THE USE OF MODIS DATA TO DERIVE ACREAGE ESTIMATIONS: A CASE STUDY IN ROSTOV REGION OF RUSSIA

S.Fritz, I.Savin, M.Massart and T.Nègre
AGRIFISH Unit, Institute for the Protection and Security of the Citizen, Joint Research Centre, European Commission, Via Fermi 1, 21020 Ispra – Steffen.Fritz@jrc.it

Abstract – The recent developments of remote sensing technology, in particular with its improved spatial and temporal resolution, open new possibilities for estimating crop acreage over larger areas. Remotely sensed data allow the estimation of crop acreage independently of sometimes biased and incomplete sub-national survey statistics. This paper presents a methodology on crop acreage estimations using the MODIS13 16-day composite NDVI product. The suitability of the different MODIS products for acreage estimations is assessed. Particular emphasis is placed on a good quality crop mask. Preliminary results for winter crop area estimates are presented.

Keywords: crop area estimates, MODIS 16-day composites, Remote Sensing, crop acreage.

1. INTRODUCTION

The general approach to crop acreage and yield estimation has been to use high resolution data such as TM/ETM, SPOT or ESR/ SAR data (Ippoliti-Ramilo *et al*., 2003). However, the disadvantage of this method is that the low frequency, the cloud cover (Doraiswamy *et al*., 2004) and the costs limit its use. Moreover, one TM/ETM scene only covers an area of approximately 32 000 square km^2 which often limits its use to small regions. There have been a number of activities in the last years on using coarser resolution data such as AVHRR for crop acreage estimations. However, due to its high geolocation error and its coarse resolution, the estimates purely based on remote sensing data were of relatively high error. More promising sensors which have a higher spatial resolution are the medium resolutions sensors MERIS and MODIS. This study focuses on the use of MODIS data which has the advantage that it is freely available and also provides composites such as the 8-day composite (MOD09Q1) and the 16-day composite (MOD13) product.

For this study MODIS 16-day (MOD13) composites were acquired for the 2000 – 2004 period over the Rostov Oblast of Russia located by the Azov Sea. The paper is divided into 5 parts. Firstly, the different MODIS products are described which are relevant for crop acreage estimations. Secondly, the study site of Rostov is described. In the third part the preparation of the crop mask is explained. In the fourth part preliminary results for the estimation of winter crop acreage are presented. In the last part further work and possible improvements of the method are discussed.

2. THE STUDY AREA: ROSTOV OBLAST

The Rostov Region consists of 55 main administrative units, 23 cities and towns, 25 settlements with town status and 43 rural districts, including 2287 rural settlements. It is situated at the southern end of the eastern European plain. It occupies a vast territory in the Lower Don river basin. The geomorphologic appearance of the Rostov region can be described as a plain divided by river valleys and gullies. The maximum altitude above sea level is 253 meters. The total area of the region is 100 800 square kilometres. The region is 470 kilometres north to south and 455 kilometres east to west. The region has a pleasant mildly continental climate. The average temperature in January is -7°C, and +23°C in July. The yearly average amount of precipitation is near 420 millimetres. The amount decreases west (650 millimetres) to east (around 400 millimetres). Most of the region's territory is farmland, mostly on highly fertile chernozem.

The Rostov Oblast is one of the most important Oblasts in Russia in terms of agriculture. Grain crops are dominant in the region and cover more than half of the total cultivated agricultural area. The region holds fourth place in the Russian Federation with regard to grain crops produced and fifth with regard to vegetable production. Winter wheat is the main type of grain. Growing corn, soy, rice, millet, buckwheat and other groats is also quite widespread. Sunflower seeds are the leading industrial crop in the Rostov Oblast and make up 20% of the total production within Russia. Furthermore gardening and grape growing play an important role.

3. THE APPROPRIATE MODIS PRODUCT

The range of MODIS products is high and it is not straightforward to decide which product is most suitable for crop acreage estimations. The most important criteria for the application presented here is to have a resolution as high as possible. This has also been shown in other studies, for example for the detecting of clear-cuts in Russia.

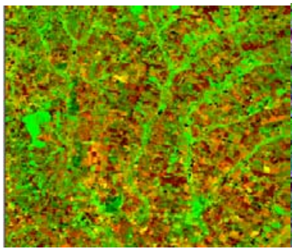
Table 1 describes the available 250m MODIS products which are relevant for crop acreage estimations. The products were available as version 3 and version 4 and it was found that version 4 is in generally higher quality. For our study the MOD13 product was used as it showed to be of sufficient quality. The big advantage is that the vegetation indices EVI and NDVI are also available and no pre-processing and compositing is required.

Therefore, the MODIS13 product was downloaded from NASA's DAAC web site (http://daac.gsfc.nasa.gov/data/) for the 200-2004 time period. One granule (covering an area of of 2300 by 2300 km) was sufficient and covered 97 percent of the Oblast Rostov.

Tab1e 1: Description of MODIS products

PRODUCT	DESCRIPTION	DISADVANTAGE
MOD02 – 1B 250m	Raw product (Top of the atmosphere radiance).	Atmospheric correction needs to be carried out
MOD09 -L2G. 250m	Daily composite atmospherically corrected and georeferenced.	Angle effects are not sufficient taken into account. The product is only useful for compositing if large off nadir angles are discarded
MOD09Q1 Version 3, Version 4	8 day composites of RED and NIR and quality indicators.	The quality depends on the way the compositing was undertaken, no control over the quality of the product.
MOD13 Version 3 Version 4	16 day NDVI and EVI product, also includes RED and NIR cannel and quality indicators of vegetation indices and bands	The quality depends on the way the compositing was undertaken, no control over the quality of the product.

In order to visually evaluate the quality of the MODIS 16 days – 250m composite it was compared with a Landsat Image over an area of 90km by 50km located in the north of the Rostov Oblast. The part of the image covers an area of approximately 4500 square kilometers. On both images natural vegetation along the river valleys, a forest area in the centre-western part and large agricultural fields can be identified. Even though on the MODIS composite certain salt and pepper effects show up, they are reasonably small and the composite is of acceptable quality (see Figure 1a and Figure 1b)



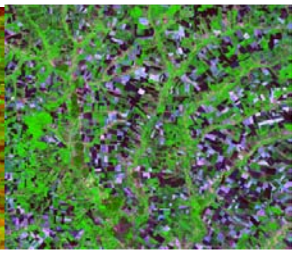


Figure 1a
MODIS 16-days composite (Version 4) 12th August 2002
NIR displayed as Green
Red – Red displayed as red

Figure 1b
Landsat ETM image
30th June 2002 (90/50km)
NIR displayed as green

3.1 EVI versus NDVI

The MODIS 16 days composite product delivers 2 different vegetation indices, the Enhanced Vegetation Index (EVI) and the traditional Normalized Difference Vegetation Index (NDVI). The enhanced vegetation index (EVI) was developed to optimize the vegetation signal with improved sensitivity in high biomass regions and improved vegetation monitoring through a de-coupling of the canopy background signal and a reduction in atmosphere influences (Gao et al., 2000). Available farm data within the study area of Rostov was used to examine the different profiles for winter crops and fallow land (see Figure 2)

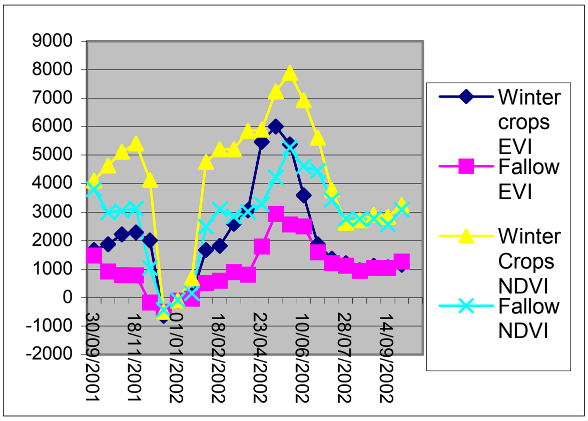


Figure 2: Difference between NDVI and EVI (factor 10000, NDVI 0.1 corresponds to 1000)

Even though the EVI might be better in suppressing background noise, it was decided to use the NDVI for this study. The reason is because the NDVI is more sensitive to lower NDVI or EVI (between 0.1 and 0.6) values which occur during the growing stage of crops (see dates of early spring on Figure 2). Separation between crops in spring time is therefore easier. Nevertheless, a study conducted using the EVI or the EVI in combination with the NDVI could be undertaken in the future and might also prove to provide useful results.

3.2 Geolocation quality

In order to use the MODIS data for acreage estimations the geolocation accuracy is an important issue. Whereas former sensors such as AVHRR were prone to a high geolocation error, the MODIS sensor has much better geolocation accuracy. According to Stroeve et al. (2005) the MODIS geolocation accuracy is better than 150 m at the 99% confidence interval. In this study a visual comparison at a specific location was undertaken. 5 random 16 day composites between 2001 – 2004 where chosen in the summer/autumn period and the geolocation was assessed visually (see Figure 3).

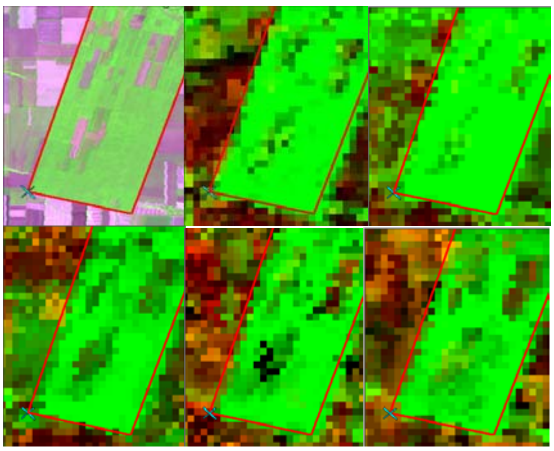


Figure 3: Geolocation accuracy: upper left- Landsat – July 2002, upper cent. – Aug. 2001, up. right – Sept. 2002, low. left – Aug. 2003, lower center Sept. 2003, low. right August 2004

The geolocation was assessed against a geo-referenced ETM image. The red border marks a forest area and the blue cross is displayed on the image as a reference point. The lighter green represents the forest area visible on the MODIS 16 days composite using the RED and NIR band available at 250m resolution. As it can be noted on Figure 3, all 5 MODIS composites show a systematic shift of around 1 pixel against the Landsat image. However the intra-geolocation error between the MODIS images is quite low. When following the red line of the forest area, a maximum of approximately one pixel shift can be identified. Therefore the MODIS 16 days composite is a suitable candidate for our study on crop acreage.

4. USE OF HIGH RESOLUTION DATA

Since data availability for the Rostov Region is quite restricted and no high resolution based land cover dataset or parcel information system was available, the additional use of ETM in combination with SRTM data was necessary. The advantages of using ETM in combination with SRTM data are several. First of all, they give a good overview of the region of interest and due to visual display capabilities of the software used in this study a simultaneous display of the Landsat ETM and the elevation data (SRTM data) was possible. In particular this method is useful to check and control the automatic classification of a crop mask and to undertake corrections based on visual interpretation. (e.g. identifying natural vegetation along river valleys). Furthermore, the ETM images can be segmented and in that way the shape of the agricultural fields can be recorded.

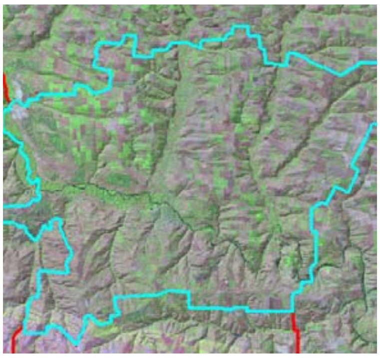


Figure 4: draped SRTM over ETM

The Landsat ETM data was downloaded from the Global Land Cover Facility web site which provides free access to ETM data. The ETM images were not all from the same year or date, but all the ETM images could be found for the summer/autumn period. In order to have a general overview mosaic for the whole area was created. The Shuttle Radar Topography Mission (SRTM) is a joint project between the National Geospatial-Intelligence Agency (NGA) and the National Aeronautics and Space Administration (NASA). The objective of this project is to produce digital topographic data the Earth's land surface. SRTM data was freely available for the study area on a 90 meter resolution (on the x-y axes).

4.1 The creation of a crop mask

In order to create an automatic classification of a crop mask an unsupervised classification of the NDVI values for the period of 3 years (2000-2003) was undertaken. Each 16 day MODIS image was used as 1 band. The general idea of this method is that a natural vegetation will show a stable profile over a certain time period, whereas the agricultural land will vary over the years due to crop rotation. The ETM scene was used to label the 50 different clusters, in that way the major crop areas of Rostov could be extracted.

4.2 Visual interpretation

Nevertheless the crop mask derived from the automatic classification proved to be insufficient and areas in particular along steeper slopes were incorrectly classified as agricultural land. Particularly in the northern zone of the Rostov Oblast this problem occurred as the area is much more hilly. In order to improve the crop mask, the TM based segments (derived from the e-cognition software) were used as baseline information to define the crop mask (see Figure 5). In that way an entirely new visual interpretation could be avoided. The crop mask was improved in the northern part of the Oblast (see black box, Figure 6). Figure 6 shows the difference between the overall crop acreage derived from the automatic and visual interpretation versus the acreage of the official statistics.

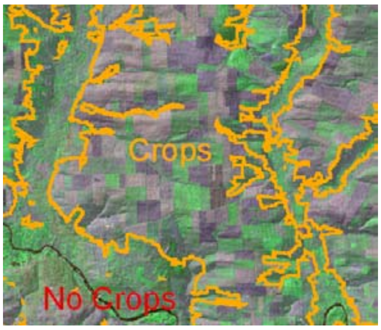


Figure 5: Creation of crop mask

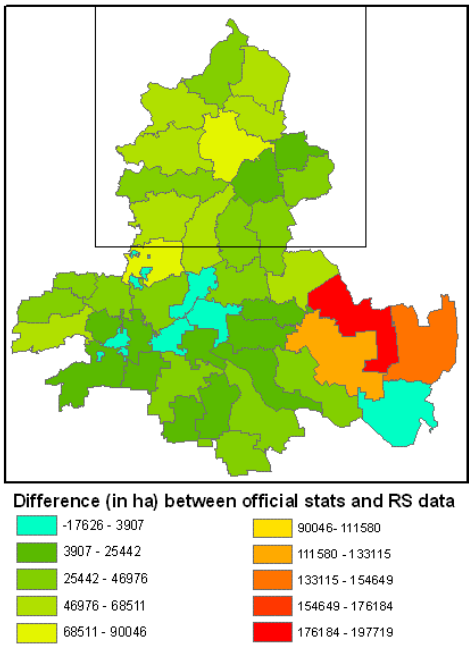


Figure 6: The sub regions of the Rostov Oblast

5. IDENTIFICATION OF WINTER CROPS

In order to have an overview of the profiles of the different crops within 2 farms, zonal statistics for each MODIS composite were derived. An automatic processing chain was implemented and the median, mean standard deviation etc. were extracted for each 16 day composite using a Python script implemented in ARC/GIS.

By using the profiles of 10 crops (1 winter crop and 9 non winter crops) which covered the highest proportion of the agricultural land of the 2 farms, two key dates for the extraction of winter crops were identified. The best dates were before/or shortly after the winter snow and one during early spring time when the other crops (not winter) are prepared for cultivation. Table 2 shows the dates of the points which were considered most suitable as the difference between the NDVI mean of the 9 crops and winter crops at that date was high and the standard deviation of the winter crop relatively low.

TABLE 2: Parameters for the extraction of winter crops

DATE [1]	WIN.[2]	STAND[3]	OTH.[4]	THRE.[5]
01/01/2001	4790	1597	2474	3632
06/03/2001	5581	1574	2724	4152
02/02/2002	5064	1448	2820	3942
22/03/2002	6198	1417	3316	4757
17/11/2002	5102	1728	2720	3911
22/03/2003	3190	795	2291	2740
21/03/2004	6183	1545	3848	5015

1) Identified key date, 2) Mean NDVI (*10000), of winter crop (310 pixels), 3) Standard deviation of winter crops for 310 pixels, 4) Mean NDVI value of other 9 non-winter crops, 5) Threshold used

The threshold was calculated by using the mean between the winter crop and the other crops. The correlation (R squared) of the estimated acreages for winter crops and the official statistics were 0.81 for the year 2001, 0.74 for the year 2002 (see Figure 7) and 0.84 for the year 2003. The Region in red on Figure 6 was discarded due to an insufficient crop mask.

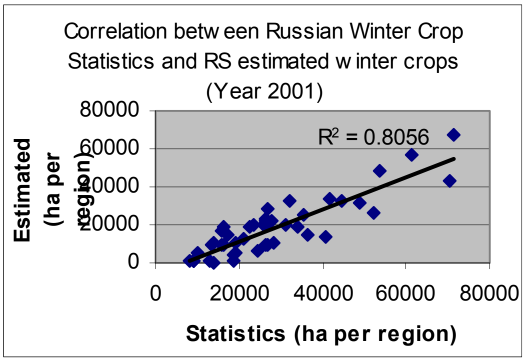


Figure 7: Correlation between estimates and statistics

6. DISCUSSION AND CONCLUSION

It has been illustrated that the MOD13 NDVI product is a suitable product for acreage estimations in the Oblast Rostov. NDVI values for a period of 4 years (2001 – 2004) were examined and analysed. In the future other MODIS products and other vegetation Indices such as the EVI and PVI could be tested and applied in this region.

The threshold values applied for extracting winter crops are based on data from 2 farms. The thresholds are chosen early in the year, which means that the method applied would allow estimations of crop acreages prior to the availability of statistical data. Nevertheless, the data from the 2 farms is quite limited. Therefore, a more general approach could be applied in a way that the thresholds are determined by more general parameters such as theoretical crop profiles based on preparation of the fields, planting time and flowering.

Preliminary results show that acreage estimation based on the MODIS composite show a high correlation with official statistical data for 50 regions within the Rostov Oblast. In order to derive good estimates for different crops a crop mask with high quality is required and even though the unsupervised classification with subsequent labelling is useful to discriminate a certain proportion of natural vegetation from crops, the visual interpretation is more accurate. The visual correction of the crop mask will be carried out for the whole of the Oblast and not only in the northern part.

The method applied is based on simple threshold values and pixel counting. In general it could be assumed that the effect of mixed pixels are cancelled out. This is however not always the case (Galliego et al, 2004) and depends highly on the spatial scale. The general rule is that crop area estimates for large regions may be improved relative to finer scales since random errors which occurred at finer scales will tend to cancel out (Lobell and Asner, 2004). Due to these uncertainties more sophisticated methods such as neural nets and linear unmixing of these crops could be applied. However these methods rely on good ground data which reinforces the need for better ground truth datasets.

These issues will be addressed in the future in order to gain more reliable estimates. Furthermore, the method will be applied to other areas in Russia namely the Oblast Moscow and Orenburg.

6. ACKNOWLEDGEMENTS

The authors would like to thank Igor Neishtadt for providing the farm data used in this study and the statistical Russian crop data. Thanks to Felix Rembold for his advice and Dementris Stathakis for his help in programming the python script.

7. REFERNCES

Doraiswamy, P.C., Hatfield, J.L., Jackson, T.J. Akhmedov, B., Pruegerb, J. and Stern, A., 2004, Crop condition and yield simulations using Landsat and MODIS, *Remote Sensing of Environment* 92, pp. 548–559

Gallego, F.J., 2004, Remote sensing and land cover estimation, *International Journal of Remote Sensing*, 25(15), pp. 3019 – 3047

Lobell, D. B., Asner, G.P., 2004, Cropland distributions from temporal unmixing of MODIS data, *Remote Sensing of Environment* 93, pp.412– 422

Ippoliti-Ramilo, G.A.; Epiphanio, J.C.N. and Shimabukuro, Y.E.[1] ,2003, Landsat-5 Thematic Mapper data for pre-planting crop area evaluation in tropical countries *International Journal of Remote Sensing*, 24 (7), pp. 1521-1534(14)

Stroeve, J., Box, J. E., Gao, F, Liang, S., Nolin, A., and Schaaf, C., 2005, Accuracy assessment of the MODIS 16-day albedo product for snow: comparisons with Greenland in situ measurements, *Remote Sensing of Environment* 94, 46–60

Gao, X., Huete, A.R., NiW., and Miura, T., 2000, Optical-biophysical relationships of vegetation spectra without background contamination, *Remote Sensing of Environment.* 74, pp.609-620.