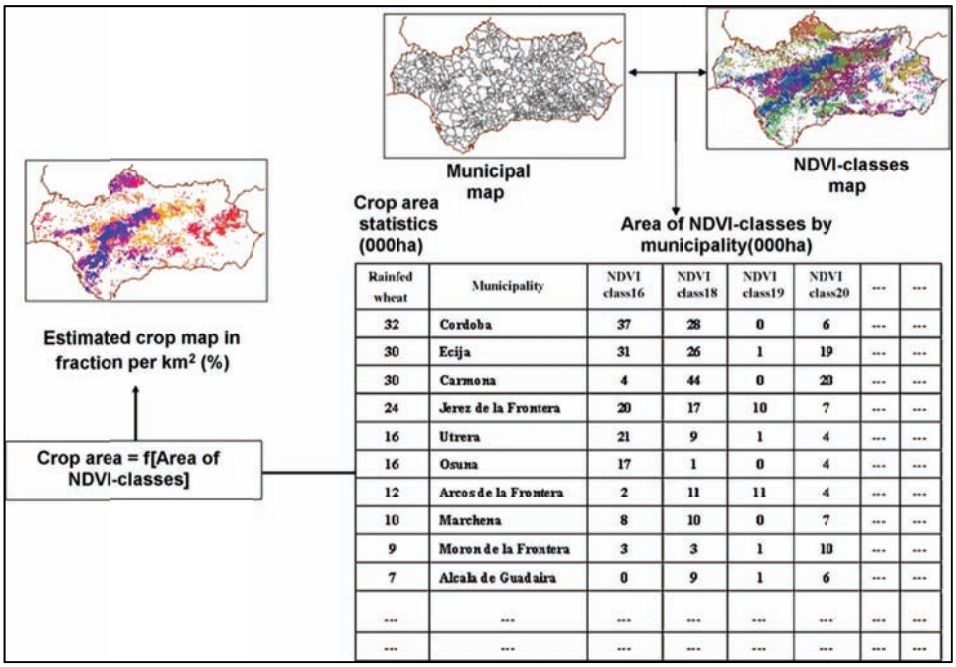
The relation between satellite information and crop characteristics is described in terms of vegetation indices which provide information on conditions of vegetation and it allows inference regarding land use/land cover. Vegetation indices have been extensively used for monitoring and detecting vegetation and land cover changes (deFries *et al.*, 1995; Liu and Kafatos, 2005). Vegetation indices are based on differential absorption, transmittance, and reflectance of energy by the vegetation in the red and near-infrared regions of the electromagnetic spectrum (Jensen, 1996). One type of spectral vegetation indices is the Normalized Difference Vegetation Index (NDVI), the ratio of near infrared (NIR) and red (R) reflectance.

NDVI is usually assumed to be broadly indicative of crop photosynthetic activity (Sarkar and Kafatos, 2004) and therefore associated with greenness and thus above-ground dry matter production (Goward and Huemmrich, 1992). The vegetation instrument onboard the SPOT satellite with four spectral bands, i.e., blue (0.43-0.47 mm), red (0.61-0.68 mm), infrared (0.78-0.89 mm) and short wave infrared (1.58-1.75 mm), at a spatial resolution of 1 km and a temporal resolution of one day, meets the requirements for vegetation mapping at continental scale. Several studies have discussed the suitability of temporal NDVI-profiles for studying vegetation phenological development, especially that of crops (Hill and Donald, 2003). The use of multitemporal images not only results in higher and consistent accuracy in mapping different classes, they are especially advantageous in areas where vegetation or land use changes rapidly. Because of their correlation with green plant biomass and vegetation cover, long temporal sequence of regularly acquired data (Hypertemporal image data),

such as NDVI time series, have been used for monitoring anomalies, drought, vegetation phenology, land cover characteristics, estimation crop yields (Agrawal *et al.*, 2001; Murthy *et al.*, 2007; Gu *et al.*, 2008) and area estimations for larger fields (Fritz *et al.*, 2008; Verbeiren *et al.*, 2008).

Crops exhibit distinctive behaviors that are captured by temporal patterns of NDVI which have strong periodic characteristics in a year so cropland can be distinguished from other vegetation types through analysis of their respective phenologies captured by the NDVI-profiles (Guo *et al.*, 2008).

The use of NDVI time series has been focused on crop production monitoring and yield forecasting rather than on mapping crops (Maselli *et al.*, 1992). The MARS project (Monitoring of Agriculture with Remote Sensing) of the Joint Research Center (JRC) has integrated use of SPOT4 images in their integrated agricultural monitoring and yield forecasting methods for Africa (Nègre *et al.*, 2004). Some studies used NDVI time series for mapping weather parameters (Creech and McNab, 2002; Champeaux *et al.*, 2004; Stöckli and Vidale, 2004). Hypertemporal images are used for mapping major land cover types and to differentiate forest, pastures and shrubs (Craig, 2001; de Bie *et al.*, 2008; Wardlow and Egbert, 2008). Their utility has yet to be explored for mapping various crops over large areas, i.e. at regional levels. The integration of satellite earth observations and ground survey data is a useful method to estimate crop acreage in small areas (Battese *et al*., 1988; Flores and Martinez, 2000). In this chapter, a properly tested improved methodology is described for preparing crop maps to be used for monitoring purposes. The method is based on ten-day temporal resolution SPOT Vegetation data to disaggregate tabular statistical data on cropped areas per administrative unit. The aim is to contribute to the development of methods for combining spatial and temporal land-use data sets using existing data sources and improved RS/GIS-based methods. In short: the extent of NDVI-classes by administrative areas is correlated with the reported area of crops in these areas and the resulting model is used to prepare individual crop maps (Figure 3- 1).



Rainfed wheat	Municipality	NDVI class16	NDVI class18	NDVI class19	NDVI class20	---	---
32	Cordoba	37	28	0	6	---	---
30	Ecija	31	26	1	19	---	---
30	Carmona	4	44	0	23	---	---
24	Jerez de la Frontera	20	17	10	7	---	---
16	Utrera	21	9	1	4	---	---
16	Osuna	17	1	0	4	---	---
12	Arcos de la Frontera	2	11	11	4	---	---
10	Marchena	8	10	0	7	---	---
9	Moron de la Frontera	3	3	1	10	---	---
7	Alcala de Guadaira	0	9	1	6	---	---
---	---	---	---	---	---	---	---
---	---	---	---	---	---	---	---

Figure 3- 1: Relating NDVI-class area with crop statistic to prepare crop maps

3.2 Materials and methods

3.2.1 Study area

Andalucia is an autonomous community in southern Spain with an area of 87,268 km^2. It comprises eight provinces and 770 municipalities and the capital is Seville in the province of Seville (Figure 3- 2). The Mediterranean climate of Andalucia is characterized by mild, rainy winters and hot, dry summers. Almeria and Seville municipalities of respective provinces have the highest average temperatures in Spain with 18.6ºC and 18.7ºC, respectively. Annual rainfall is highly heterogeneous, with a marked decreasing gradient of precipitation from west to east, and ranges from a maximum of 2000 mm to minimum of 170 mm (Font, 2000).

Agriculture plays an important role in Andalucia, with about 67% of the region utilized for agricultural purposes. Medium-sized mountains predominate in the Andalusian landscape, occupying 42% of the total surface. Consequently, 38%

of the agricultural land is mountainous with crops generally restricted to the inner valleys or to gently sloping hillsides. The major crops are grains (370, 000 ha of wheat, 29,000 of rice, 28,000 of maize), sunflower (94,000 ha), oranges (7,600 ha), olives (350,000 ha), vineyards (8,500 ha) and cotton (46,000 ha). The crops are grown on Cambisols, Regosols, Luvisols and Vertisol (FAO, 1998). Sowing (S) and harvesting (H) dates of the major crops are given in figure 3- 3.

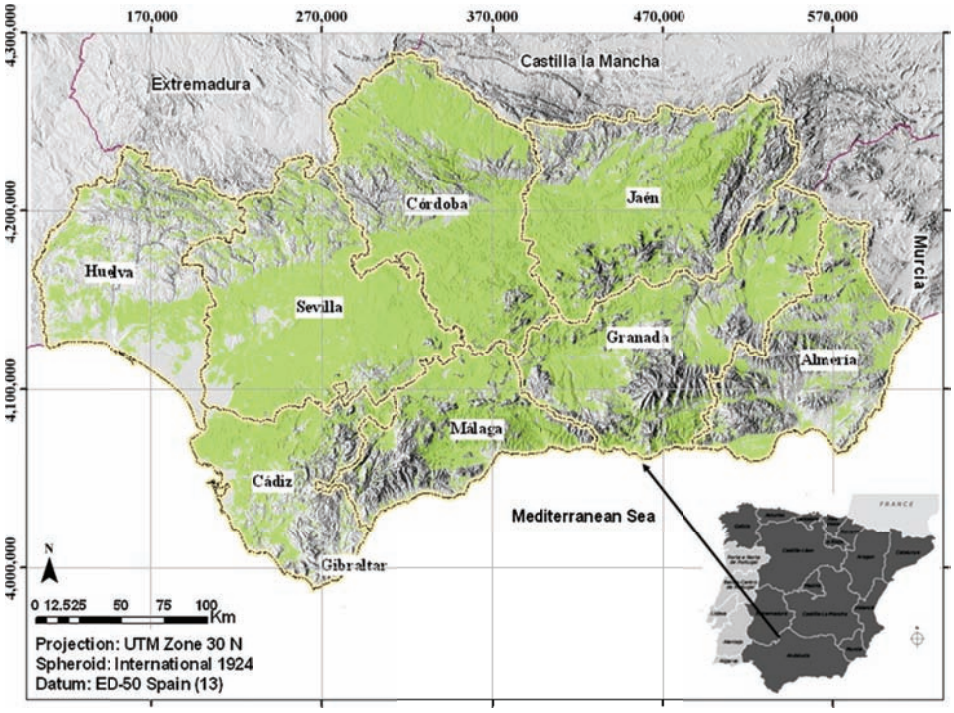


Figure 3- 2: Agricultural areas of Andalucia based on the CORINE land cover map 2000

3.2.2 Data used

Data used for the study included:

Atmospherically corrected SPOT4-5 Vegetation (VGT) Sensor's 10-day composite NDVI-images (S10 product) at 1 km^2 resolution from 1998 to 2006 obtained from www.VGT.vito.be (300 images). The first VGT sensor was launched in March 1998 on board of the SPOT4 satellite, to monitor surface

parameters with a frequency of about once a day at global basis with a spatial resolution of 1 km^2. Its specifications are well suited for terrestrial applications like land cover mapping (De Wit and Boogaard, 2001). NDVI-values used in this study are DN-Values in 0-255 format based on the formula:

$$DN = \frac{NDVI + 0.1}{0.004} \qquad \text{(Eq. 1)}$$

Crop area statistics by municipality matching the period of the NDVI-images were obtained from the Consejeria de Agricultura y Pesca (Ministry of Agriculture and Fisheries), Andalucia.

The CORINE (Coordination of Information on the Environment of the European Environmental Agency (EEA)) land cover 2000 (CLC) map (Bossard *et al.*, 2000) was used for area masks, comprising the locations where crops are certainly grown in order to get a map of major agricultural areas (Figure 3- 2). The agricultural areas comprising of arable land, permanent crops, pastures and heterogeneous agricultural areas were retained and the rest were masked out. The scale of the CLC map is 1:100,000. The accuracy of the CLC map has been reported to be greater than or equal to 85% (Martín de Santa Olalla Mañas *et al.*, 2003).

The administrative map of Andalucia, obtained from Junta de Andalucia, including the boundaries of provinces and municipalities, constructed using European 1950 Datum for Spain Zone 30 N and projected in Universal Transverse Mercator (UTM).

Plot-specific crop data, collected from 2001-2005 by the Ministry of Agriculture and Fisheries. The survey comprised 1451 randomly selected segments of 700 x 700 m. The areas with higher density of crops are sampled more intensively. i.e., additional segments are also sampled in the areas where crops dominate the land use (Figure 3- 3). The data were collected by visits to all the fields per segment. The spatial distribution of the surveyed segments was determined by dividing the territory in blocks of 10 X 10 km. Each block was sub-divided in 100 cells of 1 km^2. In each block, three cells were randomly selected for surveying. In each cell, cropped area, production of each crop per unit area, irrigation scheme, etc. of agricultural areas present in the segment of

700 x 700 m is recorded (Figure 3- 4). Out of these 1451 segments, 1424 segments coincided with agricultural areas as defined by the CORINE land cover map. These were used for validation of the produced maps.

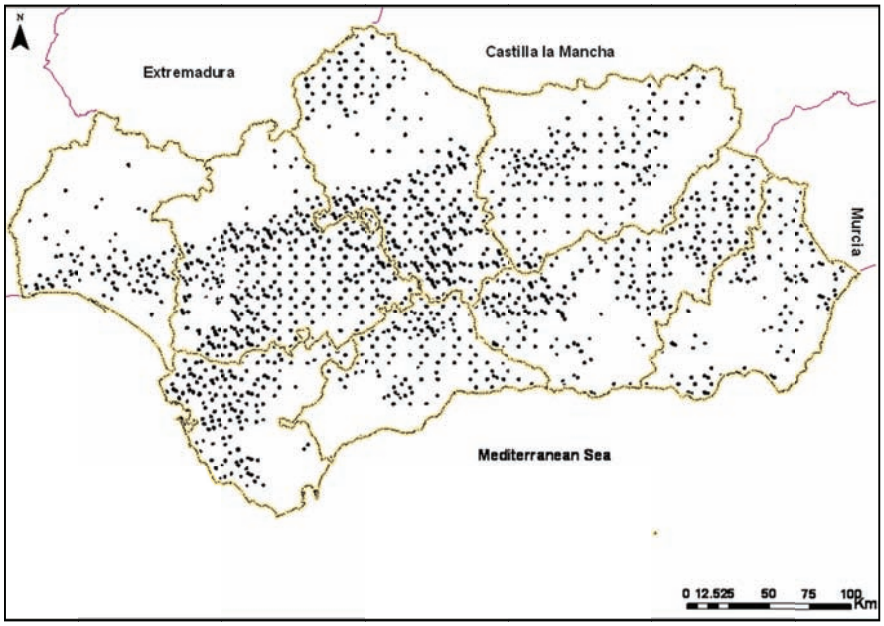


Figure 3- 3: Location of validation data set consisting of 700x700 m segments surveyed from 2001- 05

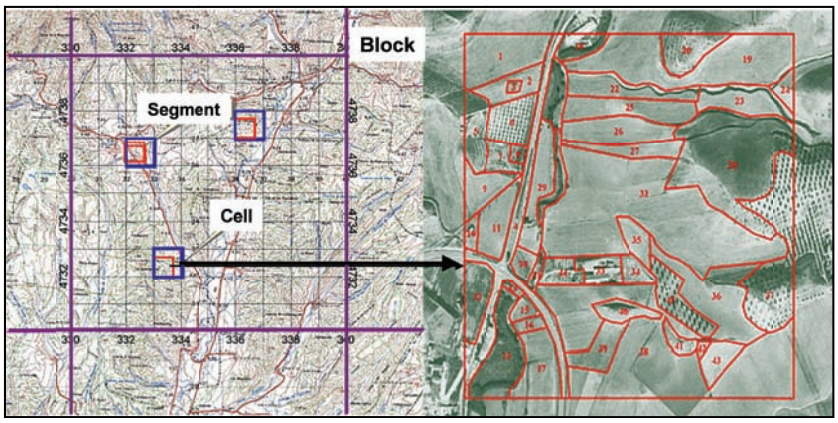


Figure 3- 4: Diagram showing area frame (segment) on the topographic map

Crop-calendar data for the study area were available on the website of the Ministry of Agriculture and Fisheries

(http://www.mapa.es/es/agricultura/agricultura.htm). These crop calendar data were verified through farmer's interviews during the field work (Figure 3- 5).

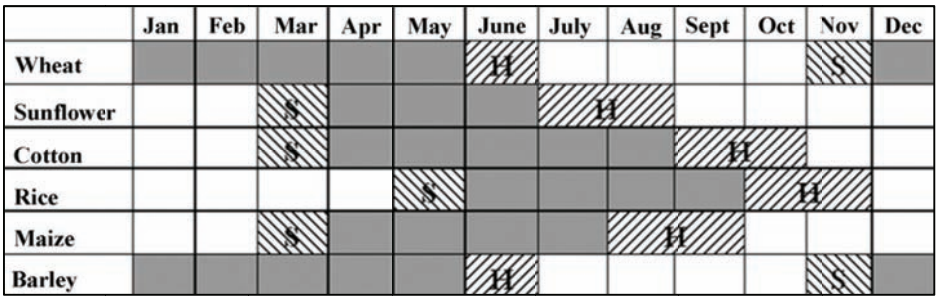


Figure 3- 5: General crop calendar for field crops in Andalucía

3.2.3 Image classification technique

The iterative self-organized unsupervised clustering algorithm (ISODATA) of the ERDAS Imagine software was used to derive spectral classes from 300 NDVI image data layers. The ISODATA procedure is iterative in that it repeatedly performs an entire classification (outputting a thematic raster layer) and recalculates statistics. Self-organizing refers to the way in which it locates clusters with minimum user input. The ISODATA clustering method uses spectral distance, as in the sequential method, it iteratively classifies the pixels, redefines the criteria for each class, and classifies again, so that the spectral distance patterns in the data gradually emerge (ERDAS, 1997). It starts from arbitrary cluster means. In each successive clustering, the means of clusters are shifted. A cluster is a group of pixels (classes) with similar spectral characteristics. The ISODATA utility repeats the clustering of pixels in the image, until either a maximum number of set iterations has been performed (50), or a maximum coverage threshold is reached (set to 1.0). Performing an unsupervised classification is simpler than a supervised classification, because the cluster signatures are automatically generated by the ISODATA algorithm. The user must predetermine the number of iterations and number of resulting clusters (classes). Separate ISODATA runs were carried out to define 10 to 100 classes. In each run the desired number of classes is produced by the ISODATA clustering Algorithm. The divergence statistical measure of distance (ERDAS,

2003; Swain and Davis 1978) between defined cluster signatures by run was used to compare the various runs. The optimal run with a clear distinguished peak in the divergence separability was selected for further study.

The NDVI-classes of Andalucia from the unsupervised classified raster map of NDVI time series after converting to polygons were masked by using the agricultural areas as defined by the CLC 2000 map. The masked map provided the NDVI-classes present in agricultural areas of Andalucia. Later on, the area of each NDVI-class for the agricultural areas in every municipality was calculated through GIS analysis by combining the classified NDVI map and the administrative map of Andalucia. The agricultural area covered by each NDVI-class in every municipality as explanatory variable and the crop statistics by municipality as dependent (a total number of 771 municipalities) were used to estimate an additive multiple linear function (Method A) that relates for a particular crop the reported cropped areas to the areas of the NDVI-classes. This was estimated through step-wise forward multiple regressions with no constant and coefficients constrained between 0.0 and 1.0 because the cropped area in a municipality can neither be in negative nor more that 100 percent of the municipality area. The model is thus:

$$Y = \sum_{i=0}^{n} b_i x_i + \varepsilon_i \qquad \text{(Eq. 2)}$$

Where

Y	= Cropped area per municipality (ha) from 2001-05
b_i	= Regression coefficient
x_i	= Area of NDVI class i per municipality (ha) from 1998-2006
n	= Number of NDVI classes
ε_i	= Residual error

By crop the function was applied on the masked NDVI-map to generate a map showing cropped fractions by map unit for various crops. The prepared maps reflect quantitatively the area status from 2001-05 for specific crops.

3.2.4 Validation of estimated crop maps

Validation of the estimated crop maps was performed by using field data obtained from the Ministry of Agriculture and Fisheries (Figures 3- 3 and 4). First, the average fractions for the years 2001 to 2005 of each crop from the segments located in agricultural areas according to the CORINE land cover map were calculated. Then these data were correlated with the estimated cropped fractions for the surveyed segments using weighted linear regression analysis by using the total area sampled over five years (2001-2005) of specific crop in each segment as weight to estimate the following equation.

$$Y = \sum_{i=0}^{n} b_i (w_i c_i) + \varepsilon_i \qquad \text{(Eq. 3)}$$

Where:

F	=	Average fractions of crops in each NDVI class per segment from 2001-05
b_i	=	Regression coefficient
c_i	=	Estimated fraction of crops per NDVI for segment i
w_i	=	Weighing factor (total area sampled over five years (2001-05) of specific crops in segment i
n	=	No. of segments
ε_i	=	Residual error

3.2.5 Direct mapping using primary field data

The average fractions of various crops (2001-2005) by segments within agricultural areas (1428 segments) as dependent variable were correlated directly with the presence and absence of segments in 45 classes in NDVI-map as explanatory variables to estimate the function (Method B) that relates average fractions of particular crop per segment in each NDVI class with the presence and absence of each segment in that particular NDVI class. For various crops the following equation was estimated through stepwise linear regression.

$$F = \sum_{i=0}^{n} b_i \delta_i + \varepsilon_i \qquad \text{(Eq. 4)}$$

Where:

F = Average fractions of crops per segment

b_i = Regression coefficient of NDVI class i

$$\delta_i = \begin{bmatrix} 1: \text{if the segment is present in NDVI class } i \\ 0: \text{if the segment is absent in NDVI class } i \end{bmatrix}$$

n = Number of NDVI classes

ε_i = Residual error term

3.3 Results

3.3.1 Image classification

To compare the signature separability across the various runs, the "divergence" statistical measure of distance between cluster signatures was used. Figure 3- 6, presents for each run, characterized by a pre-fixed number of classes, the average of all separability values between the defined classes. It shows relatively high 'average' separability values for the run with 45 classes. Thus, an unsupervised classification, using 45 classes, is statistically the best result to classify the 1-km2 NDVI-images and to stratify the area into map units (Figure 3- 7).

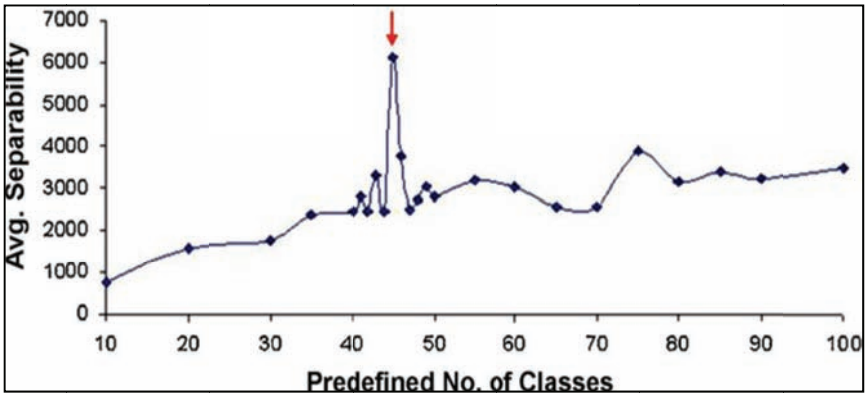


Figure 3- 6: Average divergence separability for predefined numbers of NDVI-classes

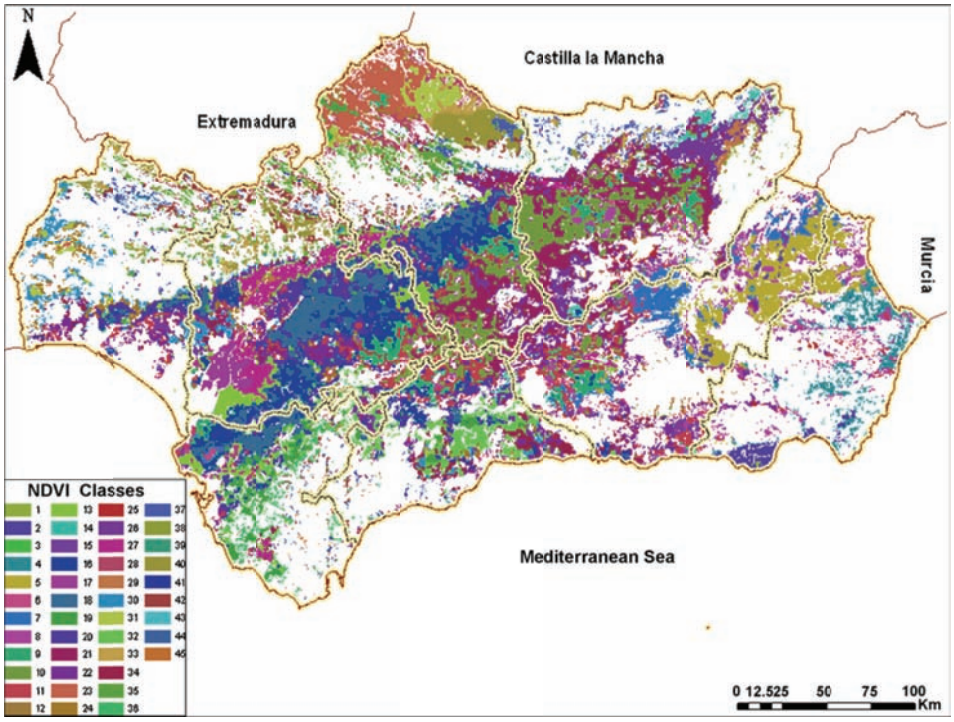


Figure 3- 7: Map of 45 NDVI-classes based on unsupervised classification of hypertemporal NDVI-images

3.3.2 Crop area maps

The occurrence of the NDVI-classes resulting from unsupervised classification with 45 classes, in agricultural areas of each municipality was calculated in hectares. Combining these results with crop statistics at municipal level (also in hectares) allows development of a matrix (part of which is shown in Figure 3-1). The areas of the 45 NDVI-classes (independent) and crop area statistics of rainfed crops (dependent) were statistically correlated through step-wise forward linear regression. The maps derived from the regression analyses are shown in figures 3- 8 to 3- 10 for rainfed wheat, rainfed sunflower and rainfed barley.

Rainfed Wheat: The coefficients derived from the stepwise forward regression, with an adjusted R^2 of 98.8 % for rainfed wheat (Table 3- 1) were used to produce the map for rainfed wheat (Figure 3- 8). Figure 3- 11 shows the temporal NDVI-profiles of selected NDVI-classes, each representing more than

a 20% area fraction of rainfed wheat per sq-km. Especially classes 18, 19 and 16 are representative for wheat.

Table 3- 1: Results of stepwise linear regression analysis for rainfed wheat

NDVI-class	Coefficient	t-Value	Sig. (%)
18	0.47	55.8	0.00
19	0.47	25.6	0.00
16	0.42	58.8	0.00
23	0.34	54.4	0.00
20	0.29	23.4	0.00
36	0.27	8.8	0.00
9	0.27	24.3	0.00
31	0.12	6.6	0.00
7	0.09	8.1	0.00
24	0.08	2.9	0.38
5	0.08	13.3	0.00
35	0.07	5.6	0.00
22	0.07	5.3	0.00
27	0.06	5.0	0.00
10	0.02	3.6	0.03

Rainfed Sunflower: The coefficients derived from the stepwise forward regression with an adjusted R^2 of 97.5% for rainfed sunflower (Table 3- 2) were used to produce the rainfed sunflower map (Figure 3- 9). Figure 3- 11 shows the temporal NDVI-profiles of NDVI-classes that represent more than 20% rainfed sunflower. Profiles 18, 20 and 16 relate to rainfed sunflower. NDVI-profiles of rained wheat and sunflower exhibit close trends because these two crops are grown on the same areas (Figure 3- 11).

Rainfed Barley: The coefficients derived from the stepwise forward regression with an adjusted R^2 of 75.7% for rainfed barley (Table 3- 3) were used to produce the rainfed barley map (Figure 3- 10).). Figure 3- 11 shows the temporal NDVI-profiles for classes having more than 10% rainfed barley. The classes considerably include other land cover types as well.

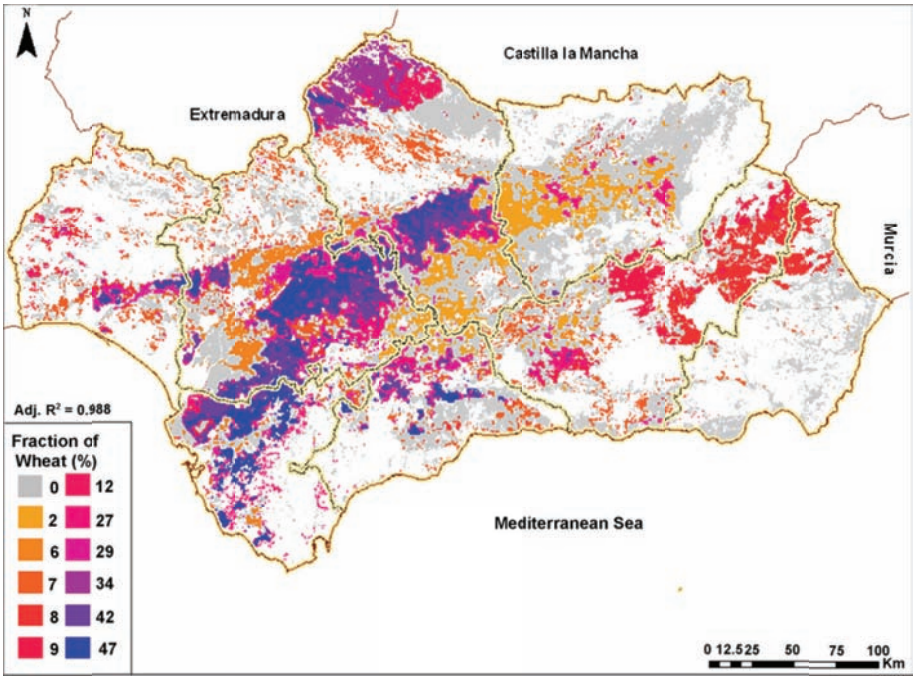


Figure 3- 8: Estimated rainfed wheat map (fractions per km^2)

Table 3- 2: Results of stepwise linear regression analysis for rainfed sunflower

NDVI-class	Coefficient	t-Value	Sig. (%)
18	0.40	50.3	0.00
20	0.22	19.5	0.00
16	0.22	33.2	0.00
9	0.15	15.3	0.00
36	0.13	5.4	0.00
24	0.11	4.1	0.00
27	0.06	4.7	0.00
7	0.03	3.2	0.14
22	0.02	2.0	5.00

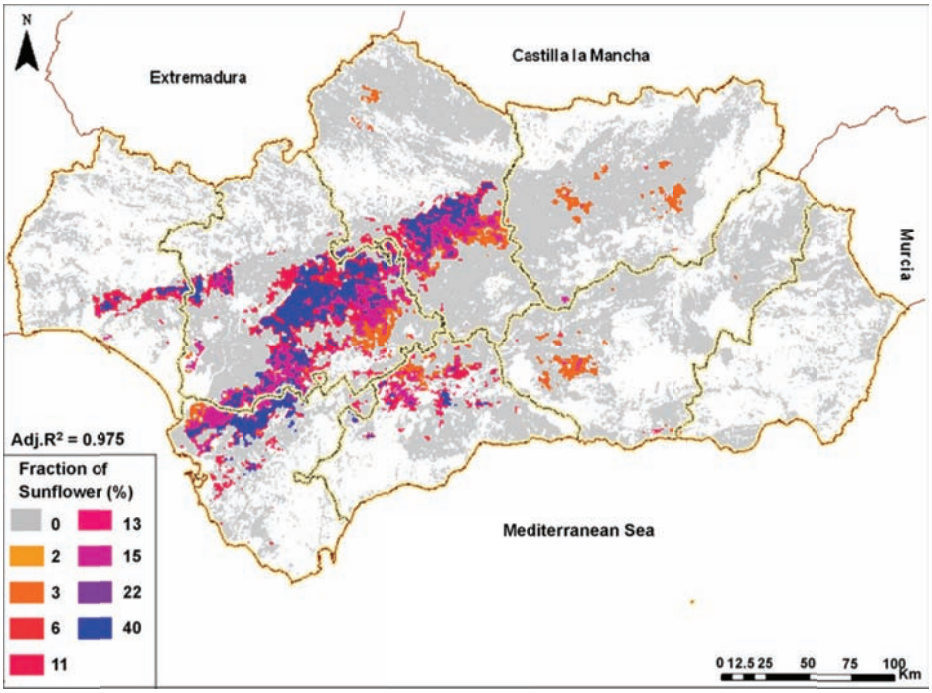


Figure 3- 9: Estimated rainfed sunflower map (fractions per km^2)

Table 3- 3: Results of stepwise linear regression analysis for rainfed barley

NDVI-class	Coefficient	t-Value	Sig. (%)
7	0.22	27.3	0.00
5	0.13	29.3	0.00
14	0.09	4.2	0.00
36	0.04	2.2	3.08
11	0.03	3.4	0.07
23	0.03	6.2	0.00
22	0.03	2.5	1.14
32	0.02	2.8	0.46
9	0.02	2.9	0.38

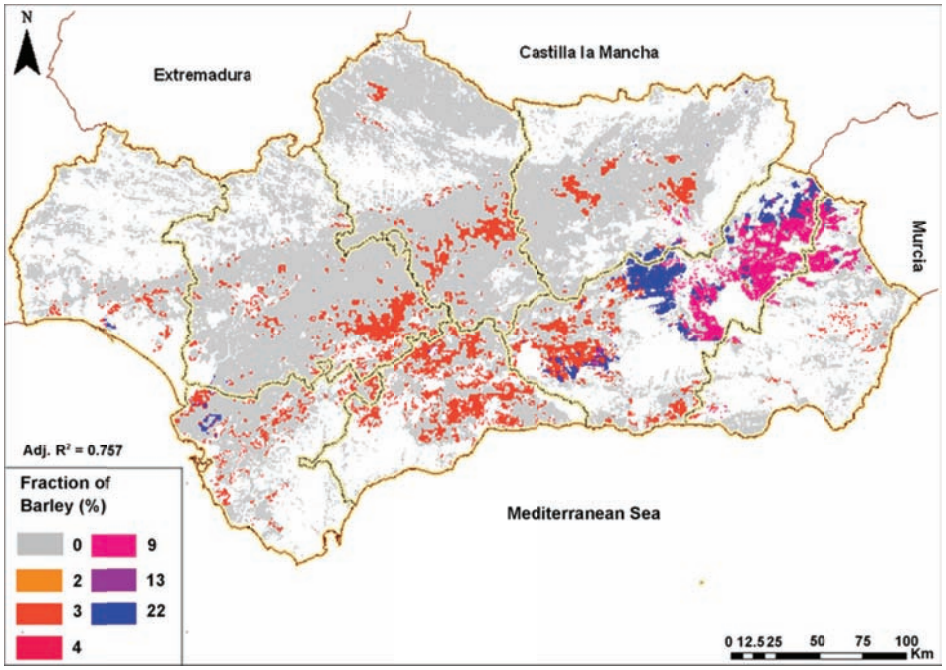


Figure 3- 10: Estimated rainfed barley map (fractions per km^2)

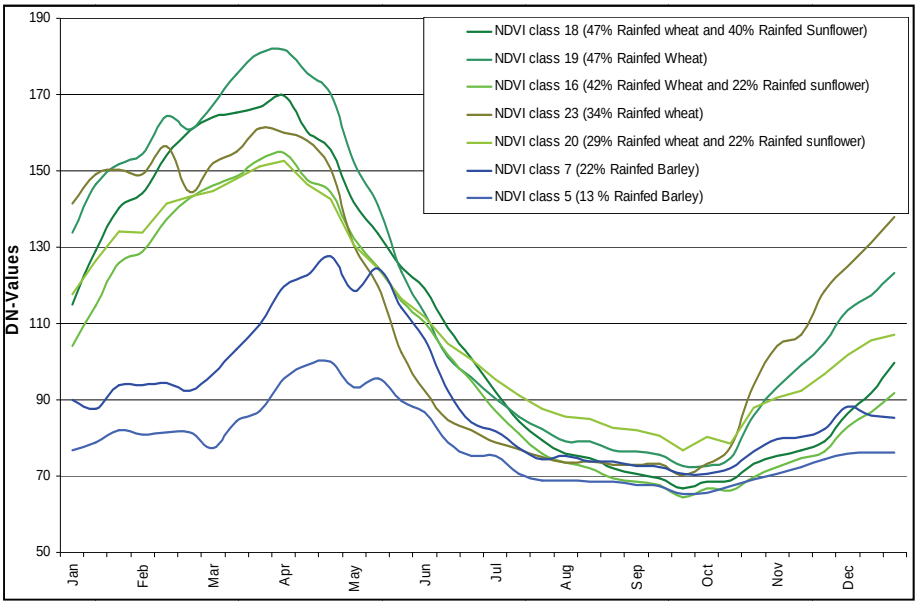


Figure 3- 11: SPOT-Vegetation temporal NDVI-profiles for classes representing studied crops (Andalucia)

3.3.3 Validation of the estimated crop maps

The average fractions of the various crops in each segment were calculated for the years 2001 till 2005. The fractions of the various crops per segment in all NDVI-classes were correlated with the fractions estimated from the reported administrative area crop statistics through weighted linear regression. The total area sampled by segment over the five years was used as weighing factor. The results showed good agreement between the actual fractions of rainfed wheat by segment and the estimated fractions presented in the NDVI-based rainfed wheat map. The rainfed wheat map explained 68% of the total variability (Adj. R^2) between sampled segments (Figure 3- 12). The variability between the segments by NDVI-class remains however high; the defined units are internally heterogeneous.

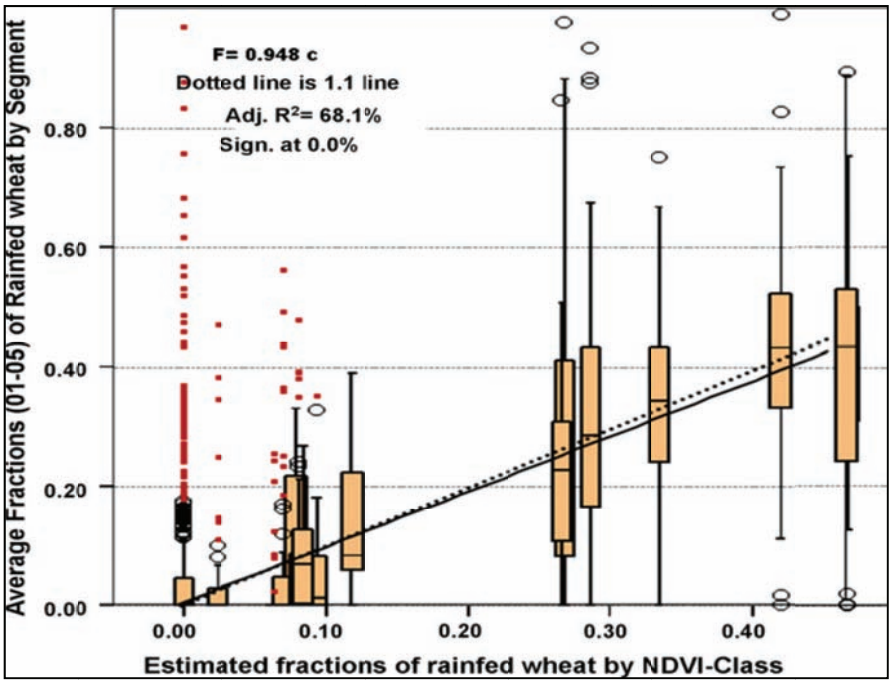


Figure 3- 12: Validation of NDVI-based rainfed wheat

The estimated wheat map has a substantial degree of generality. The NDVI-map is thus considered a small scale map. It excludes local variability within the map units delineated. The regression line is close to the 1:1 line indicating that the

estimated fractions correspond well with reality. The estimated maps of rainfed sunflower and rainfed barley explained 31% and 23% of the segments variability (Adj. R^2), respectively (Figure 3- 13and 3- 14). Equations are also presented in Figures 3- 12, 13 and14, where F is Average fraction respective crop per segment by NDVI class and c is fractions of respective crops estimated by NDVI class. For rainfed sunflower the regression line substantially deviates from the 1:1 line, the area of rainfed sunflower seems considerably over-reported by the crop statistics leading to an NDVI based map that considerably over-estimates the crop fractions. For rainfed barley, the regression line is close to the 1:1 line, but the heterogeneity within defined units is high.

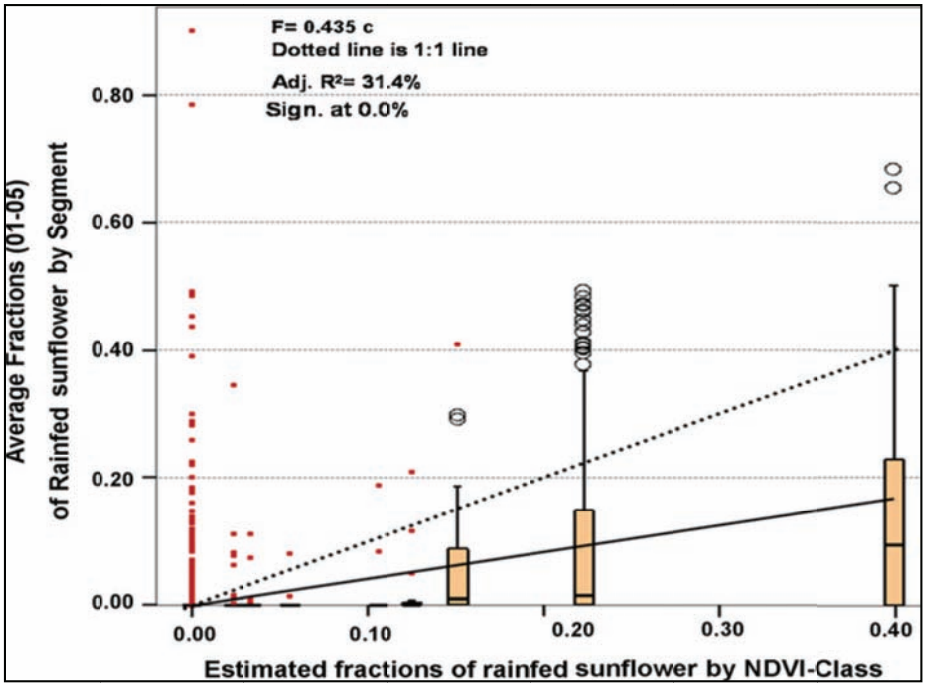


Figure 3- 13: Validation of NDVI-based rainfed sunflower map.

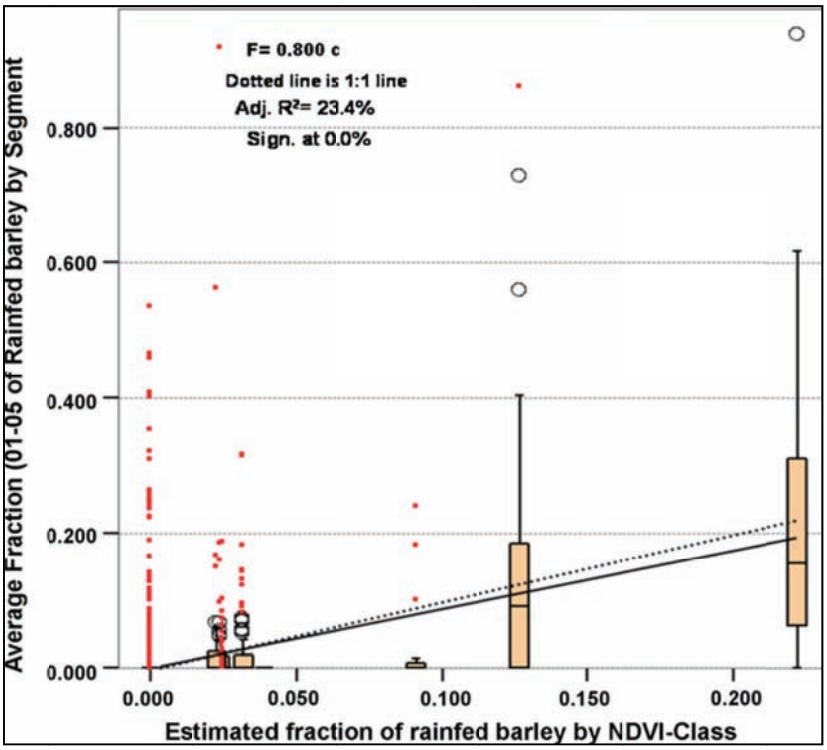


Figure 3- 14: Validation of NDVI-based rainfed barley map

3.3.4 Direct mapping using primary field data

The average fractions of crops (2001-2005) in the field segments within agricultural areas (1424 segments) were correlated with the presence and absence of segments in 45 classes in NDVI map directly to establish the functions relating cropped areas to NDVI-classes through stepwise linear regression. The resulting coefficients are given in Table 3- 4 with adjusted R^2 of 27%, 25% and 23% for rainfed wheat, rainfed sunflower and rainfed barley respectively. The relatively low values of adjusted R^2 because of the presence of thousands of segments over the five years (2001-05). These coefficients of regression analysis were used for estimating the three crop maps (% per km^2) by NDVI-class based on average fractions of crops in field segments for rainfed wheat, rainfed sunflower and rainfed barley (Figure 3- 15).

Table 3- 4: Results of the stepwise linear regression analysis using primary field data and the NDVI-map

Crop	NDVI-class	Coefficient	t-value	Sig. (%)
Rainfed Wheat	36	0.46	5.9	0.00
	16	0.44	12.4	0.00
	19	0.42	7.6	0.00
	18	0.41	10.6	0.00
	23	0.35	6.3	0.00
	20	0.30	10.5	0.00
	9	0.23	5.1	0.00
	31	0.20	2.0	4.60
Rainfed Sunflower	16	0.13	13.9	0.00
	18	0.14	13.5	0.00
	20	0.07	9.3	0.00
	9	0.06	5.0	0.00
Rainfed Barley	7	0.22	14.4	0.00
	5	0.13	14.1	0.00
	11	0.03	3.2	0.10
	32	0.03	3.1	0.20
	9	0.03	2.5	1.30

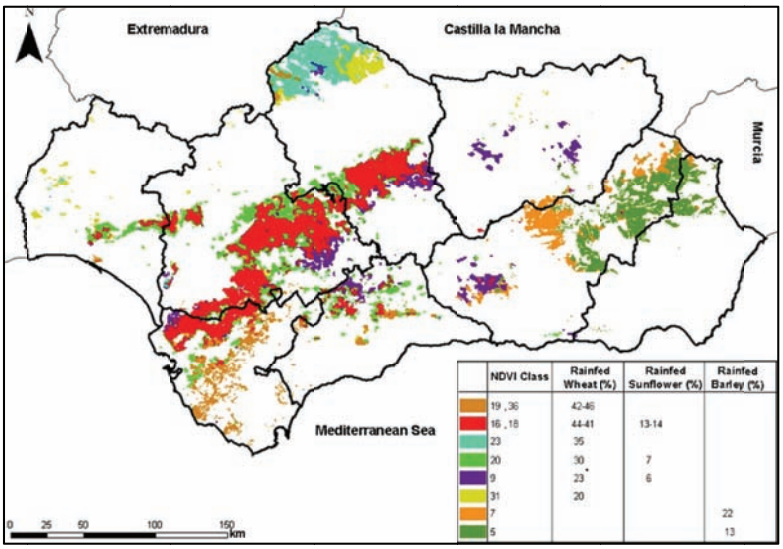


Figure 3- 15: Estimated fractions per km^2 of crops using the primary data approach

Subsequently, the coefficients derived from the analysis of NDVI-classes and crop areas from field segments (Table 3- 4) were correlated with the coefficients derived from the correlation analysis of NDVI-classes and municipal crop statistics (Tables 3- 1, 3- 2 and 3- 3) to examine the extent that the methods are in agreement with each other. The results of comparison between the two methods are given in figure 3- 16. This figure indicates once more that the official crop statistics for rainfed sunflower are considerably over-estimated, but also that fractions estimated by both methods do correspond well (high R-squares).

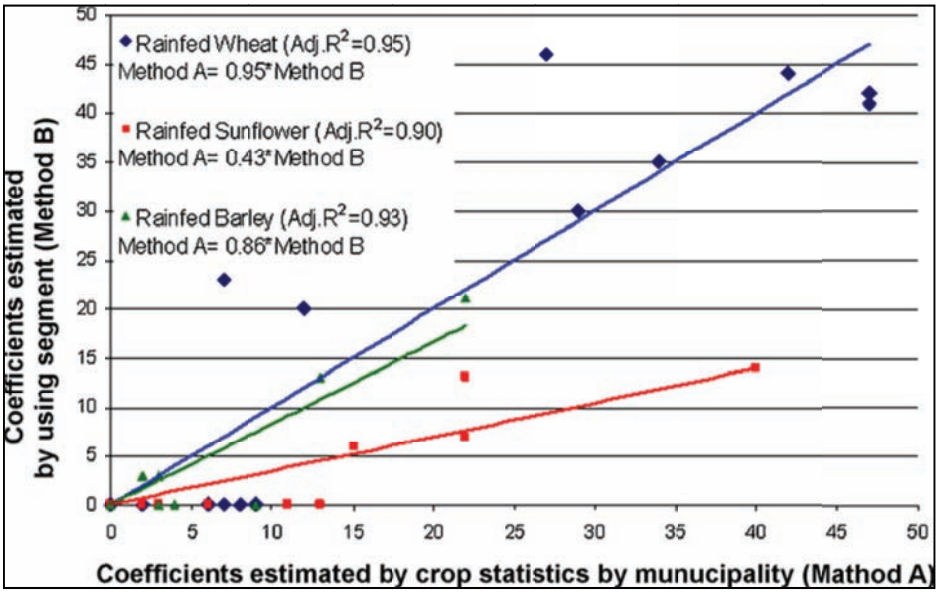


Figure 3- 16: Comparison of coefficients estimated methods A and B

3.4 Discussion

Before embarking on crop production monitoring, it is necessary to describe properly 'where-what' is being monitored, which we feel has insufficiently been recognized. The method developed in this study shows that available existing data and satellite imagery can be used to map objects of interests. Our maps remain generalized across 1 km^2 pixels, each a 'mixel', characterized by a unique land cover and land use complex. It shows clearly delineable map units for which their sizes do not seem to impair the reliability of the method described. Validation by using the primary field data clearly showed that the

maps produced have a good accuracy and that they can be used for monitoring purposes like food security and early warning and crop growth monitoring.

The long term goal of Integrated Global Observations for Land (IGOL) as described in its report (IGOL, 2007) is "enhancement of agricultural survey / monitoring capabilities through the realization of satellite observations as an integral part of the overall agriculture survey and monitoring for all countries". In order to achieve this goal the strategy which is being followed uses the combination of field survey and remote sensing observations. It has been stressed by the IGOL theme group that remote sensing can be used in agricultural surveys and its utility lies in stratification, model-based estimation in combination with ground surveys. Area frame sampling methodology is being supported by remote sensing (Tsiligirides, 1998; Pradhan, 2001). The National Agricultural Statistics Service (NASS) of the United States Department of Agriculture has used or considered since long, regression based estimators for small area crop acreage estimation with ancillary satellite data. These estimators use stratum level counts of pixels classified to crops (Bellow, 1993). Remote sensing imagery and NASS survey data have been combined to produce improved acreage estimates (Kutz *et al.*, 2005).

The method described in this chapter also suggests the usefulness of NDVI regarding the generalization and stratification of the large spatial data set by supervised grouping following the unsupervised classification. The method applied in this research makes use of analysis of NDVI time-series using unsupervised ISODATA clustering algorithm. The divergence statistics was used for determining the optimal number of NDVI-classes which were then related to reported crop statistics. The NDVI time series has the ability to capture the crop phenologies (crop calendars) and thus has a good relationship with the cropped areas as shown in this chapter. Each land use class, in this case the studied crops, has been defined by its corresponding NDVI-profiles. The scope of using NDVI time series for area frame sampling has a greater benefit that will lead to true estimates of cropped areas. The timely and accurate estimation of cropped areas is the basic element in any kind of food security endeavour.

Estimated maps of rainfed wheat, rainfed sunflower and rainfed barley are dependent on the quality of reported statistics which in case of rainfed

sunflower clearly identified a serious overestimation of reported statistics. Primary field data which is a direct approach of the method used gives better results as compared to reported statistics at municipal level because while compiling the municipal level crop statistics generalization method is applied.

The spatially explicit crop statistics helps to identify the areas where the various crops are actually grown. This supports policy makers and researcher to identify the areas of their interest and makes it easier for devising policies and strategies concerning food security. The thorough validation of the prepared crop maps suggests the additional use of higher resolution images and additional spatial information to further improve the method of generating crop maps and to capture additional spatial heterogeneity that exists at local level.

4

Integrating soil maps in a model to map crop areas using hypertemporal NDVI images and crop statistics

* This chapter is based on Khan, M.R., de Bie, C.A.J.M., van Keulen, H., Smaling, E.M.A. and Real, R., 2010. Integrating soil maps in a model to map crop areas using hypertemporal NDVI images and crop statistics. Submitted to RSE for publication

Abstract

Agricultural area statistics data are compiled and distributed in tabular form at administrative levels by governments but give no idea about the exact locations where specific crops are actually grown. Earlier presented (Khan *et al*, 2010) was a model to map and disaggregate crop area statistics using hyper-temporal NDVI images. It was proven that the NDVI data correlated very well to the spatial diversity of crops grown and cropping patterns over time. Assumed was however that the NDVI data comprised the combined influences of varying soil, terrain, weather and land use conditions. In this chapter, we tested that assumption with respect to soil-geogenesis information. Four modelling options were used to test if adding soil data to the NDVI data significantly improves model output. The first two options made use of officially reported wheat area statistics by municipality, whereas options 3 and 4 used field data on wheat. Options 1 and 3 used only NDVI data; options 2 and 4 also used soil maps. Model outputs comprised of the amount of variability explained and wheat maps. Options 1 and 2 and options 3 and 4 were compared regarding the variability explained. Results of option 1 showed that wheat statistics were significantly related with the NDVI map and explained 98% of the variability in cropped area by municipality. Option 2 that included soil information explained 99% of the variability. Use of NDVI data proved additional use of soil data redundant because it improved the amount of explained variability by only 1 %. Option 4 provided also only 1% better result than option 3 (65 versus 64% variability explained). Map validations for options 1 and 2 showed 56% and 61% variability explained respectively between the mapped information and the segment data. However, the slope of the regression line of option 1 comparing estimated versus actual cropped fractions was closer to the expected 1:1 ratio than option 2, i.e. 0.90 versus 0.85. Use of soil information to prepare a wheat map of Andalucia did not add substantially to the performance of the model solely based on NDVI data. The assumption that NDVI acts as an indicator of combined influences of varying spatial conditions is thus not proven false through this study.

4.1 Introduction

Timely availability of accurate agricultural land use/land cover maps is required by agricultural land use planners, policy makers, donor agencies and crop insurance companies for a variety of purposes (Becker-Platen, 1976; Jansen and Di Gregorio, 2004). Currently, governments and international organizations compile annual estimates of crop areas by administrative units and present those in tabular form. Detailed (mapped) information on the spatial extent of specific land uses is generally lacking. Figure 4- 1 shows reported rainfed wheat area statistics (% by municipality) masked by CORINE agricultural areas, it does not show "exactly where" in various municipalities the crop is grown. Such maps are not suitable for efficient monitoring and early warning studies on crop production (Verburg *et al*., 2002; Aalders and Aitkenhead, 2006). Preparation of conventional land use maps is both expensive and time-consuming, and their information quickly becomes outdated. This necessitates the importance of developing properly tested methods with modern techniques for mapping crop areas using existing data sources such as crop statistical and the field data.

Remote Sensing (RS) technologies combined with geographic information systems are efficient means to acquire, compile and distribute agricultural land use data. Satellite remote sensing techniques have been shown to be effective in preparing accurate land use/land cover maps and to monitor changes at regular intervals (Yang *et al*., 2007; Zhang and Zhang, 2007). International agencies, such as the Food and Agriculture Organization of the United Nations (FAO), European Space Agency (ESA), European Commission (EC), National Remote Sensing Center of China (NRSCC), United Nations Environment Programme (UNEP) and United States Geological Service (USGS) stress the importance of using remotely sensed images in combination with ground data/observations to generate necessary information for land use studies at both national and global scales (Townshend *et al*., 2008).

RS based vegetation indices have been widely used to monitor land cover changes (Stow *et al*., 2004; Focardi *et al*., 2008). Such vegetation indices are based on the differential absorption, transmittance, and reflectance of energy by the vegetation in the red and near-infrared regions of the electromagnetic spectrum (Jensen, 1996). The Normalized Difference Vegetation Index (NDVI) is the most common index used in vegetation studies. NDVI is an indicator of

greenness, and is therefore associated with photosynthetic activity and dry matter production (Cayrol *et al*., 2000; Han *et al*., 2004: le Maire *et al*., 2004).

For studying crops, the spatial and temporal resolution of the RS images is very important. Normally, annual crops are in the field for 6-7 months. Therefore, substantial information in terms of images is required. The vegetation instrument onboard the SPOT satellite has four spectral bands, i.e., blue (0.43-0.47 μm), red (0.61-0.68 μm), infrared (0.78-0.89 μm) and short wave infrared (1.58-1.75 μm), and provides information at a spatial resolution of 1 km^2 and a temporal resolution of one day. This meets the requirements for land cover mapping at regional scale.

Hypertemporal (long temporal sequence of regularly acquired) remote sensing and administrative level crop area statistics have been used to prepare crop maps and cropping pattern maps at a spatial resolution of 1 km^2 for India, Iran, Spain and Vietnam (De Bie *et al*., 2008; De Bie *et al*., 2010; Khan *et al*., 2010; Nguyen *et al*., 2010). Relating crop area statistical data averaged from 2001-05 for each crop with the unsupervised classified NDVI map has shown that the selected crops were significantly related to specific NDVI-based map units (Khan *et al*., 2010). Each 1 km^2 pixel characterizes a specific land cover and land use mosaic. It was assumed that NDVI data comprise the combined influences of varying soil, terrain, weather, climate and land use conditions. The hypertemporal NDVI images were used to stratify the study areas and afterwards linked with current spatial information to produce land use maps. However, landscapes are efficient units to divide the earth's surface into areas that share common management practices or resource opportunities and limitations. Cropping patterns and cultivation practices are conditioned by the gradient and configuration of the landscape (Schoeneberger and Wysocki, 2002).

Agricultural land use studies must include studies of land in combination with the management and operational activities, since decision-making on the use of land resources depends on many factors such as climate, soil, water availability, slope, relief etc. Soil is an important component of land that has a strong impact on crop performance and thus land is allocated to various crops in accordance with its soil characteristics. This necessitates to properly testing the explanatory behavior of NDVI data in the presences of other predictors such as soil,

weather, climate and land use conditions. Amongst others, soil characteristics such as water holding capacity and fertility play an important role in decision making about agricultural land use, because they affect crop productivity (De Bie, 2000; Arshad and Martin, 2002). How to use the land resources depends on the accessibility of the necessary information on factors such as climate, soil, water, or socio-economic factors (FAO, 1994; Barahona and Iriarte, 2001; Vandermeulen *et al*., 2009). Such data have been used for downscaling information on agricultural land use (Kempen *et al*., 2007).

This follow up chapter aims to test if the NDVI data reflect the combined influences of varying soil, terrain, weather, climate and land use conditions with a specific focus on soil information. Therefore, soil data were included in the method of mapping agricultural land cover/use developed by Khan *et al*. (2010). The method used to test the hypothesis if NDVI data also reflect soil information is based on 10 daily Maximum Value Composite (MVC) SPOT VEGETATION data, soil types and soil geomorphic types. In the study tabular statistical data on cropped areas for rainfed wheat per administrative unit, averaged for the period 2001-2005 to make rainfed wheat maps valid for the period 2001-2005. The overall aim is to further test the developed method for combining spatial and temporal land-use data sets using existing data sources and improved RS/GIS-based methods.

4.2 Materials and Methods

4.2.1 Study Area

Andalucia is the southernmost community of mainland Spain with an area of 87,268 km^2 comprising eight provinces and 770 municipalities (Figure 4- 1). The Mediterranean climate of Andalucia is characterized by mild rainy winters and hot dry summers. Almeria and Sevilla municipalities of the respective provinces have the highest average annual temperatures in Spain with 18.6 and 18.7 ºC, respectively. The average annual temperature of Andalucia as a whole is above 16 ºC. Overall annual rainfall is highly variable and ranges from a maximum of 2000 mm to a minimum of 170 mm (Font, 2000).

About 70% of Andalucia is utilized for agricultural purposes. Medium-sized mountains dominate the landscape, occupying 42% of the total area.

Consequently, 38% of the agricultural land is mountainous with crops generally restricted to the inner valleys or to gently sloping hillsides. The major annual and tree crops are wheat, olives, sunflower, cotton, rice, maize, grapes and oranges. Rainfed wheat dominates the field crops grown in Andalucia (Khan *et al*, 2010). The major soil types in Andalucia are Cambisols, Regosols, Luvisols and Vertisols (FAO, 1998), and the major soil geomorphogenetic classes are Structural, Denudative and Fluvio-Colluvial (de la Rosa *et al*., 2009).

In figure 4- 1, rainfed wheat areas of Andalucia are presented. These areas are based on crop statistics data (2001-05) masked by agricultural areas as defined by CORINE land cover with map of Spain in the inset.

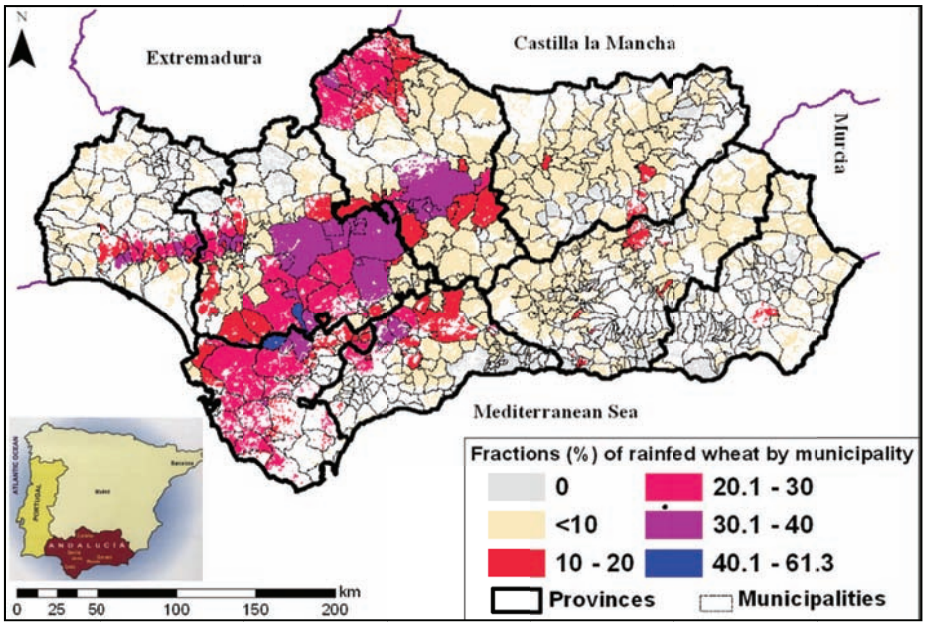


Figure 4- 1: Description of Study area

4.2.2 Data used

Data used in the study included:

- Geo-referenced and de-clouded SPOT4 and SPOT5 Vegetation (VGT) Sensor's 10-day MVC'S (Holben, 1986) NDVI images (S10 product) at 1

km^2 resolution from the first decade[1] of April 1998 to the last decade of July 2006 as obtained from www.VGT.vito.be [27 + (36 * 7) + 21 =300 images]. De-clouded means that pixels with a 'good' radiometric quality for bands 2 (red; 0.61-0.68 μm) and 3 (near IR; 0.78-0.89 μm) were retained by using the supplied quality record by pixels. The pixels having shadow, clouds and uncertainty were removed and labeled as 'missing'. The first VGT sensor was launched in March 1998 on the platform of the SPOT4 satellite. Its specifications are well suited for terrestrial applications like land cover mapping (Zhou *et al*., 2009). The NDVI values used in this research are digital number (DN) values in 0-255 format. Data conversion, made for easy handling of the data, is based on the following formula:

$$DN = \frac{NDVI + 0.1}{0.004} \qquad \text{(Eq. 1)}$$

All 300 combined NDVI images were classified to obtain an NDVI map representing differences in land cover / land use types existing in Andalucía.

- Crop (rainfed wheat) area statistics per municipality (averaged from 2001-2005), obtained from the Ministry of Agriculture and Fisheries, Andalucía. The crop statistics data is generated on the basis of areas reported by farmers at the time of application for subsidies.

- The CORINE (Coordination of Information on the Environment of the European Environmental Agency (EEA)) land cover 2000 (CLC) map (scale 1:100,000) was used to extract agricultural area.

- The administrative map of Andalucía, including the boundaries of provinces and municipalities.

- Segments data, collected from 2001-2005 by the Ministry of Agriculture and Fisheries. The survey comprised 1451 segments of 700 m x 700 m, distributed evenly over the agricultural areas of study area (Figure 3- 3). All agricultural fields present in each segment were digitized and annual data were collected for all agricultural fields per segment through field visits. Cropped area, production per crop, irrigation regime, management practices

[1] Decade is understood in this chapter as a 10-day period

performed by farmers, etc were recorded. In total 1428 segments coinciding with agricultural areas as defined by the CORINE land cover map were used for validation of the produced maps. All segments were not sampled every year. Many segments only partially covered agricultural areas. The total extent of agricultural areas sampled per segment per year was calculated. The extent of segments by NDVI classes and by soil units was also calculated.

The aforementioned data is same as used by (Khan *et al*., 2010).

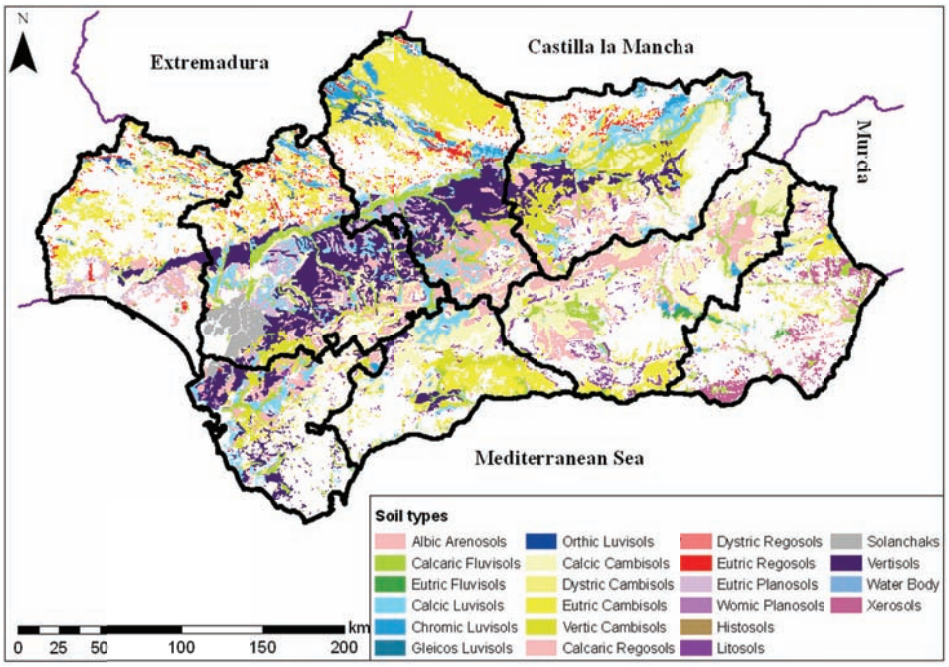


Figure 4- 2: Soil types map of Andalucia, Spain

- Soil type map of Andalucía (1:400,000) was obtained from the Consejo Superior de Investigaciones Científicas (CSIC). The map contains information on morphological and analytical soil data; the legend is presented according to the FAO soil classification system (Figure 4- 2).

- Soil geomorphic map of Andalucía (1: 400,000) was obtained from the Junta de Andalucía. The legend is based on a soil morphogenesis

classification (Figure 4- 3). In geomorphogenesis both physiographic and geomorphic dynamics are considered (Krol *et al*., 2007).

All geo-datasets were projected in Universal Transverse Mercator (UTM) system (Zone 30N) using the European 1950 Datum (Spain).

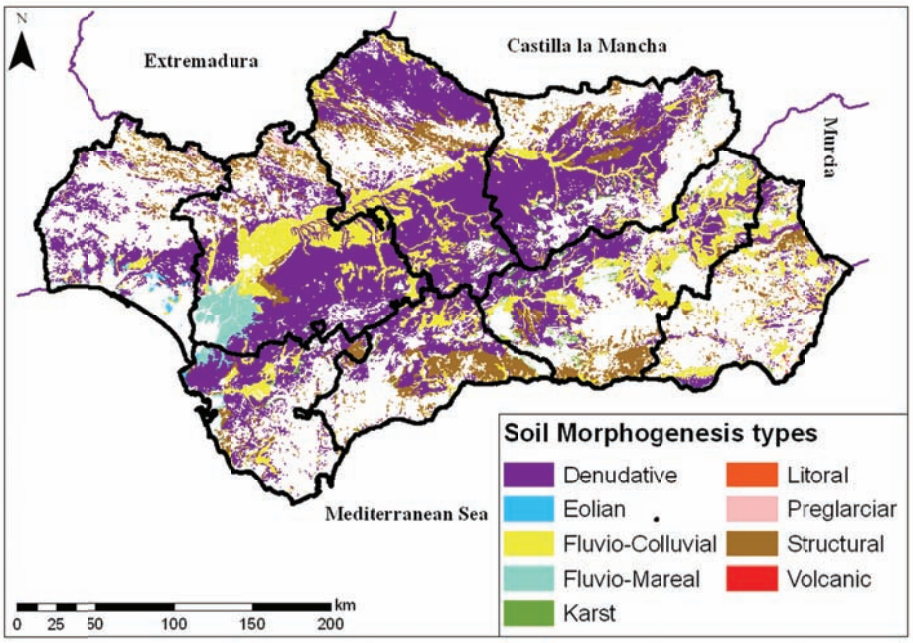


Figure 4- 3: Soil geomorphology map of Andalucia, Spain

4.2.3 Image classification

The stack of 300 NDVI images was classified to obtain the optimum number of NDVI classes to stratify the study area into mapping units. The iterative self-organized unsupervised clustering algorithm (ISODATA) of the ERDAS Imagine software, based on minimum spectral distance, was used to identify spectral clusters from the NDVI images (ERDAS, 1997; 2003). A cluster is a group of pixels (classes) with similar spectral characteristics. The algorithm begins by treating the entire data set as one cluster and identifies a number of natural spectral clusters after performing a predefined number of iterations. The ISODATA is self organized because it requires minimum user inputs. The ISODATA utility repeats the clustering of pixels until either a pre-defined maximum number of iterations has been performed (set to 50), or a maximum

convergence threshold is reached (set to 1.0). This convergence value shows that the utility will stop processing as soon as 100 % of the pixels stay in the same cluster between iterations. Performing an unsupervised classification is simpler than a supervised classification, because the cluster signatures are automatically generated by the ISODATA algorithm. Various ISODATA runs were carried out to identify 10 to 100 classes. In each run the desired number of classes was produced by the ISODATA clustering algorithm. The divergence statistical measure of separability (Swain and Davis, 1978; ERDAS, 2003) between defined cluster signatures per run was used to compare the various runs. The optimum run, i.e. with a distinguishable peak in divergence separability values was selected to continue this study.

4.2.4 Resultant NDVI classes by agricultural area

The NDVI classes in the agricultural areas of Andalucia were extracted from the unsupervised classified map of NDVI time series (output of Section 2.3). For this purpose we used the agricultural areas as defined in the CLC 2000 map. Areas comprising arable land, permanent crops, pastures and mixed crops were retained to construct a map of major agricultural areas. For generating the mask of agricultural areas, all parts of these polygons were considered while overlaying with the NDVI classified map. The other CLC-classes were masked out. The area of each NDVI class per municipality in these agricultural areas was calculated through overlay of the extracted NDVI map of agricultural areas and the administrative map of Andalucia.

4.2.5 Area of soil units at municipal level

The soil units were included as soil information in the method proposed for mapping crop areas using NDVI class areas and the reported crop statistics (Khan *et al*., 2010). The detailed soil parameters such as nutrient contents, water content, pH, etc., were not available to be included in the study. Soil types (according to FAO classification) and soil geomorphogenesis types (referring to the evolutionary material of the soils) were combined to generate the soil units (soil type - soil geomorphology combinations) for the study area. The areas of the various soil units per municipality in agricultural areas were calculated by combining this map after extraction of agricultural areas as defined in CLC 2000 map with the administrative map.

4.2.6 Mapping rainfed wheat areas using municipal statistics

The areas covered by each NDVI class and by each soil unit and the rainfed wheat area statistics per municipality were used in the regression model to estimate the areas cropped to rainfed wheat from (i) the areas of the NDVI classes only (Option 1 as shown by Eq. 2) and (ii) the areas of the NDVI classes combined with the soil units (Option 2 as shown by Eq. 3).

$$Y = \sum_{i=1}^{n} b_i x_i + \varepsilon_i \qquad \text{(Eq. 2)}$$

Where,

Y = Average rainfed wheat area (ha) per municipality from 2001-05

b_i = Regression coefficient for NDVI class i per municipality

x_i = Average area (ha) of NDVI class i per municipality from 1998-2006

n = Number of NDVI classes

ε_i = Residual error

$$Y = \sum_{i=1}^{n} b_{1i} x_{1i} + \sum_{j=1}^{m} b_{2j} x_{2j} + \varepsilon_{ij} \qquad \text{(Eq. 3)}$$

Where,

Y = Average rainfed wheat area (ha) per municipality from 2001-05

B_{1i} = Regression coefficient for NDVI class i

x_{1i} = Average area (ha) of NDVI class i per municipality from 1998-2006

n = Number of NDVI classes

b_{2j} = Regression coefficient for soil unit j

x_{2j} = Area (ha) of soil unit j per municipality

m = Number of soil units

ε_{ij} = Residual error

The equations were estimated through forward step-wise multiple linear regression at 95% confidence interval with no constant. The coefficients were constrained between 0 and 1.0, because the cropped area can neither be negative nor exceed the total area available. The stepwise multiple regression analyses were continued until the adjusted R^2 did not increase more than 1% in successive runs. Jackknife test was applied to the results of equation 2 to confirm the importance of soil units as predictors of rainfed wheat areas by municipality. For this purpose every single predictor was tested with an equation comprising all other predictors except the concerned predictor. The jackknife test yields the relative importance of each predictor by checking how much explained variability is decreased when a predictor is not used in the regression analysis (Efron and Gong, 1983; Lobo *et al*., 2002).

4.2.7 Validation of rainfed wheat maps based on municipal area statistics using segments data (2001-2005)

Validation of the estimated rainfed wheat maps was performed by using segments data for 2001 to 2005 (Figure 4- 2). The estimated rainfed wheat fractions from the maps (results of options 1 and 2) for the surveyed segments were compared to the average rainfed wheat fractions (2001 to 2005) of those segments. Since not all segments were surveyed every year and also many segments were only partially covered by the agricultural fields, a relative weight was used for this comparison.

A weighted linear regression analysis was performed by using the total area of agricultural fields sampled over five years for each segment as weighting factor. This allowed for assigning a relatively higher weight to those segments that were sampled more. It was observed that each surveyed segment was covered by a single NDVI class and a single soil unit. Subsequently, equations 4 and 5 were estimated by weighted linear regression for validation of the rainfed wheat maps produced using options 1 and 2, respectively.

$$F = \sum_{i=1}^{n} b_i \, (w_i c_i) + \varepsilon_i \qquad \text{(Eq. 4)}$$

Where,

F = Average actual fraction of rainfed wheat in each NDVI class per segment from 2001-05

b_i = Regression coefficient for NDVI class located in segment i

c_i = Estimated fraction of rainfed wheat per NDVI class for segment i

w_i =Weighing factor (total area sampled over five years (2001-05) of rainfed wheat in segment i)

n = Number of segments

ε_{ij} = Residual error

$$F = \sum_{i=1}^{n} b_{1i} \, (w_{1i} c_{1i}) + \sum_{i=1}^{n} b_{2j} \, (w_{2j} c_{2j}) + \varepsilon_{ij} \qquad \text{(Eq. 5)}$$

Where,

F = Average actual fraction of rainfed wheat per segment separately for each NDVI class and for each soil unit from 2001-05

b_{1i} = Regression coefficient for NDVI class located in segment i

c_{1i} = Estimated fraction of crops per NDVI class for segment i

w_{1i} = Weighing factor for NDVI classes (total area sampled over five years (2001-05) of specific crops in segment i)

b_{2j} = Regression coefficient for soil unit located in segment i

c_{2j} = Estimated fraction of crops per soil unit for segment i

w_{2j} = Weighing factor for soil units (total area sampled over five years (2001-05) of rainfed wheat in segment i)

n = Number of segments

ε_{ij} = Residual error

4.2.8 Direct mapping using segments data

The average fraction of rainfed wheat (2001-2005) per segment within the agricultural areas (1428 segments) as dependent variables were correlated directly with the NDVI-map alone and with the NDVI along with soil unit maps, to establish the functions to estimate the average fractions of rainfed wheat from (i) the NDVI classes only (Option 3; Eq. 6) and (ii) the NDVI classes along with the soil units (Option 4; Eq. 7) in combination with segments data.

$$F = \sum_{i=1}^{n} b_i \delta_i + \varepsilon_i \qquad \text{(Eq. 6)}$$

Where,

F = Average fraction of rainfed wheat per segment from 2001-2005

b_i = Regression coefficient of NDVI class i

$$\delta_i = \begin{bmatrix} 1: \text{if the segment is located in NDVI class } i \\ 0: \text{if the segment is not located in NDVI class } i \end{bmatrix}$$

n = Number of NDVI classes

ε_i = Residual error term

$$F = \sum_{i=1}^{n} b_{1i} \delta_{1i} + \sum_{j=1}^{m} b_{2j} \delta_{2j} + \varepsilon_{ij} \qquad \text{(Eq. 7)}$$

Where,

F = Average fractions of crops per segment from 2001-2005

b_{1i} = Regression coefficient of NDVI class i

$$\delta_{1i} = \begin{bmatrix} 1: \text{if the segment is located in NDVI class } i \\ 0: \text{if the segment is not located in NDVI class } i \end{bmatrix}$$

b_{2j} = Regression coefficient of soil unit j

$$\delta_{2j} = \begin{bmatrix} 1: \text{if the segment is located in soil unit j} \\ 0: \text{if the segment is not located in soil unit j} \end{bmatrix}$$

n = Number of NDVI classes

m = Number of soil units

ε_{ij} = Residual error term

The equations were estimated through forward step-wise multiple regression for the average fraction of rainfed wheat per segment, as explained in Section 2.6.

4.3 Results

4.3.1 Image Classification

Average divergence, a statistical measure of distance between cluster signatures, was used to compare the signature separability across the various runs. Figure 3- 6 presents the average of separability values between the defined classes for all runs. It shows 45 clearly separable NDVI-classes from the 300 NDVI images. Thus, an unsupervised classification, using 45 classes, appears statistically the optimal way to classify the 1-km^2 NDVI image series and to stratify the study area into map units. The possible reasons for the peak in divergence statistics are; the 1 km^2 resolution of the SPOT NDVI images where each pixel represents a mix of land cover/ land use which are clearly separable in case of homogenous land cover of the study area (Andalucía, Spain) at this scale. We applied the same classification on a heterogeneous area in Nizamabad, India where a clear peak could not be identified and to overcome this constraint we did supervised grouping after the unsupervised classification (De Bie *et al*, In press). Further, the pixels with clouds were not cleaned before the classification and were not used in the classification.

Subsequently, the NDVI classes present in agricultural areas were extracted and the respective NDVI-class areas (in ha) per municipality were tabulated. In figure 4- 4, the 15 major NDVI classes in terms of area with in agricultural areas are presented.

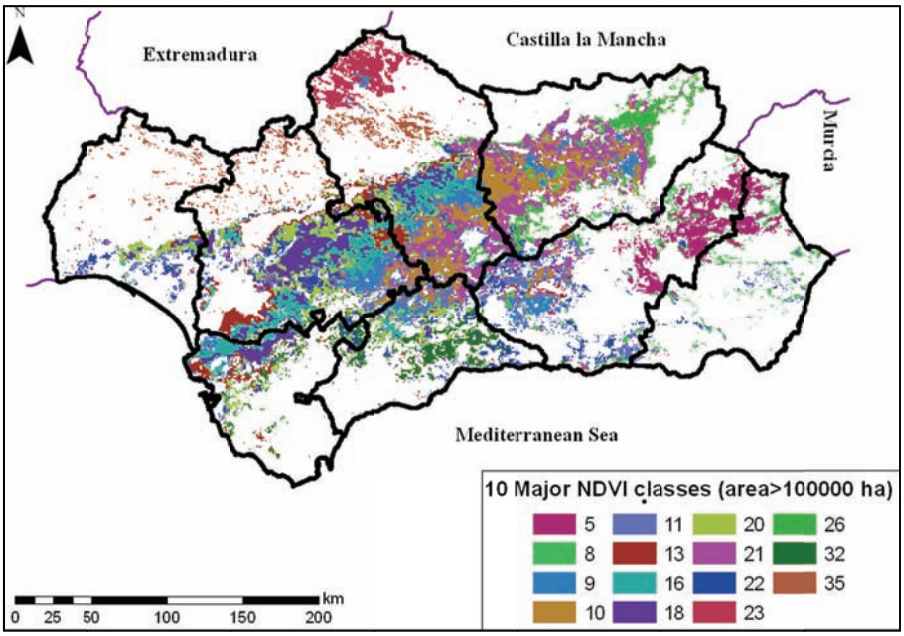


Figure 4- 4: NDVI classes map based on unsupervised classification

4.3.2 Soil units

By combining soil type and soil geomorphology, 150 soil units were identified in Andalucia, 30 are larger than 15.000 ha within the CLC agricultural zones. In terms of area, Eutric Cambisols-Structural (1,707,444 ha), Calcaric Regosols-Denudative (733,258 ha), Calcic Cambisols-Denudative (693,488 ha), Eutric Regosols-Structural (628,479 ha) and Vertisols-Denudative (607,966 ha) soil units dominate.

4.3.3 Rainfed wheat maps derived from municipal area statistics

A large matrix was generated by combining areas of derived NDVI classes and soil units with rainfed wheat area statistics at municipal level (in ha). The areas per municipality of the 45 NDVI classes and soil units (as independent) and reported rainfed wheat areas per municipality (as dependent) were statistically processed through step-wise forward linear regression as explained in Section 2.6.

The regression coefficients of estimated equations 2 and 3 to model rainfed wheat areas are reported in Table 4- 1 and Table 4- 2 respectively. These coefficients were used to make two rainfed wheat maps (Figures 4- 5 and 4- 6).

Table 4- 1: Step-wise linear regression analysis results using only NDVI class areas as predictors of rainfed wheat areas

Predictor	Coefficient	t-value	Sig. (%)	Adj. R^2 when included (%)
NDVI-18	0.47	42.9	0.00	74.8
NDVI-16	0.41	44.2	0.00	84.7
NDVI-23	0.35	45.5	0.00	92.1
NDVI-19	0.59	30.5	0.00	95.3
NDVI-9	0.28	20.3	0.00	96.9
NDVI-20	0.31	19.7	0.00	98.0

Table 4- 2: Step-wise linear regression analysis results using NDVI class areas and areas of soil units as predictors of rainfed wheat area

Predictor	Coefficient	t-value	Sig. (%)	Adj. R^2 when included (%)
Vertisols-Denudative	0.11	8.8	0.00	76.4
NDVI-23	0.37	41.0	0.00	89.6
NDVI-20	0.36	20.0	0.00	93.0
NDVI-16	0.42	32.0	0.00	95.4
NDVI-19	0.58	25.2	0.00	97.0
NDVI-18	0.33	21.8	0.00	99.1

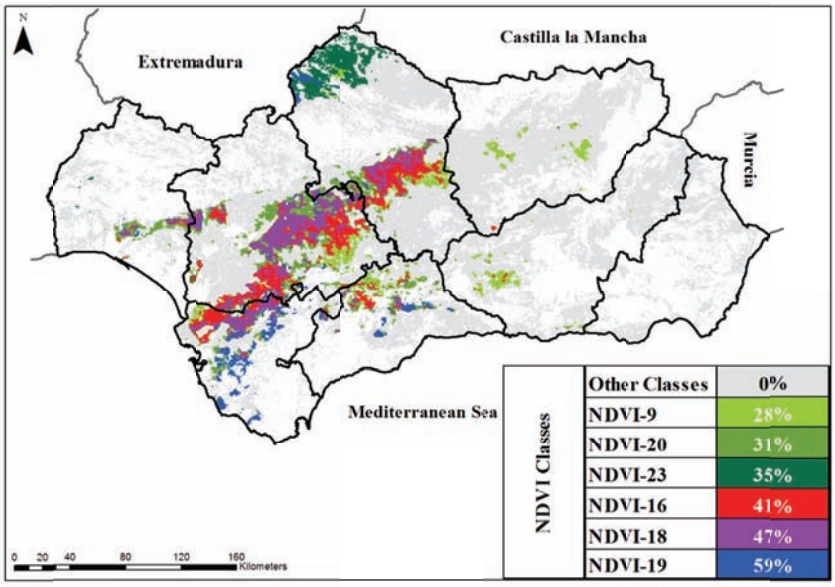


Figure 4- 5: Estimated fractions of rainfed wheat (% per km^2) based on option 1

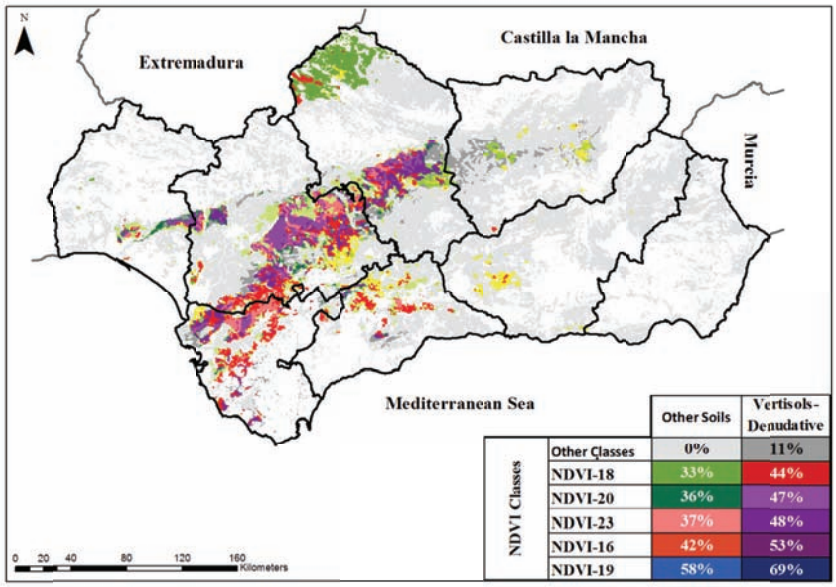


Figure 4- 6: Estimated fractions of rainfed wheat (% per km^2) based option 2

The results of jack-knife (Figure 4- 7) test show that NDVI-class 23 is the most important predictor.

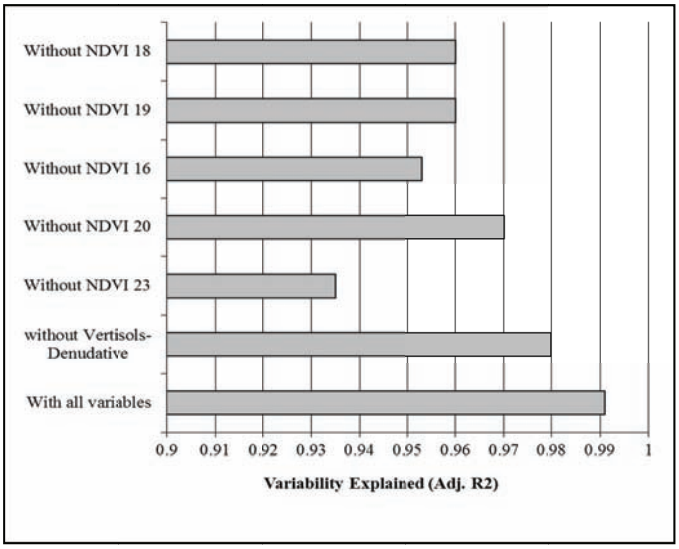


Figure 4- 7: Results of the jack-knife analysis

When the NDVI-classed 23 was removed in the jackknife procedure, the explained variability reached to minimum value. The soil data performed poorly, because when they were dropped in the jackknife procedure the explained variability only diminished by 1% (Figure 4- 7).

4.3.4 Validation of the rainfed wheat maps derived from municipal area statistics

The average fractions of rainfed wheat in the segments sampled from 2001 to 2005 were used to validate the area fractions estimated from the reported crop statistics per municipality. The results showed a highly significant positive correlation between the actual fractions and the estimated fractions of rainfed wheat per segment (Figures 4- 8 and 4- 9).

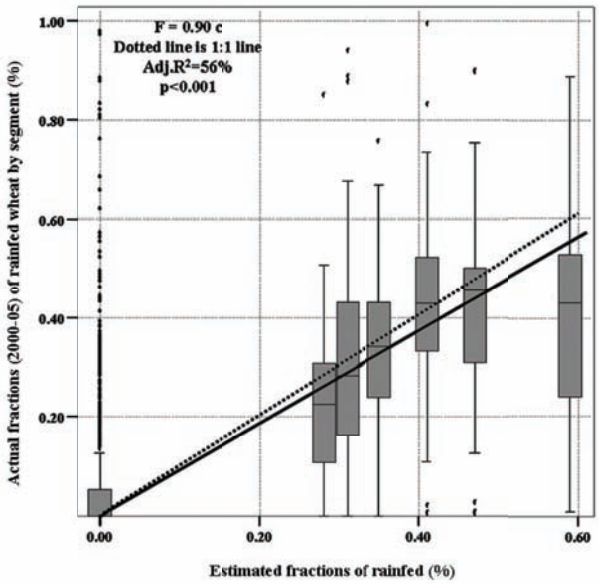


Figure 4- 8: Actual versus estimated fractions of rainfed wheat presented in Figure 4- 5

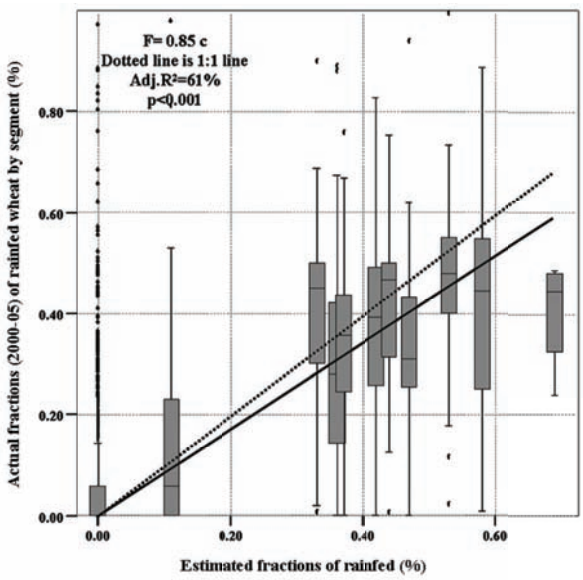


Figure 4- 9: Actual versus estimated fractions of rainfed wheat presented in Figure 4- 6

The calculated regression equations are presented in Figures 4- 8 and 4- 9, where F is average actual fraction of rainfed wheat in each NDVI class per segment from 2001-05 and c is estimated fraction of rainfed wheat by NDVI (Figure 4- 8) and per NDVI classes along with soil units per segment (Figure 4- 9). The equations suggest that the rainfed wheat areas per municipality are 10 to 15% higher than the segment data. The box plots in figures 4- 8 and 4- 9 show that though the variability within the NDVI units is higher but still the regression lines are close to 1:1 line. Therefore, it can be deducted that these maps can be used at regional scales. The equations explained 56% and 61% of the total variability among segments, respectively; hence, the estimated wheat maps exhibit a substantial degree of generalization and are not of use at local level.

4.3.5 Direct mapping using segments data

The average area fractions of rainfed wheat (2001-2005) in field segments within agricultural areas (1428 segments) were correlated with the 45 NDVI classes (Table 4- 3 and Figure 4- 10) and with the NDVI classes along with soil units (Table 4- 4 and Figure 4- 11) to establish the associated statistical relationships.

Table 4- 3: Step-wise linear regression analysis using the NDVI classes as predictor of rainfed wheat areas (option 3)

Predictor	Coefficient	t-value	Sig. (%)	Adj. R^2 when included (%)
NDVI-18	0.42	24.1	0.00	36.7
NDVI-20	0.29	23.5	0.00	50.7
NDVI-19	0.41	16.3	0.00	57.5
NDVI-23	0.32	11.7	0.00	60.9
NDVI-9	0.22	11.0	0.00	64.0

Table 4- 4: Step-wise linear regression using the NDVI and soil units as predictor of rainfed wheat areas (option 4)

Predictor	Coefficient	t-value	Sig. (%)	Adj. R^2 when included (%)
Vertisols-Denudative	0.08	7.3	0.00	22.9
NDVI-16	0.40	25.3	0.00	33.5
NDVI-20	0.27	21.4	0.00	42.6
NDVI-18	0.38	21.2	0.00	53.0
NDVI-19	0.40	16.2	0.00	59.3
NDVI-23	0.32	11.9	0.00	62.7
NDVI-9	0.20	10.2	0.00	65.2

Table 4- 5 represents relevant sample frequency information.

Table 4- 5: Area (1000 ha) and sample frequency (No. of segments) by predictors of rainfed wheat areas

NDVI Class	Other soil units	Vertisols-Denudative
NDVI-9	330 (33)	81 (13)
NDVI-16	435 (41)	228 (37)
NDVI-18	385 (38)	186 (25)
NDVI-19	210 (28)	41 (3)
NDVI-20	887 (95)	178 (22)
NDVI-23	193 (31)	0
Others	7159 (1000)	481 (62)

Since the soil unit Vertisols-Denudative and NDVI class 23 do not coincide in the segment data, the combination is indicated in Figure 4- 11 as NA.

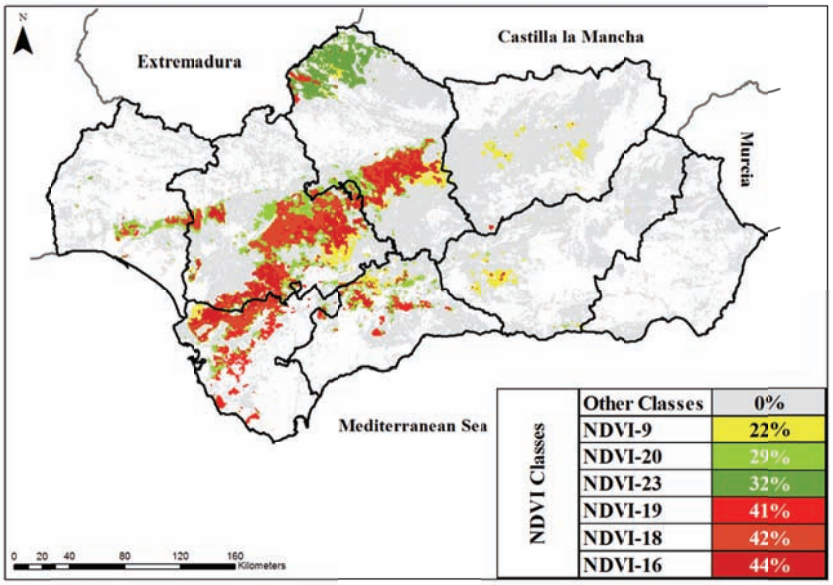


Figure 4- 10: Estimated rainfed wheat map of Andalucia (fractions km^2) based on options 3

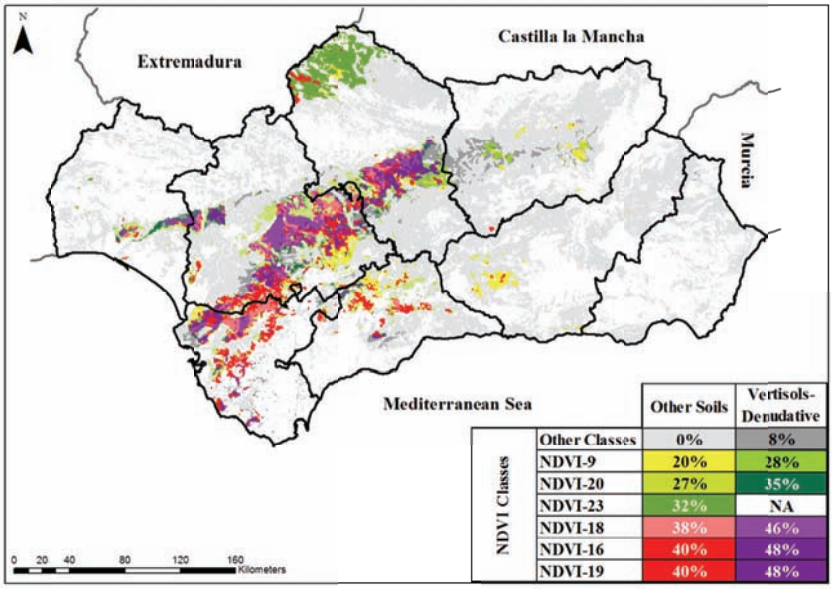


Figure 4- 11: Estimated rainfed wheat (fractions km^2) based on options 4

4.4 Discussion

The present study evaluated the assumption that the NDVI data comprise the combined influences of varying soil, terrain, weather, climate and land use conditions for disaggregation and mapping of crop area statistics. Four modeling options were used to test this assumption with a specific focus on soil information. The results of rainfed wheat areas modeled by municipality (Options 1 and 2) show that almost all variability was explained by the NDVI classes data (Adj. R2 = 98%). In both analyses, the NDVI classes selected to estimate the rainfed wheat area are almost identical. Adding the soil data increased the explanatory value by only 1%. Noteworthy is that Table 4- 2 (results of Option 2) shows that the first step in the multiple regression selected the soil unit as the strongest predictor; its explanatory power of 76%, however, reduced to 1% after NDVI classes entered the equation. The results of the jack-knife analysis (Figure 4- 7), applied to establish the relative importance of each predictor, also indicated that the NDVI data are better predictors of rainfed wheat areas than the soil data. Clearly, the NDVI predictors sufficiently reflected the landscape variability related to differences in soil units, thus rendering the use of the soil map unnecessary. This strongly supports our assumption that NDVI data comprise the combined effects of varying land characteristics.

The rainfed wheat maps produced by using segments data (direct mapping procedure, referred as options 3 and 4) also show that almost all the variability was explained by the NDVI classes data. In options 3 and 4 only 64% and 65% of the variability is explained respectively. The present study shows that NDVI data provided sufficient results for mapping crop areas and the use of soil units marginally increased the explained variability. This makes the mapping exercise simple and fast. In more complicated methods, crop areas statistics (at level 2 of the so-called Nomenclature of Territorial Units for Statistics (NUTS 2) regions) have been downscaled to the level of the homogenous mapping units (HSMU), comprising environmental characteristics (climate, soil properties, land cover, etc.) by statistical regression models (Leip *et al*., 2008).

Validation of the prepared rainfed wheat maps (options 1 and 2) on the basis of the average fractions of rainfed wheat in the segments data revealed that 56% of the variability was explained by NDVI class data, and 61% following the

addition of soil units. Validation also suggested that the official rainfed wheat statistics data are 10-15% higher than the actual situation (Figures 4- 8 and 4- 9). The results of rainfed wheat map produced by using option 3 were aggregated at NUTS level 3 and compared with the officially published crop statistical data (Khan *et al*., 2010). The aggregated results showed a very good agreement with the reported data (R2 = 98 %). Further the aggregated results of option 3 also show a good agreement with the the crop statistical data at municipal level (R2 = 91%). Therefore it may be inferred that the option 3 which uses NDVI data in combination with segments data is the "method of choice". This indicates that the prepared rainfed wheat maps have a substantial degree of generality, as a result of using a 1 km^2 pixel size of the hypertemporal RS images. Each 1 km^2 pixel of produced NDVI based rainfed wheat map is related to a mix of rainfed wheat and other land cover.

As the variability within the NDVI classes is high due to 1 km^2 resolution of SPOT NDVI images at local levels the accuracy of generated wheat map is low. Wheat maps produced by using options 1 and 2 show that each NDVI class related to wheat areas also represent other land covers/land use. Therefore, it can be deducted that the major reason of low variability explanation is that at local level the accuracy of generated wheat maps because of high local heterogeneity in land cover and use. Such land use maps can further be improved by using higher resolution imagery. The other reason is that 1-km^2 resolution NDVI images partially match the 700 m x 700 m segments data.

Validation shows their accuracy (85-90 %, coefficients of figures 4- 8 and 4- 9) to be quite comparable to crop areas estimated by using different approaches. Verbeiren *et al*., 2008 reported accuracy of winter wheat areas estimated by applying linear mixture model (R^2 =0.39) and Neural Networks (R^2 =0.86) on SPOT-Vegetation time series. In another study, the reported Kappa coefficients are 0.69 to 0.74 for individual crop type mapping by using an eco-region stratification approach on MODIS NDVI data (Shao *et al*., 2010). Wardlow and Egbert in 2008 used a hierarchical crop mapping protocol by applying a decision tree classifier to NDVI time series data collected over the growing season and reported classification accuracy of 84% for the summer crop map. The use of our unsupervised classification method to make mapping units proved in comparison very successful, though results of options 1 and 2 depend upon the quality of the used crop statistical data by municipality.

4.5 Conclusions

We conclude that NDVI data are suitable for mapping crop areas in combination with crop statistical data by municipality or with segments data. Options 1 and 2 depend upon the quality of crop statistics data by municipality. Thus selected "best" map which only relies on segment data of rainfed wheat (Figure 4- 10) can be used as an input in crop monitoring methods to improve production assessments. Knowledge on other land cover fractions (non-rainfed wheat) per NDVI class can then be a requirement that can be satisfied, using the same methodology. Such maps can be further improved by using higher spatial resolution hypertemporal images. The explanatory value can be further improved by using the NDVI map in the area frame sampling method. Similar behaving groups of NDVI classes can function as strata for various related land cover/land use types. Though soil data seemed relevant to explain the rainfed wheat area, the use of NDVI class areas rendered their use unnecessary. NDVI alone explained a substantial proportion of the variability in the rainfed wheat areas. This makes the exercise of mapping crop areas faster, easier and accurate. Further, such maps can be regularly updated because of the hypertemporal NDVI images.

5

[*]Comparing a crop growth model driven by remotely sensed data with the European Crop Growth Monitoring System, agricultural statistics and primary field data

[*] This chapter is based on M. R. Khan, V. Venus, C.A.J.M. de Bie, E.M.A. Smaling and H. van Keulen (in preparation). Comparing a crop growth model driven by remotely sensed data with the European Crop Growth Monitoring System, agricultural statistics and primary field data.

Abstract

This study aimed at evaluating the performance of a crop growth model driven by remotely sensed data (C*f*-Water) that estimates actual crop yields at 1-km^2 resolution. Evaluation included (i) comparing the output of C*f*-Water at regional scale (province) with the output of an operational crop growth model, CGMS (Crop Growth Monitoring System), of the European Union's Monitoring Agriculture with Remote Sensing (MARS) program and with published agricultural statistics and (ii) accuracy assessment of the output of C*f*-Water using primary field data. CGMS only reports calculated water-limited and potential crop yields at NUTS-1 scale (group of counties/communities level) and after time-trend adjustments, actual crop yields at NUTS-0 scale (country level). Using the trend adjustment logic, CGMS estimates at NUTS-3 (province) scale were derived for the required comparison. The C*f*-Water model has lower data requirements than CGMS which requires also soil and historical yield data. For 2001, comparison of the estimated actual rainfed wheat production of Andalucía, Spain at NUTS-3 scale (province) by C*f*-Water and the estimated actual rainfed wheat production by CGMS with published agricultural statistics showed for C*f*-Water excellent agreement (R^2 = 98%; RSME = 16 Mg**)** and for CGMS good agreement (R^2 = 67 %; RSME= 41 Mg). The accuracy assessment of C*f*-Water estimates using primary field data comprising 334 segments of 700 x 700 m showed excellent agreement (Adj. R^2 = 98%). We conclude that C*f*-Water has a very high potential to support food security studies. However, before recommending incorporation in an operational system, C*f*-Water needs to be tested for additional years, crops and regions.

5.1 Introduction

Reliable and timely assessment of production of the main food crops is required so that appropriate decisions can be made in time to ensure food security. Food security exists "when all people, at all times, have physical and economic access to sufficient, safe and nutritious food to meet their dietary needs and food preferences for an active and healthy life" (FAO, 1983; World Food Summit, 1996). Simplified, it can be defined as the sufficiency of food available in any given year in any given region for the people living in that region during that time. The number of people living in a region can be determined through a census and other social data sets with a fair degree of accuracy (Matras, 1973; United Nations, 1973; Keyfitz, 1976; Elvidge *et al.*, 1997). Reliable and relevant information on crop production systems then becomes a key factor in decision making and defining strategies related to food security. Timely information on crop production is also important for organizations working in the field of crop insurance, agricultural marketing, pricing and trading.

The first component of an efficient monitoring system for crop production is to properly describe "where-what" is grown, before embarking on the "how much" is harvested question. Thus, first crop area maps are required, followed by maps of crop yield estimates for the crop-producing areas. Georeferenced land use information facilitates meaningful analyses. Crop area maps have been prepared by using hypertemporal SPOT NDVI images (S-10 product) and crop statistical data (De Bie *et al.*, 2010; Khan *et al.*, 2010; Nguyen *et al.*, 2010). Spatially explicit crop statistics help to identify the areas where the various crops are actually grown. Such information supports policy makers and researchers to identify the areas of their interest and makes it easier to devise policies and strategies concerning food security.

The use of remote sensing for crop monitoring started in 1974 with an initial focus on data for crop condition monitoring. The United States Department of Agriculture (USDA), the National Oceanic and Atmospheric Administration (NOAA), the National Aeronautics and Space Administration (NASA) and the United States Department of Commerce (USDC) carried out the "Large Area Crop Inventory Experiment (LACIE)" program (Leamer *et al.*, 1975; Roberto, 1993). From 1980 to 1986, these institutes carried out the "Redirected from Agriculture and Resources Inventory Surveys through Aerospace

(AgRISTARS)" program. Following that program, a global scale operational crop monitoring system was developed in 1986. The system not only established a crop condition assessment and production prediction for many crops (such as wheat, rice, maize, soybean and cotton) in the United States, but also monitored the main food producing countries in the world such as the former USSR, Canada, Mexico, Argentina, Brazil, China, India and Australia (NASA, 1984; Hogg, 1986).

The United Nation's Food and Agriculture Organization (FAO) developed its method to monitor crop conditions at global scale with GIEWS (Global Information and Early Warning System on Food and Agriculture). The system used 10-day composite NOAA-AVHRR NDVI data, which were pre-processed with the WINDISP software (Kileshye Onema and Taigbenu, 2009).

In the late 1960s, following a long period of experimental and empirical work in agricultural research in combination with statistical analysis, the first dynamic simulation models were developed (Van Ittersum *et al*., 2003). Subsequently, in the 1970s the development of crop growth simulation models started (Jame,1992; Bouman *et al*., 1996). Since then, a wide variety of crop models has been developed all over the world to serve many different purposes. DSSAT (Decision Support System for Agrotechnology Transfer) was developed in the USA by IBSNAT (International Benchmark Sites Network for Agrotechnology Transfer) (Jones *et al.,* 2003). APSIM (Agricultural Production system SIMulator) modeling framework was developed by APSRU (Agricultural Production Systems Research Unit) in Australia (Keating *et al.,* 2003; Thorburn *et al*., 2010). In the Netherlands, at Wageningen, the late C. T. de Wit started the work on crop growth modeling at the Department of Theoretical Production Ecology of Wageningen Agricultural University. One of the most widely used ' products' of the 'Wageningen School' (Bouman *et al*., 1996) is the WOFOST (World Food Studies) crop growth simulation model which is the core of the Monitoring Agriculture with Remote Sensing (MARS) program of the European Union implemented by the Joint Research Center (JRC'; Van Diepen *et al*., 1989; Reidsma *et al*., 2009). Other recent models developed in Wageningen are SUCROS (Simple Universal Crop growth Simulator; Van Laar *et al*., 1997) and ORYZA (a crop growth model for rice; Bouman *et al*., 2001). The Crop Growth Monitoring System (CGMS) of JRC, developed in the MARS program, is based on the WOFOST crop growth simulation model (Supit *et al*.,

1994; Boogaard *et al*., 1998). Monitoring results of CGMS have been applied in the Common Agricultural Policy (CAP) of the European Union such as for establishment of agricultural subsidies and verification of farmers' declarations (MacDonald and Hall, 1980). CGMS has been widely applied to estimate crop production at various regional levels (Lăzar *et al*., 2009; Reidsma *et al*., 2009).

Information from remote sensing observations can effectively be integrated into crop modeling methodologies. Such data have been used in crop models for regional yield assessment (Roebeling *et al.*, 2004). An appropriate crop model and careful application of input information derived from satellite-based observations can be highly beneficial in regional crop yield assessments, because satellite-based inputs can relatively be obtained, considering the amount of time and labor that regional level data collection requires (Doraiswamy *et al.* 2005). Remote sensing techniques can be used in calibration and validation procedures through supply of input data for spatial applications of crop growth simulation models (Jongschaap, 2006). Remote sensing can provide instantaneous information on important crop state variables at regional scale as a basis for assessing crop production.

An experimental model (C*f*-Water) that operates on the basis of satellite-derived observations is under development at the Faculty of Geo-Information Science and Earth Observation (ITC), University of Twente. This chapter aims to test the capabilities of C*f*-Water by comparing its production estimates and those of CGMS with published agricultural statistics and primary field data in Andalucía, Spain.

5.2 Material and Methods

5.2.1 Study Area

Andalucía is located between 36º and 38º 44' NL, in the warm-temperate region. It is the southernmost community of mainland Spain, with an area of 87,268 km^2 comprising eight provinces and 770 municipalities. It is the most populous (8,285,692 inhabitants in 2009) of the seventeen autonomous communities of Spain. The Mediterranean climate of Andalucía is characterized by mild rainy winters and hot dry summers. Average annual temperature of Andalucía as a whole is above 16 ºC. Overall annual rainfall is highly variable, with a marked

decreasing gradient of precipitation from west to east, and ranges from a maximum of 2000 mm to a minimum of 170 mm.

About 70% of the land area of Andalucía is utilized for agricultural purposes. Medium-sized mountains dominate the landscape, occupying 42% of the total area. Consequently, 38% of the agricultural land is mountainous with crops generally restricted to the inner valleys or to gently sloping hillsides. The main agricultural system is dryland (rainfed) farming of cereals and sunflower in the vast countryside of the Guadalquivir valley. Geographically focused cultivation of barley and oats takes place in the high plains of Granada and Almeria. The major annual crops are wheat, sunflower, cotton, rice and maize, in addition to olives, grapes and oranges.

The Nomenclature of Units for Territorial Statistics (NUTS) classifies the EU member states according to three spatial scales. Thus, the territory of Spain is classified for statistical purposes as NUTS-1 (groups of autonomous communities), NUTS-2 (individual autonomous communities) and NUTS-3 (provinces of the autonomous communities) as shown in table 5- 1.

Table 5- 1: NUTS regions of Spain relevant to the study area

CODE	LABEL	NUTS LEVEL
ES	Spain	0
ES6	SUR	1
	Subdivision of Sur region	
ES61	Andalucia	2
ES62	Region of Murcia	2
ES63	Autonomous city Ceuta	2
ES64	Autonomous city Melilla	2
	Subdivision of Andalucia	
ES611	Almeria	3
ES612	Cadiz	3
ES613	Cordoba	3
ES614	Granada	3
ES615	Huelva	3
ES616	Jaen	3
ES617	Malaga	3
ES618	Sevilla	3

Following this classification, the southern communities, known as Sur, are combined at NUTS-1 with code ES6. The Sur region is subdivided into four at NUTS-2 level with codes ES61-ES64. The eight provinces of Andalucia at NUTS-3 level are coded from ES611-ES618 (Table 5- 1). In this chapter, the estimates of the models are evaluated at NUTS-3 level (provinces of the community of Andalucia).

5.2.2 Crop Growth Monitoring System (CGMS)

The Crop Growth Monitoring System (CGMS) is used by the European Commission's MARS program. The core of CGMS is the WOFOST crop growth simulation model which is combined with a Geographical Information System (GIS) and a yield prediction routine (Boogaard *et al*., 2002). CGMS comprises three main components (Figure 5- 1):

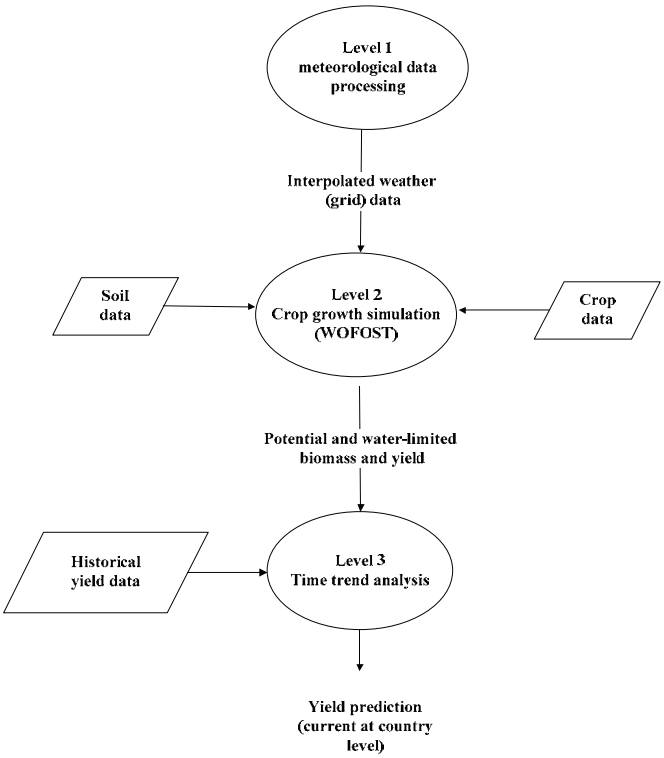


Figure 5- 1: Schematic overview of yield prediction in CGMS
(Modified after: http://supit.net/main.php?q=aXRlbV9pZD02Mg==)

Followings are the three levels as identified in figure 5- 1

1 Interpolation of meteorological data for the whole territory to a square grid of 50 x 50 km.

2 Simulation of crop growth for the whole territory

3 Statistical evaluation of the results (time trend analysis)

The first level in Figure 5- 1 is the weather system in which the weather data are collected, corrected and interpolated to the grid centre. The interpolated data are introduced in WOFOST at the second level, where crop growth simulation takes place. In addition to the interpolated weather data, crop characteristics and soil information are needed as input for WOFOST. Execution of the model results in water-limited yield estimates for the whole grid that are corrected with a time trend analysis by using historical yield data at NUTS-1 scale.

WOFOST simulates phenological development, leaf area development and aboveground dry matter accumulation of annual field crops from emergence (or sowing) to maturity in daily time steps, based on daily weather data, soil properties and crop characteristics. Crop growth rate depends on daily net CO_2 assimilation rate, calculated as a function of intercepted light, which is determined by the level of incoming radiation and the leaf area of the crop. From absorbed radiation and the photosynthetic characteristics of single leaves, the daily rate of potential gross photosynthesis is calculated (Boogaard *et al*., 2002). The assimilates, after subtraction of respiration, are partitioned over the various plant organs, i.e. leaves, roots, stems and storage organs. WOFOST simulates crop production in two production situations (potential and water-limited (Van Ittersum and Rabbinge, 1997). Potential yield of a crop is only dependent on weather (solar radiation and temperature) and crop characteristics (Boogaard *et al*., 1998). Water-limited yield is dependent on weather (solar radiation, temperature, rainfall, humidity and wind speed) and soil physical characteristics.

In the CGMS system, the area of interest is divided into homogeneous units, EMU's (Elementary Mapping Units), to each of which WOFOST is applied. Crop growth in each EMU is simulated based on soil characteristics, grid weather data, crop characteristics and crop calendar. EMU simulation results are

aggregated to NUTS-2 scale, after which the time trend analysis is performed to produce independent forecasts aggregated at NUTS-0 scale for the main crops (De Koning *et al*., 1993).

Input data

a) **Weather data:** Daily grid weather data, i.e. minimum and maximum temperature, wind speed (at 10 m height), vapour pressure, rainfall and global radiation or sunshine hours, are generated through interpolation of daily weather data from weather stations. Additional environmental characteristics (Van Diepen *et al*., 2004) such as daily evaporation from a free water surface (E0), evaporation from wet bare soil (ES0) and evapotranspiration for a reference crop (ET0) (Allen *et al*., 1998) are calculated for each weather station in CGMS. These characteristics are also interpolated during the grid weather generation.

b) **Soil data:** The soil database used in CGMS is the Soil Geographical Data Base of Europe (SGDBE) at scale 1:1,000,000, containing information on maximum rooting depth, crop-specific suitability and water holding capacity.

c) **Crop characteristics:** To characterize the crop (variety)**,** about 40 crop parameters are used by WOFOST in CGMS.

d) **Crop calendar:** Crop calendar data, containing average sowing and harvest dates and the distribution of crops (showing which crop is grown in a particular location) is linked to the grid system.

e) **Historical yield data:** Historical data on planted area, yield and production at NUTS-0 scale, used in the time trend analysis, are obtained from national statistical agencies of the EU member states.

5.2.3 The C*f*-Water model

The C*f*-Water model used in this study (Venus and Rugege, 2004), is adapted from modified algorithms of WOFOST documented by Driessen and Konijn (1992). The model has been programmed to improve the production situation

analysis for regional application, by including satellite-derived parameters to estimate canopy heating[2].

In the crop growth simulation model, C*f*-Water, first the actual gross rate of assimilation is calculated from leaf area and radiation. Gross assimilate production is then partitioned over leaves, stem, root and storage organs as a function of development stage of the crop. Next, maintenance respiration losses are calculated for each plant organ and subtracted from gross assimilate allocations to obtain net assimilate available for growth. The available net assimilates are then multiplied by organ-specific 'conversion efficiency' coefficients to obtain the increments in dry organ masses (Rugege, 2002; Venus and Rugege, 2004).

The C*f*-Water model allows simulation of actual crop production/yield as a function of radiation and temperature, and compounded constraints to crop growth as reflected in the temperature difference between the canopy and ambient temperature. Thus, C*f*-Water takes the following form:

Production/Yield = *f* (radiation, temperature, C_3/C_4[3], canopy heating)

The canopy is heated by incident radiation and part of the absorbed energy is dissipated by transpiration (Barros, 1997; Kalluri and Townshed, 1998). Incoming radiation available for heating the canopy is set equal to net intercepted radiation minus the energy needed for assimilation and for vaporization of water lost in actual transpiration. The instantaneous difference between air temperature and canopy temperature is approximated from the sensible heat component of the energy balance equation. The 'water sufficiency coefficient' (C*f*-Water) is then calculated (Equation 1). C*f*-Water is used to estimate the relative rate of gross assimilation and represents the relative sufficiency of available water Venus and Rugege, 2004).

$$Cf\text{-Water} = TR_{act} / TR_{max} \qquad (1)$$

where:

TR_{act} is actual transpiration rate

[2] Program was written by Mr. Valentijn Venus

[3] C_3/C_4 refers to the photosynthetic mechanism characteristic for the crop simulated

TR_{max} is maximum (potential) transpiration rate

Or

$$Cf\text{ - Water} = \left[\frac{INTER - \frac{[\Delta T * VHEATCAP]}{AERODR}}{LATHEAT * TRO * CFLEAF * TC} \right]$$

where

INTER is net radiation intercepted by the canopy

ΔT	is temperature difference between canopy temperature and air temperature [K]
VHEATCAP	is volumetric heat capacity [J m^{-3} K^{-1}]
AERODR	is aerodynamic resistance to heat transfer [s m^{-1}]
LATHEAT	is latent heat of vaporisation ([2.46 * 10^6 J kg^{-1}])
CFLEAF	is ground cover fraction of the actual canopy [0-1]
TR_0	is potential transpiration rate (Penman) of the canopy [kg m^{-2} s^{-1}]
TC	is 'actual turbulence coefficient'

Hence, actual crop growth rate follows from instantaneous measurements or derivations of canopy and ambient temperatures. On this basis, assimilation is adjusted and actual crop growth rate is calculated. Note that the value of C*f*-Water calculated in this way directly takes into account the combined effects of all yield-limiting and yield-reducing factors (stress due to water scarcity, water logging, nutrient shortages or excesses, pests, diseases, pollutants, etc.). In WOFOST, a serious error propagation component, i.e. the coarse resolution of all soil-based data is included, which is excluded in the C*f*-Water model.

Figure 5- 2 presents an example of the calculated dry matter dynamics of rainfed wheat, sown at 180 kg/ha and germinating on 1st of November in Andalucía, Spain during 2000-01.

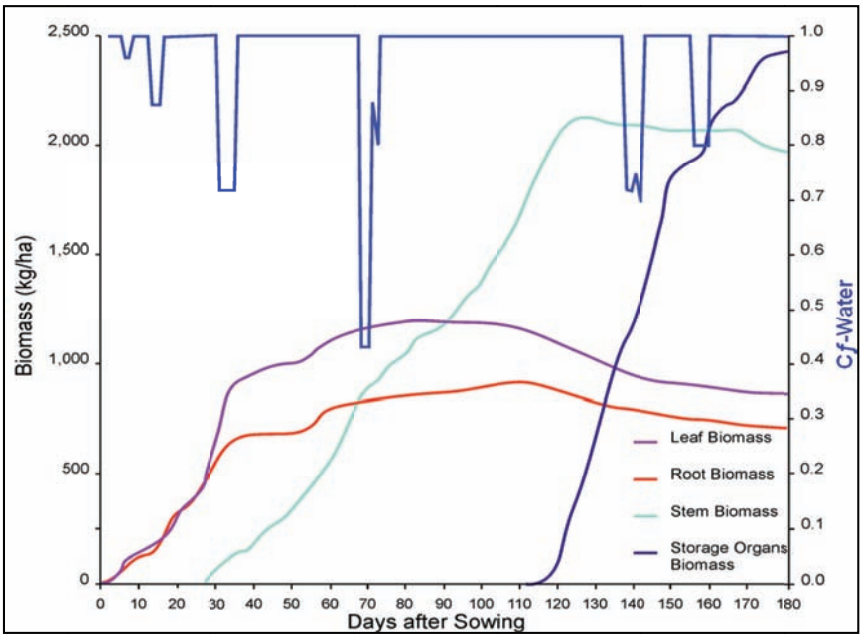


Figure 5- 2: Dynamics of simulated dry matter accumulation of rainfed wheat with the Cf-Water (blue line)

Figure 5- 2 shows that the production of leaves, stems and roots has priority early in the growth cycle. Vegetative growth declines after the mid-season stage, when newly formed assimilates are entirely allocated to the storage organs and maintenance respiration declines, as due to senescence the weight of leaves, stems and roots declines, indicating that crop maturity is approaching. The rate of biomass accumulation decreases when Cf-Water is below 1. During such periods, a variety of yield-limiting and yield-reducing factors (Van Ittersum and Rabbinge, 1997) may affect crop growth.

Input data

a) **Weather data:** Daily maximum and minimum temperature, E0 and ET0 calculated from interpolated weather data obtained from weather stations.

b) **Thermal remote sensing data**: In this study, MODIS Land Surface Temperatures (LST) are used

c) **Crop characteristics:** 25 crop characteristics are used in the model

d) **Management data:** Sowing date, germination date and seeding rate

A brief overview of the similarities and differences between CGMS and the C*f*-Water model is presented in Table 5- 2.

Table 5- 2: Comparison of CGMS and C*f*-Water

CGMS	C*f*-Water
Crop growth simulation is based on daily weather data, soil properties and crop characteristics.	Crop growth simulation is based on daily weather data and crop characteristics.
Soil properties are taken into account.	Soil properties are not taken into account.
Remote sensing data are not used for yield estimation.	Remote sensing data from thermal bands are used for yield estimation.
Evapotranspiration is calculated from interpolated meteorological data from the weather stations	Evapotranspiration is calculated from the surface temperature from remote sensing data and ambient air temperature
Potential and water-limited yields are estimated	Actual yield is estimated
Adjustments are made through time-trend analysis based on historical yield data	Time-trend analysis is not required.
Estimated crop yields (kg/ha) are made available at country level through AGRI4CAST.	Estimated crop yields (kg/ha) are generated at a 1 km^2 resolution grid

5.2.4 Comparison of rainfed wheat production estimates from CGMS and C*f*-Water

Rainfed wheat production estimates from CGMS and C*f*-Water are compared at NUTS-3 level (provinces). Estimated production of rainfed wheat was aggregated to provincial level on the basis of the rainfed wheat area map of Andalucia (Khan *et al.*, 2010; Figure 5- 3), constructed on the basis of hypertemporal SPOT NDVI images (1998-2006) and the rainfed wheat fractions reported in the primary field data (segments).

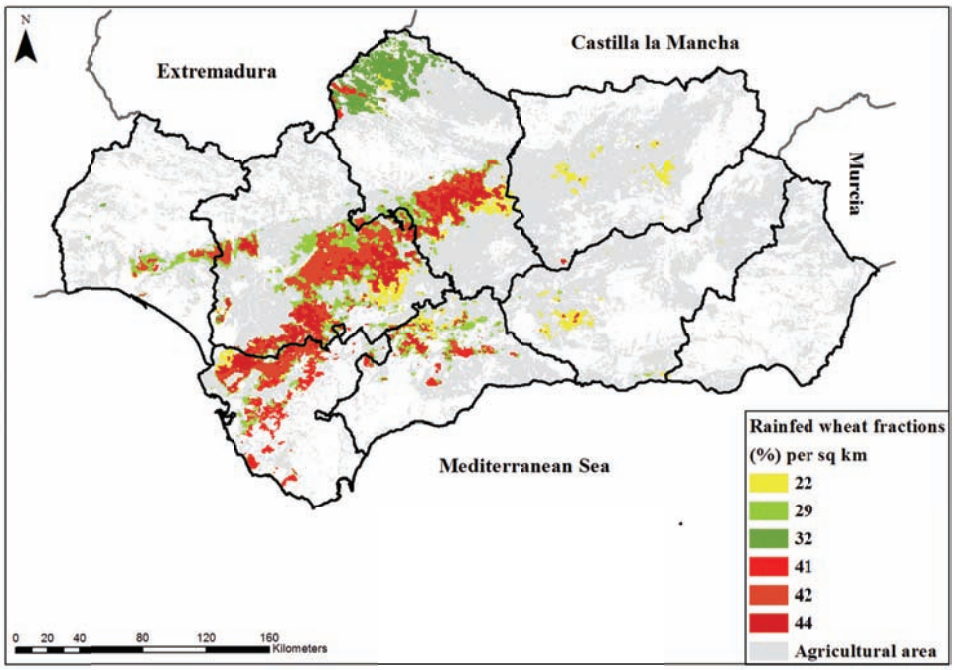


Figure 5- 3: Map of rainfed wheat area of Andalucia based on NDVI and segment data (2001-05) [Source: Khan *et al.*, 2010] with province boundaries (NUTS-3)

Data preparation of actual production estimates by CGMS at NUTS-3 level: Estimated potential and water-limited yields by CGMS for a 5 x 5 km grid (2001) were obtained from the Ministry of Agriculture and Fisheries, Andalucia. For that purpose, CGMS was applied for major crops grown in Andalucia in collaboration with JRC from 1990 till 2007. After up-scaling yield

data per grid to total production at NUTS-3 level, adjustments were introduced through time trend analysis using historical rainfed wheat production data (1990-2000). Through multiple linear regressions a second degree polynomial function was used to adjust for the time trend and to convert water-limited production to actual production, as suggested by Dennett *et al*., 1980; Vossen, 1990, 1992; Palm and Dagnelie, 1993; Supit, 1997 and JRC, MARS STAT 2004.

Wheat yield and production estimates by Cf-Water (2001)

Rainfed wheat yields for Andalucia were simulated for 2001 (Figure 5- 4). Seed rate was set to 180 kg/ha on the basis of farmers' interviews held in 2006-07 and published literature (López-Bellido *et al*., 2005; López-Bellido *et al*., 2006; monthly statistical bulletins available at http://www.mapa.es/es/estadistica/pags/publicaciones/BME/introduccion.htm). Emergence date was set to 1st of November 2000. The model was run only for the areas where rainfed wheat is grown, derived from the rainfed wheat map of Andalucia (khan *et al*., 2010: chapter 4, figure 4-10).

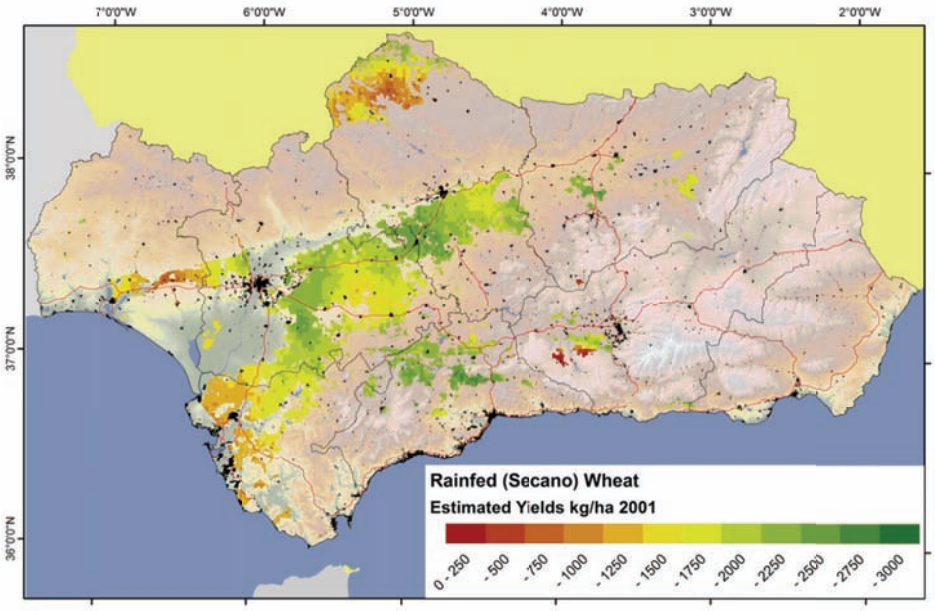


Figure 5- 4: Estimated rainfed wheat yield map of Andalucia for 2001

Estimated production of rainfed wheat for 2001 by C*f*-Water was aggregated to provincial level on the basis of the rainfed wheat area map of Andalucia at NUTS-3 (province level).

Reported agricultural statistical data on crop yields

Annual agricultural statistical data on crop yields were obtained from the Ministry of Agriculture and Fisheries, Andalucia. Each year, the data for various crops are complied and uploaded on the website

(http://www.mapa.es/es/estadistica/pags/anuario/introduccion.htm; see also Table 5- 3). Data for the year 2001 were used to compare the output of CGMS and C*f*-Water at NUTS-3 scale.

Table 5- 3: Rainfed wheat production (Tg[1]) in Andalucia (1990 - 2001)

NUTS LEVEL	Region	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001
NUTS-3	Almeria	11	8	7	7	1	1	3	3	3	1	5	5
	Cadiz	252	331	323	270	144	48	252	157	219	45	213	241
	Cordoba	372	356	273	118	245	83	382	227	218	30	390	362
	Granada	37	34	31	21	12	2	24	18	18	3	17	32
	Huelva	20	68	39	38	60	46	73	74	59	26	52	58
	Jaen	60	51	45	26	31	1	42	25	23	23	30	58
	Malaga	59	70	60	32	28	13	80	26	36	11	61	68
	Sevilla	329	697	453	132	237	100	593	348	447	35	439	525
NUTS-2	Andalucia	1,140	1,615	1,231	644	760	295	1,447	880	1,024	176	1,207	1,347

([1]Tg = Teragram: 10^9; Source: Anuario de Estadísticas Agrarias y Pesqueras de Andalucía, 1990-2001 obtained from the website of the Ministry of Agriculture and Fisheries, Andalucia)

5.2.5 Accuracy assessment of the outputs of C*f*-Water at field level

Estimated rainfed wheat yields of C*f*-Water (Figure 5- 4) for the surveyed segments were compared to the rainfed wheat yields (2001) reported for those segments. Since not all segments were surveyed each year and many segments are only partially covered by agricultural fields, a weighted linear regression analysis was performed, using the total area sampled in each segment as

weighing factor. This procedure leads to assigning a relatively higher weight to segments that were sampled more frequently.

Primary field data

The primary field data for the year 2001, collected by the Ministry of Agriculture and Fisheries, Andalucia were used for accuracy assessment of the output of C*f*-Water. The survey comprised segments of 700 x 700 m for 2001 (Figure 5- 5), coinciding with agricultural areas as defined by the CORINE land cover map (Bossard *et al*., 2000). All agricultural fields in each segment were digitized and data were collected for all agricultural fields per segment through field visits. The total number of segments sampled in the agricultural areas was 334.

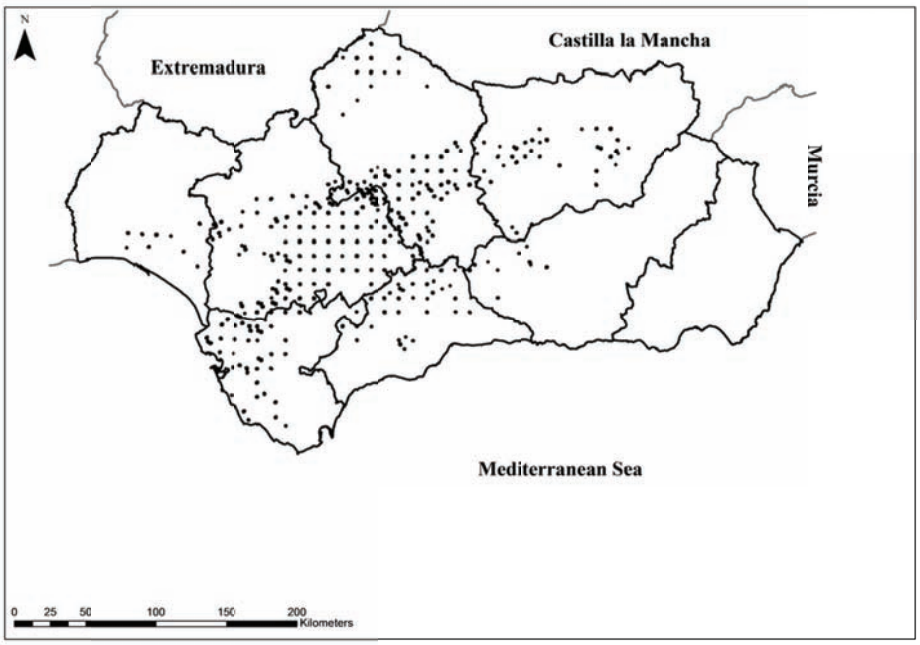


Figure 5- 5: Validation data set consisting of 700 x 700 m segments in the rainfed wheat area of Andalucia, Spain (2001)

5.3 Results

5.3.1 Comparison of CGMS output (after time trend analysis) and results of C*f*-Water at NUTS-3 scale with agricultural statistical data

Figure 5- 6 shows the comparison of rainfed wheat production estimated by CGMS and C*f*-Water for 2001 with the reported official rainfed wheat production data (Table 5- 3). Data from the province of Almeria were omitted from the validation, because of the very low production and the small rainfed wheat area (<300 ha) compared to the other provinces, as shown in the map of rainfed wheat (Khan *et al*., 2010).

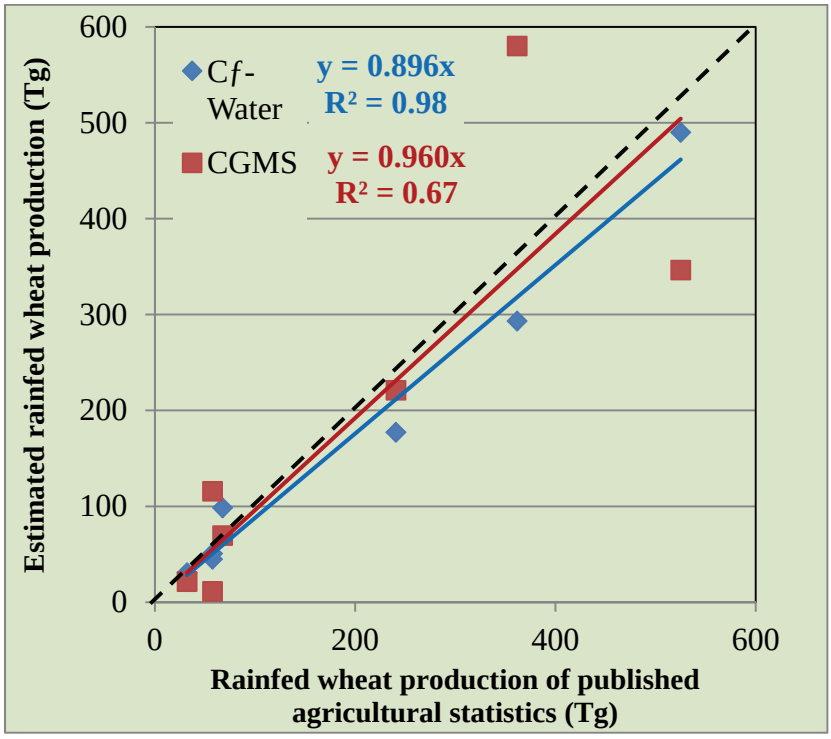


Figure 5- 6: Comparison of estimated yields (kg/ha) with observed yields at NUTS-3 scale

Estimated rainfed wheat production by both, CGMS and C*f*-Water, shows good agreement with the reported values at NUTS-3 scale (R^2 = 0.98 and 0.67 for

CGMS and C*f*-Water, respectively), while the regression lines for both models are close to the 1:1 line. The root mean square error (RMSE) of the production estimates at NUTS-3 scale of C*f*-Water is 16 Mg (9% of the mean) and that of CGMS 41 Mg (21% of the mean).

5.3.2 Accuracy assessment of the results of the C*f*-Water model with primary field data

Comparison of yields estimated with the C*f*-Water model with observed yields at segment scale (Figure 5- 7) shows good agreement, with a regression line close to the 1:1 line.

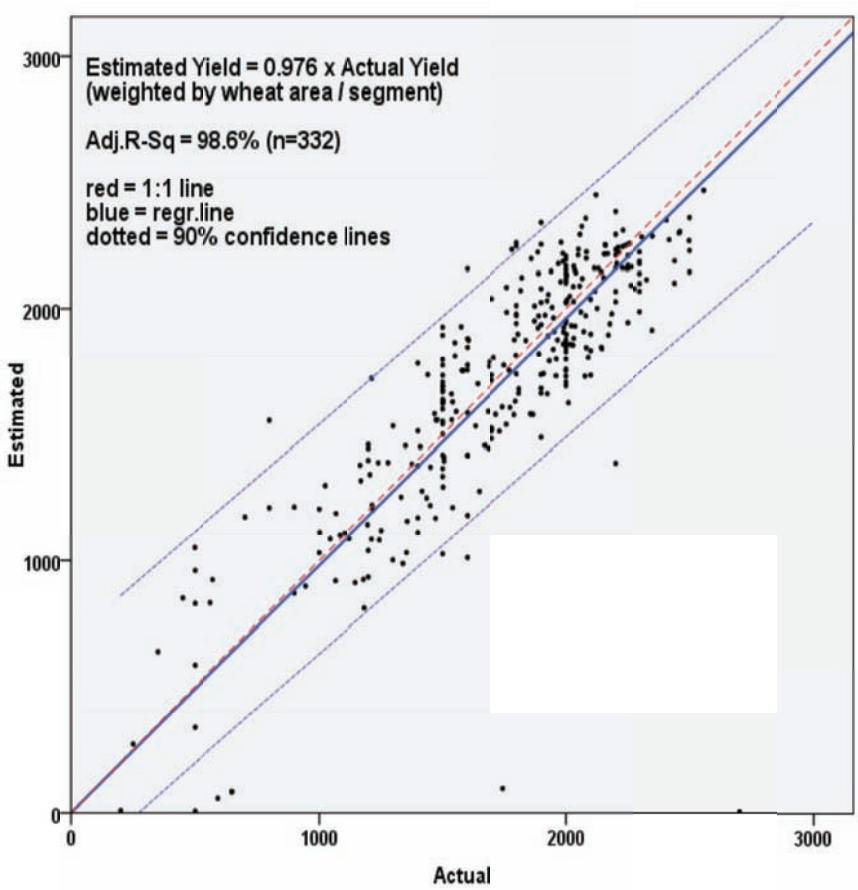


Figure 5- 7: Accuracy assessment of estimated yields (C*f*-Water; kg/ha) on observed yields in segments

Almost all the points are situated within the 90% confidence interval; however, the estimated yields vary 500 kg/ha around the observed yields at segment scale. Overall RSME is 16 kg/ha (1% of the mean)

5.4 Discussion and conclusions

The aim of the study was to evaluate the outputs of a spatial crop growth model (C*f*-Water) that uses remote sensing-based input data. The evaluation was performed at NUTS-3 scale, on the basis of output of the CGMS system and agricultural statistical data and at field scale on the basis of primary field data. At NUTS-3 scale, rainfed wheat production, estimated by the C*f*-Water model showed good agreement with published statistical data, better than that estimated with the CGMS system. A possible reason for this better agreement could be that contrary to the C*f*-Water model, in CGMS soil characteristics are used at a very coarse scale and the outputs are compared at sub-national level. De Wit and van Diepen (2008) compared the output of CGMS executed using remotely sensed input data, with information from EUROSTAT at national scale. They reported R^2-values of 0.93 to 0.97 for Spain, i.e., at NUTS-0 scale.

Agreement between estimated yield by C*f*-Water and CGMS is more close than that reported by Mo *et al*. (2005) who used a remotely sensed input data in a crop growth model (SVAT) and reported an R^2 of 0.57 between simulation results and statistical data. Our results are also better than those of Bai *et al*. (2010) who coupled satellite-derived solar radiation with temperature data from weather stations in a crop growth model to estimate the yield potential of maize and reported R^2-values of 0.62 to 0.90 in comparison with experimental data.

Comparison of results of C*f*-Water at field level with observed yields reported in segment data (R^2=0.98) showed close agreement. These results are more favorable than those reported by Duchemin *et al*. (2008). They used a crop growth model in combination with remotely sensed data for monitoring of wheat production and reported a correlation coefficient of 0.69 with experimental data at field level.

Our results suggest that the C*f*-Water model can be applied at regional scale (provincial or municipal) to estimate the production of crops with lower data requirements than CGMS, without a correction through time trend adjustments.

On the basis of comparison of the results of the C*f*-Water model with the results of CGMS, reported agricultural statistical data and field data, we conclude that the model can potentially be used in food security studies. In principle, regional production of food crops can be calculated by combining crop yield maps (Figure 5- 4) with a crop area map (Figure 5- 3) and an administrative map. Currently, the arable areas or other land cover maps are used to aggregate the results of crop growth models at administrative levels. De Wit and van Diepen (2008) stressed the importance of updated crop area maps instead of using only the arable areas as reported in regional statistical data sets such as EUROSTAT. However, the model has only been applied in a single case study, thus before incorporating the model into an operational systems; more extensive testing of the model is required for different crops, years and regions.

6

Users' perspective on available land use data and the generated outputs; a top-down valorisation approach

Abstract

In this chapter, the users' opinion regarding available land use datasets and the generated outputs of previous chapters of the thesis (chapters 3, 4 and 5) is incorporated. This was done by requesting the researchers working on agricultural land use mapping and monitoring in Andalucia, Spain to provide their opinion through an online questionnaire. The respondents were asked about the availability, updating frequency, documentation (meta data), spatial level of detail and usefulness of the available datasets and the generated outputs of chapters 3-5 (crop area and crop yield estimation). Thirty two researchers were requested to respond out of which 21 responded with a response rate of 66%. The respondents were having different research interests such as land use mapping (47%), land use monitoring (33%), policy related work in regional organizations (10%) and management of agricultural companies (10%). The responses obtained were analysed to appraise the problems faced by the respondents regarding the available land use datasets. The respondents were already using land use maps (50% of the respondents), specific crop maps (12.5%), agricultural statistical data (37.5%), CORINE land cover map (62.5%) and the primary survey data (12.5%). Analysis of the responses showed that the respondents were not satisfied about the availability, updating frequency and spatial level of detail of the land use data. Further, 72% of the respondents were satisfied with the spatial level of detail, the type of land use data and updating frequency of the generated rainfed wheat map. However, 28% of the respondents preferred to have more spatial level of detail and showed interest in having high resolution specific crop maps. Respondents had reservations about the quality and accuracy assessments of the available yield data owing to the use of different methods to collect these data. Responses about the generated rainfed wheat yield map revealed that 70% of the respondents were satisfied with the spatial level of detail in the yield map but they preferred to have such information prior to harvesting, whereas, our method provides yield estimation after harvesting of the crops.

6.1 Introduction

Timely and accurate information on areas of crops and estimates of their production is needed for many purposes such as food security, land use planning, crop insurance, marketing of agricultural commodities and agricultural research. Trade organizations are interested in information on food production in various regions to decide 'what to import' from 'where' and vice versa. Timely information on agricultural land use in terms of both the area and the production is also required for price control and management of agricultural markets. Without such reliable information, problems of disruptions in food supply, artificial price hikes, and problems in implementation of policies are expected to occur. Land use decisions aim at better management of natural resources and environment. It is necessary to evaluate the outcome of such decisions that whether these are responsive to needs both at national and international levels (Young, 1998; Dore *et al*., 2001; Schlamadinger *et al.*, 2007) For such evaluation, up to date and accurate land use information is required.

Regional to global land use databases constructed from the data collected by different countries face technical problems such as inconsistent use of definitions of land use, lack of harmonization and differences in the methods used for inventorying (George and Nachtergaele, 2002). Also, there is a paucity of regional to global land use datasets. In many cases the available information is incomplete or unreliable (FAO, 2005). Moreover, the quality of available information is quite variable and often presents a confused mixture of land use and land cover categories (De Bie, 2000).

The available land use information often lacks the required level of user needs. This limitation is due to the conflicting purposes of land use data collection, differences in methods used to collect land used data and limitations in these methods (chapter 1). Further, the available land use information often lacks the required level of accuracy and is incompatible with other data sets. Therefore, the assessment of available land use information is an essential requirement for sustainable management of natural resources, food security and related studies (FAO, 2002; FAO, 2004).

International organizations such as the Food and Agriculture Organization (FAO), the European Union (EU), the United Nations Environment Programme (UNEP) and the United States Department of Agriculture (USDA) are now providing assistance to countries to build updated and accurate land use databases (Cohen and Shoshany, 2002; Townshend *et al*., 2008). FAO supports developing countries in enhancing, processing and distribution of local agricultural land use data.

The scarcity of data, substandard data quality at all scales and the lack of common data exchange formats and protocols occur everywhere with an exception of few developed countries (Sombroek and Antoine, 1994; Lepers *et al*. 2005; Ramankutty *et al*. 2007). The problem becomes more severe owing to poor communication between data producers/suppliers, information technology and users. Stakeholders report that the effective use of GIS technology is constrained by the limited adequacy of data on land use systems (Dalal-Clayton and Dent 1993; Zeijl-Rozema *et al*. 1997; FAO and UNEP, 2002; Dietz, 2003). The problems identified by users of land use data were reported by De Bie, (2000) which are summarized in table 6- 1. The constraints were recorded at selected (sub-) national institutes in a number of developing countries and in four European countries.

To overcome these limitations, effective coordination is needed among potential data users of global land use data to identify, define and harmonize their data needs. It also asks for consistent land use classification systems at national and international level.

In the present study, available land use information is assessed and information on desired characteristics of land use data is collected by conducting a survey through an online questionnaire. The questionnaire was presented to thirty two researchers working in the fields of agricultural mapping and monitoring in the Ministry of Agriculture and Fisheries, Andalucia to ask their opinion about the existing Land use/ land cover (LULC) information and their desired data qualities. The questionnaire comprised of both closed and open questions to obtain the user's judgment about LULC data sets.

Table 6- 1: Constraints associated with effective use of land use system information as reported by stakeholders (Modified after De Bie, 2000)

Data Aspect	Problem	Observation
Availability	Not accessible	Often occurs
	Limited	Regularly occurs
	Restricted	Often occurs
Format	Inconsistent	Often occurs
	Data integrity problems	Often occurs
	Units used vary from region to region	Regularly occurs
Quality	Lack of uniformity	Often occurs
	No accuracy assessment	Regularly occurs
Documentation	Not available	Sometimes occurs
	Incomplete	Sometimes occurs
	Poor	Often occurs
Geo-referencing	Absent	Sometimes occurs
Cost	Expensive	Regularly occurs
Update	Poor update frequency	Regularly occurs
Coordination	Users are not involved in the surveys	Often occurs
	Poor coordination between organizations with the mandate of producing such information	Often occurs

6.2 Methods

Experts/user consultation was performed to solicit (i) the opinion of respondents about the already available land use data and (ii) what is their opinion about the generated crop area maps and crop yield estimates (outputs of chapter 3, 4 and 5 of this thesis). The respondents were also asked to explicitly mention possible reasons of disagreements with the provided outputs of the thesis in order to incorporate the suggestions/ requirements of the users for future recommendations.

6.2.1 Study sample and procedure

Data were collected using a population of thirty two researchers working in the field of agricultural mapping and monitoring in the Ministry of agriculture and fisheries, Andalucia. Before developing the questionnaire two meetings were held in the Ministry of agriculture and fisheries, Andalucia in order to understand the terms and definitions used. This was done to ensure that the respondents understand the statements of the questionnaire.

6.2.2 Available land use data sets

The following land use data sets are used by the researchers which can be grouped into following five categories based on the land use information presented by these land use data sets.

Agricultural Land use maps include those maps which contain the information about general land cover/land use classes and do not differentiate various crops grown in Andalucia.

- Land Use map (Usos del Suelo), at 1/1,000,000 scale. It includes classes such as forests, arable land mixed with grass land, arable land, pastures etc.

- Map of vegetation land cover/land use in Andalusia in 1999 (mapa de usos y coberturas vegetales de Andalucía, 1999), at 1/1,50,000 scale. The map of vegetation land cover/land use in Andalusia was updated in 2007

  (Source: Consejería de Medio Ambiente. Junta de Andalucía).

- Map of irrigated areas (Mapa de las zonas de regadío). (Source: http://www.juntadeandalucia.es/agriculturaypesca/sigregadios/servlet/regadios)

- Map of ecological assessment of natural resources (Mapa de evaluación ecológica de los recursos naturales), at 1:1,000,000. (1987).

Specific crop type maps are those maps which present information about the spatial distribution of specific crops. Such maps are prepared to meet the requirement of European commission. These maps include:

- Map of CAP crops declared, updated each year at a scale of 1/50,000 by using the data obtained for subsidy claims of the framers and LPIS cartography (SIGPAC; http://sigpac.mapa.es/fega/visor/).

- Other thematic maps of agronomic land use such as rice area map, olive area maps, other field crop maps and fruit trees maps such as olive trees maps (SIG-Oleícola), vineyard maps (SIG-Vitivinícola), citrus map (SIG-Citrícola) for the years 1980, 1987 and 2000.

CORINE land cover map is the Coordination of Information on the Environment of the European Environmental Agency (EEA)) land cover map of 2000. The CORINE land cover 2000 (CLC2000) was produced by photo-interpretation of Landsat ETM+ images. The scale of the CLC map is 1:100,000. The researchers in Andalucia, Spain use it as baseline information, i.e., to find locations of major agricultural areas and statistical sampling schemes for crop area estimations (Gallego and Bamps, 2008).

Primary segments data are plot-specific crop data, collected annually (twice) by the Ministry of Agriculture and Fisheries. The survey comprises of randomly selected segments of 700 m × 700 m. The data are collected by visits to all the fields per segment. The spatial distribution of the surveyed segments is determined by dividing the whole territory in blocks of 10 km × 10 km. Each block was sub-divided into 100 cells of 1 km^2. In each block, three cells were randomly selected for surveying. In each cell, cropped area, production of each crop per unit area, irrigation scheme, etc. of agricultural areas present in the segment of 700 m × 700 m is recorded (Figure 6- 1).

Agricultural Statistics Data are the annual statistical data available in tabular format, comprising the information on cultivated area and production of crops at provincial level (available online) and at municipal level (available in the department of statistics, Ministry of agriculture and fisheries, Andalucía).

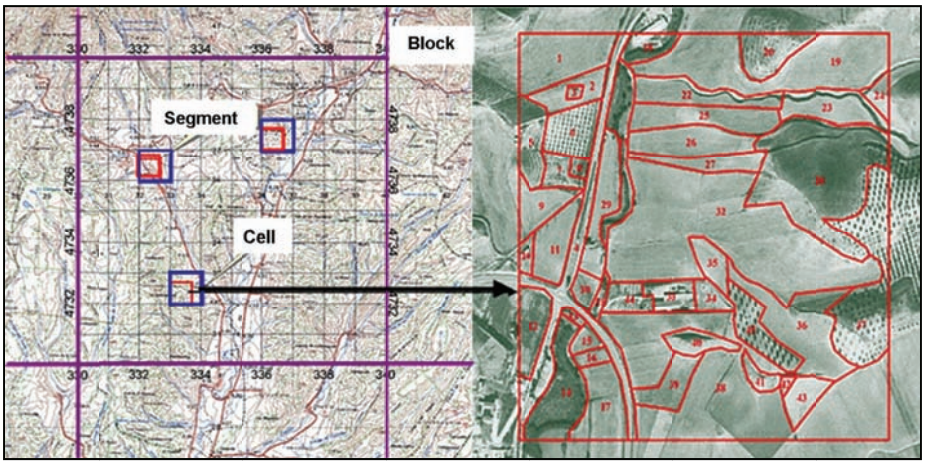


Figure 6- 1: Diagram showing area frame (segment) on the topographic map

6.2.3 Provided land use information (generated outputs of this thesis)

The generated outputs of this thesis were also provided to the users in order to know their opinion. These outputs were:

***Rainfed wheat area map*:** This map was produced by using mapping technique developed by the authors (chapter 2-4). The methods developed produce land use data on their spatial and temporal dynamics. The use of hypertemporal imagery (long term regularly acquired) provides options for location specific monitoring of crops (Khan *et al*., 2010). Further, the methods developed in (chapters 2-4) also provide the options to detect land use change over time (De Bie *et al*., 2010). The details on the data used and methods are included the appendix-1 (Q 29).

Estimated rainfed wheat yield map: This map was produced by using a remotely sensed data driven crop growth model (output of chapter 5). The crop growth model used has fewer data requirements (Appendix 1, Q 44). The model can be run at any regional level (provincial or municipal) to estimate the production of crops.

A brief description of the generated rainfed wheat yield map is included the appendix 1 (Q 45).

6.2.4 Measures

Five data sources as described in section 2.2 were provided in the questionnaire (Appendix 1 question 5). In order to obtain the information on all the data sources which are being used, the respondents were also provided option to indicate other land use data available to them.

All items in the questionnaire regarding the characteristics of different land use data used a 5 point scale with anchors of 1= strongly disagree and 5 = strongly agree.

6.2.5 Statistical analysis

Responses were analyzed using quantitative analysis of the close ended questions (descriptive statistics and graphical representation of data). Further, qualitative data obtained from responses of open ended questions were also summarized and presented.

6.3 Results and discussion

Twenty-one out of a total of 32 questionnaires were returned, giving a response rate of 66%. The respondents were having different research interests such as land use mapping (47%), land use monitoring (33%), policy related work in regional organizations (10%) and management of agricultural companies (10%). The respondents having their research interest in land use mapping, monitoring and policy related work in regional organizations working under the Ministry of Agriculture and Fisheries, Andalucía. Their main activities are farm management, research and knowledge transfer, systems and information technology, analysis of food production, rural development and international cooperation. For their work, they use various land use data as described in section 2.2. The respondents belonging to management of agricultural companies were the managers of agricultural cooperatives. They also use information about cropped areas and production for decision making.

The respondent’s opinions on the available land use data and the outputs of chapters 3-5 of the thesis is given below.

6.3.1 Land use data available to the respondents

The respondents identified 5 major types of land use data available to them from different sources (Figure 6- 2). None of the users mentioned other land use data available to them.

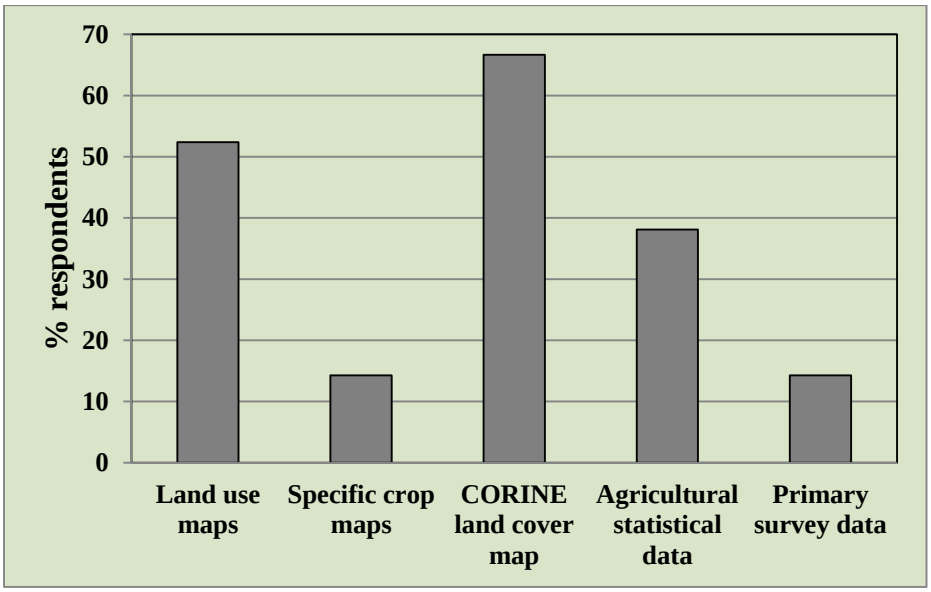


Figure 6- 2: Land use data already used by the respondents

6.3.2 Opinion of the respondents about available land use data

Land use maps

A majority of respondents using the land use maps were not satisfied with the update frequency, spatial level of detail and confused mixture of land use classes present in these maps (Figure 6- 3). However, documentation was found to be sufficient as per requirements of the respondents.

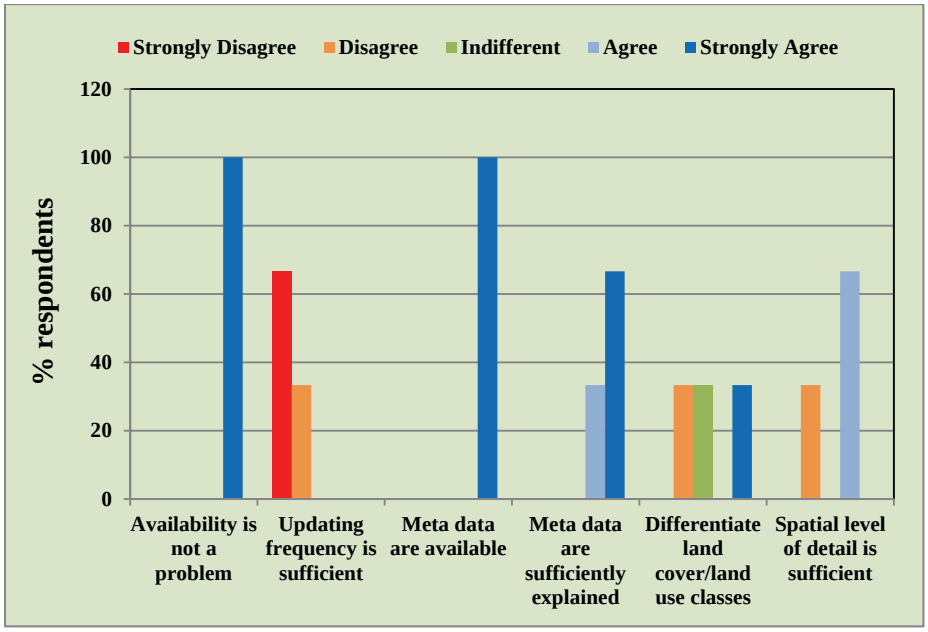


Figure 6- 3: Opinion of the respondents about available land use maps

Specific crop maps

Contrary to land use maps, a majority of respondents were less happy about the availability and update frequency when it comes to specific crop maps. However, they were satisfied with the land use presented in the legend and the spatial level of detail (map resolution) in these specific crop maps (Figure 6- 4).

Agricultural statistical data

As to agricultural statistical data, a majority of respondents had reservations about the spatial level of detail. However, most respondents were satisfied with the availability, update frequency and documentation of these data (Figure 6- 5). The respondents preferred to have these data in the forms of maps rather than just in tabular format.

CORINE land cover maps

A majority of the respondents was satisfied with the availability and documentation of the CORINE land cover map.

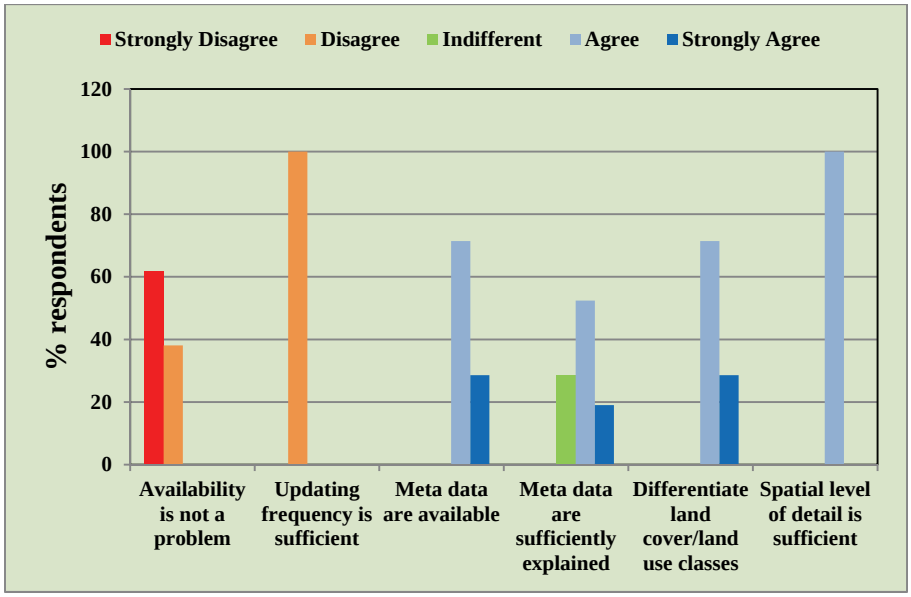


Figure 6- 4: Opinion of the respondents about available specific crop maps

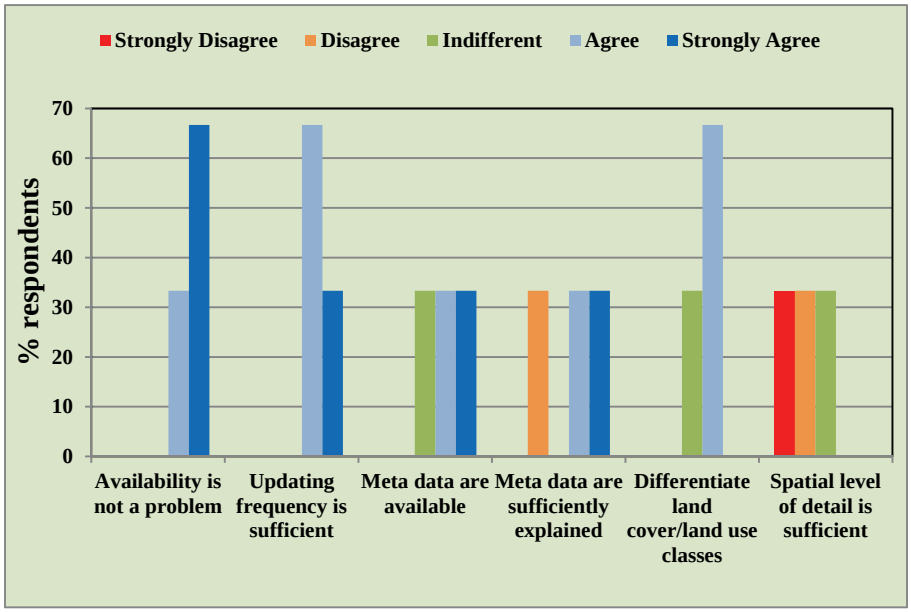


Figure 6- 5: Opinion of the respondents about available agricultural statistical data

The users of CORINE were less happy about the updating frequency and the confused mixture of land use classes and the spatial level of detail in the CORINE land cover map (Figure 6- 6).

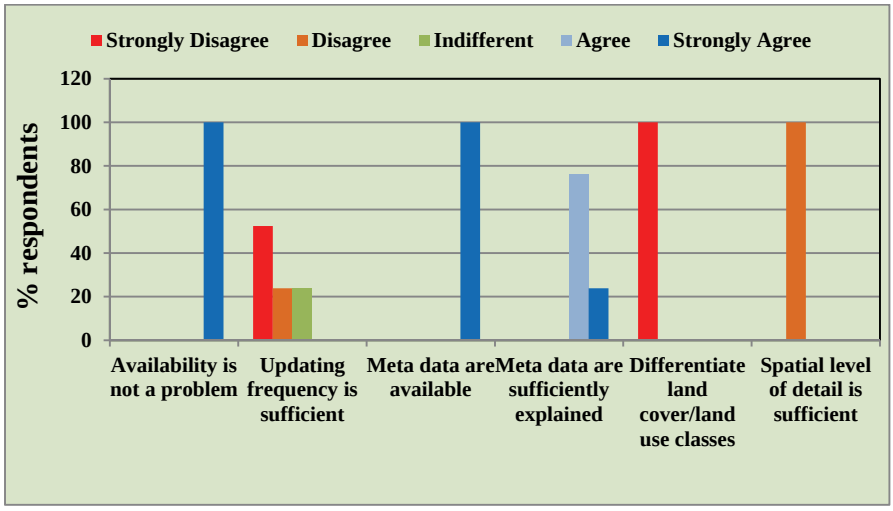


Figure 6- 6: Opinion of the respondents about CORINE land cover map

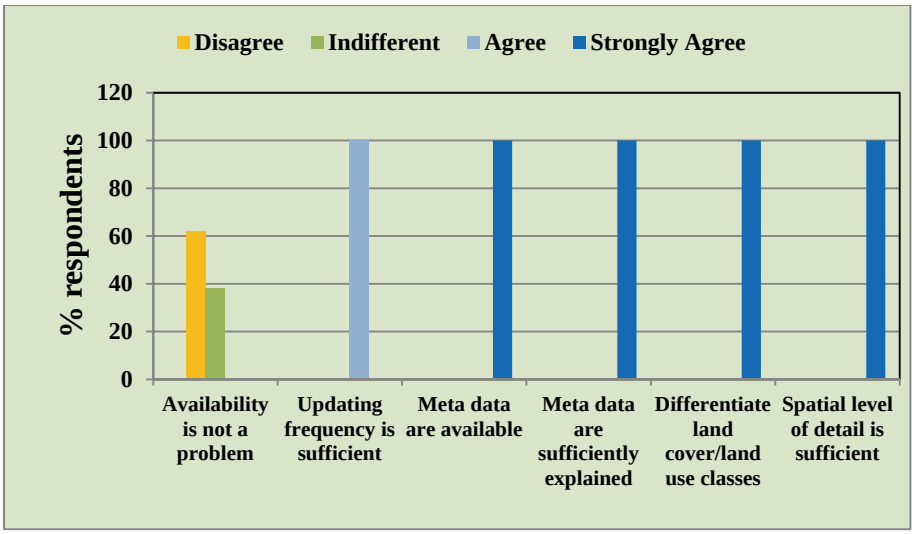


Figure 6- 7: Opinion of the respondents about the primary segments data

Primary segments data

The users of primary segments were not happy with the availability of the primary segments data. However, they were satisfied with the rest of the data aspects of primary segments data (Figure 6- 7).

6.3.2 Opinion of the respondents about the generated rainfed wheat map (output of chapter 3 and 4)

The respondents, after evaluating the provided rainfed wheat map, were satisfied with the data aspects of the outputs of chapter 3 and 4. Seventy five percent of the respondents were comfortable with the spatial level of detail, data type requirements and update frequency of the generated wheat map. They also showed their keen interest in using the developed method to produce the specific crop maps for their own areas of interest (Figure 6- 8).

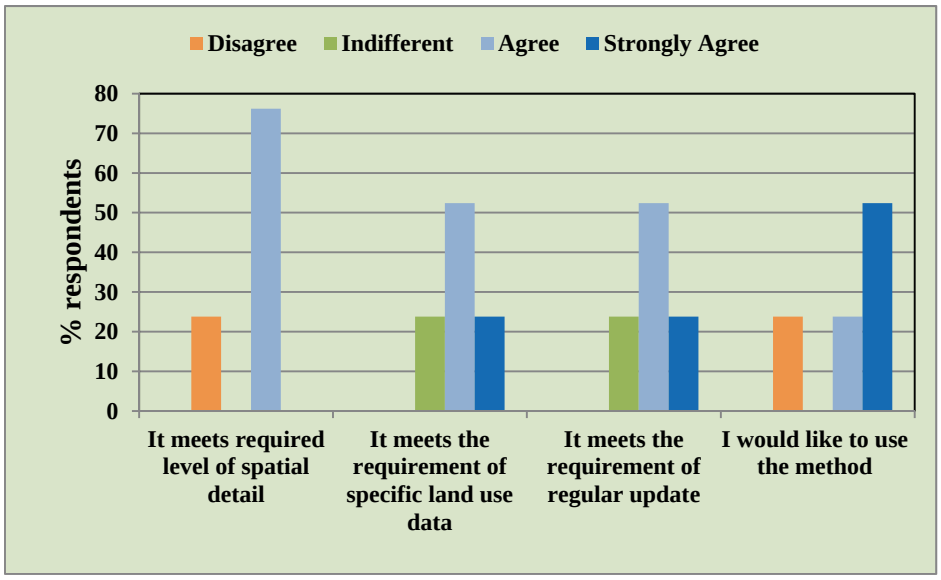


Figure 6- 8: Opinion of the respondents about the generated rainfed wheat map

While expressing the opinion on the generated rainfed wheat map (open ended part of question 30; see appendix1), the respondents showed their interest in having high resolution maps (500 m or 250 m). The respondents were highly satisfied with the update frequency of the generated rainfed wheat. Further, the

available land use data (section 2.2) are is not updated regularly and presents mixtures of agricultural land use classes. The high resolution crop maps (CAP crops) are produced by using the data obtained from the subsidy claims of farmers are not very reliable because of substantial over reporting (40%) in case of sunflower area (Khan *et al*., 2010). On the other hand, the rainfed wheat map produced can be easily updated because of using hypertemporal remote sensing inputs (NDVI). The spatial resolution of the rainfed wheat map can also be improved by using higher resolution imagery which is available from the Moderate Resolution Imaging Spectroradiometer (MODIS).

6.3.3 Opinion of the respondents about available yield data

Amongst the respondents who were using the yield data, only 29% were satisfied with the availability, 62% were satisfied with the update frequency, 55 % reported the availability of documentation (Figure 6- 9).

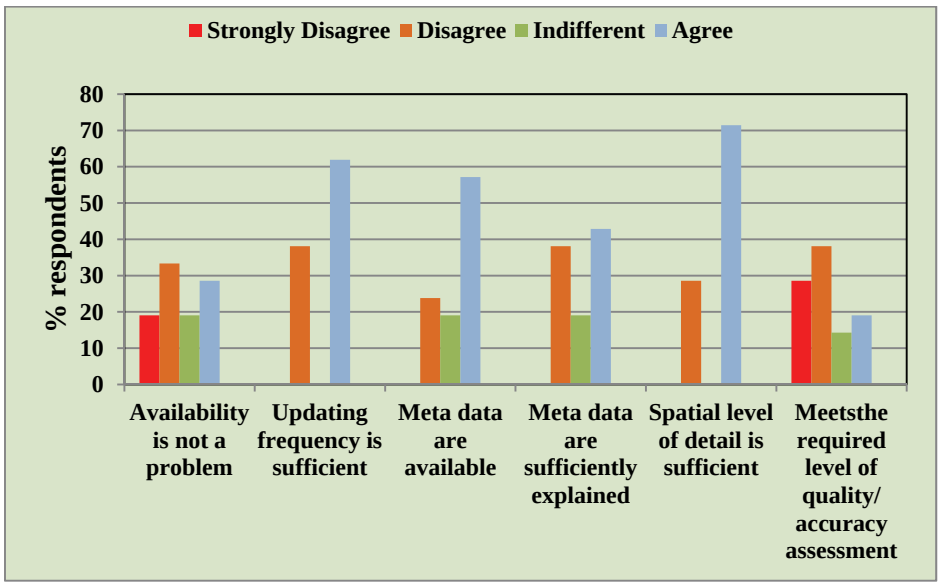


Figure 6- 9: Opinion of the respondents about the available yield data

The yield data available to the majority of the users were obtained after the survey results and the respondents showed the reservations because of using different methods such as farmers reporting, counting and weighing fruits, estimation by an expert (as a rough guess) and machine harvests etc.

6.3.4 Opinion of the respondents about the generated yield map of rainfed wheat (output of chapter 5)

The respondents after evaluating the provided rainfed wheat map were satisfied with the data aspects of the outputs of chapter 3 and 4. 76% of the respondents was comfortable with the spatial level of detail, accuracy assessment and quantitative yield estimation. 52% showed their keen interest in using the developed method to produce the specific crop maps for their own areas of interest (Figure 6- 10). Whereas, 25 % were of the opinion they would like to see the accuracy assessment after calibration as per authors' suggestion (chapter 5).

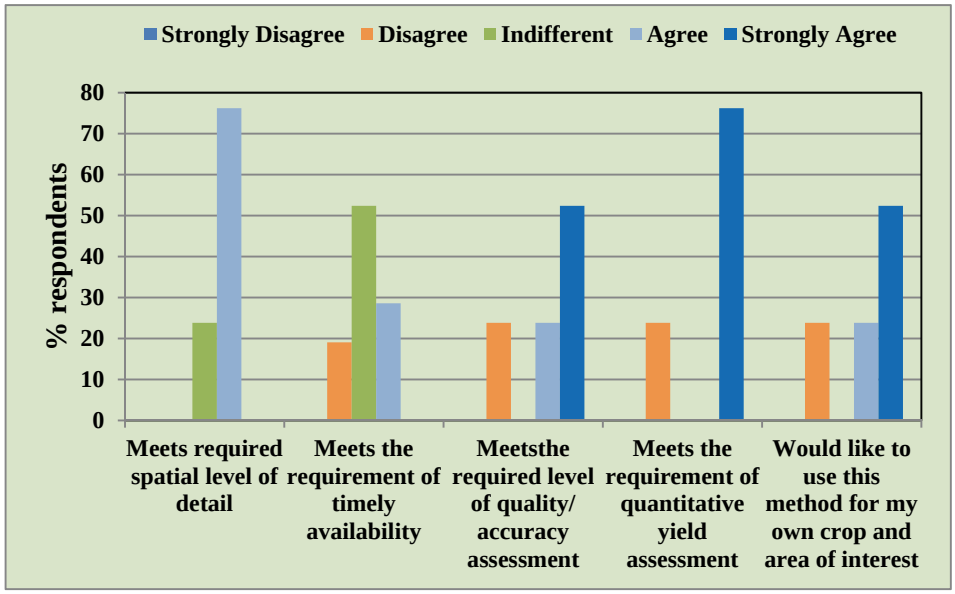


Figure 6- 10: Opinion of the respondents about the generated rainfed wheat yield map

While responding to the open ended part of question 46 (see appendix 1), the respondents appreciated the accuracy assessment and the presentation of quantitative yields as mapped information. However, the respondents showed their concern for detailed calibration of the model used in the research. Further, it was also expressed that the yield estimates are required before harvesting of the crops.

6.4 Conclusions

It is concluded that:

- Amongst the problems identified in the data aspects of available land use data ‘easy availability’ (70 % of respondents) and ‘update frequency’ (70 % of respondents) are the most common. Further, thematic agricultural land use data (specific crop maps) are highly demanded by the respondents, of which 68% declared not to be satisfied with the capability of available land use data to properly differentiate land use/land cover classes.

- Geo-referenced and high resolution maps are desired characteristics of the land use data as expressed by the respondents.

- Remote sensing techniques should be used in combination with field observations to enhance the update frequency and availability of land use data.

- The products of chapter 3-5 are of interest to users. They were keen to use the methods developed for crop area maps (75%) and crop yield maps (70%). The generated land use data has the potential of further improvement by using high resolution imagery e.g. MODIS 500 m and 250 m images and further testing the method of yield estimation for other crops.

- The method developed (output of chapters 2, 3 and 4) can help the researchers to improve the statistical sampling method in Andalucia in particular and in Europe in general, by including the NDVI based strata in the stratification. Such maps can also be used as input for crop monitoring, crop yield estimation, agricultural land use planning and crop insurance purposes. Further, the estimated yield map (output of chapter 5) has an excellent potential for the researchers and policy makers working in the field of crop monitoring. The method developed in chapter 5 can be used to estimate crop yield and production at any administrative level, i.e., municipal or provincial.

- The major limitation of this research is small number of respondents (21).

Acknowledgement

We are highly indebted to Mr. Alfredo Stoelo of the Junta de Andalucía for his support in obtaining the responses of the online questionnaire and also for facilitating the personal meeting with the experts/users of land use data.

7

General Discussion

7.1 Introduction

The overall aim of the research described in this thesis is to extend and improve the existing toolbox to describe agricultural land use. Currently, this is largely made up of maps with scales or legends that are inappropriate for crop-level assessments. Annual updates are not common either, forcing staff to rely on the land use of one particular year for too long. Yield assessments are often also restricted to predetermined sampling frames, without clear tools to scale up the results. The final products of such exercises are often tabular data sets for administrative units, providing area and production totals, but without any notion of the spatio-temporal explicitness of agricultural land use in that unit, and a general lack of insight into data accuracy. The research in this thesis follows a systems approach to map and monitor agricultural land use using a combination of remote sensing, GIS, agricultural statistical data and crop modelling.

The research had two focal points. In three chapters (2-4), the results are described of efforts to make hypertemporal NDVI imagery a useful instrument for collecting agricultural land use data, improving both the spatial and temporal explicitness. Case study areas include Spain and India. Also, ancillary environmental variables such as soil types and geomorphology, were tested on their contribution to explain the spatial distribution of agricultural areas. The second focus was on comparing the outputs of a new crop growth model (C*f*-Water) that also makes extensive use of remote sensing, with the well-known Crop Growth Monitoring System (CGMS) of the European Union's Monitoring Agriculture with Remote Sensing (MARS) project. The aim was to investigate whether improvements as to the predictive value of crop growth models were possible, at different spatial scales. Finally, a qualitative validation was done by interviewing professionals working with agricultural land use maps and data bases, also investigating to what extent they thought they could benefit from the research products developed in this thesis.

This chapter firstly synthesizes the main findings of the aforementioned research topics along with the practical relevance of these findings. Secondly, the limitations of the research are described and lastly, recommendations for future research have been put-forth.

7.2 Agricultural land use mapping

In chapters 2, 3 and 4, SPOT4 and SPOT5 Vegetation (VGT) Sensor's 10-day composite hypertemporal NDVI-images (S10 product) at 1 km^2 resolution were used to produce agricultural land use information in the form of crop calendar and crop area maps. In chapter 2, hypertemporal NDVI images of Nizamabad District, India, were linked with the existing land cover maps. The latter lacked precise location of crops, and was based on an image of the 1994-5 season. After linking the classified map of the stack of hypertemporal images with the existing land cover map and then with the crop statistical and crop calendar data, we could appropriately define 'what' is grown 'where' and 'how much' is grown 'there'. In chapter 3, specific crop area maps were generated by using the classified hypertemporal images of Andalucia, Spain in combination with the crop statistical and primary field data. The NDVI data explained 77% - 98% of the variability. The estimated crop maps in chapter 3 also showed good agreement with the primary field data. In chapter 4, soil and geomorphology data were included in the exercise of mapping crop areas to find out whether the NDVI data reflect their influence. The contribution turned out to be minimal, showing that NDVI alone reflects almost all the crop variability. The results of series of unsupervised classifications of the SPOT images helped to stratify the study area into an optimum number of classes on the basis of minimum and average divergence. 'Visual supervised grouping' of the annual averaged classes produced by the unsupervised ISODATA algorithm was performed to repair possible deficits in classification that were due to the absence of a very clear single peak in the divergence values.

This method shows promise for application where similar kinds of uncertainties can arise due to, for example, missing satellite data as a result of cloud cover. Supervised grouping after unsupervised classification helps to eliminate the annual variations within the original classes without any loss of important detail. The method developed resulted in a map that represents the period of NDVI images (1998-2002 in case of Nizamabad, India as shown in chapter 2). The classification procedure adopted has the capability to capture the changes in land use. The temporal behavior of NDVI class profiles which are relevant to any land use/land cover depicts the changes on ground. Figure 7- 1 shows NDVI profile 19 along with reported production over the rainfed wheat growth period for 2003-2006. The 2005 was a bad year as reported by MARS bulletin

2005 and markedly lower peaks of NDVI profile 19 for 2005 confirms that the method of classification is capable of capturing the land cover features of such a year.

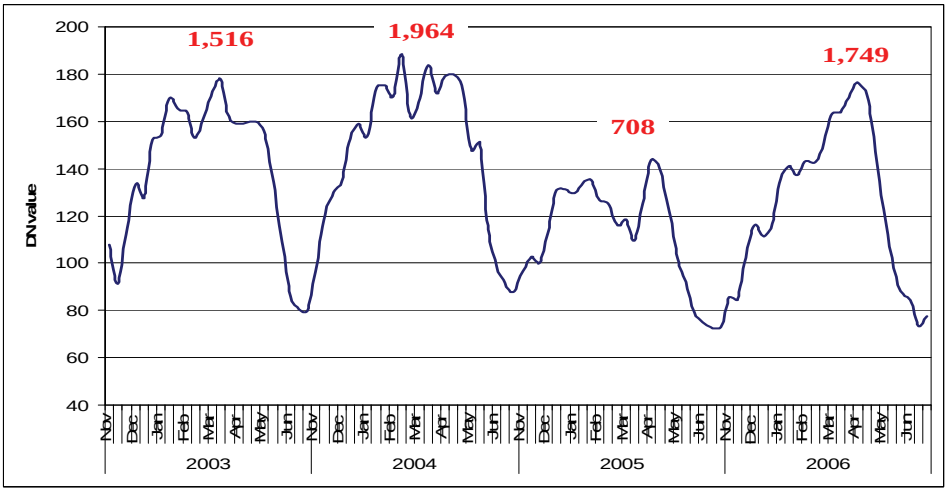


Figure 7- 1: Average annual NDVI profiles for rainfed wheat along with the annual production of rainfed wheat (2003-2006). Red figures show the production (1000 tons) of Rainfed wheat reported by Ministry of Agriculture and Fisheries, Andalucia

The land use maps produced in chapter 2-4 have the potential to support a broad range of applications in agriculture, such as early warning in relation to food security, yield gap analysis and regional to global assessment of agricultural productivity. Some of these applications are described below:

Monitoring of land use modifications

In chapter 2, we used the results as an example of monitoring of land use modifications over time. Figure 2- 9, for example, shows the irrigated areas where rice is the dominant crop during the *Rabi* season. The annual profiles of relevant NDVI classes indicate a decline in NDVI (DN values) starting in 2001 and expanding one year later (2002). These declines represented a substantial number of fields that had not been cultivated during *Rabi* because of unreliable availability of electricity to run their water pumps.

Location specific crop monitoring

Our method enables the monitoring of changes in crop over years by comparing annual NDVI classes for selected map units and the crop calendar information. Once NDVI map units have been delineated for a district or a larger geographical region, continuous monitoring of their specific performance over time can be related to the information presented in the legend. Stratified monitoring - based on units with a uniform known land cover and land use - allows the preparation of more specific early warning bulletins. The anomalies in the NDVI classes in time and duration of their state below predefined thresholds can be monitored continuously. This, in turn, could effectively lead to a continuous monitoring and prediction of various crop conditions at different scales – from pixel to region levels.

Provision of input data required by regional crop growth modeling approaches

The estimated crop maps can be used as an important input in the regional crop growth models for estimation of crop yields. De Wit and van Diepen (2008) reportedly had to aggregate their simulation results to regions for comparison with yield statistics available from EUROSTAT, using only 'arable' areas instead of wheat areas. Crop area maps can also be used in models such as Common Agricultural Policy Regionalised Impact (CAPRI) modeling system (Kempen *et al*., 2007; Britz and Leip, 2009).

Provision of input data required by drought monitoring

The method illustrated in chapters 2-4 can potentially be incorporated into RS/GIS based drought monitoring systems like Famine Early Warning System Network (FEWSNET), US Drought Monitor (www.drought.unl.edu) and the Southern African Development Community's (SADC) Regional Remote Sensing Unit Drought Monitoring Center (Thenkabail *et al*., 2004; Smakhtin *et al*., 2005; Justice and Becker-Reshef, 2007). Including the land use maps generated in chapters 2-4 in such systems will take the outputs of these systems beyond drought monitoring at administrative unit levels to agricultural drought monitoring which is experienced by crops over severe water shortage periods.

Improving the statistical sampling methods for better estimates of crop areas

The comparison of crop maps generated directly by using the segments data and the agricultural statistical data at municipal level shows that hypertemporal SPOT NDVI images can be used to improve statistical methodologies for estimation of crop areas. This holds for agricultural ministries (at country level) and for international organizations such as the European Union (EU), the Food and Agriculture Organization (FAO) of the United Nations and the United States Geological Survey (USGS) (IGOL, 2006; Townshend *et al*., 2008). Hypertemporal NDVI data provide the opportunity to improve the current stratified sampling methods for statistical data collection and up-scaling these data to administrative divisions. This can now be done by using the similar-behaving NDVI classes (results of supervised grouping after unsupervised classification) as a stratum for sampling.

The methods developed in chapters 2, 3 and 4 have meanwhile also been tested in Vietnam (Nguyen *et al*., 2010). This shows that it is a replicable approach for land use data acquisition.

Further, the remote sensing data (Hypertemporal NDVI images) is freely available from SPOT, Moderate Resolution Imaging Spectroradiometer (MODIS) and MEdium Resolution Imaging Spectrometer (MERIS) satellite systems.

Limitations of the research on land use mapping

The land use maps produced by using the methods developed in chapters 2-4 are based on hypertemporal NDVI images at 1 km^2 spatial resolution. Each pixel of a produced NDVI based land use map is also related to a mix of other land cover. Therefore, such maps should be considered as ‘small-scale’ land use maps valid at regional scales. Further, the accuracy of produced land use map is low at local scales.

The accuracy of produced specific crop area maps is dependent on the quality of crop statistical data. This issue was dealt by using the segments data directly for mapping crop areas (in chapters 3 and 4).

7.3 Comparing crop growth models using agricultural statistics and field data

In chapter 5 the outputs of two spatial crop growth models were compared with a specific aim to test a new (still under development) crop growth model [C*f* (Water)]. Currently, input data requirements in most operational crop growth models are substantial, which makes their use problematic, particularly in low data environments which is the case in many developing countries. This makes it necessary to look for models with fewer data requirements, but still capable of producing accurate estimates of crop yield/production.

For 2001, comparison of the estimated actual rainfed wheat production at NUTS-3 scale by C*f*-Water and the estimated actual rainfed wheat production by CGMS with published agricultural statistics showed an excellent agreement for C*f*-Water and a good agreement for CGMS (R^2 = 67 % and 98 %; RSME= 16 Mg and 41 Mg, Figure 5- 6), respectively. The accuracy assessment of C*f*-Water estimates using primary field data comprising 334 segments of 700 x 700 m showed excellent agreement (Adj. R^2 = 98%). On the basis of these results we conclude that C*f*-Water has a potential to be incorporated in food security studies. Two points should be noted. First, the C*f*-Water model has lower data requirements than CGMS which requires also soil and historical yield data. Second, the outputs of CGMS are potential and water limited yields which are then adjusted by a time-trend analysis to produce actual yields. Whereas, C*f*-Water directly generates the actual crop yield estimates (Table 5.2). However, the model has only been applied in a single case study, thus before incorporating the model into operational systems; more extensive testing of the model is required for different crops, years and regions. Further research is needed to explore the extraction of sowing and harvesting time on regional scales to be used in the model which could not be explored in this thesis.

Practical relevance of the outputs

The outputs of chapter 5 (estimated crop yields) can be used along with the outputs of chapters 2-4 (crop area maps) for assessing the production at regional scales. By combining the estimated rainfed wheat area map (Figure 4- 10) and the estimated rainfed wheat yield map (Figure 5- 4) with the administrative map

of Andalucia, we generated the map of rainfed wheat production in Andalucia for 2001 (Figure 7- 2).

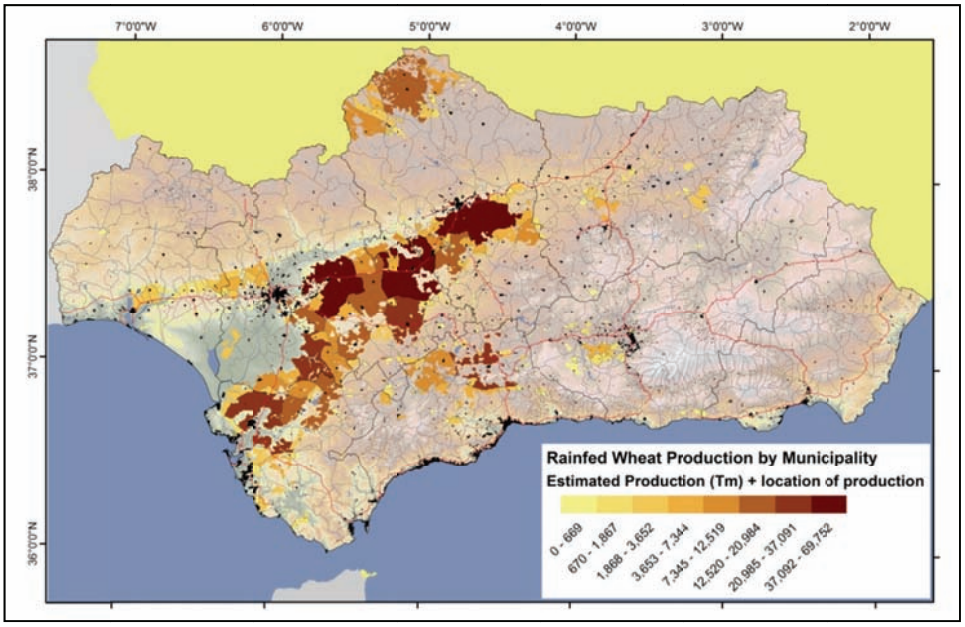


Figure 7- 2: Estimated production of rainfed wheat, Andalucia, Spain (2001)

Figure 7- 2, shows the production of rainfed wheat in 2001 in the agricultural areas per municipality. Such information can be used to know the production of food in various municipalities. Such data are only provided at country level (NUTS-0) by CGMS through MARS bulletins (at http://mars.jrc.ec.europa.eu/mars/Bulletins-Publications) or only at NUTS level 3 by the Ministry of agriculture and Fisheries, Andalucia, Spain (http://www.mapa.es/es/estadistica/pags/publicaciones/BME/introduccion.htm).

Limitations of the research on crop yield estimation

The limitation in the presented research is the use of only one year's simulation results. Another limitation is that the crop yield estimated by C*f*-Water is only available after harvest. Also, crop yield estimation can be affected by cloud cover which may cause substantial percentages of missing data. In order to overcome the issue of missing remotely sensed data, the observations can be

combined from different satellites and used in crop growth models as shown by Venus and Rugege (2004).

7.4 Users' perspective on available land use data and the generated outputs

Amongst the problems identified in the data aspects of available land use data, update frequency and limited availability of thematic agricultural land use data were the most common. Since the methods developed in chapters 2, 3 and 4 make use of hypertemporal NDVI data, therefore, the demands of users pertaining to easy availability and update frequency are easily met. This was also reflected by the respondents, i.e., a majority was keen to use the methods developed for crop area maps (75%) and crop yield maps (70%). On the basis of stakeholders' opinion on existing data sources and the outputs of this thesis we can conclude that remote sensing techniques should be used in combination with field observations to enhance the update frequency and availability of land use data.

7.5 Recommendations for future research

The following recommendations could be taken on board for future research:

- The land use maps produced (chapters 2-4) can be further improved by using higher spatial resolution hypertemporal images. Hypertemporal NDVI images are available at a spatial resolution of 500 m and 250 m from MODIS which can be used to improve the spatial resolution of produced agricultural land use data. By using high resolution hypertemporal NDVI data the occurrence of the mixed pixels in the generated crop area maps can be overcome depending on the size of the fields.
- The NDVI map can be used in the area frame sampling method to improve the crop statistical data. For this purpose similar-behaving groups of NDVI classes after unsupervised classification can function as strata for various related land cover/land use types. Currently, the generalized land cover classes present in the CORINE land cover map are used as strata for statistical sampling.

- The presented Cf (Water) model needs further testing for other crops and areas for multiple years in order to verify the results obtained and to gain confidence in their general applicability.
- In the long run when other crops are included in the yield estimation method it is advisable to prepare sowing time and harvesting time maps at regional scales. This can be achieved by investigating the temporal behavior of relevant NDVI classes because in chapters 2 and 3 it was proven that the temporal behavior of NDVI match well with the crop calendars. Further, analyses of uncertainties that can arise should also be investigated.

Bibliography

Aalders, I.H. and Aitkenhead, M.J., 2006. Agricultural census data and land use modelling. Computers Environment and Urban Systems 30, 799-814.

Agrawal, S., Sarnam, S., Joshi, P.K. and Roy, P.S., 2001. Phenology based classification Model for Vegetation Mapping using IRS-WiFS. Workshop on classification methods, March 28-29, 2001, Ispra, Italy.

Alexandratos, N. (Ed.), 1995. World Agriculture: Towards 2010 - An FAO Study. Chichester, UK: John Wiley and Sons.

Allen, R.G., Pereira, L.S., Raes, D. and Smith, M., 1998. Crop evapotranspiration - Guidelines for computing crop water requirements. FAO Irrigation and Drainage paper No. 56. FAO, Rome, Italy.

APAU, 1989. National Agricultural Research Project: Northern Telengana Zone- Status Report Vol 1. Andhra Pradesh Agricultural University, Regional Agricultural Research Station, Jagitial, Karimnagar District. September 1989.

Archer, E.R.M., 2004. Beyond the “climate versus grazing” impasse: using remote sensing to investigate the effects of grazing system choice on vegetation cover in the eastern Karoo. Journal of Arid Environments 57: 381-408.

Arshad, M.A. and Martin, S., 2002. Identifying critical limits for soil quality indicators in agro-ecosystems. Agriculture Ecosystems & Environment 88, 153-160.

Bagan, H., Takeuchi, W., Kinoshita, T., Yuhai, B. and Yamagata, Y., 2010. Land Cover Classification and Change Analysis in the Horqin Sandy Land From 1975 to 2007. Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of 3, 168-177.

Bai, J., Chen, X., Dobermann, A., Yang, H., Cassman, K.G. and Zhang, F., 2010. Evaluation of NASA Satellite- and Model-Derived Weather Data for Simulation of Maize Yield Potential in China. Agronomy Journal 102, 9-16.

Barahona, E. and Iriarte, A., 2001. An overview of the present state of standardization of soil sampling in Spain. Science of the Total Environment 264, 169-174.

Baret, F. and Guyot, G., 1991. Potentials and limits of vegetation indices for LAI and APAR assessment. Remote Sensing of Environment 35, 161-173.

Barros, J.M., 1997. Quantitative Analysis of Selected Land-Use Systems. Ph.D. Thesis, Wageningen Agricultural University, Wageningen, The Netherlands. 169 pp.

Battese, G.E., Harter, R.M. and Fuller, W.A., 1988. An Error-Components Model for Prediction of County Crop Areas Using Survey and Satellite Data. Journal of the American Statistical Association 83, 28-36.

Becker-Platen, J.D., 1979. Geoscientific maps as an aid to land use and regional planning. Resources Policy 5, 71-77.

Bellow, M.E., 1993. Application of satellite data to crop area estimations at the county level. International Conference on Establishment Surveys. June 28-30, Buffalo, New York, USA, pp. 1-6.

Bhuiyan, C., Singh, R.P. and Kogan, F.N., 2006. Monitoring drought dynamics in the Aravalli region (India) using different indices based on ground and remote sensing data. International Journal of Applied Earth Observation and Geoinformation, 8(4): 289-302.

Boogaard, H.L., Van Diepen, C.A., Roetter, R.P., Cabrera, J.M.C.A., Van Laar, H.H., 1998. WOFOST 7.1; user's guide for the WOFOST 7.1 crop growth simulation model and WOFOST Control Center 1.5**.** Wageningen, SC-DLO: Techn. Doc. 52. 144 pp.

Boogaard, H., Eerens, H., Supit, I., Van Diepen, C., Piccard, I. and Kempeneers, P., 2002. Description of the MARS crop yield forecasting system (mcyfs). Technical Report 19226-2002-02-F1FED ISP NL, Centre for Geo-information, Wageningen-UR, Wageningen, The Netherlands.

Borlaug, N.E., 2000. Ending World Hunger. The Promise of Biotechnology and the Threat of Antiscience Zealotry. Plant Physiol. 124, 487-490.

Bossard, M.J., Feranec, J. and Otahel, J., 2000. CORINE Land Cover Technical Guide – Addendum 2000. Technical report No 40. European Environmental Agency (EEA), Copenhagen, Denmark.

Bouman, B.A.M., van Keulen, H., van Laar, H.H. and Rabbinge, R., 1996. The 'School of de Wit' crop growth simulation models: A pedigree and historical overview. Agricultural Systems 52, 171-198.

Bouman, B.A.M., Kropff, M.J., Tuong, T.P., Wopereis, M.C.S., Ten Berge, H.F.M. and Van Laar, H.H., 2001. ORYZA2000: modeling lowland rice. Los Baños (Philippines): International Rice Research Institute, and Wageningen: Wageningen University and Research Centre. 235 pp.

Brand S. and Malthus T.J., 2004. Evaluation of AVHRR NDVI for monitoring intra-annual and interannual vegetation dynamics in a cloudy environment (Scotland, UK). In: Proceedings of the XXth ISPRS Congress, Commission-II, 12-23 July 2004 Istanbul, Turkey. 6pp.

Britz, W. and Leip, A., 2009. Development of marginal emission factors for N losses from agricultural soils with the DNDC-CAPRI meta-model. Agriculture, Ecosystems & Environment 133, 267-279.

Budde, M.E., Tappan, G., Rowland, J., Lewis, J. and Tieszen, L.L., 2004. Assessing land cover performance in Senegal, West Africa using 1-km integrated NDVI and local variance analysis. Journal of Arid Environments 59, 481-498.

Carfagna, E. and Gallego, F.J., 2005. Using Remote Sensing for Agricultural Statistics. International Statistical Review 73, 389-404.

Carter, S.E., Fresco, L.O. and Jones, P.G. with Fairbairn, J.N., 1992. An atlas of cassava in Africa. Historical, agroecological and demographic aspects of

crop distribution. Cali, Colombia: Centro Internacional de Agricultura Tropical.

CAST, 1999. Animal Agriculture and Global Food Supply. Centre for Agricultural Science and Technology, Ames, Iowa, USA.

Cayrol, P., Chehbouni, A., Kergoat, L., Dedieu, G., Mordelet, P. and Nouvellon, Y., 2000. Grassland modeling and monitoring with SPOT-4 VEGETATION instrument during the 1997-1999 SALSA experiment. Agricultural and Forest Meteorology 105, 91-115.

Census of India, 2001. Andhra Pradesh, Provisional Population Totals. Paper 3 of 2001, Director of Census Operations. Hyderabad, India.

Champeaux, J.L., Kyung-Soo, H., Arcos, D., Habets, F. and Masson, V., 2004. Ecoclimap2: a new approach at global and European scale for ecosystems mapping and associated surface parameters database using SPOT/VEGETATION data - first results. Geoscience and Remote Sensing Symposium, 2004. IGARSS '04. Proceedings. 2004 IEEE International, pp. 2046-2049 vol.2043.

Claggett, P.R., Okay, J.A. and Stehman, S.V., 2010. Monitoring Regional Riparian Forest Cover Change Using Stratified Sampling and Multiresolution Imagery. JAWRA Journal of the American Water Resources Association 46, 334-343.

Cohen Y. and Shoshany M., 2002. A national knowledge-based crop recognition in Mediterranean environment. International Journal of Applied Earth Observation and Geoinformation 4: 75-87.

Conway, G., 1997. The Doubly Green Revolution: Food for all in the 21st century. London, UK: Penguin Books.

CPO, 2001. Handbook of Statistics- Nizamabad District, Chief Planning Officer: Nizamabad, Andhra Pradesh, India.

Craig, M.E., 2001. The NASS cropland data layer program. Third Annual International Conference on Geospatial Information in Agriculture, Denver, CO, November 5–7.

Creech, T.G. and McNab, A. L., 2002. Using NDVI and elevation to improve precipitation mapping. Preprints, 13th Conference on Applied Climatology, Portland, OR, May 13-16, 2002, American Meteorological Society, Boston, MA, 184-187. ftp://ftp.ncdc.noaa.gov/pub/data/papers/2002tc05.pdf

Dalal-Clayton, B. and Dent, D., 1993. Surveys, Plans and People. A Review of Land Resource Information and its Use in Developing Countries. Environmental Planning Issues No.2. International Institute for Environment and Development (IIED), London.

De Bie, C.A.J.M., 2000. Comparative performance analysis of agro - ecosystems. ITC Dissertation;75. ITC, Enschede, p. 232.

De Bie, C.A.J.M., Khan, M.R., Toxopeus, A.G., Venus, V. and Skidmore, A.K., 2008. Hypertemporal image analysis for crop mapping and change detection. In: ISPRS 2008 : Proceedings of the XXI congress : Silk road

for information from imagery : the International Society for Photogrammetry and Remote Sensing, 3-11 July, Beijing, China. Comm. VII, WG VII/5. Beijing : ISPRS, 2008. pp. 803-812.

De Bie, C.A.J.M., Khan, M.R., Smakhtin, V.U., Venus, V., Weir, M.J.C., Smaling, E.M.A., 2010. Analysis of multi-temporal SPOT NDVI images for small-scale land-use mapping. International Journal of Remote Sensing (in press).

DeFries, R., Hansen, M. and Townshend, J., 1995. Global discrimination of land cover types from metrics derived from AVHRR pathfinder data. Remote Sensing of Environment 54(3), 209-222.

De Groote, H. and Traoré, O., 2005. The cost of accuracy in crop area estimation. Agricultural Systems 84, 21-38.

De Koning, G.H.J., Jansen, M.J.W., Boons-Prins, E.R., Van Diepen, C.A., Penning de Vries, F.W.T., 1993, Crop growth simulation and statistical validation for regional yield forecasting across the European Community. Simulation Reports CABO-TT no 31, Wageningen Agricultural University, CABO-DLO, JRC, Wageningen/Ispra.

De la Rosa, D., Anaya-Romero, M., Diaz-Pereira, E., Heredia, N. and Shahbazi, F., 2009. Soil-specific agro-ecological strategies for sustainable land use - A case study by using MicroLEIS DSS in Sevilla Province (Spain). Land Use Policy 26, 1055-1065.

Dennett, M.D., Elston, J., Diego Q.R., 1980. Weather and yields of tobacco, sugar beet and wheat in Europe. Agricultural Meteorology 21, 249-263.

De Wit, A.J.W. and Boogaard, H.L., 2001. Monitoring of crop development and crop model optimisation using NOAA - AVHRR : towards an integrated satellite and model - based crop monitoring system in the European context. Beleidscommissie Remote Sensing (BCRS), Delft.

De Wit, A.J.W. and van Diepen, C.A., 2008. Crop growth modelling and crop yield forecasting using satellite-derived meteorological inputs. International Journal of Applied Earth Observation and Geoinformation 10, 414-425.

Dietz, T. 2003. Land Use Policy for Michigan: Views from Stakeholders. Michigan State University (http://www.michiganlanduse.org/resources/councilresources/MSU-stakeholder.pdf).

Dixon, J, Gulliver, and Gibbon, D. 2001. Farming Systems and Poverty - Improving farmers' livelihoods in a changing world. FAO and World Bank, Rome and Washington DC. 412p. Available at http://www.fao.org/farmingsystems/index_en.htm.

Doraiswamy, P.C., Sinclair, T.R., Hollinger, S., Akhmedov, B., Stern, A., Prueger, J., 2005. Application of MODIS derived parameters for regional crop yield assessment. Remote Sensing of Environment 97, 192-202.

Dore, M. Kulshreshtha, S. and Johnston, M., 2001. An Integrated Economic-Ecological Analysis of Land Use Decisions in Forest-Agriculture Fringe

Regions of Northern Saskatchewan. Geographical and Environmental Modelling 5, 159-175.

Drenge H.E. and Tucker C.J., 1988. Green biomass and rainfall in semi-arid sub-Saharan Africa . Journal of Arid Environments 15: 245-252.

Driessen, P.M., Konijn, N.T., 1992. Land Use Systems Analysis. Wageningen Agricultural University, Wageningen, The Netherlands.

Duchemin, B., Maisongrande, P., Boulet, G., Benhadj, I., 2008. A simple algorithm for yield estimates: Evaluation for semi-arid irrigated winter wheat monitored with green leaf area index. Environmental Modelling & Software 23, 876-892.

Dymond, J.R., Bégue, A. and Loseen, D., 2001. Monitoring land at regional and national scales and the role of remote sensing. International Journal of Applied Earth Observation and Geoinformation 3, 162-175.

Dyson, T., 1996. Population and Food: Global Trends and Future Prospects. London, UK: Reutledge.

Eerens H., Kempeneers P., Piccard I. and Verheijen Y., 2001. Crop monitoring and yield forecasting with NOAA-AVHRR or SPOT-Vegetation. http://www.geosuccess.net/geosuccess/documents/Dry_Matter_Productivity.pdf (accessed 8 Dec'09).

Efron, B. and Gong, G., 1983. A Leisurely Look at the Bootstrap, the Jackknife, and Cross-Validation. The American Statistician 37, 36-48.

Elvidge, C.D., Baugh, K.E., Hobson, V.R., Kihn, E.A., Kroehl, H.W., Davis, E.R. and Cocero, D., 1997. Satellite inventory of human settlements using nocturnal radiation emissions: a contribution for the global tool chest. Global Change Biology 3, 87-395.

Engelman, R., Cinotta, R. P., Dye, B., Gardner-Outlaw, T., and Wisnewski, J., 2000. People in the balance: Population and natural resources at the turn of the millenn ium. Washington, DC: Population Action International.

ERDAS, 1997. ERDAS Field Guide. 4th ed., ERDAS Inc., Atlanta GA.

ERDAS, 2003. Erdas Field Guide. 7th ed., Leica Geosystems, GIS and Mapping LLC, Atlanta GA

FAO, 1974. Commodity Review and Outlook, 1973-1974. FAO, Rome, Italy

FAO, 1983. World Food Security: a Reappraisal of the Concepts and Approaches. Director Generals Report, FAO, Rome, Italy.

FAO, 1994. Integrated Approach to the Planning and Management of Land Resources. Draft report of the UN Secretary-General on the Implementation of Chapter 10 of Agenda 21 (UNCED) to the Commission on Sustainable Development. Third Draft of Task Manager's Report. FAO/AGL, Rome. 30 pp.

FAO, 1998. World Reference Base for Soil Resources. Food and Agriculture Organization of the United Nations, Rome.

FAO, 2000. Basic texts of the Food and Agriculture Organization of the United Nations, Vols. I and II – 2000 edition. FAO, Rome, Italy.

FAO, 2002. Report of the second consultation on agricultural information management, Rome, Italy.

FAO, 2003. World Agriculture: towards 2015/2030. An FAO perspective. FAO-Earthscan, Rome/London, 432 pp.

FAO, 2004. The state of food and agriculture 2003-2004. Agricultural Biotechnology: Meeting the needs of the poor. Food and Agriculture Organization of the United Nations, Rome, 2004. 208 p.

FAO, 2005**.** Report of expert consultation on Analysis and dissemination of agricultural census and survey data, Bankok, Thailand.

FAO and UNEP. 2002. Proceedings of the FAO/UNEP Consultation on Strategies for Global Land Cover Mapping and Monitoring. Artimino, Florence, Italy, 6-8 May, 2002. FAO, Rome, Italy. 39 pp. including Artimino Declaration on establishment of a Global Land Cover Network.

Feranec, J., Hazeu, G., Christensen, S. and Jaffrain, G., 2007. Corine land cover change detection in Europe (case studies of the Netherlands and Slovakia). Land Use Policy 24, 234-247.

Flores, L.A. and Martínez, L.I., 2000. Land Cover Estimation in Small Areas Using Ground Survey and Remote Sensing. Remote Sensing of Environment 74, 240-248.

Focardi, S., Loiselle, S.A., Mazzuoli, S., Bracchini, L., Dattilo, A.M. and Rossi, C., 2008. Satellite-based indices in the analysis of land cover for municipalities in the province of Siena, Italy. Journal of Environmental Management 86, 383-389.

Font, I., 2000. Climatología de España y Portugal. Ediciones Universidad de Salamanca. Salamanca, Spain.

Fraser, R.H., Olthof, I. and Pouliot, D., 2009. Monitoring land cover change and ecological integrity in Canada's national parks. Remote Sensing of Environment, 113(7): 1397-1409.

Fresco, L.O., Stroosnijder, L., Bouma, J., and Van Keulen, H. (Eds.), 1994. The Future of the Land. Mobilising and Integrating Knowledge for Land use Options. John Wiley and Sons, London.

Fresco, L.O., Leemans, R., Turner Ii, B.L., Skole, D., Van Zeijl-Rozema, A.G., Haarmann, V., 1997. Land use and cover change (LUCC). Open science meeting proceedings. LUCC Report Series No. 1. International Project Office, Institut Cartografic Catalunya, Barcelona, Spain.

Fritz, S., Massart, M., Savin, I., Gallego, J. and Rembold, F., 2008. The use of MODIS data to derive acreage estimations for larger fields: A case study in the south-western Rostov region of Russia. International Journal of Applied Earth Observation and Geoinformation 10, 453-466.

Gallego, J. and Bamps, C., 2008. Using CORINE land cover and the point survey LUCAS for area estimation. International Journal of Applied Earth Observation and Geoinformation 10, 467-475.

George H. and Nachtergaele F.O., 2002. Land Use Data. In: R. Tateishi and D. Hastings (Eds.) 2002. Global Environmental Databases - Present Situation; Future Directions, pp 55-67. Vol.2. ISPRS, WG IV/8. ISBN 962-8226-03-7; 154 pp.

Ghulam, A., Qin, Q., Teyip, T. and Li, Z.-L., 2007. Modified perpendicular drought index (MPDI): a real-time drought monitoring method. ISPRS Journal of Photogrammetry and Remote Sensing, 62(2): 150-164.

Goerner, A., Reichstein, M. and Rambal, S., 2009. Tracking seasonal drought effects on ecosystem light use efficiency with satellite-based PRI in a Mediterranean forest. Remote Sensing of Environment 113, 1101-1111.

Gorham B., 1998. Mapping multi-temporal agricultural land use in Mississippi Alluvial valley of Arkansas. In: Proceedings of the Spatial Technologies in Agricultural and Natural Resources: SRIEG-10 Annual Conference, Global Hydrology and Climate Centre, Huntsville, USA.

Goward, S.N. and Huemmrich, K.F., 1992. Vegetation canopy PAR absorptance and the normalized difference vegetation index: An assessment using the SAIL model. Remote Sensing of Environment 39, 119-140.

Groten S.M.E.and Ocatre R., 2002. Monitoring the length of the growing season with NOAA. International Journal of Remote Sensing 23.14: 2797 - 2815.

Gu, Y., Hunt, E., Wardlow, B., Basara, J.B., Brown, J.F. and Verdin, J.P., 2008. Evaluation of MODIS NDVI and NDWI for vegetation drought monitoring using Oklahoma Mesonet soil moisture data. Geophys. Res. Lett. 35, L22401.

Guo, W.Q., Yang, T.B., Dai, J.G., Shi, L. and Lu, Z.Y., 2008. Vegetation cover changes and their relationship to climate variation in the source region of the Yellow River, China, 1990–2000. International Journal of Remote Sensing 29, 2085 - 2103.

Hafner, S., 2003. Trends in maize, rice and wheat yields for 188 nations over the past 40 years: a prevalence of linear growth. Agriculture, Ecosystems and Environment 97, 275-283.

Han, K.S., Champeaux, J.L. and Roujean, J.L., 2004. A land cover classification product over France at 1 km resolution using SPOT4/VEGETATION data. Remote Sensing of Environment 92, 52-66.

Hatfield J.L., 2001. Emerging Challenges in Land Management: An American Perspective on efficiency. In: Proceedings of the 10th Australian Agronomy Conference, Hobart. http://www.regional.org.au/au/asa/2001/plenery/1/hatfield.htm (accessed 5 Apr 2005).

Hedal Kløverpris, J., 2009. Identification of biomes affected by marginal expansion of agricultural land use induced by increased crop consumption. Journal of Cleaner Production 17, 463-470.

Hill, M.J. and Donald, G.E., 2003. Estimating spatio-temporal patterns of agricultural productivity in fragmented landscapes using AVHRR NDVI time series. Remote Sensing of Environment 84, 367-384.

Hogg, H.C., 1986. Foreword. IEEE Transactions on Geoscience and Remote Sensing, GE-2(1), 3–4.

Holben, B.N., 1986. Characteristics of maximum-value composite images from temporal AVHRR data. International Journal of Remote Sensing 7, 1417-1434.

Huang, S. and Siegert, F., 2006. Land cover classification optimized to detect areas at risk of desertification in North China based on SPOT VEGETATION imagery. Journal of Arid Environments, 67(2): 308-327.

IAWG [Inter-Agency Working Group], 2000. Guidelines for National FIVIMS. Background and Principles. IAWG Guideline Series, No. 1, FAO, Rome.

IEA, 2007. World Energy Outlook. OECD/IEA, Paris, France.

IFPRI, 2001. Biotechnology for Developing Country Agriculture: Problems and Opportunities. (Pinstrup-Andersen, P. and Pandya-Lorch, R., Eds.). Washington DC, USA: International Food Policy Research Institute.

IGOL, 2006. Report 3rd Meeting of the Theme Team of the Integrated Global Observations of the Land (IGOL), Beijing, China, 27 Feb.-1 March 2006.

IGOL, 2007. Developing a strategy for Global Agricultural Monitoring in the framework of Group on Earth observations (GEO) workshop report. Available at http://www.fao.org/gtos/igol/docs/meeting-reports/07-GEO-AG0703-Workshop-Report-nov07.pdf.

Ingram, J.S.I., Gregory, P.J. and Izac, A.M., 2008. The role of agronomic research in climate change and food security policy. Agriculture, Ecosystems & Environment 126, 4-12.

Jame, Y. W., 1992. Modelling soil moisture for agricultural water management. Pages 69–88. *in* F. J. Eley, R. Granger and L. Martin eds. Soil moisture modelling. Environment Canada, Ottawa, Canada.

Janick, J., 1974. Plant science: an introduction to the world of crops. San Francisco, USA: Freeman

Jansen, L.J.M. and Di Gregorio, A., 2003. Land-use data collection using the "land cover classification system": results from a case study in Kenya. Land Use Policy 20, 131-148.

Jansen, L.J.M. and Di Gregorio, A., 2004. Obtaining land-use information from a remotely sensed land cover map: results from a case study in Lebanon. International Journal of Applied Earth Observation and Geoinformation 5, 141-157.

Jensen, J.R., 1996. Introductory digital image processing : a remote sensing perspective. Prentice-Hall, London.

Ji, L. and Peters, A.J., 2003. Assessing vegetation response to drought in the northern Great Plains using vegetation and drought indices. Remote Sensing of Environment, 87(1): 85-98.

Jones, J.W., Hoogenboom, G., Porter, C.H., Boote, K.J., Batchelor, W.D., Hunt, L.A., Wilkens, P.W., Singh, U., Gijsman, A.J., Ritchie, J.T., 2003. The DSSAT cropping system model. European Journal of Agronomy. 18, 235-265

Jongschaap, R.E.E., 2006. Integrating crop growth simulation and remote sensing to improve resource use efficiency in farming systems. Ph.D. Thesis, Wageningen University, Wageningen, The Netherlands.

JRC, MARS STAT, 2004. Methodology of the Mars Crop Yield Forecasting System. Vol 4. *European Commission EU Report 21291 EN/1-4*. Luxembourg, Luxembourg. Available at: http://agrifish.jrc.it/marsstat/Crop_Yield_Forecasting/METAMP/.

Julien, Y. and Sobrino, J.A., 2009. The Yearly Land Cover Dynamics (YLCD) method: An analysis of global vegetation from NDVI and LST parameters. Remote Sensing of Environment, 113(2): 329-334.

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

Justice C.O., Townshend J.R.G., Holben B.N. and Tucker C.J., 1985. Analysis of the phenology of global vegetation using meteorological satellite data. International Journal of Remote Sensing 6: 1271-1318.

Kalluri, S.N.V. and Townshed, J.R.G., 1998. A Simple Single layer Model to Estimate Transpiration from Vegetation using Multi-spectral and Meteorological Data. International Journal of Remote Sensing 19, 1037-1053.

Keating, B.A., Carberry, P.S., Hammer, G.L., Probert, M.E., Robertson, M.J., Holzworth, D., Huth, N.I., Hargreaves, J.N.G., Meinke, H., Hochman, Z., McLean, G., Verburg, K., Snow, V., Dimes, J.P., Silburn, M., Wang, E., Brown, S., Bristow, K.L., Asseng, S., Chapman, S., McCown, R.L., Freebairn, D.M. and Smith, C.J., 2003. An overview of APSIM, a model designed for farming systems simulation. European Journal of Agronomy 18, 267-288.

Kempen, M., Heckelei, T. and Britz, W., 2005, An econometric approach for spatial disaggregation of crop production in the eu, University of Bonn, Institute for Agricultural policy, market Reserach and Economic sociology; working paper presented at the EAAE Seminar, Parma, 3-5 February 2005 .

Kempen, M., Heckelei, T., Britz, W., Leip, A., and Koeble, R., 2007 Computation of a European Agricultural Land Use Map – Statistical Approach and Validation. Technical Paper, Institute for Food and

Resource Economics, Bonn, (available at: http: //www.ilr1.uni-bonn.de/agpo/rsrch/dynaspat/publicat.htm).

Keyfitz, N., 1976. Applied Mathematical Demography, New York: John Wiley and Sons.

Keyzer, M.A., Merbis, M.D., Pavel, I.F.P.W. and van Wesenbeeck, C.F.A., 2005. Diet shifts towards meat and the effects on cereal use: can we feed the animals in 2030? Ecological Economics 55, 187-202.

Khan, M.R., de Bie, C.A.J.M., van Keulen, H., Smaling, E.M.A. and Real, R., 2010. Disaggregating and mapping crop statistics using hypertemporal remote sensing. International Journal of Applied Earth Observation and Geoinformation 12, 36-46.

Kileshye Onema, J.-M., Taigbenu, A., 2009. NDVI-rainfall relationship in the Semliki watershed of the equatorial Nile. Physics and Chemistry of the Earth, Parts A/B/C 34(13-16), 711-721.

Knorn, J., Rabe, A., Radeloff, V.C., Kuemmerle, T., Kozak, J. and Hostert, P., 2009. Land cover mapping of large areas using chain classification of neighboring Landsat satellite images. Remote Sensing of Environment 113, 957-964.

Krol, B.G.C.M., Rossiter, D.G. and Siderius, W., 2007. Ontology-based multi-source data integration for digital soil mapping. In: (Lagacherie, P., MacBratney, A.B. and Voltz, M., Eds.) Digital soil mapping: an introductory perspective. Developments in Soil Science 31, Elsevier, Amsterdam, pp. 119-133

Kutz, F. W., Morgan, J. M., Willis, III, P. and Mueller R., 2005. The Towson/NASS 2002 Agricultural Cropland Project. Paper presented at the 2005 Maryland Environmental
Public Health Tracking Summer Conference: Analysis of EPHT Data, August 22, Baltimore, Maryland

Latifovic, R., Zhu, Z.-L., Cihlar, J., Giri, C. and Olthof, I., 2004. Land cover mapping of North and Central America--Global Land Cover 2000. Remote Sensing of Environment 89, 116-127.

Launay, M. and Guerif, M., 2005. Assimilating remote sensing data into a crop model to improve predictive performance for spatial applications. Agriculture, Ecosystems & Environment 111, 321-339.

LAZĂR, C., Bettina, B., Fabio, M., Daniela, L. A., 2009, Adaptation of WOFOST model from CGMS to Romanian conditions. Journal of Plant Development 16, 97-102.

Leamer, R.W., Weber, D.A., Wiegand, C.L., 1975. Pattern recognition of soils and crops from space. Photogrammetric Engineering 41, 471–478.

Ledwith M., 2000. Land cover classification using SPOT Vegetation 10-days composite images – Baltic Sea catchment basin. GLC2000 Meeting, Ispra, Italy, April 18-22.

le Maire, G., Francois, C. and Dufrene, E., 2004. Towards universal broad leaf chlorophyll indices using PROSPECT simulated database and

hyperspectral reflectance measurements. Remote Sensing of Environment 89, 1-28.

Leip, A., Marchi, G., Koeble, R., Kempen, M., Britz, W. and Li, C., 2008. Linking an economic model for European agriculture with a mechanistic model to estimate nitrogen and carbon losses from arable soils in Europe. Biogeosciences 5, 73-94.

Lepers, E., Lambin, E.F., Janetos, A.C., DeFries, R., Achard, F., Ramankutty, N. and Scholes, R.J., 2005. A synthesis of information on rapid land-cover change for the period 1981-2000. Bioscience 55, 115-124.

Liu, X. and Kafatos, M., 2005. Land-cover mixing and spectral vegetation indices. International Journal of Remote Sensing 26, 3321-3327.

Lobell, D.B., Asner, G.P., Ortiz-Monasterio, J.I. and Benning, T.L., 2003. Remote sensing of regional crop production in the Yaqui Valley, Mexico: estimates and uncertainties. Agriculture, Ecosystems & Environment 94, 205-220.

Lobo, J.M., Lumaret, J.P. and Jay-Robert, P., 2002. Modelling the species richness distribution of French dung beetles (Coleoptera, Scarabaeidae) and delimiting the predictive capacity of different groups of explanatory variables. Global Ecology and Biogeography 11, 265-277.

López-Bellido, L., López-Bellido, R.J., Redondo, R., 2005. Nitrogen efficiency in wheat under rainfed Mediterranean conditions as affected by split nitrogen application. Field Crops Research 94, 86-97.

López-Bellido, L., López-Bellido, R.J., López-Bellido, F.J., 2006. Fertilizer nitrogen efficiency in durum wheat under rainfed Mediterranean conditions: Effect of split application. Agronomy Journal 98, 55-62.

Loveland, T.R., Reed, B.C., Brown, J.F., Ohlen, D.O., Zhu, Z., Yang, L. and Merchant, J.W., 2000. Development of a global land cover characteristics database and IGBP DISCover from 1 km AVHRR data. International Journal of Remote Sensing 21, 1303 - 1330.

Luck, G.W., 2007. The relationships between net primary productivity, human population density and species conservation. Journal of Biogeography 34, 201-212.

Lutz, W., Sanderson, W. and Scherbov, S., 1997. Doubling of world population unlikely. Nature 387, 803-805.

Ma, H.Y., Rae, A., Huang, J.K. and Rozelle, S., 2004. Chinese animal product consumption in the 1990s. Australian Journal of Agricultural and Resource Economics 48, 569-590.

MacDonald, R.B. and Hall, F.G., 1980. Global Crop Forecasting. Science 208, 670-679.

Maggi M. and Stroppiana D., 2002. Advantages and drawbacks of NOAA-AVHRR and SPOT-VGT for burnt area mapping in a tropical savannah ecosystem. Canadian Journal of Remote Sensing 28.2: 231-245.

Martín de Santa Olalla Mañas. F, F.J. Cabanero Soria and A. Artigao Ramirez, Validation of the CORINE land cover database in a pilot zone under

semi-arid conditions in LaMancha (Spain), Cybergeo European Journal of Geography, p. 235.

Maseli F., 2004. Monitoring forest conditions in a protected Mediterranean coastal area by analysis of multiyear NDVI data. Remote Sensing of Environment. 89: 423-433.

Maselli, F., Conese, C., Petkov, L. and Gilabert, M.-A., 1992. Use of NOAA-AVHRR NDVI data for environmental monitoring and crop forecasting in the Sahel. Preliminary results. International Journal of Remote Sensing 13, 2743 - 2749.

Matras, J., 1973. Population and Societies, Englewood Cliffs, NJ: Prentice-Hall.

Matthews, R.B. and Stephens, W., 2002. Crop-soil simulation models: Applications in Developing Countries. CAB International, Wallingford, UK.

Mayaux, P., Bartholomé, E., Fritz, S. and Belward, A., 2004. A new land-cover map of Africa for the year 2000. Journal of Biogeography 31, 861-877.

McNairn, H., Champagne, C., Shang, J., Holmstrom, D. and Reichert, G., 2009. Integration of optical and Synthetic Aperture Radar (SAR) imagery for delivering operational annual crop inventories. Isprs Journal of Photogrammetry and Remote Sensing 64, 434-449.

Mo, X., Liu, S., Lin, Z., Xu, Y., Xiang, Y. and McVicar, T.R., 2005. Prediction of crop yield, water consumption and water use efficiency with a SVAT-crop growth model using remotely sensed data on the North China Plain. Ecological Modelling 183, 301-322.

Morgan, W.B. and Solarz, J.A., 1994. Agricultural crisis in Sub-Saharan Africa - Development constraints and policy problems. Geographical Journal 160, 57-73.

Murakami T., Ogawa S., Ishitsuka N., Kumagai K. and Saito G., 2001. Crop discrimination with multi-temporal SPOT/HRV data in the Saga Plains, Japan. International Journal of Remote Sensing 22.7: 1335-1348.

Murthy, C.S., Sai, M.V.R.S., Kumari, V.B. and Roy, P.S., 2007. Agricultural drought assessment at disaggregated level using AWiFS/WiFS data of Indian Remote Sensing satellites. Geocarto International 22, 127-140.

Nachtergaele, F.O.F., 2000. Soil Resources Information. In "Global environmental databases - Present situation; future directions". Tateishi, R and Hastings, D. (Eds.). ISPRS Working Group IV/6 (1996-2000). Geocarta International Centre, Hong Kong, China. pp. 157-177.

NASA, 1984. AgRISTARS: Agriculture and Resources Inventory Surveys through Aerospace Remote Sensing, Research Report, JSC-18920, fiscal year 1983, National Aeronautics and Space Administration, Lyndon B. Johnson Space Center, Houston, Texas, 74 pp.

Nègre, T., Rembold, F., Rojas, O., 2004, Integrated agricultural Monitoring and Yield Forecasting for Eastern Africa: the JRC Approach. Proceedings of the JRC-FAO Workshop: Crop and Rangeland Monitoring in Eastern

Africa for Early Warning and Food Security, Nairobi 28-30 January 2003.

Nguyen, T., T. H., De Bie, C.A.J.M., Ali, A., Smaling, E.M.A. and Hoanh, C.T., 2010. Mapping the irrigated rice cropping patterns of the Mekong delta, Vietnam through hyper-temporal SPOT NDVI image analysis. International Journal of Remote Sensing (in press).

Nicholls, T., 2010. Mitigate food loss to feed more people right now. Nature 466, 920-920.

Nicholson, S.E. and Farrar, T.J., 1994. The influence of soil type on the relationships between NDVI, rainfall, and soil moisture in semiarid Botswana. I. NDVI response to rainfall. Remote Sensing of Environment 50, 107-120.

NRSA, 1995. Integrated Mission for Sustainable Development, Integrated Survey Group, National Remote Sensing Agency, Hyderabad.

Oetter, D.R., Cohen, W.B., Berterretche, M., Maiersperger, T.K. and Kennedy, R.E., 2001. Land cover mapping in an agricultural setting using multiseasonal Thematic Mapper data. Remote Sensing of Environment 76, 139-155.

Palm, R. and Dagnelie, P., 1993. Tendance générale et effets du climat dans la prévision des rendements agricoles des différents pays des C.E. Publication EUR 15106F. Luxembourg, Luxembourg: Office for Official Publications of the EU.

Pradhan, S., 2001. Crop area estimation using GIS, remote sensing and area frame sampling. International Journal of Applied Earth Observation and Geoinformation 3, 86-92.

Prenzel, B., 2004. Remote sensing-based quantification of land-cover and land-use change for planning. Progress in Planning 61, 281-299.

Pruitt, E.L., 1979. The Office of Naval Research and Geography. Annals of the Association of American Geographers 69, 103-108.

Ramankutty, N., T. Hertel, H.-L. Lee, and Rose, S.K., 2007. Global Agricultural Land Use Data for Integrated Assessment Modeling, in Human-Induced Climate Change: An Interdisciplinary Assessment, edited by M. Schlesinger, *et al*., pp. 252-265, Cambridge University Press, New York.

Ran, Y.H., Li, X. and Lu, L., 2010. Evaluation of four remote sensing based land cover products over China. International Journal of Remote Sensing 31, 391-401.

Rao, S.V.C.K., 1995. Cropping Pattern Analysis using Temporal IRS-LISS II Data – A Case Study of Madnur Watershed, Nizamabad District, AP. Agriculture and Soil Division, IIRS, NRSA, Dehradun.

Ray, S.S. and Dadhwal, V.K., 2001. Estimation of crop evapotranspiration of irrigation command area using remote sensing and GIS. Agricultural Water Management 49, 239-249.

Reidsma, P., Ewert, F., Boogaard, H. and van Diepen, K., 2009. Regional crop modelling in Europe: The impact of climatic conditions and farm characteristics on maize yields. Agricultural Systems 100, 51-60.

Rembold F., 2004. Application – MARS-FOOD crop monitoring for Food Security in Eastern Africa. http://www.iao.florence.it/documentation/glcn/app1.pdf (accessed 8 Dec'04).

Richards, J. A., 1986. Remote Sensing Digital Image Analysis. Springer-Verlag, Berlin, Heidelburg, New York, London, Paris, Tokyo, pp. 206-225.

Ringrose S., Vanderpost C. and Mattheson W., 1996. The use of integrated remotely sensed and GIS data to determine the causes of vegetation cover change in southern Botswana. Applied Geography 16: 225-242.

Roberto, B., 1993. On the use of NDVI profiles as a tool for agricultural statistics: The case study of wheat yield estimate and forecast in Emilia Romagna. Remote Sensing Environment 45, 311-326.

Roebeling, R., Van Putten, E., Genovese, G. and Rosema, A., 2004. Application of Meteosat derived meteorological information for crop yield predictions in Europe. International Journal of Remote Sensing 25(23), 5389–5401.

Runge, E.C.A. and Odell, R.T., 1958. The Relation between Precipitation, Temperature and the Yield of Corn on the Agronomy South Farm, Urbana, Illinois. Agronomy Journal 50, 448-454.

Rugege, D., 2002. Regional analysis of maize-based land use systems for early warning applications. ITC Dissertation no. 87. Wageningen University, Wageningen, The Netherlands. 122 pp.

Sachs, J., Remans, R., Smukler, S., Winowiecki, L., Andelman, S.J., Cassman, K.G., Castle, D., DeFries, R., Denning, G., Fanzo, J., Jackson, L.E., Leemans, R., Lehmann, J., Milder, J.C., Naeem, S., Nziguheba, G., Palm, C.A., Pingali, P.L., Reganold, J.P., Richter, D.D., Scherr, S.J., Sircely, J., Sullivan, C., Tomich, T.P. and Sanchez, P.A., 2010. Monitoring the world's agriculture. Nature 466, 558-560.

Sarkar, S. and Kafatos, M., 2004. Interannual variability of vegetation over the Indian sub-continent and its relation to the different meteorological parameters. Remote Sensing of Environment 90, 268-280.

Sarris, A.H. and Taylor, L., 1976. Cereal stocks, food aid and food security for the poor. World Development 4, 967-976.

Schlamadinger, B., Bird, N., Johns, T., Brown, S., Canadell, J., Ciccarese, L., Dutschke, M., Fiedler, J., Fischlin, A., Fearnside, P., Forner, C., Freibauer, A., Frumhoff, P., Hoehne, N., Kirschbaum, M.U.F., Labat, A., Marland, G., Michaelowa, A., Montanarella, L., Moutinho, P., Murdiyarso, D., Pena, N., Pingoud, K., Rakonczay, Z., Rametsteiner, E., Rock, J., Sanz, M.J., Schneider, U.A., Shvidenko, A., Skutsch, M., Smith, P., Somogyi, Z., Trines, E., Ward, M., Yamagata, Y., 2007. A

synopsis of land use, land-use change and forestry (LULUCF) under the Kyoto Protocol and Marrakech Accords. Environmental Science & Policy 10, 271-282.

Schnittker, J.A., 1973. The 1972-73 Food Price Spiral. Brookings Papers on Economic Activity 1973, 498-507.

Schoeneberger, P.J. and Wysocki, D.A., 2002. A Geomorphic Description System (Version 3.1); National Soil Survey Center, Natural Resources Conservation Service, USDA, Lincoln, NE.

Searchinger, T., Heimlich, R., Houghton, R.A., Dong, F., Elobeid, A., Fabiosa, J., Tokgoz, S., Hayes, D. and Yu, T.-H., 2008. Use of U.S. Croplands for Biofuels Increases Greenhouse Gases Through Emissions from Land-Use Change. Science 319, 1238-1240.

Sellers P.J., 1985. Canopy reflectance, photosynthesis and transpiration. International Journal of Remote Sensing **6**: 1335-1372.

Shao, Y., Lunetta, R.S., Ediriwickrema, J. and Liames, J., 2010. Mapping Cropland and Major Crop Types across the Great Lakes Basin using MODIS-NDVI Data. Photogrammetric Engineering and Remote Sensing 76, 73-84.

Shaw, D.J., 2007. World food security. A history since 1945. Basingstoke, United Kingdom: Palgrave MacMillan.

Smakhtin, V., Thenkabail P. and Gamage N., 2005. Developing on-line near-real time drought monitoring system for southwest Asia. Submitted for publication in the monograph "Crop growth monitoring and yield forecasting in Central Asia", prepared by the European Commission.

Sombroek,W.G. and J.Antoine, 1994. The use of geographic information systems in land resource appraisal. Outlook on Agriculture, London.

Souza, C., Firestone, L., Silva, L.M. and Roberts, D., 2003. Mapping forest degradation in the Eastern Amazon from SPOT 4 through spectral mixture models. Remote Sensing of Environment 87, 494-506.

Stöckli, R. and Vidale, P.L., 2004. European plant phenology and climate as seen in a 20-year AVHRR land-surface parameter dataset. International Journal of Remote Sensing 25, 3303 - 3330.

Stow, D.A., Hope, A., McGuire, D., Verbyla, D., Gamon, J., Huemmrich, F., Houston, S., Racine, C., Sturm, M., Tape, K., Hinzman, L., Yoshikawa, K., Tweedie, C., Noyle, B., Silapaswan, C., Douglas, D., Griffith, B., Jia, G., Epstein, H., Walker, D., Daeschner, S., Petersen, A., Zhou, L.M. and Myneni, R., 2004. Remote sensing of vegetation and land-cover change in Arctic Tundra Ecosystems. Remote Sensing of Environment 89, 281-308.

Supit, I., Hooijer, A.A. and Van Diepen, C.A. (editors), 1994. System description of the WOFOST 6.0 Crop Growth Simulation System. Publication EUR 15956 EN, Luxembourg, Luxembourg: Office for Official Publications of the European Union.

Supit, I., 1997. Predicting national wheat yields using a crop simulation and trend models. Agricultural and Forest Meteorology 88, 199-214.

Swain P.H. and Davis S.M., 1978. Remote Sensing: The Quantitative Approach, McGraw Hill Book Company, New York, 375 p.

Thenkabail, P.S., Gamage M.S.D.N. and Smakhtin V.U., 2004. The use of remote sensing data for drought assessment and monitoring in Southwest Asia. Research report 85. IWMI International Water Management Institute, Colombo, Sri Lanka; 25pp.

Thompson, L.M., 1962. An Evaluation of Weather Factors in the Production of Corn. Report No. 12T, Center for Agricultural and Rural Development, Iowa State University, Ames, USA.

Thorburn, P.J., Biggs, J.S., Collins, K. and Probert, M.E., 2010. Using the APSIM model to estimate nitrous oxide emissions from diverse Australian sugarcane production systems. Agriculture Ecosystems & Environment 136, 343-350.

Tilman D., Cassman K.G., Matson P., Naylor R. and Polasky S., 2002. Agricultural Sustainability and Intensive production practices. Nature 418(6898): 671-677.

Townshend, J.R., Latham, J., Arino, O., Balstad, R., Belward, A., conant, R., Elvidge, C., Feuquay, J., El Hadani, D., Herold, M., Janetos, A., Justice, C.O., Liu Jiyuan, Loveland, T., Nachtergaele, F., Ojima, D., Maiden, M., Palazzo, F., Schmullius, C., Sessa, R., Singh, A., Tschirley, J., and Yamamoto, H., 2008. Integrated Global Observations of the Land: an IGOS-P Theme. IGOL Report No. 8, GTOS 54.

Tsiligirides, T.A., 1998. Remote sensing as a tool for agricultural statistics: a case study of area frame sampling methodology in Hellas. Computers and Electronics in Agriculture 20, 45-77.

Uchida S., 2001. Discrimination of Agricultural Land use using multi-temporal NDVI data. In: Proceedings of 22nd Asian Conference of Remote Sensing. Singapore, Centre for Remote Imaging, Sensing and Processing (CRISP). 5-9 Nov. Vol 2: 813-818.

Unganai L.S. and Kogan F.N., 1998. Drought Monitoring and Corn Yield estimation in Southern Africa from AVHRR Data. Remote Sensing of Environment 63: 219-232.

United Nations, 1973. Determinants and Consequences of Population Trends. New York: United Nations.

United Nations, 1974. World Food Conference. The World Food Problem. Proposals for National and International Action. November 5th–16th, Document E/CONF. 65/4, UN, New York, USA. 187 pp.

United Nations, 2000. Millennium Declaration. UN General Assembly Resolution 55/2, New York, USA: UN.

United Nations, 2005. World Population Prospects: The 2004 Revision. UN Population Division, New York available at

http://www.un.org/esa/population/publications/ WPP2004/World_Population_2004_chart.pdf.

Ustin, S.L. and Gamon, J.A., 2010. Remote sensing of plant functional types. New Phytologist 186, 795-816.

Vandermeulen, V., Gellynck, X., Van Huylenbroeck, G., Van Orshoven, J. and Bomans, K., 2009. Farmland for tomorrow in densely populated areas. Land Use Policy 26, 859-868.

Van Diepen, C.A., Wolf, J., Vankeulen, H. and Rappoldt, C., 1989. WOFOST - A SIMULATION-MODEL OF CROP PRODUCTION. Soil Use and Management 5, 16-24.

Van Diepen, C.A., Boogaard, H.L., Supit, I. and LAZĂR, C., Orlandi, S., Van der Goot, E., Schapendonk, A.H.C.M., 2004. Methodology of the MARS crop yield forecasting system. Vol. 2 agrometeorological data collection, processing and analysis. Publication EUR 21291 EN/2, Luxembourg, Luxembourg: Office for Official Publications of the European Union.

Van Ittersum, M.K. and Rabbinge, R., 1997. Concepts in production ecology for analysis and quantification of agricultural input-output combinations. Field Crops Research 52, 197-208.

Van Ittersum, M.K., Leffelaar, P.A., Van Keulen, H., Kropff, M.J., Bastiaans, L. and Goudriaan, J., 2003. On approaches and applications of the Wageningen crop models. European Journal of Agronomy 18, 201-234.

Van Keulen, H., Van Laar, H.H. and Rabbinge, R., (Eds.) 2008. 40 years theory and model at Wageningen UR. Wageningen University and Research Centre, Wageningen, The Netherlands.

Van Laar, H.H., Goudriaan, J. and Van Keulen, H., 1997. SUCROS97: Simulation of crop growth for potential and water-limited production situations. As applied to spring wheat. Quantitative Approaches in Systems Analysis No. 14. AB-DLO/GTPE, Wageningen, The Netherlands. 52 pp. + Ann.

Venus, V. and Rugege, D., 2004. Combined use of polar orbiting and geo-stationary satellites to improve time interpolation in dynamic crop models for food security assessment. In: ISPRS 2004: Proceedings of the XXth ISPRS congress: Geo-imagery bridging continents, 12-23 July 2004, Istanbul, Turkey. Comm. VII PS, WG VII/2. pp. 212-219.

Verbeiren, S., Eerens, H., Piccard, I., Bauwens, I. and Van Orshoven, J., 2008. Sub-pixel classification of SPOT-VEGETATION time series for the assessment of regional crop areas in Belgium. International Journal of Applied Earth Observation and Geoinformation 10, 486-497.

Verburg, P.H., Soepboer, W., Veldkamp, A., Limpiada, R., Espaldon, V. and Mastura, S.S.A., 2002. Modeling the spatial dynamics of regional land use: The CLUE-S model. Environmental Management 30, 391-405.

Von Braun, J., 2010. Strategic body needed to beat food crises. Nature 465, 548-549.

Vossen, P., 1990. Modèles agrométéorologiques pour le suivi des cultures et la prévision de rendements des grandes régions des Communautés Européennes. In: Proceedings of the Conference on The Application of Remote Sensing to Agricultural Statistics, Varese, Italy, 10–11 Octobre 1989, Publication EUR 12581 EN Luxembourg, Luxembourg: Office for Official Publications of the EU. pp. 75–84.

Vossen, P., 1992. Forecasting national crop yields of EU countries: the approach developed by the Agricultural Project. In: Proceedings of the Second Conference on The Application of Remote Sensing to Agricultural Statistics, Belgirate, Italy, 26–27 November 1991, Publication EUR 14262 EN Luxembourg, Luxembourg: Office for Official Publications of the EU. pp. 159–176.World Food Summit, 1996. Rome Declaration on World Food Security**.** FAO, Rome.

Walker P.A. and Mallawaarachchi T., 1998. Disaggregating Agricultural Statistics using NOAA-AVHRR NDVI. Remote Sensing of Environment 63: 112-125.

Wardlow, B.D. and Egbert, S.L., 2008. Large-area crop mapping using time-series MODIS 250 m NDVI data: An assessment for the U.S. Central Great Plains. Remote Sensing of Environment 112, 1096-1116.

Weiss J.L., Gutzler D.S., Coonrod J.E.A. and Dahm C.N., 2004. Long-term vegetation monitoring with NDVI in a diverse semi-arid setting, central New Mexico, USA." Journal of Arid Environments 58(2): 249-272.

Wessels, K.J., De Fries, R.S., Dempewolf, J., Anderson, L.O., Hansen, A.J., Powell, S.L. and Moran, E.F., 2004. Mapping regional land cover with MODIS data for biological conservation: Examples from the Greater Yellowstone Ecosystem, USA and Pará State, Brazil. Remote Sensing of Environment 92, 67-83.

Williams, D.L., Goward, S. and Arvidson, T., 2006. Landsat: Yesterday, today, and tomorrow. Photogrammetric Engineering and Remote Sensing 72, 1171-1178.

Wood, S., Sebastian, K. and Scherr, S.J. 2000. Pilot analysis of global ecosystems- Agroecosystems. World Resources Institute, Washington DC. 110 pp.

World Food Summit, 1996. Rome Declaration on World Food Security**.** FAO, Rome, Italy.

Yang, C., Everitt, J.H., Fletcher, R.S., and Murden, D., 2007. Using high resolution QuickBird imagery for crop identification and area estimation. Geocarto International, 22, 219 - 233

Young, A. 1998. Land Resources, Now and for the Future. Cambridge University Press. ISBN 0521-785596 (paperback). 319 pp.

Zeijl-Rozema, A.E. Van, A.iglesias, P.Pilesjo and D.Tantram, 1997. LUCC Report Series No. I. Institut Cartografie de Caialunya. Barcelona. 1997. 143 p.

Zhang, J.X. and Zhang, Y.H., 2007. Remote sensing research issues of the National Land Use Change Program of China. Isprs Journal of Photogrammetry and Remote Sensing 62, 461-472.

Zhou, H.J., Van Rompaey, A. and Wang, J.A., 2009. Detecting the impact of the "Grain for Green" program on the mean annual vegetation cover in the Shaanxi province, China using SPOT-VGT NDVI data. Land Use Policy 26, 954-960.

Appendix 1

1 Personal Information

Please note that the questions marked with (*) are compulsory to be filled in

Name: Organization: City/Town:

Country: Email Address:

***2 Please mention your research interest/ nature of work**

Research in the field of agricultural land use mapping

Research in the field of agricultural land use monitoring

Research/ policy related work in government organization

Research/ policy related work in global/regional level organization

Research/ policy related work in an agricultural company

Other (please specify)

3 Do you use agricultural land use information?

Yes No

4 Do you require agricultural land use maps/information, especially the specific crop type maps?

Yes No

5 What kind of agricultural land use maps/ information do you require?

Agricultural Land use maps Specific crop type maps

Agricultural statistical data Primary survey data

Other (please specify)

6 What is your required update frequency of agricultural land use maps/information?

One year More than one year but less than 5 years

More than 5 years Other (please specify)

7. The required spatial level of detail is:

(Please select from 7.a to 7.c)

7.a. If provided as aggregated data by administrative region:

NUTS 1 (Country level) NUTS 2 (Region/province level)

NUTS 3 (Municipal/district level) Field level

7.b. If presented as raster maps:

1000 m 500 m 250 m 100 m 10 m

7.c. If provided at the following spatial levels:

Agro-ecological zones Soil units Land use zones/units

Other (please specify)

8 Which type of agricultural land use information do you use?

Agricultural Land use maps

Specific crop type maps

CORINE land cover map

Agricultural Statistics Data

Other (please specify)

9. Please mention the following information about the land use maps you use:

Name: Source: Description:

10. Update frequency (number of years) of the land use maps:

One year 2-4 years 5 years > 5 years

11. The spatial level of detail in these land use maps is: Please select from 11.a to 11.c

11.a. If provided as aggregated data by administrative region:

NUTS 1 (Country level) NUTS 2 (Region/province level)

NUTS 3 (Municipal/district level) Field level

11.b. If presented as raster maps:

1000 m 500 m 250 m 100 m 10 m

11.c. If provided at the following spatial levels:

Agro-ecological zones Soil units Land use zones/units

Other (please specify)

12. Your opinion regarding aforementioned available agricultural land use information source(s)

Please select one for each statement from Strongly Disagree, Disagree, Indifferent, Agree and Strongly Agree

Their availability is not a problem

Their updating frequency is sufficient

Their documentation (Meta data) is available
Their documentation is sufficiently explained
These differentiate required land cover/land use classes
Their spatial level of detail is sufficient
Other remarks about such data (please specify)

13. Please mention the following information about the specific crop maps you use:

Name: Source: Description:

14. Update frequency (number of years) of the specific crop maps:

One year 2-4 years 5 years > 5 years

15. The spatial level of detail in these specific crop maps is: Please select from 15.a to 15.c

15.a. If provided as aggregated data by administrative region:

NUTS 1 (Country level) NUTS 2 (Region/province level)
NUTS 3 (Municipal/district level) Field level

15.b. If presented as raster maps:

1000 m 500 m 250 m 100 m 10 m

15.c. If provided at the following spatial levels:

Agro-ecological zones Soil units Land use zones/units
Others (please specify)

16. Your opinion regarding aforementioned available specific crop maps

Please select one for each statement from Strongly Disagree, Disagree, Indifferent, Agree and Strongly Agree

Their availability is not a problem
Their updating frequency is sufficient
Their documentation (Meta data) is available
Their documentation is sufficiently explained
These differentiate required land cover/land use classes
Their spatial level of detail is sufficient
Other remarks about such data (please specify)

17. Your opinion regarding CORINE land cover map

Please select one for each statement from Strongly Disagree, Disagree, Indifferent, Agree and Strongly Agree

Its availability is not a problem
Its updating frequency is sufficient
Its documentation (Meta data) is available
Its documentation is sufficiently explained
It differentiate required land cover/land use classes
Its spatial level of detail is sufficient

18. Please mention the following information about the agricultural statistical data you use:

Name: Source: Description:

19. Update frequency (number of years) of the agricultural statistical data you use:

One year 2-4 years 5 years > 5 years

20. The spatial level of detail in these agricultural statistical data you use:

20.a. If provided as aggregated data by administrative region:

NUTS 1 (Country level) NUTS 2 (Region/province level)
NUTS 3 (Municipal/district level) Field level

20.b. If presented as raster maps:

1000 m 500 m 250 m 100 m 10 m

20.c. If provided at the following spatial levels:

Agro-ecological zones Soil units Land use zones/units
Other (please specify)

21. Your opinion regarding aforementioned available agricultural statistical data you use:

Please select one for each statement from Strongly Disagree, Disagree, Indifferent, Agree and Strongly Agree

Their availability is not a problem
Their updating frequency is sufficient

Their documentation (Meta data) is available
Their documentation is sufficiently explained
These differentiate required land cover/land use classes
Their spatial level of detail is sufficient
Other remarks about such data (please specify)

22. Please mention the following information about the primary survey data you use:

Name: Source: Description:

22. Update frequency (number of years) of the primary survey data you use:

One year 2-4 years 5 years > 5 years

23. The spatial level of detail in these primary survey data you use: Please select from 23.a to 23.c

23.a. If provided as aggregated data by administrative region:

NUTS 1 (Country level) NUTS 2 (Region/province level)
NUTS 3 (Municipal/district level) Field level

23.b. If presented as raster maps:

1000 m 500 m 250 m 100 m 10 m

24.c. If provided at the following spatial levels:

Agro-ecological zones Soil units Land use zones/units
Other (please specify)

24. Your opinion regarding aforementioned available agricultural statistical data you use:

Please select one for each statement from Strongly Disagree, Disagree, Indifferent, Agree and Strongly Agree

Their availability is not a problem
Their updating frequency is sufficient
Their documentation (Meta data) is available
Their documentation is sufficiently explained

These differentiate required land cover/land use classes
Their spatial level of detail is sufficient
Other remarks about such data (please specify)

25. Do you require additional data set(s) containing agricultural land use information?

Yes No

26. What kind of agricultural land use maps/ information do you require?

Agricultural Land use maps Specific crop type maps
Agricultural statistical data Primary survey data
Other (please specify)

27 What is your required update frequency of agricultural land use maps/information?

One year More than one year but less then 5 years
More than 5 years Other (please specify)

28. The required spatial level of detail is:
(Please select from 28.a to 28.c)

28.a. If provided as aggregated data by administrative region:

NUTS 1 (Country level) NUTS 2 (Region/province level)
NUTS 3 (Municipal/district level) Field level

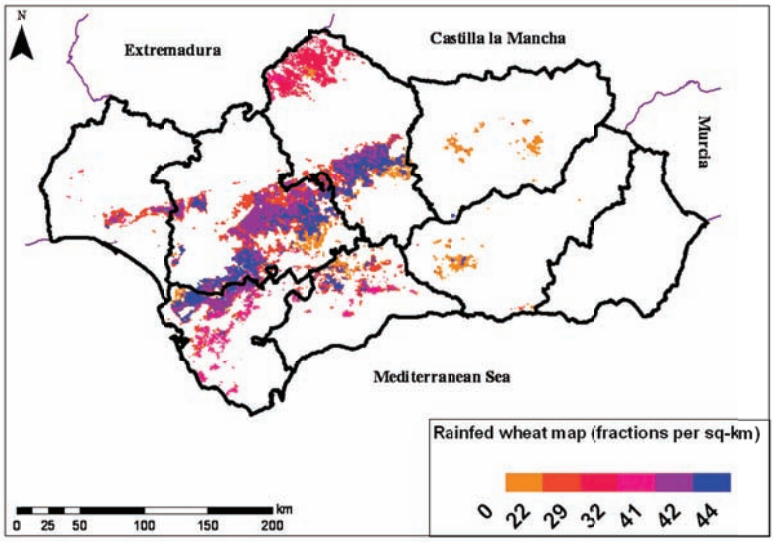
28.b. If presented as raster maps:

1000 m 500 m 250 m 100 m 10 m

28.c. If provided at the following spatial levels:

Agro-ecological zones Soil units Land use zones/units
Other (please specify)

29 Following is the rainfed wheat map produced by us.
Please consult the BOX below map for details
specify your opinion regarding this map in question 30:



The rainfed wheat map of Andalucia (fractions km2 from 2001-05)
Map characteristics:

Map characteristics	**Details**
Description	It is an example of hypertemporal mapping techniques developed by us.
Data used	**Remote sensing data:** 10 daily Maximum Value Composite (MVC) SPOT VEGETATION NDVI images (1998-2006) **Primary segments data:** segments data. 1428 segments (700mX700m)
Map resolution	The spatial resolution of the map is 1 km^2
Availability	At the end of harvesting season (when the data on crop areas is available)
Update frequency	Annually
Validation	**This map is produced by directly using the Field data (primary segments data).** The other rainfed wheat map produced by using NDVI images and rainfed wheat area statistical data by municipality shows a 90 % agreement with the primary segments data

30. Please give your opinion on the map shown above:

Please specify the reason of your (dis)-agreement in the space provided at the end of this question

It meets my required level of spatial detail
It meets my requirement of specific land cover/ land use information
It meets my required level of quality/ accuracy assessment
It meets my requirement of regular update
I would like to use the method to produce specific crop maps of my own interest
Other (please specify)

Please specify the reason of your (dis)-agreements

31. Do you use yield data such as estimated or reported yield data of specific crops?

Yes No

32. Do you use yield data such as estimated or reported yield data of specific crops?

Yes No

33. Which type(s) of crop yield data do you require?

Crop yield estimated by a crop growth model Agricultural statistical data
Primary survey data

34. When do you require crop yield estimates?

At the end of harvesting Season
One month prior to the end of harvesting Others (please specify)

35. The required spatial level of detail is:
(Please select from 35.a to 35.c)

35.a. If provided as aggregated data by administrative region:

NUTS 1 (Country level) NUTS 2 (Region/province level)
NUTS 3 (Municipal/district level) Field level

35.b. If presented as raster maps:

1000 m 500 m 250 m 100 m 10 m

35.c. If provided at the following spatial levels:

Agro-ecological zones Soil units Land use zones/units
Other (please specify)

Please mention what kind of crop yield estimates are available to you (For example the model/method used and source of this information)

36. Please mention the following information about the land use maps you use:

Name: Source: Description:

37. Update frequency (number of years) of the land use maps:

One year 2-4 years 5 years > 5 years

38. The spatial level of detail in these land use maps is: Please select from38.a to 38.c

38.a. If provided as aggregated data by administrative region:

NUTS 1 (Country level) NUTS 2 (Region/province level)
NUTS 3 (Municipal/district level) Field level

38.b. If presented as raster maps:

1000 m 500 m 250 m 100 m 10 m

39.c. If provided at the following spatial levels:

Agro-ecological zones Soil units Land use zones/units
Other (please specify)

40. Your opinion regarding aforementioned available yield estimates

Please select one for each statement from Strongly Disagree, Disagree, Indifferent, Agree and Strongly Agree

Its availability is not a problem
Its updating frequency is sufficient
Its documentation (Meta data) is available
Its documentation is sufficiently explained
Its spatial level of detail is sufficient
It meets my required level of quality/ accuracy assessment

41. Do you require additional crop yield data?

Yes No

42. Which type(s) of additional crop yield data do you require?
Crop yield estimated by a crop growth model Agricultural statistical data
Primary survey data

43. When do you require these additional crop yield estimates?
At the end of harvesting Season
One month prior to the end of harvesting
Others (please specify)

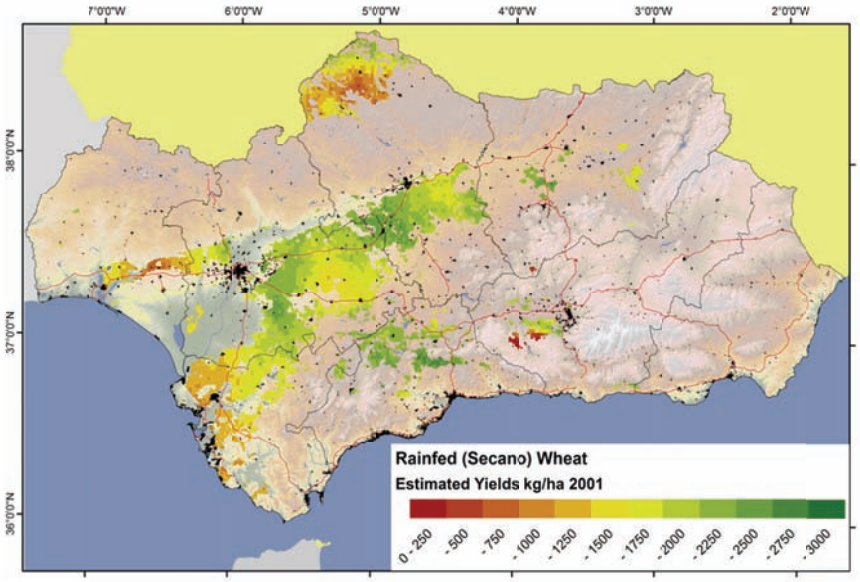
44. The required spatial level of detail is:
(Please select from 44.a to 44.c)

44.a. If provided as aggregated data by administrative region:
NUTS 1 (Country level) NUTS 2 (Region/province level)
NUTS 3 (Municipal/district level) Field level

44.b. If presented as raster maps:
1000 m 500 m 250 m 100 m 10 m

44.c. If provided at the following spatial levels:
Agro-ecological zones Soil units Land use zones/units
Other (please specify)

45. Following is the crop yield map prepared by us:



The estimated rainfed wheat yield map of Andalucia (2001)
Please consult the following BOX for details and specify your opinion regarding this map in question 46:

characteristics	Details
Description	It is based on a crop growth model with improved regional incorporating remote sensing data The model estimates actual yield as a function of **available light, temperature, photosynthesis and compounded constraints reflected by canopy heating.**
Data used	**Remote sensing data:** MODIS LST PRODUCT **Weather Data:** Daily max. and min. temp., Relative humidity, Evapotranspiration **Crop information** **Management:** seeding rate and sowing date.
Spatial resolution	It is a raster map with 1 km^2
Availability	At the end of crop growing season
Update frequency	Annual
Validation	Validated by using actual rainfed wheat yields (322 segments of 700 m X 700 m). The accuracy of estimated yield maps is 92%.

46. Please give your opinion on the map shown above:

Please specify the reason of your (dis)-agreement in the space provided at the end of this question

It meets my requirement of spatial level of detail

It meets my requirement of timely availability

It meets my required level of quality/ accuracy assessment

It meets my requirement of quantitative yield assessment

I would like to use this method for my own crop and area of interest

Please specify the reason of your (dis)-agreements

Thank you for your time!

Note: The questionnaire was designed in a way that only the relevant questions were displayed on the screen of the respondents. For this purpose, jump statements were used.

Summary

Policy makers, responsible for food security and land use planning, need accurate and timely information on crop areas and production at regional level. Such information should be regularly updated to monitor changes in land use and agricultural production. The value and relevance of such information substantially increases when it is spatially explicit. Currently, governments and international organizations such as the Food and Agriculture Organization of the United Nations (FAO), the EUROpean STATistical Office (EUROSTAT) and the United States Geological Service (USGS) compile and distribute statistical data on crop areas and production. Such data are often tabular data sets pertaining to administrative units, providing area and production totals, but without any indication of the spatial distribution of agricultural land use in that unit. Consequently, such information is insufficient for monitoring of crop production and for food security studies.

The objective of this study was to develop methods to quantitatively monitor and map crop production systems, including assessment of a prototype crop growth model that uses remotely sensed input data. The research had two focal points, namely land use mapping and crop yield estimation. We developed methods to generate agricultural land use maps and crop yield maps by combining remote sensing, Geographical Information Systems (GIS), crop statistical data and crop models. Hypertemporal SPOT-Vegetation (VGT) Sensor's 10-day composite Normalized Difference Vegetation Index (NDVI) images (S10 product), at 1 km^2 resolution, were used to disaggregate agricultural land use data in the form of crop maps showing fractions of area cropped by grid cell. The NDVI images were used to stratify the study areas (Nizamabad, India and Andalucía, Spain) in mapping units reflecting homogeneity in space and time.

The classification results of SPOT-Vegetation NDVI images of Nizamabad showed that the study area can be stratified in 11 distinct mapping units. These mapping units were then related to an existing land cover map compiled from high resolution images, reported crop areas by sub-district, and crop calendar information. The existing land cover map only reports, at high spatial detail (1: 50 000), the location of cropped fields for 1994-95. The NDVI-derived map on the other hand, was based on a number of images acquired between 1998 and

2002, and therefore reflected spatial and temporal variability. This NDVI-derived map, after linking with crop areas and crop calendar information, depicts *what* is grown *where*. Thus an improved map was generated, containing baseline information on both land cover and land use.

For Andalucía, the areas of the NDVI classes per municipality were correlated with the reported cropped areas per municipality to disaggregate the crop statistics. Relating statistical data on areas cropped per municipality with the NDVI-based unit map showed that the selected crops were significantly related to specific NDVI-based map units. The results were validated by using primary field data. These data were collected by the Spanish government from 2001 to 2005 through grid samples of 700 m × 700 m within agricultural areas. In this part of the research, a properly tested methodology for preparing crop area maps, useful for monitoring crop performance, was developed.

However, it was assumed that the NDVI data fully express the combined influences of varying soil, terrain, weather and land use conditions. This assumption was tested with respect to soil types and soil-geomorphological information. It was shown that the additional use of soil data increased the explained variability by only 1 %. Validity of the assumption made was also confirmed by applying the jack-knife procedure to establish the relative importance of each predictor of crop areas. The assumption that NDVI can serve as an indicator of the combined influences of varying spatial conditions was thus confirmed.

After defining the methods of generating agricultural land use information, we proceeded with the evaluation of the output of a crop growth model (Cf-Water) driven by remotely sensed data that estimates actual crop yields at a 1-km^2 resolution. The evaluation was performed by (i) comparing the output of Cf-Water at regional level (province) and the output of an operational crop growth model, CGMS (Crop Growth Monitoring System), of the European Union's Monitoring Agriculture with Remote Sensing (MARS) program with published agricultural statistics and (ii) accuracy assessment of the output of Cf-Water using primary field data. Note that the CGMS, after time-trend adjustments, only reports generalized estimates of actual crop yields at NUTS-0 (country) level, after time-trend adjustments. The Cf-Water model has lower data requirements than CGMS which requires also soil properties and historical yield

data. Our results demonstrate that C*f*-Water has a high potential to support food security studies; it showed excellent agreement with both primary field data and published statistical data. However, the model has only been applied in a single case study, thus before recommending the model as an operational system, more extensive testing is required for different crops, years and regions.

To get an impression of the relevance of the study, the opinion of potential users was solicited on available land use datasets and the generated outputs of the thesis (crop area and yield maps). Analysis of the responses showed that the respondents were not satisfied with the updating frequency and the spatial level of detail of the land use data already available to them. On the other hand, 72% of the respondents were satisfied with the spatial level of detail, the type of land use data and the optional updating frequency of the generated crop maps. Responses to questions on the generated rainfed wheat yield map revealed that 70% of the respondents were satisfied with the spatial level of detail in the yield map, but that they would prefer to have such information available prior to harvesting, whereas our method only provides yield estimates after harvesting. The products generated in Chapters 3-5 are of interest to potential users. They were keen to use the methods developed to generate of crop area maps (75%) and crop yield maps (70%).

On the basis of results described in this thesis, we conclude that use of hypertemporal NDVI data is highly suitable to map crop areas through data-mining of existing statistical data. Though soil data seemed relevant to explain the extent and location of the rainfed wheat area, the use of NDVI class areas rendered their use unnecessary. NDVI data can be incorporated in area frame sampling methods to compile more efficient and accurate statistical data on crop areas. Hypertemporal image analysis techniques can be used either to disaggregate crop statistics and map crop areas for homogenous farming systems (like commercial farming) or to prepare cropping pattern maps (for areas where mixed cropping systems prevail, such as Africa and Asia). We demonstrated that the crop area and crop yield maps generated in the thesis, can improve crop production assessments at regional scales.

Samenvatting

Beleidsmakers, verantwoordelijk voor voedselzekerheid en landgebruik aansturing, hebben belang bij goede en tijdige informatie op regionaal niveau betreffende geteelde arealen en behaalde opbrengsten. Regelmatige vernieuwing van deze informatie ondersteunt bewaking van veranderingen in landgebruik en productie niveaus. De waarde en relevantie van zulke informatie neemt behoorlijk toe als het ruimtelijk expliciet is. Regeringen, internationale organisaties zoals de Food and Agriculture Organisation van de verenigde naties (FAO), the European Statistical Office (EUROSTAT), en de United States Geological Service (USGS), produceren en verspreiden momenteel statistische data betreffende arealen en behaalde producties. Deze zijn vaak in tabelvorm per administratieve eenheid maar missen de informatie waar per eenheid de gewassen werkelijk worden geteeld. Als gevolg zijn zulke data minder geschikt voor bewaking van gewasproductieniveaus, zoals nodig voor voedselzekerheidstudies.

Het doel van deze studie was om methodes te ontwikkelen die kwantitatief gewasarealen kunnen bewaken via een cartografisch proces en om een prototype gewasgroeimodel te testen dat gebruikt maakt van satellietbeelden. Het onderzoek had twee zwaartepunten: landgebruikkartering en de raming van gewasopbrengsten. Bestudeerde methodes die gewas- en opbrengstkaarten genereren zijn ontwikkeld door gebruik te maken van satellietbeelden, geografische informatie systemen, gepubliceerde gewasstatistieken, en gewasmodellen. Geïntegreerde dagelijkse beelden (hyper-temporal) van de SPOT-Vegetatie (VGT) satelliet tot 10-daagse producten (S10 product) die de Normalized Difference Vegetation Index (NDVI) representeren op een 1 km^2 resolutie, zijn gebruikt om landbouw statistische data in tabelvorm te splitsen naar een gewaskaart die per cel (1 km^2) de geteelde areaal-fractie weergeeft. De NDVI beelden zijn gebruikt om een studiegebied (Nizamabad, India; Andalucia, Spanje) te stratificeren in kaarteenheden die homogeen zijn in ruimte en tijd.

De classificatie resultaten van de SPOT-Vegetatie NDVI-beelden van Nizamabad lieten zien dat het gebied opgedeeld kon worden in 11 duidelijk verschillende kaarteenheden. Deze eenheden werden daarna gerelateerd aan een bestaande vegetatiekaart, gebaseerd op hoge-resolutie satellietbeelden, met

gemelde gewasareaal statistieken op sub-district basis, en met teeltkalenders. De bestaande vegetatiekaart vermeldt met een hoog ruimtelijk detail (1:50.000) specifiek de locaties van gebruikte akkers in 1994-95, terwijl de kaart gebaseerd op een serie van NDVI- beelden (1998 – 2004) zowel ruimtelijke als de temporele variaties representeert. Na het koppelen van de NDVI kaart met de areaal-statistieken en de teeltkalenders, is een verbeterde kaart geproduceerd die bovendien aangeeft waar ieder gewas is geteeld met additionele achtergrond informatie t.a.v. bestaande vegetatie en landgebruik.

De areaaldata per NDVI kaarteenheid en dorp in Andalucia zijn gecorreleerd met gerapporteerde statistieken per dorp en gewas betreffende het totaal geteeld areaal, om de statistieken te verdelen per 1 km^2 cel. Deze oefening liet zien dat de bestudeerde gewasstatistieken significant gecorreleerd zijn met de NDVI kaarteenheden. De resultaten zijn bekrachtigd door gebruik te maken van primaire veldgegevens. Die gegevens zijn verzameld door de Spaanse overheid door jaarlijks (2001-05) 700x700m blokken binnen landbouwgebieden volledig te inventariseren. In dit deel van het onderzoek werd de methode om gewaskaarten te produceren verder ontwikkeld en uitvoerig statistisch getest.

De methode is gebaseerd op de veronderstelling dat NDVI beelden volledig de gecombineerde verschillen in bodems, landschap, weer en landgebruik in zich bergen. Ten aanzien van bodems en geomorfologie is deze veronderstelling verder bestudeerd. Additioneel gebruik van bodem gerelateerde kaarteenheden verbeterde de hoeveelheid verklaarde variabiliteit door het NDVI model slechts met 1%. De Jack-Knife procedure, die de relatieve inbreng per parameter binnen het model test, bevestigde deze bevinding. NDVI kan dus gebruikt worden als een indicator die de gecombineerde verschillen in ruimtelijke condities weergeeft.

Na het bepalen van een methode om gewaskaarten te genereren werd het onderzoek voortgezet met het evalueren van resultaten van een gewasgroeimodel (Cf-Water), dat op basis van satellietbeelden schattingen maakt van de reële behaalde opbrengst per 1 km^2 cel. Na opschaling van de geschatte opbrengsten door Cf-Water naar provincie niveau zijn deze data vergeleken met de resultaten van het operationele model CGMS (Crop Growth Monitoring System) van het European Union Monitoring Agriculture with Remote Sensing (MARS) programma, en met gepubliceerde

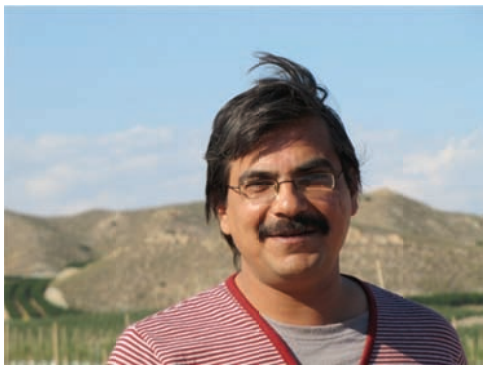
opbrengstgegevens. Ook zijn Cf-Water resultaten vergeleken met primaire veldgegevens, verzameld door de Spaanse overheid. Te vermelden is dat CGMS, na een aanpassing via trendanalyse, slechts gegeneraliseerde statistieken op landniveau rapporteert. Cf-Water heeft een lagere gegevensbehoefte dan CGMS, waarvoor ook bodem- en historische opbrengstgegevens nodig zijn. De resultaten lieten de hoge potentie van Cf-Water voor gebruik in voedselzekerheidsstudies zien; hoge correlaties met veldgegevens en met gepubliceerde statistieken werden gevonden. Het model is echter eenmalig getest en voordat het model aanbevolen kan worden voor operationeel gebruik zijn studies met andere gewassen, in andere jaren en andere gebieden nodig.

Om een beeld te krijgen van de relevantie van dit onderzoek voor potentiële gebruikers, is hun mening gevraagd over momenteel beschikbare informatie en het nut van de in dit onderzoek geproduceerde gewas- en opbrengstkaarten. De gebruikers bleken niet tevreden met de updatefrequentie en met het ruimtelijk detail van bestaande landgebruiksinformatie, terwijl 72% van de gebruikers wel tevreden bleek met de in dit onderzoek geproduceerde kaarten, vooral met het ruimtelijke detail, het type informatie, en de potentiële updatefrequentie. Wel uitten 70% van de gebruikers de wens dat gewasproductiekaarten al ruim voor de oogst beschikbaar gesteld worden; de huidige methode heeft echter nog geen voorspellingsroutines. De gemaakte producten bleken interessant voor potentiële gebruikers; 75% wil de methode om gewaskaarten te maken gebruiken en 70% die van opbrengstkaarten.

Op basis van resultaten beschreven in deze thesis is de conclusie dat het gebruik van hypertemporal NDVI beelden zeer geschikt is om gewaskaarten te produceren met hergebruik van gepubliceerde gewasstatistieken. Hoewel bodemkaarten verklaren waar regenafhankelijke tarweproductie plaatsvindt, bleek door gebruik van NDVI beelden incorporatie van deze relatie in het model overbodig. Gebruik van NDVI beelden voor het maken van een veldmeting plan (‘area frame’ design) zal de uitvoeringsefficiëntie en de kwaliteit van beoogde gewasareaalschattingen verbeteren. Gebruik van verwerkingstechnieken betreffende hypertemporal beelden kan worden ingezet voor het splitsen van landbouwstatistieken in tabelvorm naar een gewaskaart. Dit geldt zowel voor commerciële homogene landbouwgebieden als voor het in kaart brengen van gebieden waar een ruimtelijk en sequentieel patroon van

gewassen geteeld wordt. Dit laatste komt vaak voor in Azië en Afrika. Duidelijk is dat gebruik van verbeterde gewas- en productiekaarten, opbrengstschattingen op regionaal niveau ten goede komt.

Biography



Mobushir Riaz Khan was born in Faisalabad, Pakistan on 27 November 1974. In 1991, he completed his higher secondary education from his home town. Having always been interested in agriculture, he chose to extend his education in this direction and went to the University of Agriculture Faisalabad, Pakistan. At the end of 1991, he completed his M.Sc. (Hons.) in Agriculture. To get field knowledge in agriculture, he joined the Agro-service industry in the same year and managed the territory of mixed cropping systems in the province of Punjab. During his stay, there, he delivered "on and off farm" trainings to farmers and the company's field staff for pest management and crop production. Later on, in November 2002 he joined the PMAS-University of Arid Agriculture as Lecturer in the Faculty of Crop and Food Sciences. Where he taught various courses in the department of Entomology and supervised Bachelor and Masters level students. In 2004, he was honoured by the Vice Chancellor of the University and was appointed as his Technical Staff Officer. He assisted the Vice Chancellor in the meetings and seminars. He conducted many international seminars and workshops. In 2005, he got the training on Decision Support System for Agro-technology Transfer (DSSAT) at the Chiang Mai University, Thailand with the collaboration of University of Georgia, USA. In 2006, he joined the department of Natural Resources in the then International Institute for Geo-Information Science and Earth Observation (ITC), Enschede together with Wageningen University and Research Centre, the Netherlands to pursue his doctoral studies. There, he worked in the theme of Food security and environmental sustainability (FSES) as a Ph.D. research student. In 2008, he was selected as Assistant Professor in PMAS-University of Arid Agriculture. He focused his research work on the integration of RS and GIS technologies to study crop production systems. This thesis is the outcome of this research/study work.

ITC Dissertation List

http://www.itc.nl/Pub/research/Graduate-programme/Graduate-programme-PhD_Graduates.html

PE&RC PhD Education Certificate

With the educational activities listed below the PhD candidate has complied with the educational requirements set by the C.T. de Wit Graduate School for Production Ecology and Resource Conservation (PE&RC) which comprises of a minimum total of 32 ECTS (= 22 weeks of activities)

Review of literature (5.6 ECTS)
- Review of crop monitoring capabilities (current and past and spatial crop growth modelling approaches

Writing of project proposal (4.5 ECTS)
- Quantitative mapping and monitoring of crop production systems

Post-graduate courses (5 ECTS)
- Learning IDL for building expert applications in ENVI; ITC (2009)

Invited review of (unpublished) journal (1 ECTS)
- Change detections using hypertemporal NDVI data; Remote Sensing of Environment

Deficiency, refresh, brush-up courses (2.8 ECTS)
- Principles and applications of GIS and remote sensing

Competence strengthening / skills courses (5.8 ECTS)
- Scientific writing course; ITC (2008)
- Technical writing and editing; UT (2009)

PE&RC Annual meetings, seminars and the PE&RC weekend (2.9 ECTS)
- ITC-IPC PhD weekend (2009)
- ITC-IPC PhD weekend (Resource person) (2010)
- ITC PhD research day (oral and poster presentations) (2010)

Discussion groups / local seminars / other scientific meetings (7.3 ECTS)
- Biodiversity and fragmentation group discussion (NRS PhD) (2007-2011)
- 139th "Themadag" of the NBV, *Past and Future of Land Evaluation* (2007)

International symposia, workshops and conferences (3.4 ECTS)
- AGRO 2010, the XIth ESA Congress (2010)
- Presentation at the Central Offices of Consejería de Agricultura y Pesca (Ministry of Agriculture and Fisheries); Seville, Spain (2009)
- Presentation at the departamento Biologia animal, Universidad de Malaga, Spain (Department of Animal Sciences, University of Malaga, Spain) on Hypertemporal image analysis and its results for Andalucía (2007)