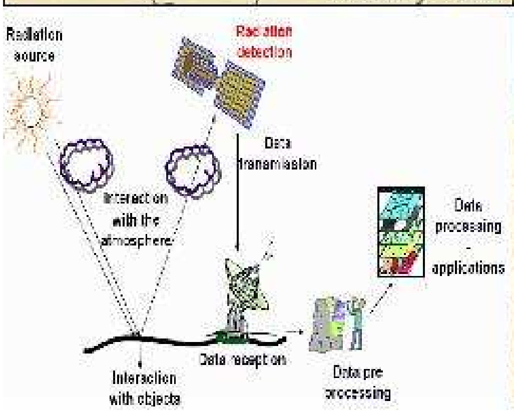
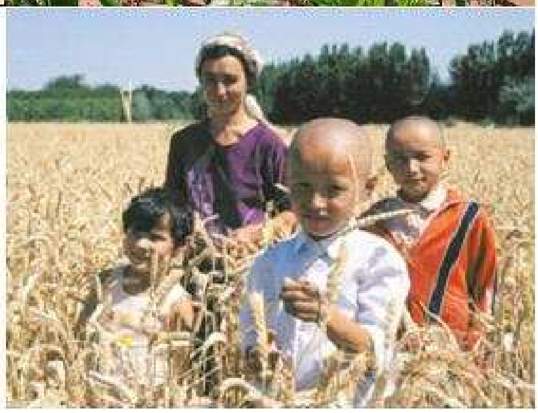
Centre for Geo-Information

Thesis Report GIRS-2008-13

Prediction of wheat yield in Uzbekistan by using the CGMS model and SPOT-Vegetation data

Bahtiyor Pulatov





July, 2008

WAGENINGEN UNIVERSITY
WAGENINGEN UR

Prediction of wheat yield in Uzbekistan by using the CGMS model and SPOT-Vegetation data

Bahtiyor Pulatov

Registration number: 85 05 23 673 090

Supervisors:

Dr. Ir. Jan Clevers
Ir. Hendrik Boogaard
Ir. Kees van Diepen

A thesis submitted in partial fulfillment of the degree of Master of Science at Wageningen University and Research Centre,
The Netherlands.

July, 2008
Wageningen, the Netherlands

Thesis code number: GRS-80436
Thesis Report: GIRS-2008-13
Wageningen University and Research Centre
Laboratory of Geo-Information Science and Remote Sensing

Acknowledgements

First and foremost I offer my sincerest gratitude to my supervisors Dr. Ir Jan Clevers, Ir. Hendrik Boogaard and Ir. Kees van Diepen, who have supported me through my thesis work with their patience and knowledge. Without their guidance and persistent help this work would not have been possible. From them I have learned how it is to be thoughtful, logical and accurate. One simply could not wish for better supervisors.

I am grateful to study advisor Willy ten Haaf for his advice and arrangement of conditions during my study.

Thanks to all my friends for their helps and supports during my thesis study.

Finally, I thank my parents for supporting me throughout all my studies at Wageningen University.

Abstract

Wheat yield forecast under difficult environmental conditions in Uzbekistan is receiving considerable attention from governmental agencies, commodity traders and farmers. An objective of this study is to investigate how we can best predict wheat yield early in the season in Uzbekistan. The approach used in this study is based on a crop growth simulation model which is able to quantify the effect of weather conditions on crop growth. Focus was on selection which indicators either from CGMS or SPOT-Vegetation data can be the best predictors for an explanation of year to year variation of wheat yield in Uzbekistan. The Crop Grow Monitoring System (called CGMS) provides a timely, accurate, synoptic and objective estimation for crop growth conditions and issues yield forecasts at regional and national level using remote sensing data. After CGMS and SPOT-Vegetation data comparison the results showed that indicators maximum NDVI and maximum DMP which are driven from remote sensing SPOT-Vegetation data are performing the best at regional and at national level.

Table of contents:

List of Figures

List of Tables

Abbreviations

.

AVHRR	Advanced Very High Resolution Radiometer, LR-sensor on-board of the NOAA-satellites
B	Blue
CGMS	Crop Growth Monitoring System, the combination of an agrometeorological crop growth simulation model WOFOST, a database, and a yield prediction routine.
CNES	Centre National d'Etudes Spatiales
CORINE	Coordination of Information on the Environment
DMP	Dry Matter Productivity (kg DM / ha / day), RS vegetation indicator derived with the Monteith approach
DSSAT	Decision Support System for the Agrotechnology Transfer
ECMWF	European Centre for Medium Range Weather Forecast
EUROSTAT	Statistical Office of the European Commission
EC	European Commission
fAPAR	Fraction of Absorbed Photosynthetically Active Radiation
GIS	Geographical Information System, software for storage of geographical data, mostly in vector format
Green Seeker	Ground based sensors provide real time data, day or night, regardless of weather conditions
iNDVI	integrated Normalised Difference Vegetation Index
JRC	Joint Research Centre of the European Commission (at Ispra, Italy)
LAI	Leaf Area Index
LR	Low-Resolution imagery (pixel of 1 km), parcel structure invisible, mixed pixels
MCYFS	Mars Crop Yield Forecasting System, an important instrument of the MARS-STAT team to assess the crop yields
MARS	Monitoring Agriculture with Remote Sensing, EU-JRC program started in 1998
NDVI	Normalised Difference Vegetation Index, RS-indicator for the amount of standing vegetation
NIR	Near infrared range of the spectrum, roughly from 780 nm to 1300nm

NOAA	Series of near-polar satellites monitored by the US National Oceanographic and Atmospheric Administration
R	Red
RMSFE	Root Mean Square of Forecast Error
RMSE	Root Mean Square Error
PAR	Photosynthetically Active Radiation
PRESS	Predicting Error Sum of Square Residuals
PYB	Potential Above Ground Biomass
PYS	Potential Storage Organs
RS	Remote Sensing: earth observation with imaging sensors on-board of space/airborne platforms
SPOT	Système Pour l'Observation de la Terre
USDA	United States Department of Agriculture
WOFOST	World Food Studies crop growth model, a crop growth simulation model
WYS	Water Limited Storage Organs
WYB	Water Limited Above Ground Biomass

1. Introduction

1.1. Background

Nowadays, agriculture in Uzbekistan is developing, especially in wheat production. Since many people live in the countryside, they grow wheat in order to make their own bread at home. Uzbekistan provides an excellent example of the shift in agricultural land use in Central Asia. After independence in 1991 Uzbekistan introduced winter wheat in irrigated areas largely for reasons of food security. Wheat is a relatively new crop. In the past, a big part of the wheat consumed in Uzbekistan was imported either from Russia or neighboring Kazakhstan. Production of wheat has increased from 0.61 million tons in 1991 to 3.07 million tons in 1997. Wheat production in Uzbekistan for 2001/02 is estimated at 3.4 million tons, comparable to 3.6 million in 2000/01, and 3.7 million in 1999/2000. The area under wheat was about 1.4 million hectares. Wheat area dropped by 15 percent in 2000/01 to 1.2 million hectares, and remained at that level for 2001/02, but the reduction was due mainly to a decrease of previously irrigated 150,000 hectares of land from a marginal production area. Preliminary official statistics indicate that the sown area of 2002/03 in Uzbekistan for winter grains, over 90 percent of which is winter wheat, remained stable at roughly 1.2 million hectares. Conditions have been beneficial for winter crops, with above-normal precipitation per month. Wheat is considered to be a strategic crop by the government and its production has been subject to extensive government controls. Families in Uzbekistan historically have been enlarged (4-5 children in rural areas) and a sharp decline in the life standards made bread a main staple for Uzbek families. While low food prices benefited consumers, they discouraged farmers from increasing production and encouraged harvest diversion to the neighboring countries where prices were somewhat higher (USDA, 2002). In order to improve abilities in making decisions, it is necessary to develop or introduce a yield prediction model in Uzbekistan. In the present thesis the European yield prediction system was tested for its potential to serve as a basis for a national crop yield prediction system in Uzbekistan.

The European Commission launched the Monitoring Agriculture with Remote Sensing (MARS) project in 1988 in order to predict crop yields, crop areas and crop productions. The project was led by the Joint Research Centre (JRC) of the EC, located in Ispra, Italy. One of the actions of the JRC-MARS projects was to create the Crop Growth Monitoring System (CGMS), which is a combination of a crop growth simulation model (WOFOST), using as inputs meteorological information, a detailed soil map, parameters for crop and spatial crop information.(Boogaard et al., 2002). The objective of this European crop yield forecasting system was to supply foreseen information to the DG-Agriculture Outlook group on the development and growth conditions of crops across Europe. The system relies on a Pan-European agro-meteorological method of analysis (Vossen and Rijks, 1995). After ten years of research and development in co-operation with Member States and a pre-operational-phase, following a Council/Parliament decision in 1999, it is running in an operational context, called Mars Crop Yield Forecasting System (MCYFS). The system consists of an ensemble of methodologies and tools providing early information on crops across the agricultural season.

In order to predict yields using meteorological information, the Crop Growth Monitoring System (CGMS) was created for Europe. Currently, the CGMS model is also used to predict crop yields in Central Asia. However, this is still a prototype. Because of the importance of wheat for Uzbekistan, the adaptation of this model at a national level in Uzbekistan will provide a tool to predict the wheat production in the following years. The main objective of this model is to generate country-wide information during the cropping season on the state of the crops and yield outlook. The approach is based on crop growth simulation to quantify the effect of weather conditions on crop growth and on statistical analysis of many years of crop model output and observed regional yield figures.

1.2. Introduction of CGMS

The yield prediction system used in the European Union also known as the Crop Growth Monitoring System (CGMS) has been adapted to run for Uzbekistan. The main objective of CGMS is to monitor agricultural season conditions over the whole of the European Union (EU) and neighboring countries, and to make quantitative within season yield forecasts at regional and national scale for specific crops. CGMS uses daily meteorological observations from a network of some 2000 weather stations to estimate crop status under rainfed (water-limited) and irrigated (potential) conditions in the course of the growing season and to estimate final crop yield at the end of the season.

Three levels can be distinguished in the CGMS model:

− The first level is the weather system. Historical and actual weather data are collected, corrected and subsequently interpolated to the grid centre. The CGMS system as used by the European Commission is based on a 50 x 50 km grid. Other CGMS systems may use a different grid size. Historical, actual and interpolated meteorological data are stored in a database.

− The second level is the crop simulation. The soil database and the crop knowledge database are used to identify areas where a given crop can possibly grow. Then crop growth is simulated on a daily basis using the interpolated meteorological data from level 1. Crop variables like leaf area index (LAI), above ground biomass and storage organs are simulated. Parallel to the crop growth, the soil water balance is simulated, which provides soil moisture status and quantification of drought effects on crop growth.

− The third level is the statistical analysis of simulated yield and observed historical yields over some 8 to 15 years as a basis for yield prediction in the current year. The focus is on yield per hectare. Statistics on crop yields, in combination with statistics on planted area, and production volume are used at level 3. The main component of this level is a regression model consisting of three terms: a mean yield figure, a trend and a residual term (see chapter 3). The residual term is the one determined by the current growing conditions and this one is estimated using simulated crop variables by CGMS. The regression parameters are determined by means of an analysis of historical multi-annual time series of observed and simulated yield data. When this regression model is used in prediction model, the simulated yield is used as yield predictor, and is also called yield indicator. In stead of simulated yield also other variables can be used as yield indicator, e.g. variables derived from remote sensing such as a vegetation index.

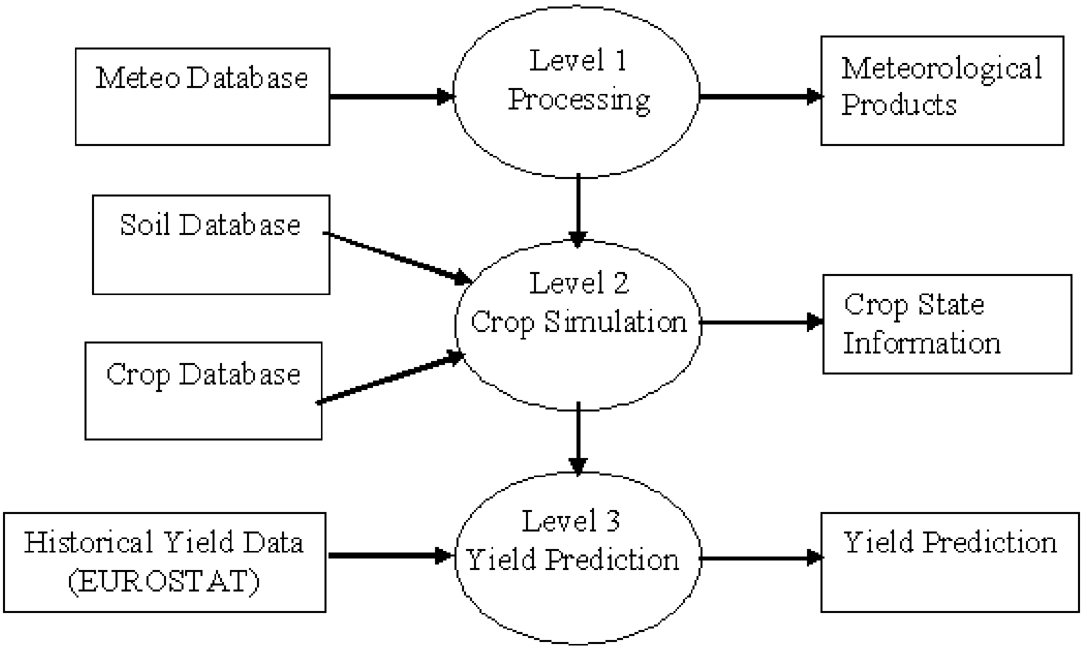


Figure 1.1 ***Schematic overview of the Crop Growth Monitoring System in the EU (Guo, 2004).***

1.3. Problem Description

From the first days of independence of Uzbekistan, the wheat production has increased to 3.7 million tons between the years 1991-2000. During the years 2000 and 2001 wheat production decreased to approximately 3.6 million tons. The reason of the change in the wheat production can be explained by the changes in the weather in Uzbekistan. From an economic point of view, wheat production is vulnerable to weather changes which may damage or harm the wheat production and could be a barrier to extent productivity of wheat. Wheat is considered to be a strategic crop by the government and its production has been subject to extensive governmental controls.

Description of CGMS

The CGMS model is already used by the MARS projects to predict crop yields in Uzbekistan. For this prediction, the wheat yield statistics from Uzbekistan are used in the regression model of CGMS level 3 (described above). The analysis of the outputs of the level 2 of the CGMS model adapted to Uzbekistan for wheat and the indicators of SPOT-Vegetation may help to investigate and understand the causes of the changes in the wheat production. At the moment, it is not known which indicators can be used to explain the year to year variation. These indicators can be the indicators obtained from CGMS, like biomass at a certain moment during the season, or an indicator based on remote sensing images, like the NDVI obtained from SPOT-Vegetation or a combination of both types of indicators. Major problem at the moment is the question which indicator can best be used in the regression model and how well the indicators can be used to predict wheat yield in Uzbekistan at regional and national level.

The prediction of the wheat yield and the reasons of the changes as a result of the regression model may help the government to determine its plans strategically and help the farmers to reduce the risks which they may face while producing wheat. In the first place it would give the government a better insight early in the year on the need for import or export of cereals at national scale, and the need for redistribution of produce from surplus regions to deficient regions. For the planning for the coming years, a better understanding of yield reducing factors such as drought and heat may help to make better planning of land use for rainfed and irrigated agriculture, and to make better decisions on allocation of scarce irrigation water over crops and over regions, on the designation of preferred irrigation areas, and on the balance between purely rainfed conditions, and partial and full irrigation.

1.4. Research objective

The main research objective is to investigate whether indicators from CGMS or SPOT-Vegetation data can be used in a regression model to explain year to year variation of wheat yield in Uzbekistan during the growing season well before harvest.

1.5. Research questions

- Which indicator from CGMS can best be used for explaining yield statistics of wheat in Uzbekistan at both oblast and national level?
- Which indicator from SPOT-Vegetation data can best be used for explaining yield statistics of wheat in Uzbekistan at both oblast and national level?
- Which level (regional or national) can be explained best by indicators in CGMS statistical tool?
- Should we distinguish rainfed and rainfed plus irrigated lands in the analysis?

Sub - research questions

- What are the indicators of CGMS related to yield statistics of wheat in Uzbekistan at oblast level?
- What are the indicators of CGMS related to yield statistics of wheat in Uzbekistan at national level?
- What are the indicators of SPOT-Vegetation data related to yield statistics of wheat in Uzbekistan at oblast level?
- What are the indicators of SPOT-Vegetation data related to yield statistics of wheat in Uzbekistan at national level?
- What are the advantages and disadvantages of distinguishing rainfed and rainfed plus irrigated lands in the analysis?
- What are the statistical wheat yield data to be collected for Uzbekistan?

- How can the prediction of wheat production be described in terms of a regression model?
- What is the result of the regression model by using the indicators of the CGMS model and SPOT vegetation?

1.6. Overview of the report

Chapter one of this report comprises an introduction about the general background and an overview of the context. Description and definition of the problem is also main part of this chapter. The objective of this study and research questions as well as sub-research questions are covered in this chapter. The second chapter deals with a review of the relevant literature and data collection in Uzbekistan. The third chapter describes the methodologies followed in order to achieve the research objectives. Results of this study are presented in chapter four and discussed in chapter five. Final chapter will conclude and give recommendation for future work.

2. Literature review

2.1. Yield prediction

The explanation of inherent crop yield variability is due to variations in the topography, local water conditions, management, soil conditions, etc. In the past, agricultural research has been dedicated mainly to the yield prediction issue of deriving local response functions to various fertilizers, pesticide and other application rates. Difference in microclimate has only been addressed in a few attempts.
Yield forecasts are of high interest to the wheat production industry in order to allow the most convenient purchasing policy of raw materials. There are two possibilities to predict crop yield. First way has been used for many centuries mainly by farmers all over the world. Each farmer knows about the yield and how much they can get next year if they sow their crop at a certain time in the season. Therefore, they can predict the amount of yield before the harvest time, based on knowledge and experience of yearly wheat production. This type of prediction is called field-scale (Schwab et al., 1996), but this type is not accurate, since it is qualitative but not quantitative. It is just a word of farmers, they make decisions by themselves, because they are harvesting and sow each year their own crops.

Second way of wheat yield prediction is by simulation of the future using different models. There are different models which can be used to make wheat yield prediction. Here two different models used for crop yield prediction will be discussed in the next paragraphs.

The Decision Support System for Agrotechnology Transfer (DSSAT) model is a microcomputer software program combining crops soil and weather databases. This model is simulating multi-year outcomes of crop management strategies (Clarke et al., 1996). Since crop growth factors are so complex, it is difficult to isolate them. For example, if local soil conditions, microclimate, intermittent air deficiency and many more factors caused a delay in crop development at early stages in some zones within the field, the following crop development will always be affected by early stage differentiation (Clarke et al., 1996). The model DSSAT does not adequately reflect this early-state crop variability and will therefore falsely predict the final crop yield. If instead simple observations can be taken that adequately reflect the spatially differing crop development, an important additional input can be given to the model, and the following calculations will be based on a spatially differentiated pattern. Conceptually this is what the farmer does when he takes decisions for managing locally varying crop stands that were for example affected by locally unfavorable growing conditions during seasonal time. However, the model cannot reflect to precision biomass production of the farmer due to the complexity of the various processes. Moreover, if conditions for precise crop growth simulation can be improved, we provide ample opportunities for using computer simulation models in scenario calculations for environmentally and economically sustainable management decisions for Uzbek Government.

Another approach is called GreenSeeker (http://nue.okstate.edu). The GreenSeeker hand-held sensor can be used to estimate actual wheat or corn grain yield using the Normalized Difference Vegetation Index (NDVI) values from the Nitrogen Rich Strip (can tell you how much N the environment delivers as N mineralized from soil organic matter and that deposited in the rainfall) compared to the farmer practice, and knowing the date when the wheat was planted (Solie et al., 2005). Essentially, the NDVI value from the hand-held sensor outputs “total biomass.” From data which were collected between January and March (regardless of when the wheat was planted), we can estimate “biomass produced per day.” This value is used to predict the wheat grain yield obtainable. With this GreenSeeker we can accurately predict the yield (http://nue.okstate.edu/Hand_Held/Hand_Held_Pictures.htm). This type of research is possible but it takes time to go to the field and measure all yields. This is a disadvantage of this apparatus. Nowadays there are a lot of new technologies producing different types of prediction models for different types of crops. One of those models is the CGMS.

2.2. Remote Sensing

Remote sensing offers great potential for regional production monitoring and estimates. It provides a timely, accurate, synoptic and objective estimation of crop growth conditions and it can improve yield forecasts at a range of spatial scales. In agricultural planning and policy making, knowledge of crop yield at an early stage is very important at both national and regional level. Moreover, to estimate and predict crop yield it is necessary to monitor the growth of the crop during the growing season. Time series of the Advanced Very High-Resolution Radiometer (AVHRR) Normalized Difference Vegetation Index (NDVI) show greenness values on a daily global basis and these have been used for crop yield forecasting since the 1980s.

Many researchers have done much work on the AVHRR satellite. One of this research's had been used for malting barley yield forecasts in South-Western Germany. For this approach, multisensor and multitemporal Remote Sensing data and ancillary meteorological, agrostatistical, topographical and pedological data were used as input data for prediction models (Braun et al., 2004).

Also in France yield forecast had been performed using the information of NOAA-AVHRR/NDVI and CORINE land cover data (CORINE -Coordination of Information on the Environment). The EU established CORINE in 1985 to create pan-European databases on land cover, biotopes (habitats), soil maps and acid rain. Yield indicators were used in time series at a regional scale. A high correlation was observed with wheat yield. The advantages of this method are that results confirm better wheat yield prediction (Genovese et al., 2004).

NDVI is an expression of contrasting reflectance between red and near-infrared regions of a surface spectrum (Rouse and Telesca, 2006). NDVI is one of the most common indexes to determine vegetation growth, like daily global datasets have been developed into 10-day maximum NDVI composite datasets to minimize cloud cover. The latter data set provides a time series of uniform, nearly cloud free composites.

In the past, many scientists have utilized remote sensing techniques to assess agricultural yield, production, and crop condition. First, they identified a relationship between the NDVI and crop yield using experimental fields and ground-based spectral radiometer measurements. Final grain yields were found to be highly correlated with accumulated NDVI (a summation of NDVI between two dates) around the time of maximum greenness (Tucker et al., 1980). In another experimental study remotely sensed data were used to predict wheat yield 3–4 month before harvest in Uzbekistan. These early experiments identified relationships between NDVI and crop response, paving the way for crop yield estimation using satellite imagery (Kastens et al., 2005). Based on the studies described, for the purpose of crop yield forecasting, longer time series of NDVI imagery are preferred to shorter ones. The objective in general is to use historical yield information and historical NDVI imagery to devise a thorough and strong statistical procedure for obtaining early to mid-season crop yield forecasts. The techniques are described in the following sections and can be applied to any region and for any crop that possesses sufficient historical yield information and corresponding time series of NDVI imagery (Kastens et al., 2005).

2.3. Linking remote sensing and crop growth models

Satellite remote sensing has always been an important cornerstone of the MARS-project and remote sensing is a component of the MARS vegetation monitoring and yield forecasting operations. It has been used widely for early prediction of crop production on the basis of the CGMS model. The CGMS model uses only meteorological information and it does not take the advantage of remote sensing for crop monitoring and yield prediction into account.

There are three methods which can be distinguished to link remote sensing indicator data with CGMS model. First method is called “model calibration”. This model calibration is used to calibrate a crop growth model on time series of remote sensing measurements. Bouman (1992b) developed a procedure in which remote sensing indicators were linked to crop growth models, so that canopy reflectance for wheat was simulated by linking a crop growth model with a canopy reflectance model. The crop growth model was then calibrated to match simulated values of canopy reflectance to measured values of reflectance for the wheat (Clevers and Leeuwen, 1994).

A second method can be distinguished to link optical remote sensing data with crop growth models. This method is called “model initialization”: crop values are estimated from optical remote sensing and put into a growth model as input or forcing function (Steven et al., 1983). Mostly crop parameters that have been used successfully so far are measures for the fractional light interception by the canopy. However, other parameters can also be used as input in more elaborate growth models.

A third method can be distinguished to combine remote sensing with crop growth models. It is called “model reset”. An example is to introduce LAI values, estimated from SPOT images (Clevers et al., 2002). A reset of the LAI data was performed at the date of each SPOT observation in the growing season. Such a reset means that simulated LAI is replaced by the estimated LAI value obtained from remote sensing information. After reset at a SPOT observation date, the model simulated the LAI development until the next SPOT observation date, when again a reset was done. The LAI values were simulated until the date of the last SPOT observation. Then, LAI values were estimated from the SPOT data and linearly interpolated from these estimates for the dates between the SPOT observation dates. Final observation of the last SPOT data were simulated by LAI values until the end of the growing season.

2.4. SPOT VEGETATION

The French government has undertaken the development of the Systeme Pour l'Observation de la Terre (SPOT) program. (http://www.agrecon.ca berra.edu.au/products/Satellite_Imagery/Spot_Veg/Spot_Veg.htm). The SPOT program was conceived by the Centre National d'Etudes Spatiales (CNES) and has developed into an international program with ground receiving stations and data distribution outlets in more than 30 countries. The Vegetation Instrument system is the result of a co-operation between the European Union, France, Sweden, Belgium and Italy. It aims at ensuring a regional and global monitoring of the continental biosphere and its crops. High quality remote sensing imagery are ready to use and available to end-users in near-real time. Vegetation optical instrument, a wide field of view sensor on board of the SPOT-4 earth observation satellite launched in March 1998 (http://www.spot-vegetation.com/), operates in four spectral bands:

- **Blue (B)**, mainly to perform atmospheric corrections.
- **Red and Near Infrared (R and NIR),** sensitive to the vegetation's photosynthetic activity and cell structure.
- **Short Wave Infrared (SWIR),** sensitive to soil and vegetation moisture content.

Its 1 km spatial resolution is nearly constant across the whole 2,250 km corridor it covers, which means that there is almost no distortion on image edges (Vegetation website, 2008).

Many research works have been carried out in these last years to investigate the contribution of remote sensing data in yield prediction and crop monitoring. In particular the possibilities of integrating information derived from radiometric measurements into agrometeorological models show promising results.

In the paper of Roijas (2007) spectral indices deduced from visible and near-infrared remote-sensing data have been extensively used for crop characterization, biomass estimation, and crop yield monitoring and forecasting. For example, the government of Kenya has reported the development of an operational agrometeorological yield model for maize using spectral index such as the NDVI (Normalized Difference Vegetation Index is a simple numerical indicator that can be used to analyze remote sensing measurements, typically but not necessarily from a space platform, and assess whether the target being observed contains live green vegetation or not) derived from SPOT-Vegetation, and meteorological data obtained from the European Centre for Medium-Range Weather Forecast (ECMWF) model. A statistical regression model has been developed for large scale growing areas in Kenya. The advantages of the results reported in Kenya are that this model is operational for yield prediction (Roijas, 2007).

Another study described different indicators in terms of 'rainfall variability' and 'vegetation dynamics' and their performance was compared within two study areas in different climate zones which are generally both characterized by high rainfall variability in West and East Africa. The indicators were determined for 14 stations in Kenya and 10 stations in Benin considering different land cover types. Final investigation was undertaken to identify the main driving forces for the integrated NDVI (iNDVI or sum-NDVI) which is assumed to be a good indicator for seasonal variation of different types of land (Budde et al. 2004).

The data presented in a research on Alaska AVHRR NDVI Derived Phenological Metrics contain multi-temporal phenological data of Alaskan vegetation for 1991 – 2000 (Alaska web source, 2008). These multitemporal data were produced using seasonally truncated, composited Advanced Very High Resolution Radiometer (AVHRR) derived Normalized Difference Vegetation Index (NDVI) data. In this above mentioned report (Alaska, 2008) many metrics were obtained: maximum NDVI is the maximum measurable NDVI recorded during the year and is normally associated with the peak of green vegetation during the growing season. Mean NDVI is the maximum NDVI value obtained for each recording period during the growing season divided by the total number of periods. These two metrics are based on greenness values and have been shown applicable in some situations (http://agdc.usgs.gov/data/usgs/erosafo/ak_avhrr/ak_avhrr_ndvi.html, 2008).

3. Methodology

3.1. Study Area

The study area for the prediction model which I am going to adapt is in Central Asia, east of the Caspian Sea, directly south of Kazakhstan, north of Turkmenistan, and on the western borders of Tajikistan and Kyrgyzstan. Geographic coordinates are 41 00 N, 64 00 E. The climate of landlocked Uzbekistan is continental, with hot summers and cool winters. Summer temperatures reach 40°C, averaging 32°C. Winter temperatures reach –38°C, averaging –23°C. Rainfall averages vary between 100 millimeters per year in the northwest and 800 millimeters per year in the Tashkent region. Precipitation falls mainly in the winter and spring. Land use of Uzbekistan is arable land: 10.51%, permanent crops : 0.76%, and other: 88.73% (2005). Nowadays, shrinkage of the Aral Sea is resulting in growing concentrations of chemical pesticides and natural salts; these substances are then blown from the increasingly exposed lake bed and contribute to desertification; water pollution from industrial wastes and the heavy use of fertilizers and pesticides is the cause of many human health disorders; increasing soil salination; soil contamination from buried nuclear processing and agricultural chemicals, including Dichloro-Diphenyl-Trichloroethane (DDT).

Figure 3.1 Location of study area, Uzbekistan.
Source, https://www.cia.gov/library/publications/the-world-factbook/print/uz.html

3.2. Wheat

The crops of Uzbekistan are mainly wheat and cotton. This thesis is focusing on two types of wheat: winter wheat and spring wheat production. The crop calendar is starting from September, October and harvest in June, July. Winter wheat is starting to be sown from the middle of September, usually in October after harvesting the cotton depending on the climate. Spring wheat is starting to be sown from the middle of March, sometimes at the end of April depending on the climate if it's hot or cold and till end of August, so they can have good yields. For this crop is used the data which was collected from Uzbekistan. These are rainfed and rainfed plus irrigated for wheat. "Rainfed" in this case is the total sown area on rainfed areas without irrigation and the "rainfed plus irrigated" is wheat sown on irrigated areas. In Uzbekistan has also two types of wheat sowing in rainfed areas about 1,2 million hectares and rainfed plus irrigated wheat which is about 0,1 million hectares. This data is for all regions in Uzbekistan, which is called national level in the CGMS model.

3.3. Overview of Mars Crop Yield Forecasting System (MCYFS)

The quantitative forecast level of the MCYF system aims to provide the most likely, precise, accurate, scientific and independent forecast for the main crop yields. It works at national level in the European Union, taking into account the effect of the weather, and provides figures for final yield already during the growing season as early as possible. To reach this aim, crop yield forecast procedures, which combine different input data (such as historical yield statistics, weather indicators, simulated crop indicators, remote sensing based vegetation indices) are applied. In the MCYF system, the crop yield is subdivided into three components: mean statistical yield, trend analysis and residual variation (Vossen, 1992, Odumodu and Griffits, 1980).

$$Y_T = \overline{Y} + f(T) + e \quad \text{(eq. 3.1)}$$

where:

Y_T	: observed yield in year T	[ton.ha^{-1}]
$\overline{Y}$	: mean yield	[ton.ha^{-1}]
$f(T)$	: technological trend as a function of time	[ton.ha^{-1}]
e	: residual, not explained by trend	[ton.ha^{-1}]

It is assumed that the interacting effects of climate, soil, management and technology, determine the mean statistical yield. Observed national, regional and sub-regional yields show a trend in time. The trend analysis is mainly due to long term economic and technological dynamics such as increased fertilizer application, improved crop management method, new high yielding varieties. The third component, the residual variation, is considered to be the variation amongst years (Dennett and Diego, 1980). This part should be explained by indicators. Indicators can be the biomass or the storage organs as simulated by a crop growth model (like CGMS) or a vegetation index obtained from satellite imagery. A regression model (fig. 3.2 – 3.3) is used as a tool to look at indicators, which can explain year to year statistical variation at regional and national level. Either the outputs of the CGMS indicators or from SPOT-Vegetation or both will be the inputs of the regression model (CGMS Statistical Tool), which is used for wheat yield prediction.

3.4. Flowchart

The general working methodology of this study followed the scheme indicated by the conceptual model in Fig 3.2 and Fig 3.3. There are generally two steps: first step is for regional level rainfed and rainfed plus irrigated (Fig 3.2, Fig 3.3), second step is for national level also rainfed and rainfed plus irrigated (Fig 3.2, Fig 3.3). Each conceptual model has an input and output. As an example, in Fig 3.2 and Fig 3.3 the statistical tool has an input from statistical (historical) data collected from Uzbekistan as well as from two indicators which later on will be described briefly. SPOT - Vegetation and CGMS indicators are inputs for the CGMS Statistical Tool. Regression analysis compares the results which were output from the CGMS or SPOT-Vegetation data with the actual yield data. Final results will be yield forecast after analyzing indicators from SPOT-Vegetation and CGMS results.

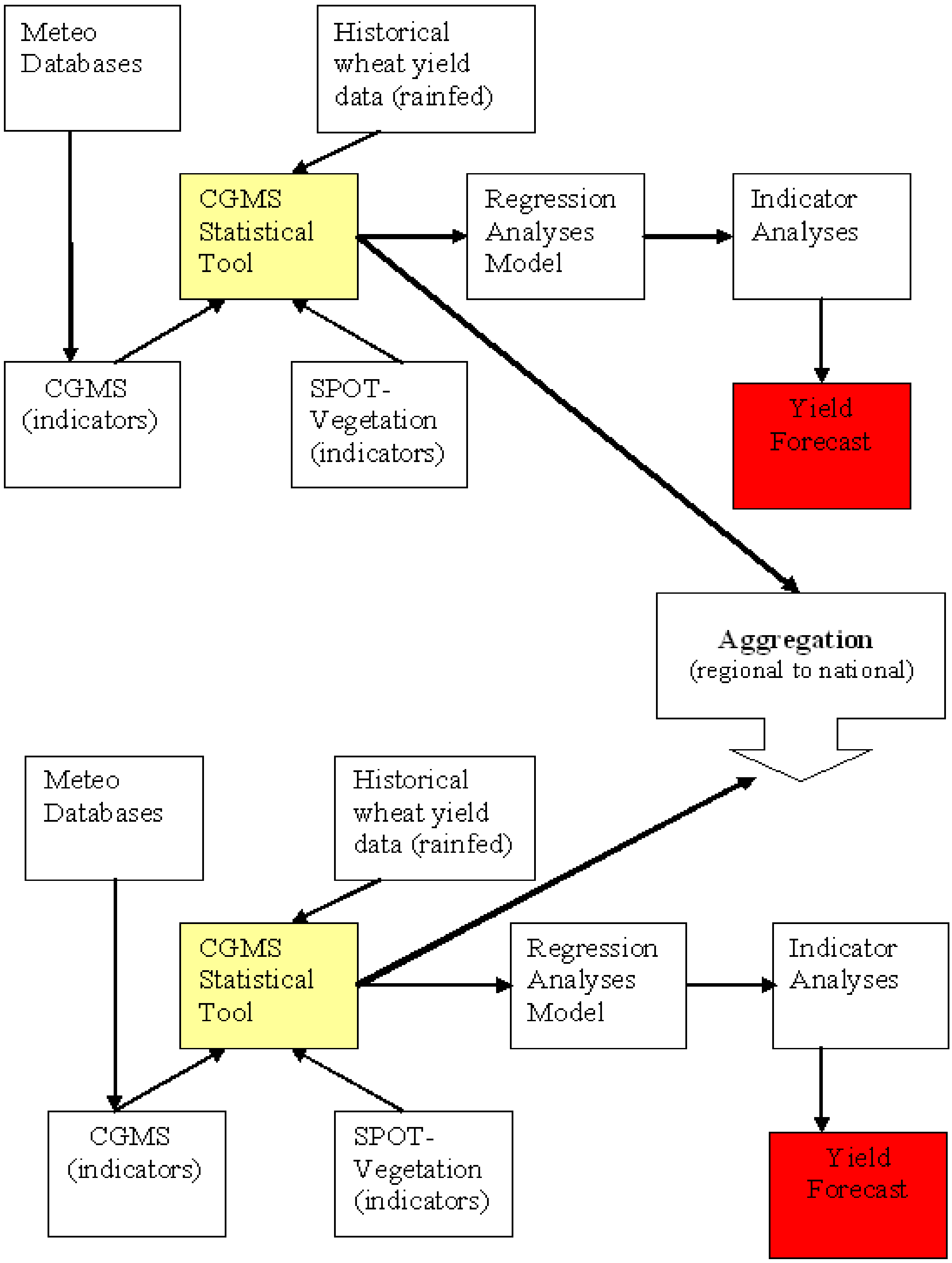


Figure 3.2 Schematic overview of CGMS_Bahtiyor at regional to national level “rainfed”

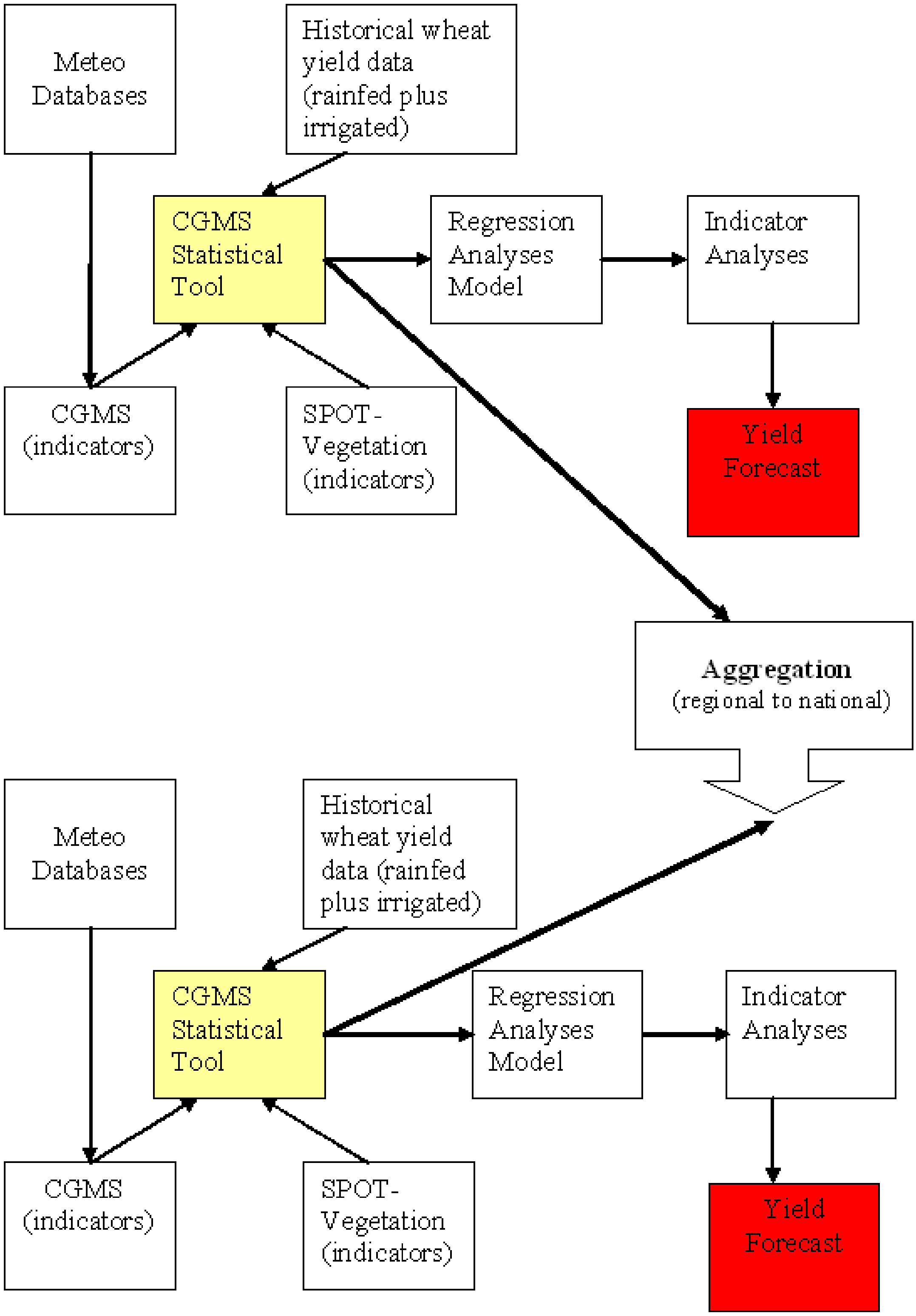


Figure 3.3 Schematic overview of CGMS_Bahtiyor at regional to national level "rainfed plus irrigated"

3.5. Regression Model analysis

1. In order to run the model, statistical data from Uzbekistan are needed. These data were collected by State Statistics of Uzbekistan for 12 years of wheat production. This will be the main input in as data to process. The yield statistics are input into the regression model as independent variable and the predictors are the dependent variables. There are two different statistical wheat yield data; one of which is irrigated lands (sown on irrigated areas) called "Rainfed plus irrigated" and the other one is non-irrigated land type (total sown area non irrigated areas) called "Rainfed". If you look at the table 4.1 and 4.2 in the results chapter it explains the differences between these two types of land Rainfed and Rainfed plus irrigated area. Below in section 3.6 it explains the results which indicator links to which area.

2. The CGMS model is run with actual meteorological data of Uzbekistan as input. Meteorological data provided by meteorological statistics of ECMWF are input for the model. The ECMWF (European Centre for Medium Range Weather Forecast) is one of the world's leading numerical modelling centres. It operates a global circulation model and runs 10 day forecasts on it each day. To evaluate the initial state of the atmosphere a data assimilation system is integrating observations from ground stations, radiosondes, satellites and many other sources. Special techniques bring these observations in balance with meteorological equations to form a physically valid state of the atmosphere. These assimilation data are used to produce meteorological and derived agro-meteorological parameters. Meteorological parameters required as input for the crop model are precipitation, maximum and minimum temperature, global radiation and evapotranspiration. These meteorological parameters should be available on a daily basis.

3. Second part of the flowchart will be the input for the CGMS Statistical Tool in terms of indicators from the CGMS and SPOT-Vegetation data. For the CGMS indicators will be **biomass** and **storage organs** at a certain day during the growing season. For the SPOT-Vegetation these are Dry Matter

Productivity (DMP) and NDVI indicators. The predictors from the CGMS model and SPOT-Vegetation will be tested in the regression model at both levels to investigate which one predicts best the final yield.

4. A trend line represents a trend, the long-term movement in time series data after other components have been accounted for. Trend lines are used in prediction analysis to show changes in data over time. In this thesis it tells whether a wheat yield production have increased or decreased over the period of 1998 to 2006. A trend line is drawn through a set of data points (yield production), but more properly their position and slope is calculated using statistical techniques like linear regression.

5. The residuals from a fitted model are defined as the differences between the response data and the fit to the response data at each predictor value.

$$e = Y_i - \hat{Y}_i \qquad \text{(eq. 3.2)}$$

Residual = field data – fitted yield

Where, $\hat{Y}_i$ is the estimated yield and Y_i is observed yield in year. Mathematically, the residual for a specific predictor value is the difference between the response value Y_i and the predicted response value $\hat{Y}_i$ (Montgomery et al., 2001). The RMSFE is an aggregation over a given period of time (T) consisting of years according to the following formula:

$$RMSFE = \frac{\sqrt{\sum_{i=1}^{T} e_i^2}}{T} \qquad \text{(eq. 3.3)}$$

The RMSFE takes only positive values, and gives an average size of the forecast error over the given period.

6. Regression analysis is concentrating on which indicators are giving the best result in a certain region and certain time of the season. The output from the regression analysis will show which indicators can best be used to explain year to year variation, either from SPOT-Vegetation or from CGMS. To know which indicators are best you need to look at the RMSE (Root Mean Square Error for prediction error). RMSE shows how accurate the model is predicting. It has to be lower than the trend. The trend from the model is also showing yield forecast and it is used to compare results from SPOT-Vegetation and from CGMS indicators. When there are no indicators shown in the model that are suitable for a certain region for a certain time in the season, then it is possible to use the trend as a representative for this region. Root mean squared error for prediction (RMSE) is defined as the difference between the i^{th} response and the predicted value for the i^{th} response based on a model fit to the remaining observations. This is sometimes called the PRESS (Predicting Error Sum of Squares residual) or the leave one out residual (Montgomery et al., 2001). The root mean squared error of prediction is the root of the mean value of all the squared prediction errors. The trend model which gives the best explanation in time will determine both CGMS and RS indicators. If the model shows that CGMS or RS indicators have higher result than the trend, then in this case trend will be representative for this result.

7. Depending on the results, the whole analysis can be done at the oblast scale giving a regression model per oblast as best one. It can also be done at the national level, giving one best regression model for the whole country. To be able to conduct the analysis at national level, an aggregation method should be used**. A**ggregation method will be used to aggregate from regional level to a national level. This work was done separately from the model. The aggregation method defined in equation (3.4) is applied in this research. The wheat yield can be explained with different indicators at different levels (regional or national).

$$\overline{Y} = \frac{\sum_{j=1}^{n}(A_j * Y_j)}{\sum_{j=1}^{n} A_j} \qquad \text{(eq. 3.4)}$$

Where: $\overline{Y}$: is the average value

$A_{j:}$ Area for region j

Y_j Yield for region j

j Each region

n maximum number of regions

8. After the analysis was done at both oblast and national level, the best output can be determined considering the results depending on the level. The results show the yield prediction of this model.

3.6. CGMS Indicator analysis (biomass and storage organs)

Yield forecasting in CGMS takes place at the level of European Union (EU) countries by searching for a relationship between CGMS crop indicators aggregated to national level and the crop yield statistics available from the European Statistical Office (EUROSTAT). It is assumed that crop yield for a region can be divided into three factors: mean yield, multi-annual trend (or technology trend) and residual variation (Vossen, 1992). CGMS uses a running time series of yield statistics of at least 9 years to determine a linear technology trend assuming that the trend is stable over this period. The time series of crop simulation results are then used to explain the residual deviation. The system selects the best predictor out of four CGMS crop indicators (simulation results): potential yield biomass, potential yield storage organs, water-limited yield biomass and water-limited yield storage organs. Crop yield forecasting takes place by using the crop simulation results to predict the deviation from the technology trend already at the beginning of the growing season. More information on the crop yield forecasting algorithm can be found in (Boogaard et al., 2002).

CGMS_WYS: Uses as predictors the simulations of the total weight of the storage organs (grains) for wheat under water limited conditions (rainfed)

CGMS_WYB: Uses as predictors the simulations of the total weight of the above ground biomass for wheat under water limited conditions (rainfed)

CGMS_PYS: Uses as predictors the simulations of the total weight of the storage organs (grains) for wheat under potential conditions (rainfed plus irrigated)

CGMS_PYB: Uses as predictors the simulations of the total weight of the above ground biomass for wheat under potential conditions (rainfed plus irrigated)

3.7. SPOT-Vegetation indicator analysis (NDVI - DMP)

In this study an operational approach was developed using time series of Normalized Difference Vegetation Index (NDVI) and Dry Matter Productivity (DMP) derived from SPOT- Vegetation using the crop yield forecasting tool in Uzbekistan during a nine-year period (1998-2006). The starting point is a series of geographically congruent and periodic (mostly 10-daily) images over the area of interest. The NDVI-values vary from 0.15 for bare soils to ±0.80 for full green vegetations, with all gradations in-between.

$$NDVI = (\lambda Nir - \lambda Red) / (\lambda Nir + \lambda Red) \qquad \text{(eq. 3.5)}$$

where Red and Nir stand for the spectral reflectance measurements acquired in the red and near-infrared regions, respectively. These spectral reflectances are themselves ratios of the reflected radiation over the incoming radiation in each spectral band individually; hence they take on values between 0.0 and 1.0. By design, the NDVI itself thus varies between -1.0 and +1.0.

NDVI indicators may be used for estimating deviations from the year to year yield trend (Zhang et al., 2003). Focus in this study will be on relating NDVI indicators to yield residuals. To predict wheat yield indicators should be selected that have good correlation between the residuals from the trend and indicators (NDVI or DMP) from the SPOT-Vegetation. SPOT-Vegetation data were used as a basis for calculation of remote sensing indicators for crop growth. There is a second indicator available that has been derived from SPOT-Vegetation data: DMP it was used to see whether it has different results from NDVI. DMP based on SPOT-Vegetation data: DMP. It provides information about conversion of global radiation, applying the Monteith approach. Monteith (1972) provided a most general formulation for the “Dry Matter Productivity” (DMP, in kg DM/ha/day), that is the increase in dry matter biomass on a daily base

$$DMP_1 = R_1 * 0.48 * fAPAR_1 * \varepsilon(T_1) * 10000 \qquad \text{(eq. 3.6)}$$

R_1 [J/m2/day] is the incoming wave short radiation of the sun (200-3000 nm), which is composed on the average for 48% of PAR (Photosynthetically Active Radiation: 400-700 nm), and $fAPAR_1$ [-] is the PAR fraction absorbed by the green vegetation. The efficiency term $\varepsilon(T_1)$ [kg DM/*J* PAR] accounts for the conversion of this absorbed energy into biomass (radiation use efficiency) and for the losses related to the transport of photosynthetates, maintenance of the standing phytomass, etc.

A linear relationship between fAPAR and NDVI has often been demonstrated in literature (William and Myneni, 1994). Hence, the fAPAR-values are estimated for each 1km-pixel from the NDVI-imagery by means of the equation fAPAR=A+B.NDVI, in which the intercept A and slope B are adapted to the sensor type. A “heuristic” calibration on Belgian data (Eerens et al., 2000) yielded the following values: A=-0.247, B=1.54 for SPOT-Vegetation.

Maximum NDVI, maximum DMP and sum NDVI are indicators, which provide information about wheat yield prediction. The maximum NDVI is derived as the highest NDVI values over the year and the maximum DMP is also derived as the highest DMP values over the year. But sum NDVI is integrated NDVI values over the year.

At last, we build our yield estimation model based on linear regression analysis between the indicators and residuals just mentioned. These are combined with trend yield, residuals from the trend and satellite based indicators to forecast yields. For the current year, we use our model in forecasting winter wheat and spring wheat production in Uzbekistan. The resulting productivity should be consistent with other existing data.

3.8. Implement CGMS Statistical Tool for Uzbekistan Government

A crop model can make a yield prediction at any time during growing season. But, if the prediction is going to be presented early in the season, then the CGMS Statistical Tool will be a very good example to implement for Uzbekistan Government for the future. Of course this must be agreed with the owner of the CGMS Statistical tool: JRC of EC. The weather conditions leading up to harvest time are unknown and are therefore a major source of uncertainty. Early in the season simulated crop indicators are still at a very early stage and the regression built on these crop indicators is not robust. Operational CGMS yield forecasts generally improve as the growing season progresses (Genovese et al., 1998). The advantage of meteorological conditions is that the actual growing conditions are used and their use as input in the crop model could provide better crop growth indicators, which in turn could improve the statistical regression for the yield prediction. If the CGMS model could provide more reliable crop yield forecasts early in the season, it will give valuable information to agricultural production managers at the Ministry of Agriculture and Water Resourses in Uzbekistan.

4. Results

This chapter is divided into two parts: the first part explains the result of table 4.1 for rainfed areas and the second part explains the results of table 4.2 for rainfed plus irrigated area both at regional and national levels.

4.1. Results for rainfed areas

Table 4.1 shows the results generated by the CGMS Statistical Tool for rainfed areas. This table shows seven indicators which are divided into two parts. First part of these results show indicators from CGMS, which are: PYB, PYS, WYB and WYS and the second part shows the SPOT Vegetation indicators which are driven by remote sensing. These indicators are: max_NDVI, max_DMP and sum_NDVI. This table also shows the different growing season of spring and winter wheat. The cropland which is shown in table 4.1 is just to see how many hectares [ha] were sown with wheat in each region.

Three examples were chosen to illustrate wheat yield prediction at regional level. Table 4.1 shows the results of the predicted wheat yield in Uzbekistan for non irrigated area both at regional and national level. For example, in Samarkand region only four indicators (PYB, PYS, max_NDVI and max-DMP) out of seven show better results of wheat yield prediction than the trend for spring wheat. These four indicators show a lower value compared to the trend value in terms of RMSE. From these four indicators, potential above ground biomass (PYB) shows best result at this period. This indicator is driven by CGMS data. Winter wheat shows only two indicators (max_NDVI, max_DMP) with better results of wheat yield prediction in Samarkand region. These two indicators are driven by SPOT-Vegetation data. From these two indicators maximum NDVI shows best result for this period.

In Surkhandarya region no indicators show good result for spring wheat. However, in chapter 3 two factors were mentioned, if the indicator does not represent as a best one for a certain region. Second possibility to explain year to year variation in Uzbekistan is the trend. In this case trend shows a lower value than the seven indicators. For that reason the trend would be a second option as representative indicator for this region, to explain year to year variation for this season. Winter wheat shows maximum NDVI as the best indicator in this period. This indicator is driven by SPOT-Vegetation data.

Table 4.1 shows no wheat yield prediction in Karakalpak region.

The results of table 4.1 show that at national level have also seven indicators, but only maximum NDVI indicator shows the best result for whole country. This indicator is derived from SPOT-Vegetation data.

Table 4.1 Results of Root Mean Square Error of prediction values (rainfed area)

code No	Name / Regions	Crop land, [ha] (1998-2006)	Wheat, season time	CGMS indicators				SPOT-Vegetation indicators				Trend
				PYB	PYS	WYB	WYS	max_NDVI	max_DMP	sum_NDVI	trend	
UZB	Uzbekistan	1255666	spring	0.21	0.28	0.22	0.37	0.24	0.26	0.28	0.23	linear
			winter	0.24	0.27	0.21	0.22	0.19	0.24	0.3	0.23	
UZB-002	Andijan	71478	spring	0.37	0.36	0.36	0.35	0.42	0.33	0.39	0.39	quadratic
			winter	0.4	0.4	0.39	0.4	0.41	0.31	0.49	0.39	
UZB-003	Bukhara	71792	spring	no	no	no	no	no	no	no	no	linear
			winter	0.36	0.36	0.33	0.36	0.41	0.41	0.42	0.31	
UZB-004	Fergana	107389	spring	0.26	0.29	0.35	0.37	0.26	0.27	0.38	0.33	linear
			winter	0.38	0.37	0.38	0.41	0.37	0.34	0.38	0.33	
UZB-005	Karakalpak	39034	spring	no	no	no	no	no	no	no	no	linear
			winter	no	no	no	no	no	no	no	no	
UZB-006	Kashkadarya	199361	spring	0.57	0.58	0.51	1.98	0.58	0.54	0.57	0.49	linear
			winter	0.42	0.39	0.45	0.37	0.34	0.44	0.53	0.49	
UZB-007	Khorezm	37157	spring	0.68	0.61	0.71	2.18	0.64	0.73	0.76	0.68	no trend
			winter	no	no	no	no	no	no	no	no	
UZB-008	Namangan	74538	spring	0.26	0.26	0.26	0.25	0.25	0.25	0.24	0.24	linear
			winter	0.27	0.31	0.29	0.27	0.25	0.26	0.27	0.24	
UZB-009	Samarkand (comb)	18379	spring	0.31	0.36	0.41	0.97	0.33	0.34	0.41	0.37	linear
			winter	0.41	0.44	0.4	0.41	0.26	0.32	0.39	0.37	
UZB-010	Surkhandarya	102384	spring	0.73	0.57	0.58	0.52	0.85	0.83	0.82	0.72	linear
			winter	0.75	0.74	0.79	0.79	0.85	0.9	0.79	0.72	
UZB-011	Syrdarya (comb)	25678	spring	0.17	0.18	0.11	0.09	0.16	0.16	0.17	0.15	linear
			winter	0.14	0.18	0.13	0.12	0.14	0.16	0.17	0.15	
UZB-012	Tashkent oblast	111963	spring	0.31	0.29	0.31	0.34	0.31	0.33	0.34	0.28	linear
			winter	0.29	0.31	0.3	0.34	0.35	0.31	0.41	0.28	

Legends of table 4. 1

Cyan	Best
Green	Good
Yellow	Sufficient
Orange	Normal
Magenta	Bad
Red	alarm trend
White	Not significant

The legend of table 4.1 shows the rank of indicators for wheat crop at regional and national level. This legend was dedicated based on the results of RMSE in the CGMS Statistical Tool. The rank was characterized based on indicator values generated in CGMS model. The lower RMSE value was characterized as the best indicator while the higher RMSE value was characterized as bad indicator. The indicator value was ranked from best to bad when its value is lower compared to the trend value. If the indicator value is higher than the trend value it indicates an alarm. The white color in legends shows not significant wheat prediction by the model.

4.1.1. Spring wheat yield prediction for 2007 in Andijan region.

In the results of table 4.1 for rainfed area, all regions have similar procedure, but this section gives a detail results for Andijan region. It has seven indicators for one region, but only one of them gives the best result. As an example of the result for table 4.1 the Andijan region will be presented. First, results for spring wheat in the Andijan region were analysed to get a prediction of yield in 2007 as shown below.

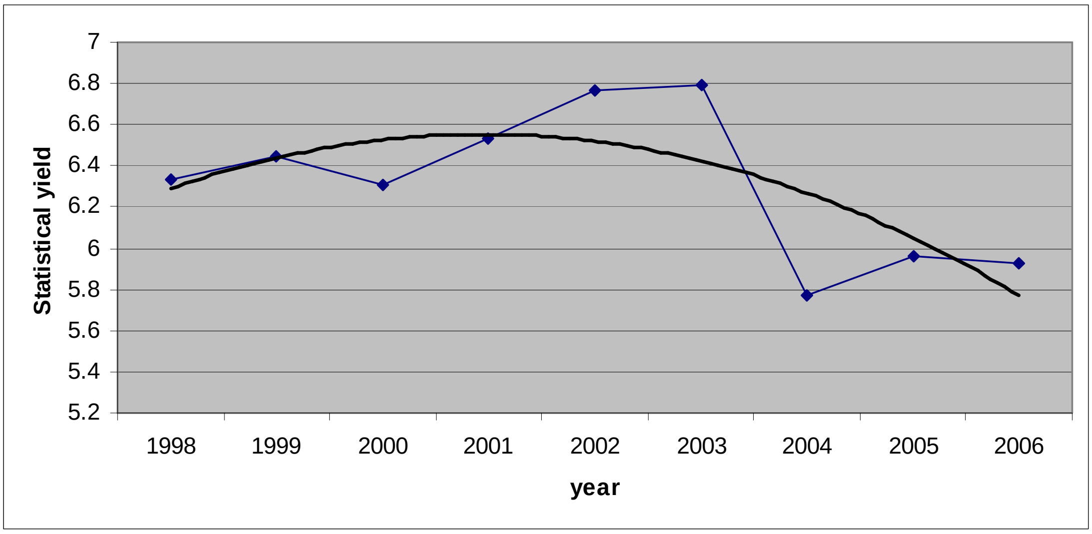


Figure 4.1 Yield Statistics of spring wheat in the Andijan region

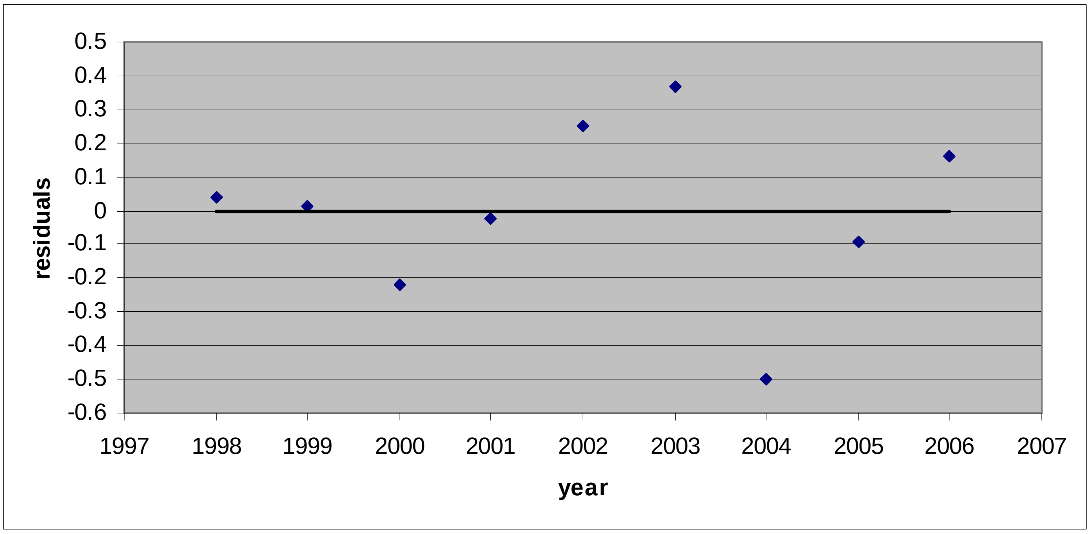


Figure 4.2 Deviation from the trend for spring wheat in the Andijan region

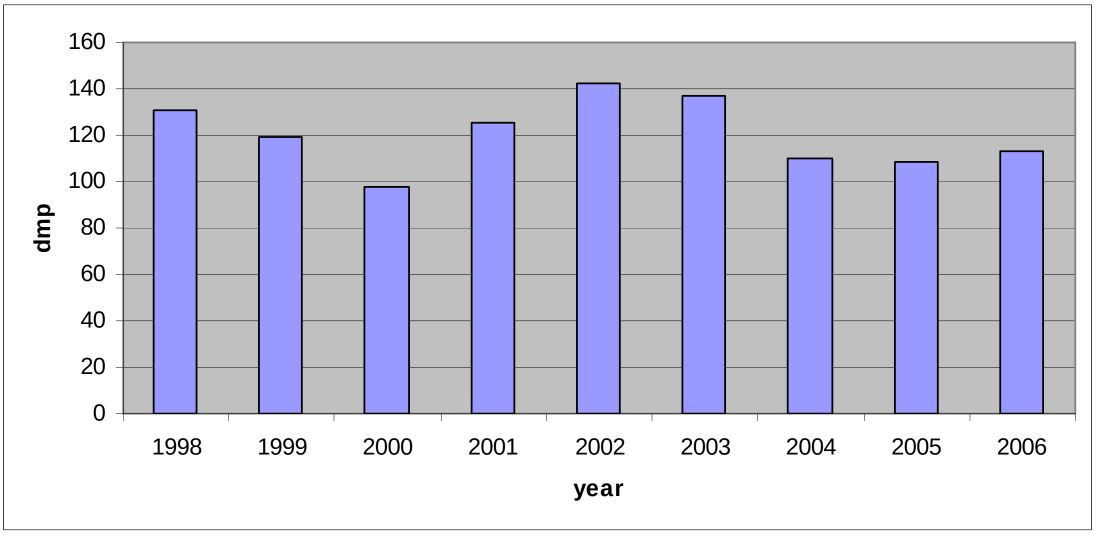


Figure 4.3 Maximum Dry Matter Productivity according to SPOT-Vegetation data in the Andijan region.

Yield figures from 1998 until 2006 in figure 4.1 show fluctuation of increasing and decreasing wheat yield for each year. Figure 4.1 also shows the fitted trend model. Of the tested functions a quadratic function performed best. Figure 4.2 shows the deviation from the trend, which is defined in terms of residual. Figure 4.2 shows the variation in residuals for each year and this variation can be compared with, for instance, the indicator dry matter productivity. Figure 4.3 shows the maximum dry matter productivity for the year 1998 to 2006 as derived from SPOT-Vegetation data by VITO according to equation (3.6). The pattern of figures 4.2 and 4.3 indicate that there is a decrease of yield from 1998 till 2000 and from 2000 it started to increase until it reaches the maximum yield in 2003. From that time the yield started to decrease till 2004, and increase again until 2006.

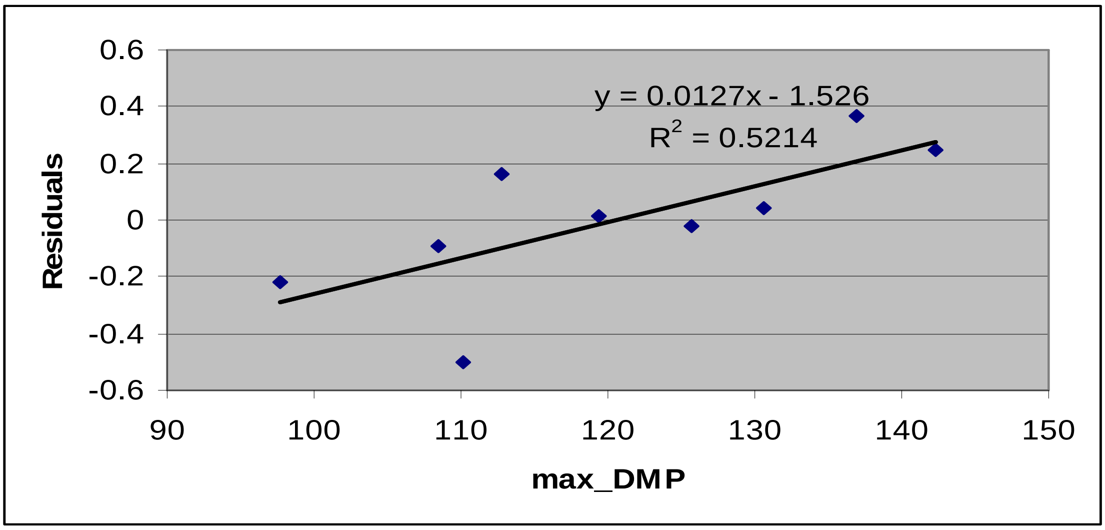


Figure 4.4 Correlation between residuals and dry matter productivity for spring wheat in the Andijan region.

Figure 4.4 shows the graph of residuals against maximum dry matter productivity. It indicates the positive correlation between the residuals and the dry matter. The R^2 is equal to 0.52 and indicates a good correlation between them, showing that this DMP may be used as an indicator for deviations from the general year to year yield trend for spring wheat in this region.

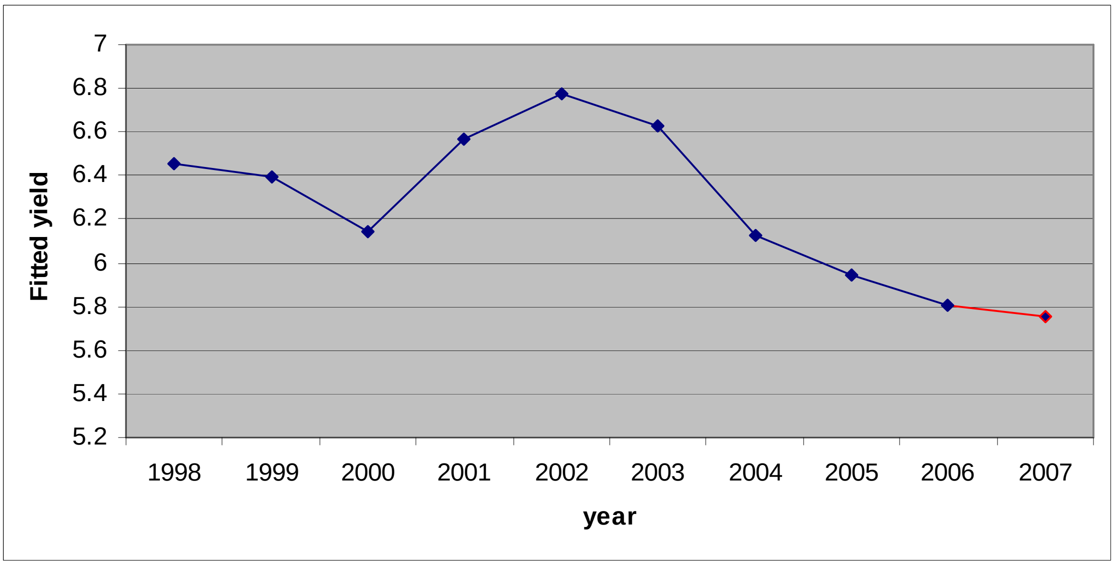


Figure 4.5 Prediction of spring wheat in the Andijan region for 2007

Figure 4.5 shows the results of the CGMS Statistical Tool. This is the yield prediction for 2007 in red color. It shows that this year wheat production will decrease.

4.1.2. Spring wheat yield prediction for 2007 in Fergana region

Results of spring wheat in Fergana region are showing two indicators as best one. Only one of them can be used as a representative result for Fergana region. This is a second example of the results for table 4.1. This example can be used as a result that explains two indicators in this case, but only one of them can be the best. These indicators are PYB and maximum NDVI. First, the indicator PYB will be analyzed for spring wheat in the Fergana region to get a prediction of yield in 2007.

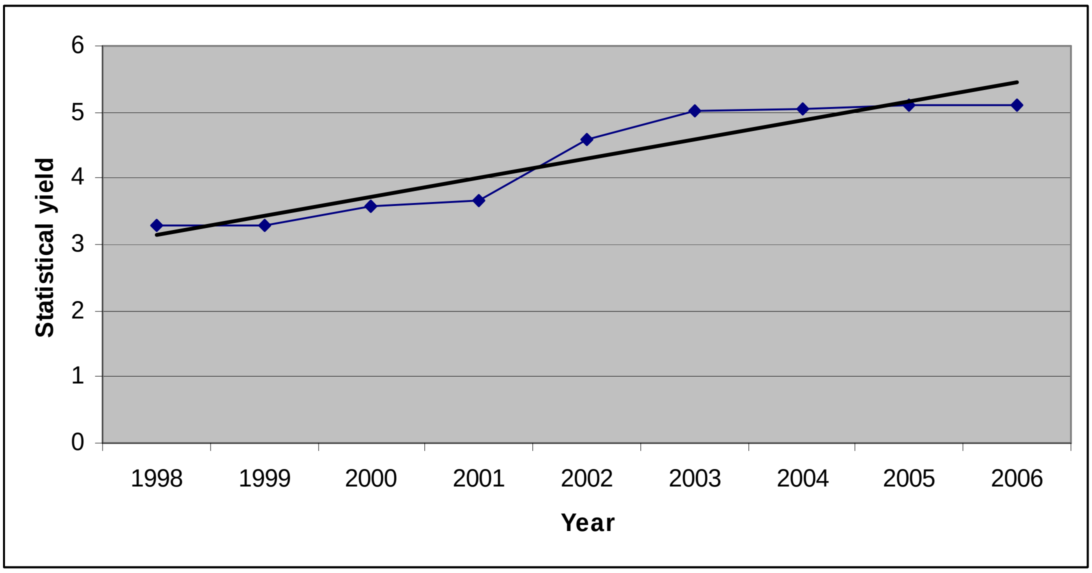


Figure 4.6 Yield statistics for spring wheat in the Fergana region

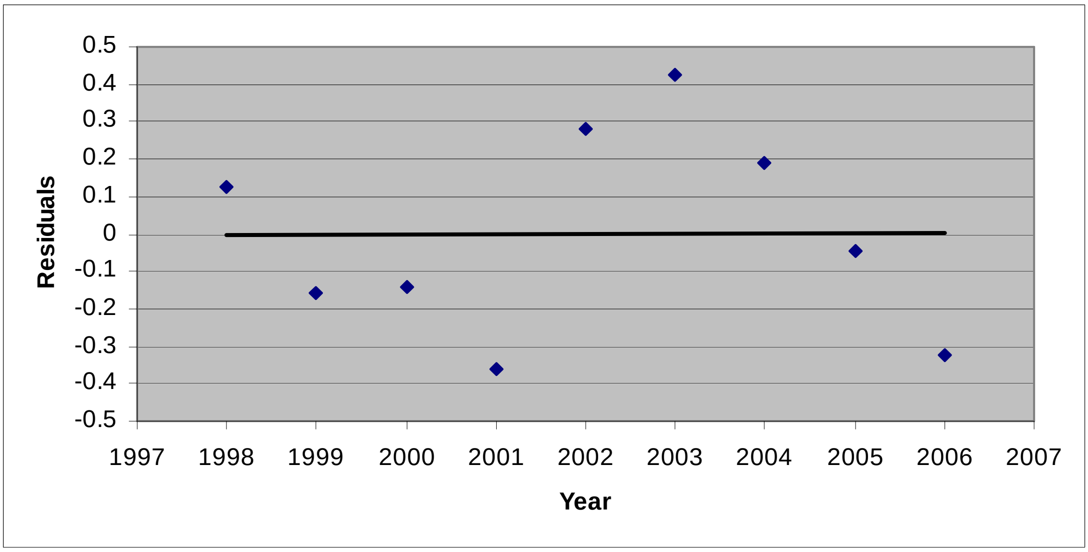


Figure 4.7 Deviation from the trend for spring wheat in the Fergana region

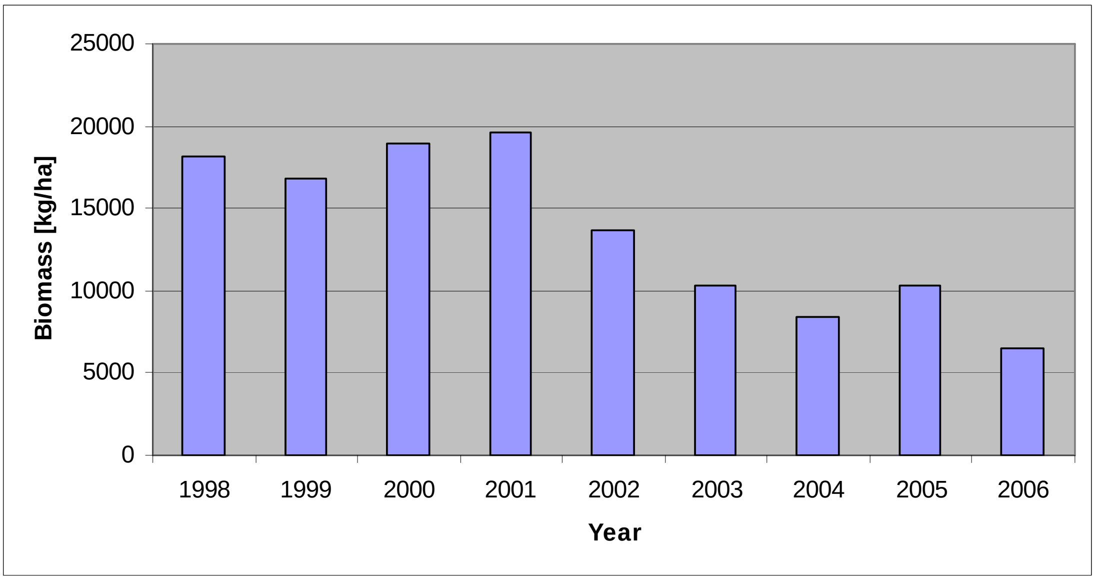


Figure 4.8 Potential above ground biomass according to CGMS model in Fergana region

Figure 4.6 shows yield statistics of collected yield from 1998 until 2006 showing fluctuation of increasing and decreasing wheat yield for each year. Figure 4.6 also shows the fitted trend model. Of the tested functions a linear function performed best. Figure 4.7 shows the deviation from the trend, which is defined in terms of residuals. Figure 4.7 shows the variation in residuals for each year and this variation can be compared with potential above ground biomass as indicator. Figure 4.8 shows the potential above ground biomass for the year 1998 to 2006 as derived from CGMS model. The pattern of figures 4.8 indicate that there is a decrease of yield from 1998 till 1999 and increase from 1999 till 2001, after this time from 2001 it started to decrease until it reaches the minimum yield in 2006, but in 2005 it had a bit better yield than in 2004 and 2006. Since the pattern of figure 4.8 does not reflect the pattern of figure 4.7, consequently, indicator (PYB) shows bad prediction. According to the results in table 4.1 this indicator (PYB) shows the best results for Fergana region. However, CGMS model does not explain wheat yield variability, but CGMS Statistical Tool facilitates the model in choosing which of them performing best.

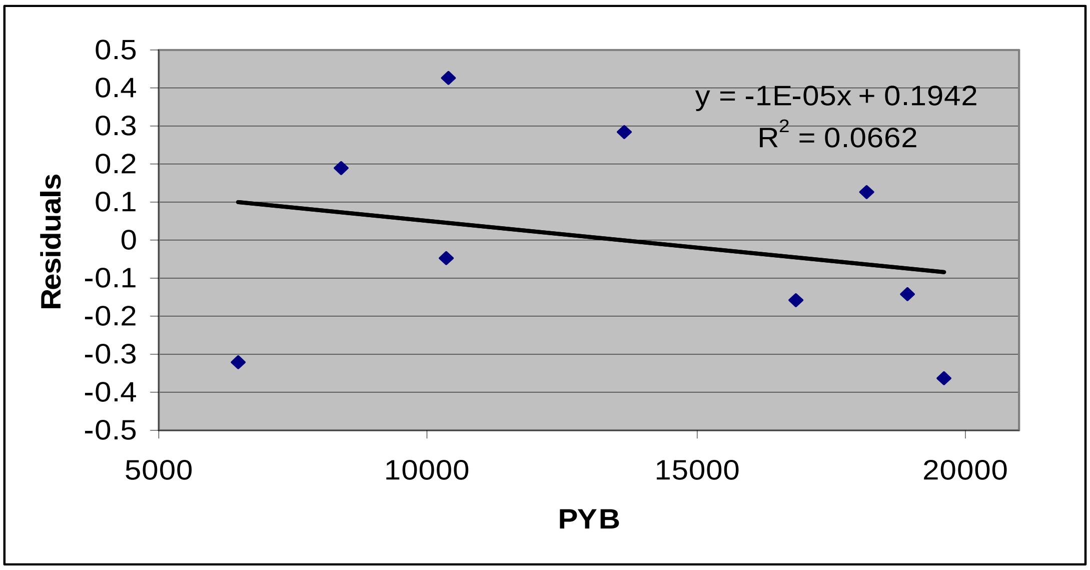


Figure 4.9 Correlation between residuals and potential above ground biomass for spring wheat in the Fergana region

Figure 4.9 shows the graph of residuals against potential above ground biomass. It indicates the negative correlation between the residuals and the above ground biomass. With R^2 equal to 0.06, it indicates poor correlation between them, which illustrates that the statistical tool may give an error for this region.

As second result the indicator maximum NDVI will be analyzed for spring wheat in the Fergana region to get prediction of yield in 2007. Previous indicator in figure 4.8 shows negative correlation between the residual and indicator PYB, which can not be used to explain the year to year variation. Following section will analyze the NDVI indicator.

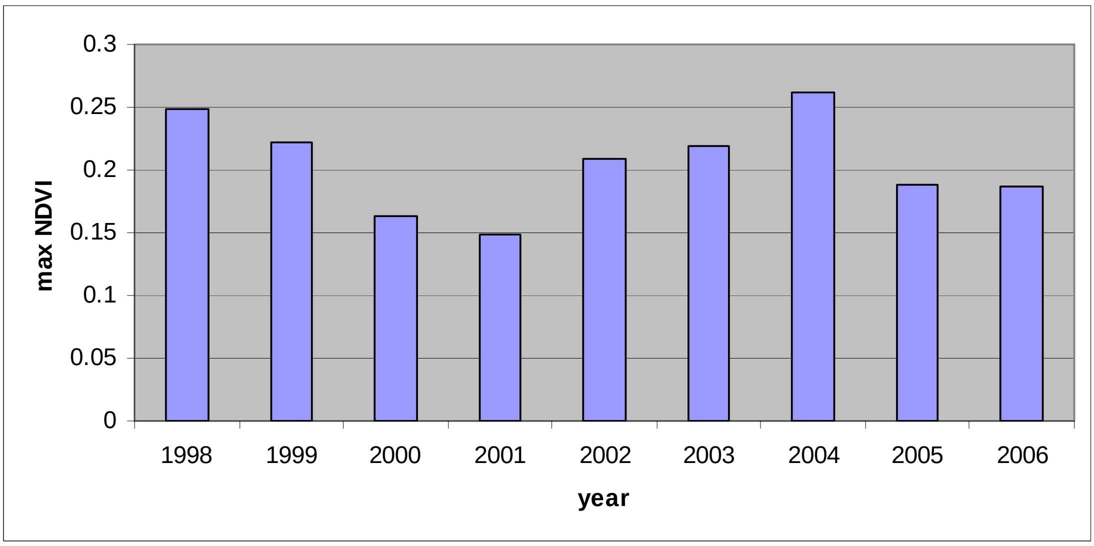


Figure 4.10 Maximum NDVI according to SPOT-Vegetation data in Fergana region.

Figure 4.10 shows the maximum NDVI according to equation (3.7) for the year 1998 to 2006 as derived from SPOT-Vegetation data. In this section the indicator NDVI is analyzed as second best indicator. Figure 4.7 showed variation in residuals for each year. The pattern of figures 4.7 and 4.10 indicate that there is a decrease of yield from 1998 till 2001 and from 2001 it started to increase until it reaches the maximum yield in 2004 and again decrease until 2006. This figure 4.10 shows a similar pattern to figure 4.7. So, this is a good indicator for this region, which can explain year to year variation.

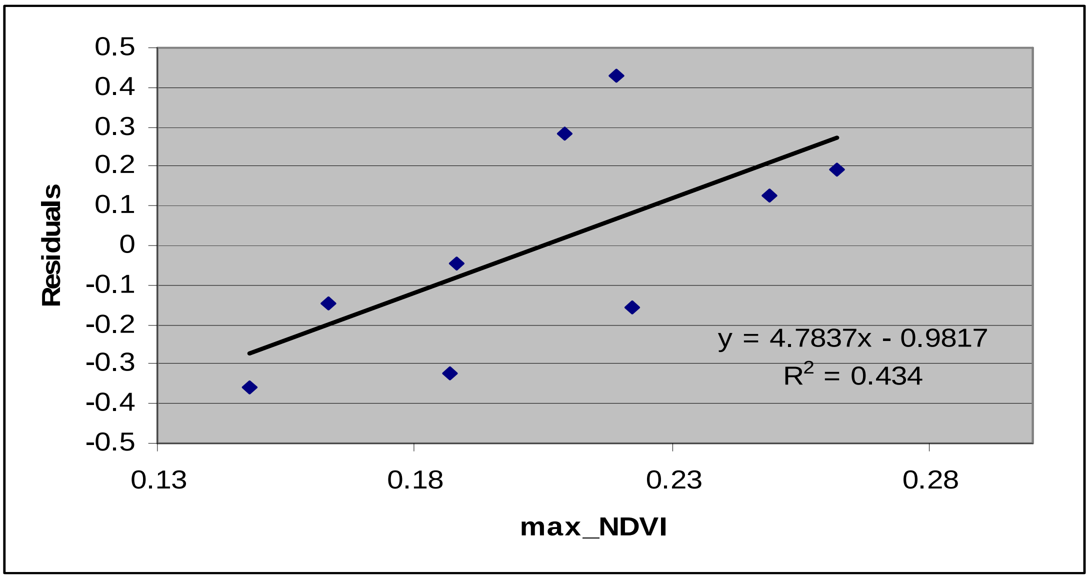


Figure 4.11 Correlation between residuals and maximum NDVI for spring wheat in the Fergana region

Figure 4.11 shows the graph of residuals against maximum NDVI. It indicates the positive correlation between the residuals and the NDVI. It shows an R^2 equal to 0.43, which indicates good correlation between them, showing that this maximum NDVI may be used as an indicator for deviations from the general year to year yield trend for spring wheat in this region.

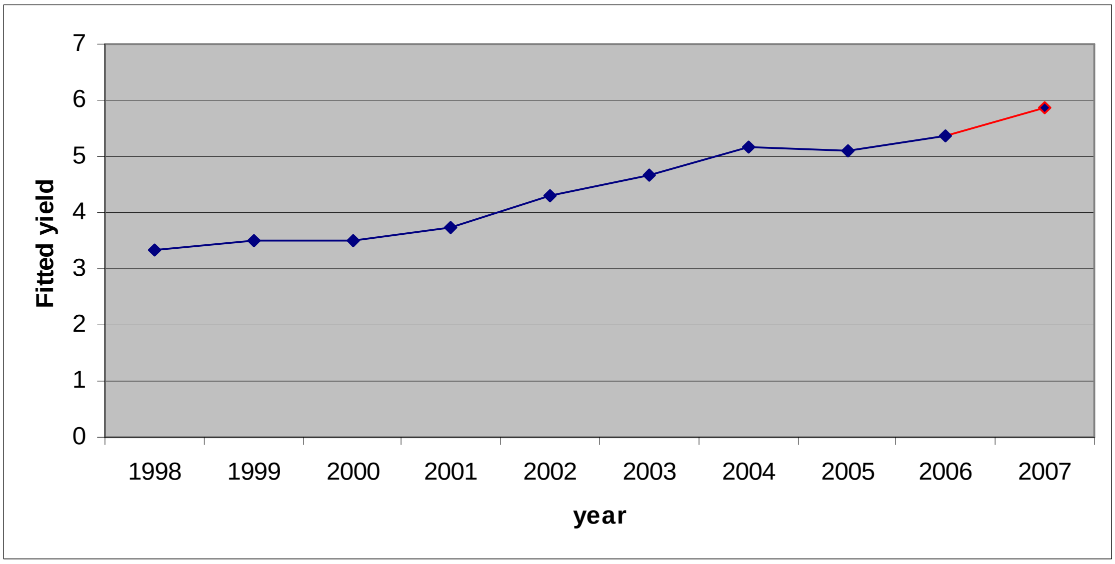


Figure 4.12 Prediction for spring wheat in the Fergana region for 2007

Figure 4.12 shows the results of the CGMS Statistical Tool. This is the yield prediction for 2007 shown in red color. It shows that this year wheat production will increase.

4.2. Results for rainfed plus irrigated areas

This part of results refers to rainfed plus irrigated areas. This section will explain the results of the table 4.2. Table 4.2 shows the results generated by the CGMS Statistical Tool for rainfed plus irrigated area. The model has seven indicators which are divided into two parts. First part of these results show indicators from CGMS, which are: PYB, PYS, WYB and WYS. The second part shows the SPOT-Vegetation indicators which are driven by remote sensing. These indicators are: max_NDVI, max_DMP and sum_NDVI. This table also shows the distinction between spring and winter wheat.

From table 4.2 as an example the results for Samarkand region were chosen. To predict the spring wheat yield the model use seven indicators given by the CGMS Statistical Tool and only four indicators show better result than the trend value for this region. These four indicators are showing a lower value compared to the trend value in terms of RMSE. But from these four indicators maximum DMP shows best result when you look at the range of table 4.2. This indicator is driven by SPOT-Vegetation data. Winter wheat shows as well seven indicators, which is given by the CGMS Statistical Tool. In this case also maximum DMP was chosen as best one for this region.

Table 4.2 Results Root Mean Square Error of predicton values (rainfed plus irrigated area)

code No	Name / Regions	Crop land, [ha] (1998-2006)	Wheat, Season time	CGMS indicators PYB	PYS	WYB	WYS	SPOT-Vegetation indicators max_NDVI	max_DMP	sum_NDVI	trend	Trend
UZB	Uzbekistan	105566	spring	0.22	0.26	0.18	0.26	0.18	0.23	0.28	**0.25**	linear
			winter	0.28	0.29	0.19	0.17	0.09	0.2	0.27	**0.25**	
UZB-002	Andijan	7145	spring	0.37	0.36	0.36	0.35	0.42	0.33	0.39	**0.39**	quadratic
			winter	0.4	0.4	0.39	0.4	0.41	0.31	0.47	**0.39**	
UZB-003	Bukhara	7179	spring	no	no	no	no	No	no	no	**no**	linear
			winter	0.36	0.36	0.33	0.36	0.41	0.41	0.42	**0.31**	
UZB-004	Fergana	10739	spring	0.26	0.29	0.35	0.37	0.26	0.27	0.38	**0.33**	linear
			winter	0.38	0.37	0.38	0.41	0.37	0.34	0.38	**0.33**	
UZB-005	Karakalpak	3902	spring	no	no	no	no	No	no	no	**no**	linear
			winter	no	no	no	no	No	no	no	**no**	
UZB-006	Kashkadarya	13849	spring	0.66	0.67	0.58	1.41	0.72	0.69	0.75	**0.61**	quadratic
			winter	0.46	0.44	0.53	0.4	0.37	0.53	0.66	**0.61**	
UZB-007	Khorezm	3716	spring	0.68	0.61	0.71	2.18	0.64	0.73	0.76	**0.68**	no trend
			winter	no	no	no	no	No	no	no	**no**	
UZB-008	Namangan	7454	spring	0.26	0.26	0.26	0.25	0.25	0.25	0.24	**0.24**	linear
			winter	0.27	0.31	0.29	0.27	0.25	0.26	0.27	**0.24**	
UZB-009	Samarkand(comb)	13759	spring	0.67	0.69	0.64	0.89	0.52	0.51	0.63	**0.65**	linear
			winter	0.68	0.64	0.64	0.58	0.57	0.43	0.59	**0.65**	
UZB-010	Surkhandarya	9802	spring	0.76	0.6	0.56	0.61	0.92	0.92	0.91	**0.79**	linear
			winter	0.84	0.83	0.89	0.88	0.92	0.96	0.94	**0.79**	
UZB-011	Syrdarya(comb)	18112	spring	0.16	0.18	0.17	0.17	0.14	0.14	0.14	**0.14**	quadratic
			winter	0.2	0.1	0.15	0.16	0.14	0.15	0.13	**0.14**	
UZB-012	Tashkent oblast	9909	spring	0.32	0.31	0.32	0.36	0.32	0.36	0.36	**0.3**	linear
			winter	0.31	0.32	0.32	0.35	0.37	0.35	0.45	**0.3**	

Legends of table 4.2

	Best
	Good
	Sufficient
	Normal
	Bad
	alarm trend
	Not significant

The legend of table 4.2 shows the rank of indicators for wheat crop at regional and national level. This legend was dedicated based on the results of RMSE in the CGMS Statistical Tool. The rank was characterized based on indicator values generated in the CGMS model. The lower RMSE value was characterized as the best indicator while the higher RMSE value was characterized as bad indicator. The indicator value was ranked from best to bad when its value is lower compared to the trend value. If the indicator value is higher than the trend value it indicates an alarm. The white color in legends shows not significant wheat prediction by the model.

4.2.1. Results of wheat yield at national level

Subsequently, the aggregation method in equation (3.4) was used to predict yield at national level. As an example of the result from table 4.2 we will illustrate only for the rainfed plus irrigated area. To start with, results of wheat yield in Uzbekistan were analysed to get a prediction of yield in 2007.

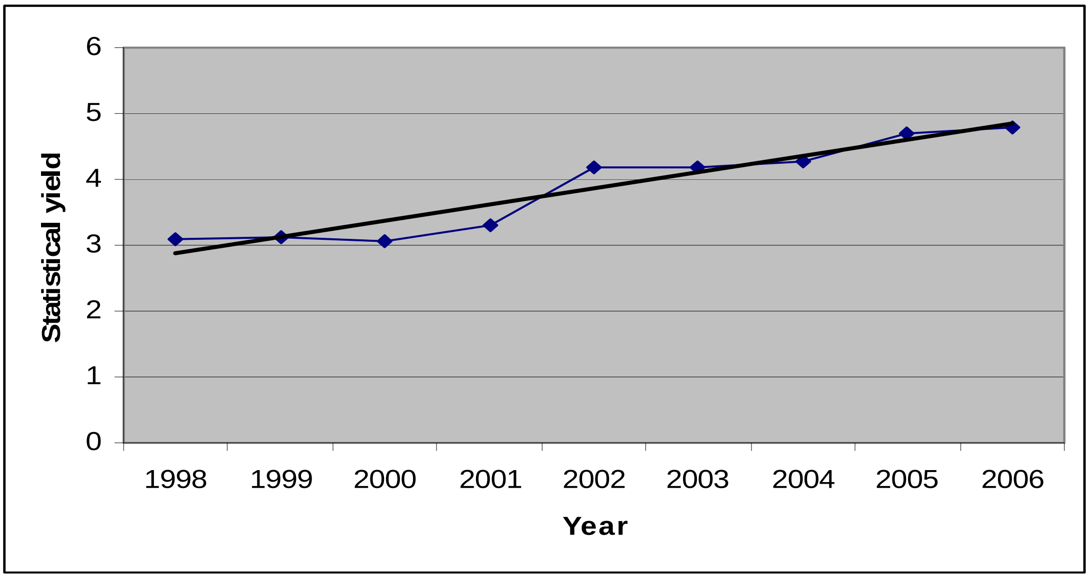


Figure 4.13 Yield statistics for winter wheat at national level

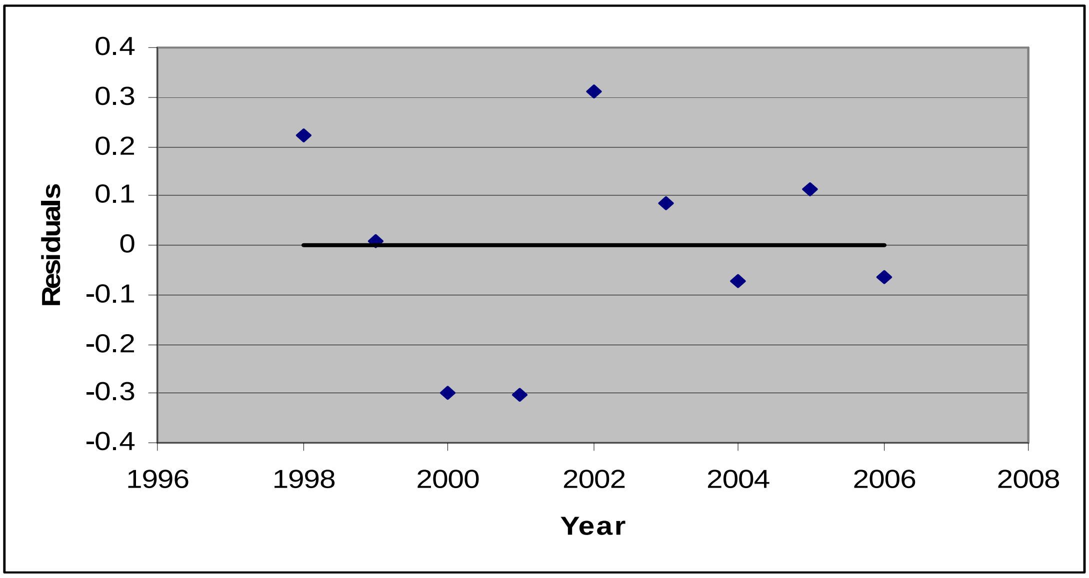


Figure 4.14 Deviation from the trend for winter wheat yield at national level

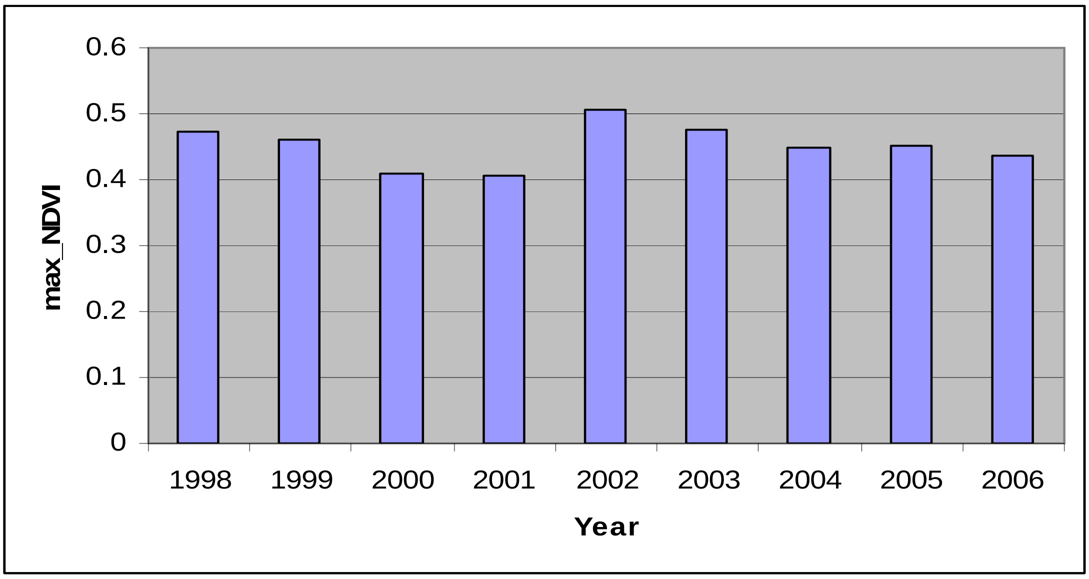


Figure 4.15 Maximum NDVI according to SPOT-Vegetation data at national level.

In this section the results of wheat yield will be analyzed at national level for rainfed plus irrigated area. Yield figures from 1998 until 2006 in figure 4.13 show fluctuation of increasing and decreasing wheat yield for each year. Figure 4.13 also shows the fitted trend model. Of the tested functions a linear function performed best. Results in figure 4.14 show the deviation from the trend, this is defined in terms of residual. Figure 4.14 shows the variation in residuals for each year and this variation can be compared with the indicator maximum NDVI. Figure 4.15 shows the maximum NDVI for the year 1998 to 2006 as derived from SPOT-Vegetation data. The pattern of figures 4.14 and 4.15 indicates that there is a decrease of yield from 1998 till 2001 and from 2001 it started to increase until it reaches the maximum yield in 2002. After 2002 the yield started to decrease till 2004, and increase again in 2005. But in 2006 the wheat trend in Uzbekistan was lower as compared to 2005.

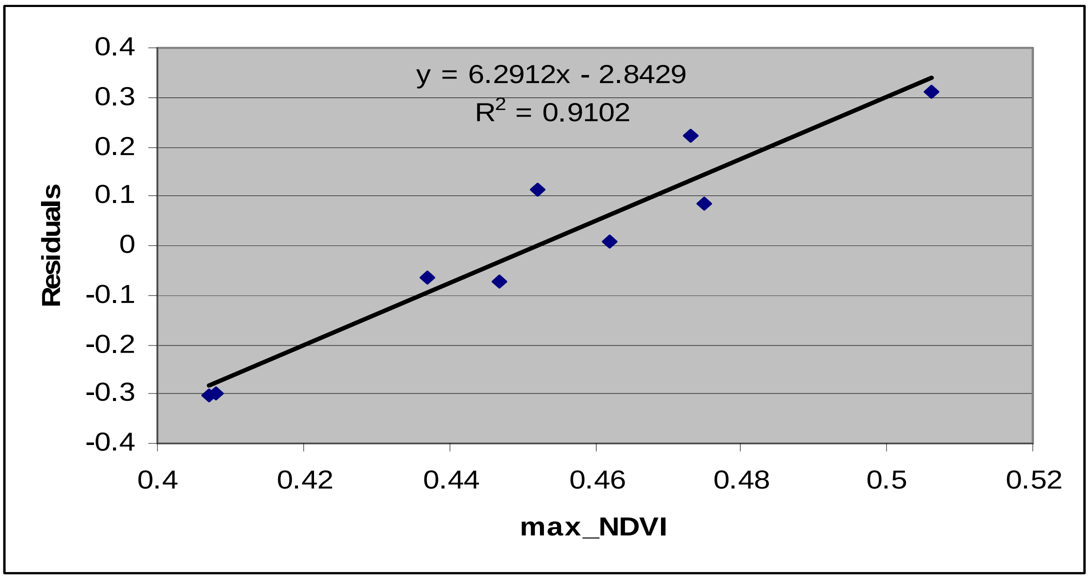


Figure 4.16 Correlation between residuals and maximum NDVI for spring wheat yield at national level

The correlation method in figure 4.16 shows the graph of residuals against maximum NDVI. It indicates the positive correlation between the residuals and the maximum NDVI. The R^2 equal to 0.91 indicates a good correlation between them, showing that this maximum NDVI may be used as an indicator for deviations from the general year to year yield trend of wheat yield at national level.

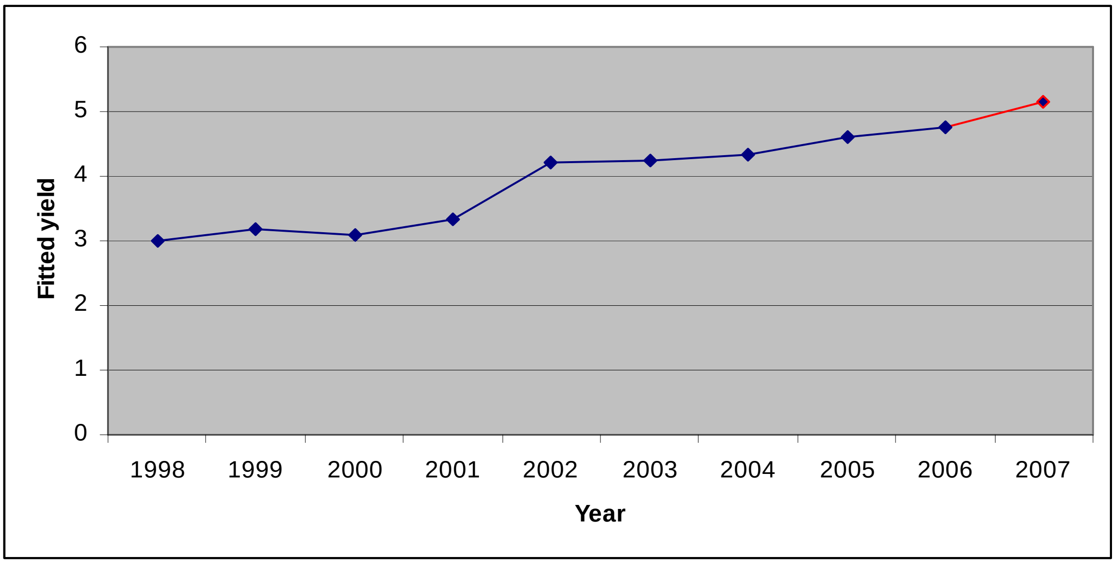


Figure 4.17 Prediction of wheat yield at national level for 2007

Figure 4.17 shows the results of the CGMS Statistical Tool. This is the yield prediction for 2007 in red color. It shows that this year wheat production will increase.

4.3. Comparison of predicted and field data yield for Andijan region at rainfed areas in 2007.

This section will compare predicted and field data yield for Andijan region at rainfed areas in 2007. In this section two different results will be analyzed in rainfed areas. As an example of this section the results in the Andijan region will be presented.

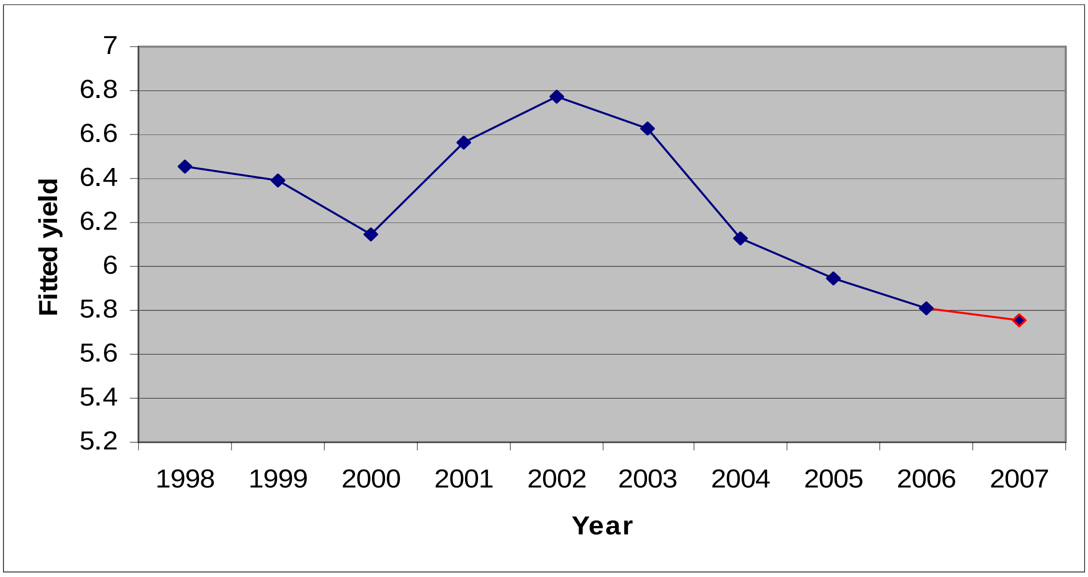


Figure 4.18 Fitted yield and model prediction using DMP indicator for Andijan region

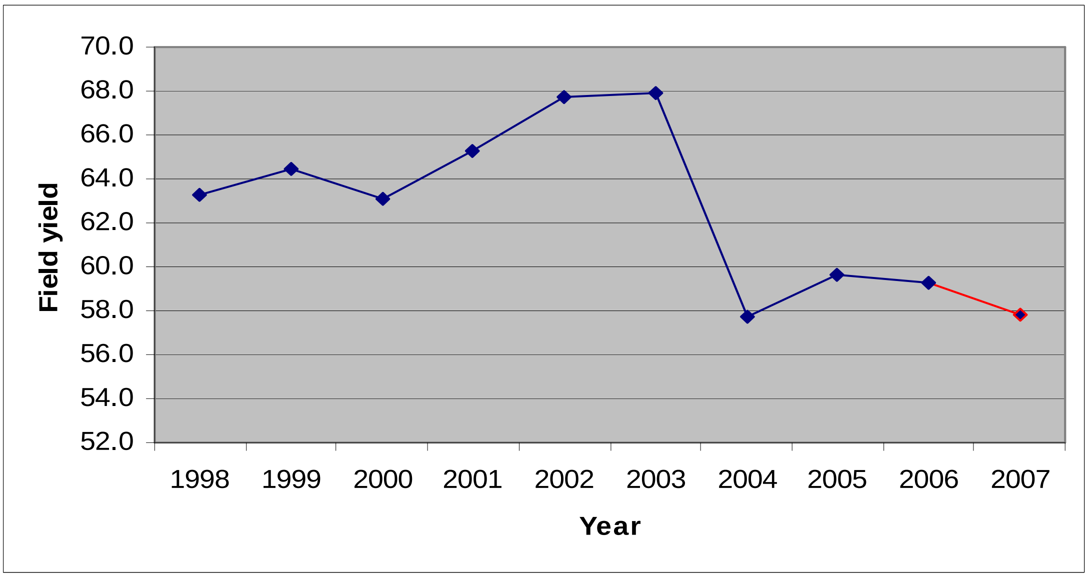


Figure 4.19 Field data collected by the State Committee Statistics of Uzbekistan for Andijan region

Figure 4.18 shows fitted value and the model predicted wheat yield for the year 2007 and in figure 4.19 the field data for the same year are shown. The field data were collected by the State Committee Statistics of Uzbekistan. Figure 4.18 and Figure 4.19 are showing the same result for 2007 in red color. This illustrates that for this region the maximum DMP indicator performed very well.

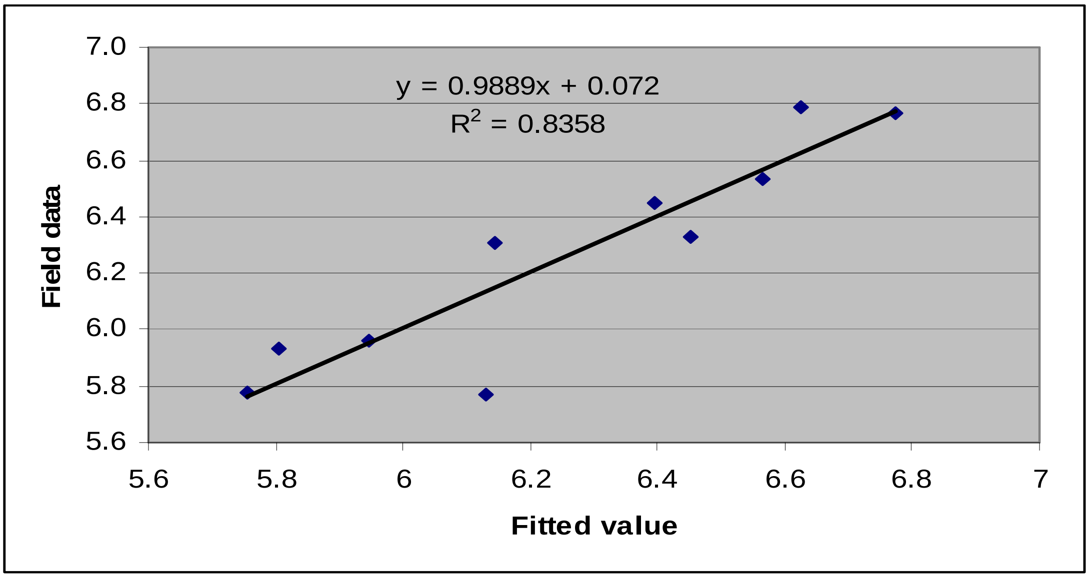


Figure 4.20 Correlation between field data and fitted wheat yield in the Andijan region

Figure 4.20 shows the graph of field data against fitted yield (predicted yield). It indicates the positive correlation between the field data and the predicted values. With R^2 equal to 0.83 it indicates very good correlation between them.

4.4. Comparison of predicted and field data yield for Fergana region at rainfed areas in 2007

In this section the results of wheat yield prediction for Fergana region will be compared with field data yield at rainfed area in 2007.

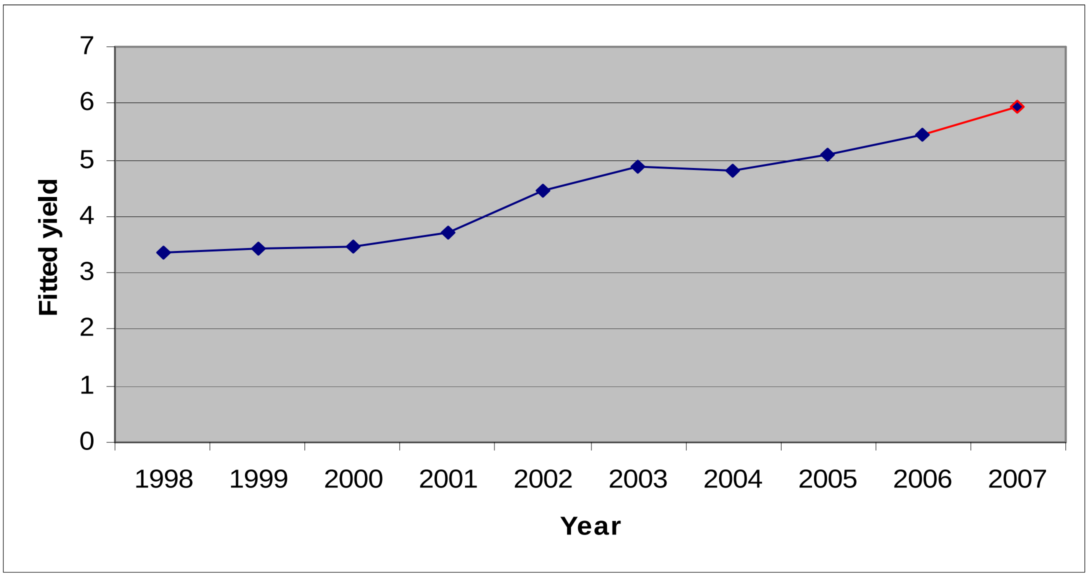


Figure 4.21 Fitted yield and model prediction using the maximum NDVI in Fergana region.

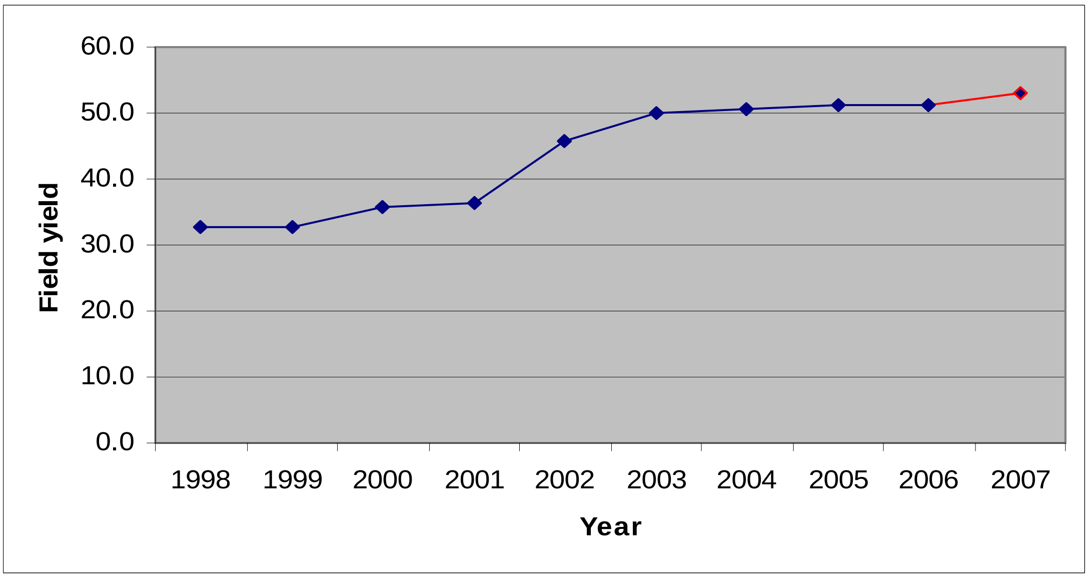


Figure 4.22 Field data collected by the State Committee Statistics of Uzbekistan for Fergana region.

Figure 4.21 shows fitted value and the model predicted wheat yield for the year 2007 and in figure 4.22 the field data for the same year are shown for Fergana region. The field data were collected by the State Committee Statistics of Uzbekistan. Figure 4.21 and Figure 4.22 are showing the same result for 2007 in red color.

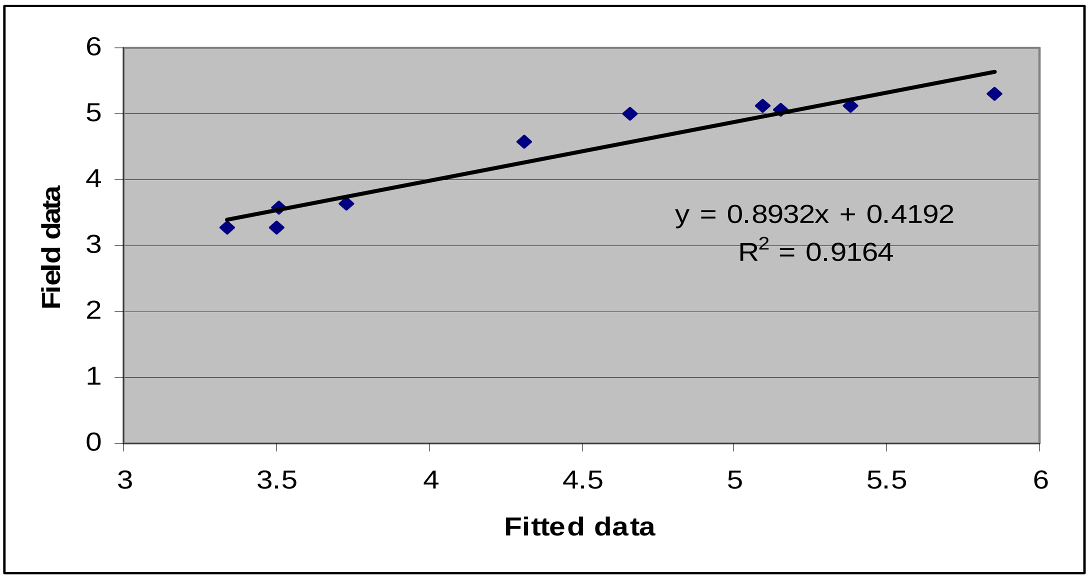


Figure 4.23 Correlation between field data and fitted wheat yield in the Fergana region

Figure 4.23 shows the graph of field data against fitted yield. It indicates the positive correlation between the field data and the predicted values. With R^2 equal to 0.92 it indicates very good correlation between them.

4.5. Comparison of predicted and field data yield at national level in rainfed plus irrigated areas in 2007

This section will compare predicted and field data yield at rainfed plus irrigated areas in 2007 at national level.

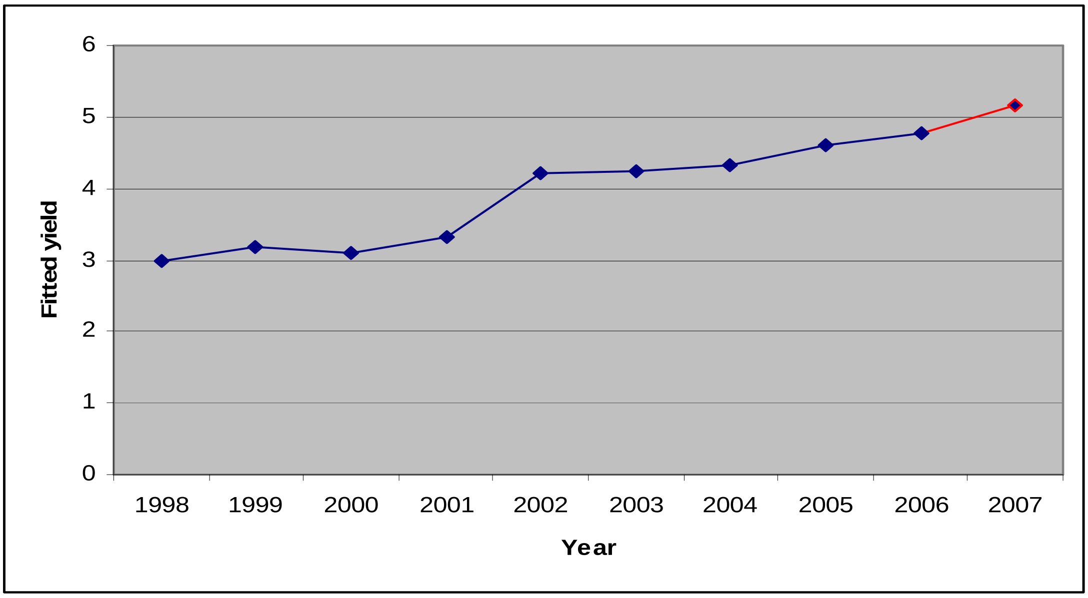


Figure 4.24 Fitted yield and model prediction using the maximum NDVI at national level

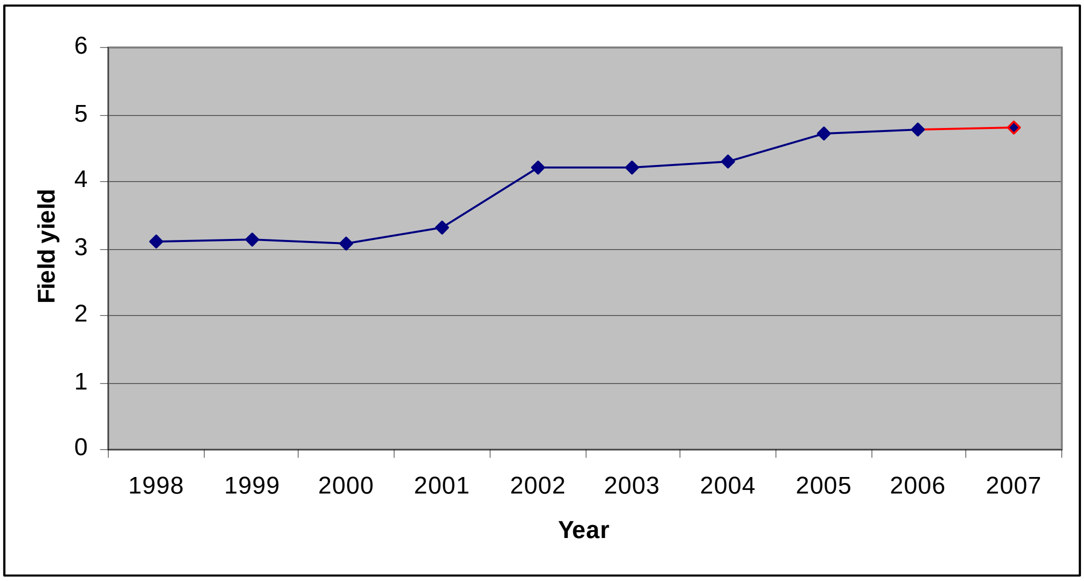


Figure 4.25 Field data collected by the State Committee Statistics of Uzbekistan at national level

Figure 4.24 shows that model predicts wheat yield for the year 2007 and in figure 4.25 field data for the same year are shown at national level. Figure 4.24 and figure 4.25 are showing the same increase of yield for 2007 in red color.

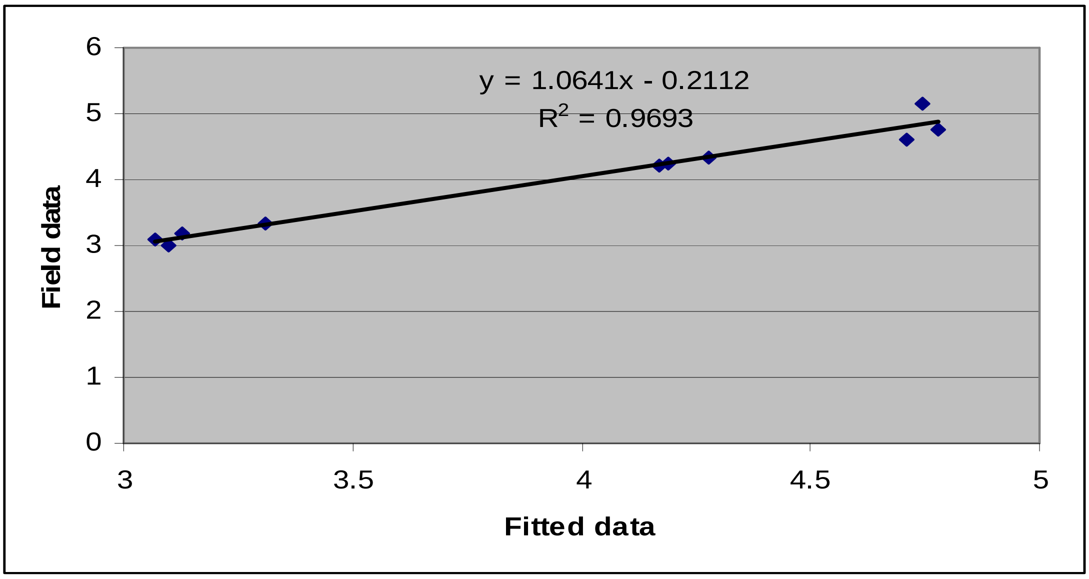


Figure 4.26 Correlation between field data and fitted yield at national level

Figure 4.26 shows the graph of field data against fitted yield (predicted yield). It indicates the positive correlation between the field data and the predicted values. With R^2 equal to 0.96 it indicates very good correlation between them.

4.6. Comparison of predicted and field data yield in rainfed areas in 2007 for all regions

This section will make a comparison for each region only for 2007. Predicted and field data yield in rainfed areas are compared.

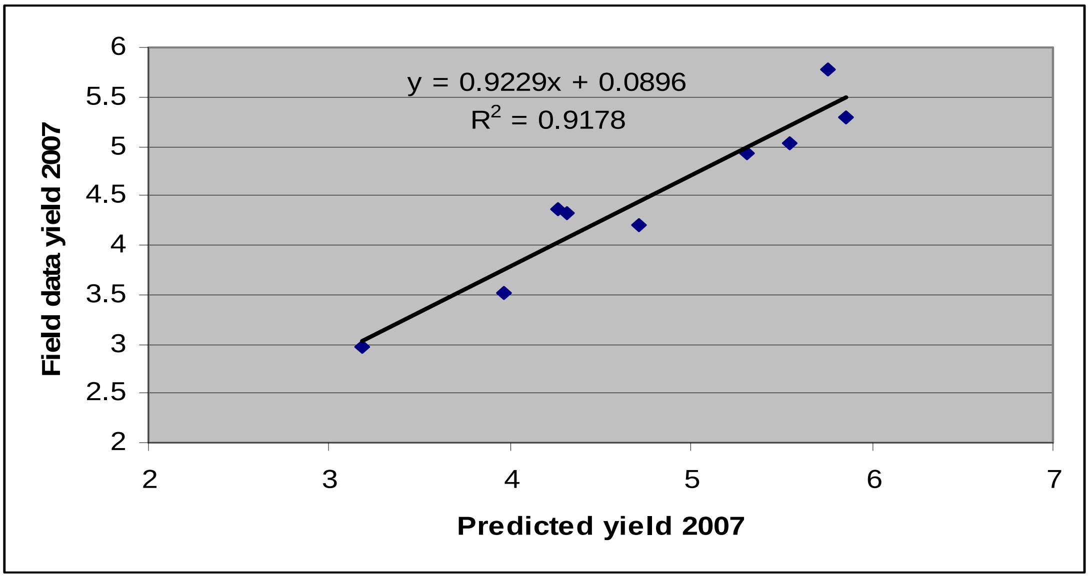


Figure 4.27 Correlation between predicted and field data yield for all rainfed regions per crop in 2007.

Figure 4.27 shows the correlation between predicted against field data yield. This correlation represents both rainfed spring and winter wheat for all regions in 2007. It indicates the positive correlation between the predicted yield and field data yield values. With R^2 equal to 0.91 it indicates very good correlation between them.

4.7. Comparison of predicted and field data yield in 2007 in rainfed plus irrigated areas for all regions.

This section makes a comparison between predicted and field data yield in 2007 in rainfed plus irrigated areas for all regions.

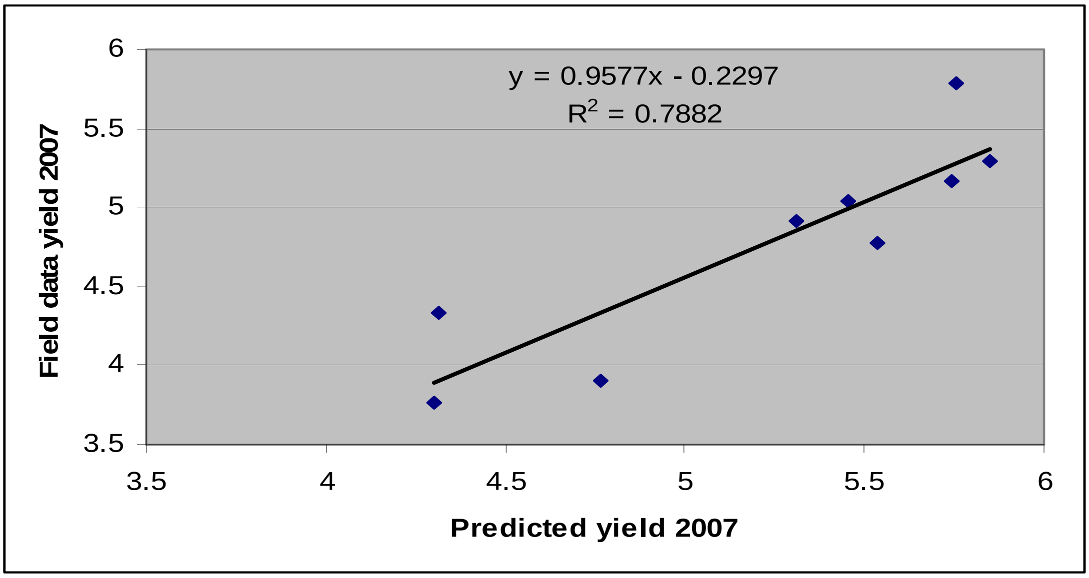


Figure 4.28 Correlation between predicted and field data yield for all rainfed plus irrigated regions per crop in 2007.

Figure 4.28 shows the correlation between predicted yields against field data yield. This correlation represents both spring and winter wheat for all regions in 2007 for rainfed plus irrigated areas. It indicates the positive correlation between the predicted yield and field data yield values. With R^2 equal to 0.78 it indicates very good correlation between them.

5. Discussion

Several simulation models are used in agriculture, which can predict yield of wheat for different regions in the world. However, they cannot predict the yield in the Central Asia well enough, especially in Uzbekistan since this country has high variety of soil and environmental conditions. The CGMS model was developed for such complicated conditions as in Uzbekistan.

5.1. Indicator analysis in rainfed area

Table 4.1 shows different indicators which were used to predict wheat yield, using the statistical model at regional and national level. Indicators maximum DMP, maximum NDVI and WYS have shown very good prediction capability of wheat yield for all regions and as well as for whole Uzbekistan. Maximum DMP and maximum NDVI indicators are driven by SPOT-Vegetation data. The WYS indicator is driven by the CGMS model and this indicator is used to predict yield on non irrigated land which is a good example for the rainfed area. The indicators max_DMP and max_NDVI were used as an example to show the results in chapter 4. Figure 4.2 and 4.3 shows how the evaluation was done based on these results. The correlation method was used in figure 4.4 between the residuals and maximum DMP to look how they are related. At the moment this indicator shows the best result for the Andijan region when compared to other indicators. However, a couple of the indicators such as PYB, PYS, WYB, maximum NDVI and sum NDVI could not predict wheat yield better than the trend for Andijan region.

A second example can be given to show that indicators which are mentioned above for Andijan region have better results than others. Fergana region was chosen for the next example. In this region the indicators PYB and maximum NDVI showed the best results. In the discussion it was mentioned that only one indicator can be the best to explain year to year variation. Looking at figures 4.6 and 4.8 and 4.10 shows how they are correlated to each other. If the indicator PYB is following the residual yield pattern from figure 4.7, it can explain the year to year variation. The correlation was shown in figure 4.9. This result shows the a negative correlation which is quite unpredictable for this region. In the results the indicator PYB was given the best result in Fergana region. However, Uzbekistan has different areas that are not flat land. It has mountains, hills and different factors which could influence wheat yield as well. These factors are not explained by the model. Also weather factors may cause CGMS not to perform in time. The most difficult factor is inside the indicator. These indicators are not showing the image or variation for the area, but only the values which are quite difficult to analyze. The maximum NDVI performs better result to explain year to year variation than PYB. The figures 4.6, 4.7 and 4.10 explain the fluctuation in yield better than the indicator PYB in this case. Figure 4.11 explains the correlation between the residuals in figure 4.7 and the indicator maximum NDVI in figure 4.10. Figure 4.11 shows a positive correlation between residuals against the maximum NDVI. In this case the maximum NDVI will be the best indicator to explain year to year variation.

The indicator sum_NDVI could not predict wheat yield for a specific region nor for the whole Uzbekistan.

5.2. Indicator analysis in rainfed plus irrigated area

Similar results were found in the rainfed plus irrigated areas with a good explanation of wheat yield by PYB, PYS, WYB, maximum NDVI and maximum DMP and on insignificant relation between the indicator sum_NDVI and wheat yield.

The indicators PYB and PYS performed better on irrigated lands and WYB on non irrigated lands. The indicators max_NDVI and max_DMP performed well for any type of landuse. From tables 4.1 and 4.2 the rainfed area shows better results than the rainfed plus irrigated area. However, at national level rainfed plus irrigated area shows the better result. One example can be introduced at national level. At this level spring and winter wheat show the maximum NDVI as best one. In figure 4.13 - 4.15 we observed that msximum NDVI can be used for spring wheat to explain year to year variation. Iin figure 4.16 it shows the high correlation between the residuals against the indicator maximum NDVI. This shows a positive correlation in this case.

Conclusion from these two tables can be drawn that the maximum NDVI and maximum DMP could predict wheat yield at both regional and at national level for rainfed or rainfed plus irrigated areas. However, some regions were not showing good results, this could be caused by intensive climate change factors.

5.3. Comparison of predicted and field data yield 2007 for Andijan region at rainfed areas in 2007

In Andijan region the predicted yield value of 2007 and field data yield have been compared for rainfed area. Figure 4.20 shows the correlation between the fitted yield and field data in Andijan region. They are positively correlated to each other with R^2 equal to 0.83. The slope is 0.98, which indicates that the variation between the predicted and observed value is small.

5.4. Comparison of predicted and field data yield for Fergana region at rainfed areas in 2007

Figure 4.23 shows the correlation between the fitted yield and the field data in Fergana region are positively correlated to each other with R^2 equal to 0.91. The slope is 0.89, which indicates the variation between the predicted and observed value is small.

5.5. Comparison of predicted and field data yield at national level at rainfed plus irrigated areas in 2007

Figure 4.26 shows the correlation between the fitted yield and the field data in Fergana region. They are positively correlated to each other with R^2 equal to 0.96. The slope is 1.06, which indicates the difference between the predicted and observed value is small.

5.6. Comparison of predicted and field data yield in rainfed areas in 2007 for all regions

Predicted and field measured yield has been compared for rainfed area in each region in 2007. Figure 4.27 shows the correlation between the predicted yield and field data yield for each region. They are positively correlated to each other with R^2 equal to 0.91. The slope is 0.92, which indicates that the variation between the predicted and field data yield is very small. This means the model fits perfect.

5.7. Comparison of predicted and field data yield in rainfed plus irrigated areas in 2007 for all regions

The comparison between the predicted and field data yield for each region per crop in 2007 is also important for rainfed plus irrigated areas. Figure 4.28 shows the correlation between the predicted yield and field data yield for each region. They are positively correlated to each other with R^2 equal to 0.91. The slope is 0.96, which indicates that the variation between the predicted and field data yield is very small. This means the model fits perfect.

6. Conclusions and Recommendation

6.1. Conclusions

The objective of this research was to investigate whether indicators from CGMS or SPOT-Vegetation data can be used in a regression model to explain year to year variation of wheat yield in Uzbekistan during the growing season well before harvest. The strategies for each model were identified and compared in order to evaluate the appropriate model for this research. The CGMS Statistical Tool was found more efficient to satisfy the objective of this research. It was therefore adopted and used to investigate which indicators show best results in explaining year to year variation.

During the set-up procedure, the questions mentioned in chapter 1 section 1.5 were answered. Below general explanation is given.

- Which indicator from CGMS can best be used for explaining yield statistics of wheat in Uzbekistan at both oblast and national level?
  The indicators potential above ground biomass (PYB), potential storage organs (PYS), water limited storage organs (WYS) and water limited above ground biomass (WYB) from the CGMS model were used at both regional and national level in rainfed and rainfed plus irrigated areas. In rainfed areas the indicators WYS and PYB performed best at regional level to explain year to year variation. But indicator PYB performed best at national level to explain year to year variation.
  However, in rainfed plus irrigated areas the PYB, PYS and WYB indicators performed best at regional level. From these three best indicators only the WYB performed best at national level to explain year to year variation.

- Which indicator from SPOT-Vegetation data can best be used for explaining yield statistics of wheat in Uzbekistan at both oblast and national level?
  The indicators maximum normalized difference vegetation index (NDVI), maximum Dry Matter Productiviy (DMP) and integrated NDVI (sum_NDVI) were used at both regional and national level in rainfed and rainfed plus irrigated areas. These indicators are derived from SPOT-Vegetation data. In

rainfed areas the indicator maximum NDVI was performing best at regional and at national level to explain year to year variation.

In rainfed plus irrigated areas the indicator maximum DMP was performing best at regional level. However, maximum NDVI performed best at national level to explain year to year variation.

- Which level (regional or national) can be explained best by indicators in CGMS statistical tool?

  The indicators maximum NDVI and maximum DMP which are derived from the SPOT-Vegetation data are performing best at regional and at national level. It is not only the regional and national level, but also in rainfed and rainfed plus irrigated areas as well. To compare two main input sources (CGMS and SPOT-Vegetation data) into the CGMS Statistical Tool, the output which gives the best model predictions will be chosen. SPOT-Vegetation data are performing best at both regional and national level and both rainfed and rainfed plus irrigated areas as compared to the CGMS model.

- Should we distinguish rainfed and rainfed plus irrigated lands in the analysis?

  This thesis work was done in rainfed and rainfed plus irrigated area. Two different results were obtained from the CGMS Statistical Tool. To distinguish the rainfed and rainfed plus irrigated area we try to compare table 4.1 and 4.2 to see which of these two areas performed best for this model. When you look at these two areas they are showing both satisfactory results. However, rainfed areas are performing better than the rainfed plus irrigated areas, when you look at the indicators shown in both tables.

The following conclusions can be drawn:

- The CGMS Statistical tool has been set up successfully with the procedures analyzing the outputs. Some regions still showed errors which have to be improved further.

- The data which were collected from Uzbekistan were has to be minimum for nine years were the minimum to predict wheat yield accurately.

- The results are showing that the prediction model fits perfectly in Uzbekistan. However, Uzbekistan has very high variety of soil and landuse planning.

- Comparison between predicted and field data yield in 2007 for all regions showed that the model worked well in such difficult environmental conditions.

6.2. Recommendations

Based on the present study, the following recommendations are identified for future consideration, as the demonstrated method is seen as a first step in assessing its applicability and efficiency at both regional and national levels.

1. One should use detailed landuse maps for the system. That could make it easy to read the values of the indicators, which it has 1X1 km resolution.

2. CGMS model can be improved use local weather data information or local landuse map.

3. For further studies, the model implementation needs to be updated as there are some bugs in the model which can be improved.

Additionally, it is reasonable to say on the basis of this thesis that exploring results for wheat yield can be used to inform the Uzbekistan Government on future yield increase or decrease.

7. Reference:

Boogaard, H. , Supit, I., K. van Diepen, C. Eerens, H., Piccard, I., Kempeneers, P., 2002. 'Description of the MARC CROP YIELD FORECASTING SYSTEM (MCYFS)". METAMP report1/3, Alterra and VITO, JRC-contract 19226-2002-02-F1FED ISP NL

Boogaard, H. , K. van Diepen, C. Lazar and G. Genovese Editors and Contributions from , I. Supit, C. Lazar, S. Orlandi, E. Van der Goot, A.H.C.M. Shapendonk, 2004. 'The Methodology of the MARS CROP YIELD FORECASTING SYSTEM" (volume 1, 2, 3, 4)

Bouman, B.A.M., 1992. Linking physical remote sensing models with crop growth simulation models, applied for sugar beet. Int. J. Remote Sensing 13:2565-2581.

Braun, M., Weissteiner, C., Kuhbauch, W., 2004. Regional Yield Predictions of Malting Barley by Remote Sensing and Ancillary Data. Journal title: Proceedings – SPIE the International Society for Optical Engineering, volume 5232, Pages 528-539.

Budde M.E., Tappan, G., Rowland, J., Lewis, J. & 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 (3): 481-498.

Clarke, N., Egli D., Ole, W., Schwab, G., 1996. "Understanding and predicting field-scale spatial variability of wheat growth and yield". Precision Agriculture: Precision Resource Management - phase II, chapter 7.3.

Clevers, J.G.P.W. and Leeuwen, H.J.C., 1994. A framework for monitoring crop growth by combining directional and spectral remote sensing information, remote sensing environ. 50:161-170.

Clevers, J.G., Oscar W.V., Jongschaap, R.E., Desprats, J.F., King, C., Prevot, L., Bruguier, N., 2002. Using SPOT data for calibrating a wheat growth model under Mediterranean conditions. Agronomie 22, 687-694.

Dennet, M.D. and Diego, R., 1980. "Weather and yield of sugar beet, tobacco and wheat in Europe", Agricultural Meteorology, 21:249:263

Genovese, G.P., Vossen, P., Terres, J.M., Rijks, D., 1998. The methodology, result and the evaluation of the MARS crop yield forecasting system. Agrometeorological applications for regional crop monitoring and production assessment, EUR 17735 EN, Joint Research Center of the European Commision, Ispra, Italy, p 67-119.

Genovese, G., Fritz, S., Bettio, M., 2004. A comparison and evaluation of performances among crop yield forecasting models based on remote sensing: results from the geoland observatory of food monitoring, ISPRS Archieves XXXVI-8/W48 Workshop proceedings: Remote sensing support to crop yield forecast and area estimates.

Kastens, J.H., Kastens, T.L., Kastens, D.L.A., Price, K.P., Martinko, E.A. and Lee, R.Y, 2005. "Image masking for crop yield forecasting using AVHRR NDVI time series imagery". Remote Sensing of Environment, Volume 99, Issue 3, 30, Pages 341-356.

Montgomery, D.C., Peck, E.A. and Vining, G.G., 2001. Introduction to Linear Regression Analysis, third edition. Wiley, New York.

Monteith, J.L., 1972. Solar radiation and productivity in tropical ecosystems. J. Applied Ecology, 19: 747-766

Odumodu, O.L, Griffits, J.F., 1980. "Some techniques for predicting winter wheat yields in major wheat producing crop districts of Texas and Oklahoma", Agricultural Meteorology, 22: 267-279

Roijas, O., 2007. Operational maize yield model development and validation based on remote sensing and agro-meteorological data in Kenya. International Journal of Remote Sensing, volume 28, Issue 17, pages 3775-3793

Rouse, L. and Telesca, L., 2006. "Quantifying intra-annual persistent behaviour in Spot-Vegetation NDVI data for Mediterranean ecosystems of southern Italy" , Remote sensing of Environment, volume 101:95-103

Schwab, G., Clarke, N., Egli D. and Ole, W., 1996. "Understanding and predicting field-scale spatial variability of wheat growth and yield". Precision Agriculture: Precision Resource Management - phase II, chapter 7.3.

Solie, J.B., Raun, W.R., Johnson, G.V. and Stone, M.L., 2005. A process for in-season fertilizer nitrogen application based on predicted yield potential. Issued, April 19, U.S. Patent: 6,880,291 B2, Steven, M.D., Biscoe, P.V. and Jaggard, K.W., 1983. Estimation of sugar beet productivity from reflection in red and infrared spectral bands, International Journal Remote Sensing 2: 117-125.

Tucker, C.J., Holben, B.N., Elgin, J.H, Jr., & McMurtrey, J.E., 1980. “Relationship of spectral data to grain yield variation”, Photogrammetric Engineering and Remote Sensing 45 (1980), pp. 657–666.

Williams, D. and Myneni, R., 1994.On the relationship between *f*APAR and NDVI. Remote Sensing of Environment, 19:200-211.

Vossen, P., 1992. Forecasting national crops yields of EC countries: *the approach developed by agriculture project*, Office of Official publications of the EU Luxembourg p159-176.

Vossen, P. and Rijks, D., 1995. Early crop yield assessment of EU countries: *the system implemented by the Joint Research Centre*, Office of Official publications of the EU, Luxembourg, p180

Zhang, F., Wu, B., Liu, C., 2003. Using time series of SPOT-VGT NDVI for crop yield forecasting. Geoscience and Remote Sensing Symposium, volume 1, Issue 21, Page 386 - 388

Web source:

1. (USDA), February, 2002, Foreign Agriculture Service, Central Asia Wheat: 2001/02 Output and 2002/03 Prospects

http://www.fas.usda.gov/pecad2/highlights/2002/02/fsu_06feb02/index.htm

2. Republic of Uzbekistan, portal of state authority, Geography: (web source)

http://www.gov.uz/en/ctx.scm?sectionId=119&contentId=1944

3. "Book of the Year 2004", Encyclopaedia Britannica, GeoHive: Global Statistics population and agriculture

http://www.xist.org/earth/ag_wheat.aspx

4. Vegetation website: May, 2008.

http://www.spot-vegetation.com/home/quick/quick.html

5. SPOT-Vegetation Instrument, AGRECON, May, 2008

http://www.agrecon.canberra.edu.au/products/Satellite_Imagery/Spot_Veg/Spot_Veg.htm

6. GreenSeeker website: Hand-Held GreenSeeker Sensors, April, 2008

http://nue.okstate.edu/Hand_Held/Hand_Held_Pictures.htm

7. Sirus Online, website: SPOT IMAGE

http://sirius.spotimage.fr/PageSearch.aspx?tab=1

8. Nitrogen use efficiency, Oklahoma State University, website, April, 2008

http://www.nue.okstate.edu/

9. Alaska report website, May, 2008

http://agdc.usgs.gov/data/usgs/erosafo/ak_avhrr/ak_avhrr_ndvi.html

10. Study area in Uzbekistan, April, 2008

https://www.cia.gov/library/publications/the-world-factbook/print/uz.html