CROP YIELD PREDICTION WITH SPOT VGT IN MEDITERRANEAN AND CENTRAL ASIAN COUNTRIES

I. Savin

AGRIFISH unit, Institute for the Protection and Security of the Citizen, Joint Research Centre, European Commission – igor.savin@jrc.it

Commission VIII, WG VIII/10

KEY WORDS: crop yield forecasting, SPOT VEGETATION, crop growth monitoring

ABSTRACT:

Some years ago the MARS-FOOD group was established to support the Food Aid and Food Security policies of the European Commission. The activities are aimed at improving methods and information on crop yield prospects. Russia, Central Asia, and non-European Mediterranean countries (MECA region), Eastern Africa (IGAD sub-region) and the MERCOSUR region in South America were selected as pilot areas. Crop growth indicators are produced based on low resolution remote sensing data, global meteorological modelling outputs (ECMWF model) and crop growth simulation models (CGMS and FAO-WSI). Crop yield forecasting is done using predictors selected from the crop growth indicators. Dekadal SPOT-VEGETATION data are used as a basis for calculation of remote sensing indicators of crop growth. The Normalized Difference Vegetation Index (NDVI) and results of Dry Matter Production modelling (DMP) applying the Monteith approach (Monteith, 1972) are used as a main source of remote sensing indicators for the MECA region. The indicators are used in aggregated for sub-national administrative unit form applying crop mask. Some indicators are derived for a network of representative points. The current dekadal indicators are compared with previous year dekadal values or with long-term average dekadal data. Additionally relative time mosaics of indicators are used as a tool for crop growth monitoring (Savin, Nègre, 2002). We analyze additionally seasonal cumulative values of indicators by comparing seasonal time profiles. As a result, near 10 remote sensing indicators can be derived for each crop for each dekad of growing season in aggregated form and the same amount for representative points. Crop yield forecasting starts from an attempt to build simple regression equation between statistical crop yield and crop growth indicators. We found that regression with high R^2 can be built for many administrative units of MECA region. During the second phase of crop yield prediction the similarity analysis is applied. The aim of analysis is to define a year-analogue for indicator time profiles. This operation is conducted mainly for the administrative units where regression analysis does not give acceptable results. The last phase is devoted to comparison of indicator's value with previous year or long-term average value. Final yield prediction is made by expert taking into consideration the results of all phases of indicators analysis. The crop yield can be predicted quantitatively based only on remote sensing indicators for many administrative units of the region. For some units only a sign of crop yield changes can be predicted. In some cases it is impossible to predict crop yield based only on remote sensing indicators. The time when crop yield prediction can be made differs from region to region. For the most part of administrative units of the region the best time for crop yield prediction is allocated near crop flowering. However, for some units the best time is shifted to earlier or to later period of crop growing season. The results of the crop growth monitoring and yield prediction are summarized in the form of agro-meteorological bulletins, issued bimonthly for Russia and Central Asia, and for the Mediterranean countries.

1. INTRODUCTION

In 2001, the MARS project started what is now called the MARS-FOOD Action aimed at giving support to the EU Food Security and Food Aid policy by improving information on crop prospects, particularly in regions of the world stricken by frequent food shortages. The main end users are the European Commission services directly involved in food aid (DG DEV, DG AIDCO and EU delegations). The activities are carried out in close collaboration with the Food and Agriculture Organization (FAO) of United Nations. After the initial development and demonstration phase (2003-2004) the developed methods and systems are now being tested on a pre-operational basis (2005-2006). One of the test areas includes non-European countries of Mediterranean basin, Russia, and Central Asian countries (MECA region).

A common problem in crop monitoring and yield forecasting in many countries of the world is generally represented by the difficulties in extending locally calibrated forecasting methods to other areas or to other scales. Several agro-meteorological and remote sensing based indicators have proven to be highly correlated with yield (Rasmussen 1997; Lewis et al., 1998; Reynolds et al. 2000) for certain crops in specific areas. Recommendations on how to use multiple regression analysis have been released by international organizations like FAO (Gommes 2001), but due to the large geographic variability of all yield indicators, no synthetic procedure is available yet for a general and simple operational yield estimation procedure, based on the data available for each single site or country.

The purpose of the method elaborated and used by MARS-FOOD (Rembold et al., 2006) is not to solve the dilemma of geographic variability in yield estimation, but to develop a simple method, which for any place of the world combines the available data for producing the best possible yield forecast. Such a processing chain is necessarily a multi-step procedure, where each step is developed case by case, depending on data availability and local climatic and agronomic indicators. The results of each step are then retained and quality flagged for computing the most likely final estimate. All intermediate results are inserted into a final matrix in a way that the user can evaluate the reliability of each step. The basic principle here is that converging results obtained by different methods reduce the error of the final estimate, while if there is no agreement they have to be used with more caution. On the other hand, a certain degree of freedom is left to the analyst in evaluating the reliability of the different estimates, leaving room to his additional knowledge of the study area. The proposed steps are not necessarily sequential but do normally follow a certain order based mainly on their complexity and the number of input data they involve. Expert knowledge is needed in the choice of indicators and in the final weighting of the results by the analyst. In this paper we describe our experience in using this approach in MECA region based on remote sensing data.

2. METHODS

Dekadal SPOT-VEGETATION data are used as a basis for calculation of remote sensing indicators of crop growth. We use the ten-day synthesis (S10) product, which is computed from all the passes on each location acquired during 10 day periods. The periods are defined from 1st to 10th, from 11th to 20th, from 21st to the end of each month. The synthesis between different passes is performed selecting the best measurement of the period defined from the following criteria: a) it does not correspond to a blind or interpolated pixel; b) it does not be flagged as cloudy in the status map; c) it does correspond to the highest value of Top of Atmosphere NDVI. For each pixel the ground surface reflectance in the four spectral bands is computed, the atmospheric correction being performed using the annotations of the primary product corresponding data. Geometric viewing conditions, reference to the date and time of the measurement, and information on the composite status map are also given. For each data set, the references of all corrections applied for calibration, atmospheric correction and geometric processing are produced. After that, clouds and snow cover masks are applied to all images.

The Normalized Difference Vegetation Index (NDVI), and results of Dry Matter Production modelling (DMP) applying the Monteith approach (Monteith, 1972) are used as a main source of remote sensing indicators for the MECA region. The indicators are used in aggregated for sub-national administrative unit form applying crop mask. The crop masks were created based on GLC2000 database (Global …, 2007), common agronomical practice knowledge, regional and countries' crop distribution paper maps, publications in scientific journals and other auxiliary information on regional crop geography. Some indicators are derived for a network of representative points. The current dekadal indicators are compared with previous year dekadal values or with long-term average dekadal data. Additionally relative time mosaics of indicators are used as a tool for crop growth monitoring (Savin, Nègre, 2002). Crop development stage-specific (relative time) NDVI mosaic is a scene of NDVI pixels, corresponding to the same crop development stage. Over the scene NDVI pixels can refer to different calendar dekads. Dates of crop phenological phases have not spatial variability within small region or for a separate pixel. In such situation the NDVI value of the current year must be compared with those of other year, which had been received for predefined crop phenological phase. In this way the NDVI values are normalized for the characteristic time of crop growth. We analyze additionally seasonal cumulative values of indicators by comparing seasonal time profiles. As a result, near 10 remote sensing indicators potentially can be derived for each crop for each dekad of growing season in aggregated form and the same amount for the representative points (fig.1).

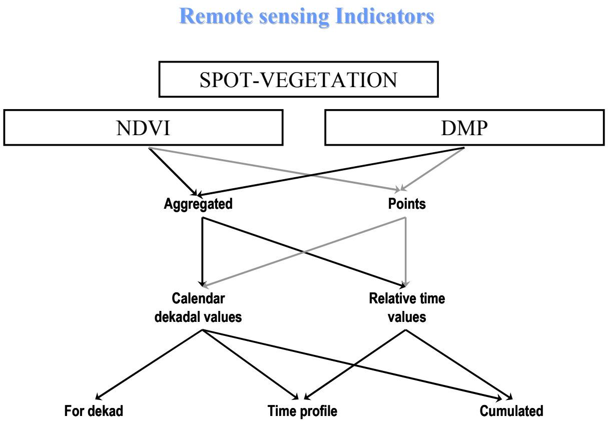

Figure 1. Diversity of potential remote sensing indicators of crop growth

Crop yield forecasting is done using predictors selected from the crop growth indicators. Crop yield forecasting starts from an attempt to build simple regression equation between statistical crop yield and crop growth indicators. We found that regression with high R^2 can be built for many administrative units of MECA region. The regression equations are updated after the end of crop growth season based on new coming crop yield statistical data. During the second phase of crop yield prediction the similarity analysis is applied. The aim of analysis is to define a year-analogue for indicator time profiles. This operation is conducted mainly for the administrative units where regression analysis does not give acceptable results. The similarity analysis was repeated 1 time every 2 months starting from the beginning of the growing season. The reliability of the results is increasing with each step until reaching a maximum on the middle of the season. The square sums of deviation between the indicator of the current and other seasons, as well as extreme deviation values were used for the definition of the year-analogue. The last phase is devoted to comparison of indicator's value with previous year or long-term average value. The results of all previous steps are resumed in a final matrix, which allows the analyst (and the users) to evaluate the quality of each single estimate. For the regression analysis the determination indices and prediction errors are retained for each of the examined indicators. Based on a final comparison of the results, by expert knowledge of the area and the qualitative use of any additional indicators (for example, analysis of local extremes of yield and indicator co-variation) available for the examined crop, the analyst finally decides a synthetic yield forecast. At the present stage, this evaluation procedure is driven mainly by expert evaluation. In the future the process could be partially automated by a higher harmonization of the single steps and a weighting system of their results.

3. RESULTS

The described above approach is used operationally for crop yield forecasting for the countries of MECA region since 2004. The approach is a part of crop yield forecasting system, where remote sensing indicators are used together with other crop growth indicators (meteorological, statistical and so on). The quality of crop yield forecasting can be observed from the table 1.

country	statistical wheat yield 2005 (t/ha) (available at FAOSTAT web site since 20.12.2005)	predicted 10.04.05	predicted 10.06.05
Algeria	1.4	1.3-1.5	1.1-1.3
Egypt	6.5	6.3-6.5	6.3-6.5
Israel	2.5	2.3-2.5	2.4-2.6
Jordan	1.4	1.3-1.5	1.3-1.5
Lebanon	2.6	2.5-2.7	2.6-2.8
Libya	0.8	0.7-0.9	0.8-1.0
Morocco	1.0	1.1-1.3	0.8-1.0
Palestine	2.3	2.2-2.4	2.2-2.4
Syria	2.5	1.9-2.1	2.3-2.5
Tunisia	1.6	1.6-1.8	1.4-1.6

Table 1. Example of wheat yield prediction for some countries from the MECA region

The significance of remote sensing indicators for crop yield prediction differs from country to country. In some regions crop yield can be predicted based only on regression analysis between yield and remote sensing indicators. In this case the R^2 of the regression models differs from 0.7-0.9 for Morocco, Algeria, Israel, Jordan, Saudi Arabia, Azerbaijan, Georgia, Iran, Uzbekistan, some regions of Russia, to 0.5-0.7 for Syria, Tajikistan, Armenia, Kyrgyzstan, and Afghanistan. Of course, the quality of regression models differs not only from country to country, but from crop to crop too (at given moment for the countries of MECA region a yield prediction is made for winter wheat, rice, maize, and potatoes).

Sometimes, analysis of similarity of the indicator behaviour between seasons can lead to receiving preliminary figures of crop yield. For example, figure 2 contains time profiles for a number of years of cumulated for the season DMP values for the areas with winter wheat in Stavropol region of Russia. As it can be deducted from the figure, two groups of profiles can be easy defined starting from the beginning of July: group with relatively high crop yield (seasons 2001/2002 and 2003/2004), and group with relatively low crop yield (other seasons). Thus, based on information in which group the time profile of the current season will be situated, it will be possible to deduct roughly the corresponding crop yield level.

The figures 3 and 4 demonstrate regions where winter wheat yield can be predicted based on only remote sensing indicators.

As can be resumed from these figures, for many administrative units the wheat yield can be predicted based only on remote sensing indicators. The usefulness of NDVI and DMP indicators depends mainly on specific of natural conditions, and varies from country to country. For some countries (for example, for Morocco and Uzbekistan) both indicators can be applied for crop yield prediction, for others – only one of these indicators can be used.

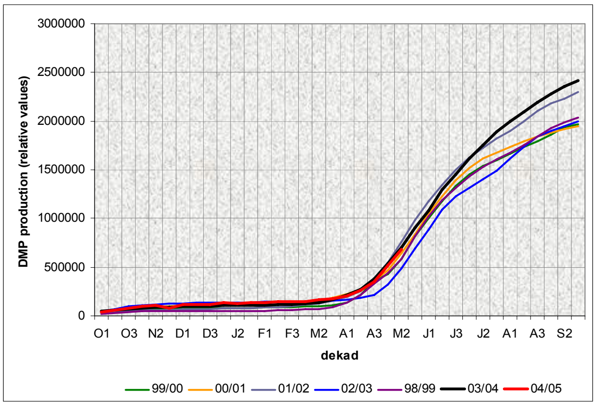

Figure 2. Time profiles of cumulated for growing season DMP values for winter wheat areas in Stavropol region of Russia (winter wheat yield (t/ha): 1999 – 2.4; 2000 – 2.3; 2001 – 2.8; 2002 – 3.3; 2003 – 2.4; 2004 – 3.5)

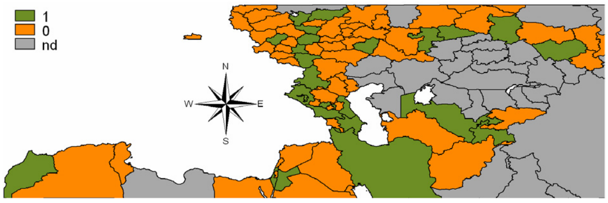

Figure 3. Administrative units within MECA region where winter wheat yield can be predicted using only NDVI based indicators (green units).

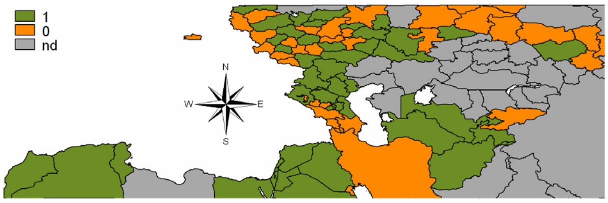

Figure 4. Administrative units within MECA region where winter wheat yield can be predicted using only DMP based indicators (green units).

Other important question is a time schedule for the yield prediction. Each step of the analysis can be implemented only at a given moment of the year, depending on statistical data availability for the previous season and on the results of the preliminary analysis of crop growth indicators. Based on analysis of crop growth predictors defined for the previous seasons, the optimal time schedule for crop yield prediction has been elaborated for each country of the region. For example, best time for each step of the yield prediction for the Maghreb countries is resumed in Table 2.

Historical yield profiles can be analyzed only when the yield statistics for the previous season are available. In the FAOSTAT database data entries are made at different times of the year, but in general, preliminary yield data become available in December (for the previous season). Thus, the first step of the yield prediction procedure (analysis of historical yield profiles) can be implemented not earlier than December.

	Step of the analysis:				
	statistical yield time series	early yield indicator analysis	regression	similarity analysis	final prediction
Morocco	36*	13	10	13	13
Algeria	36*	1	7	13	13
Tunisia	36*	4	7	13	13

* - *dekad of the previous year*

Table 2. Earliest dekad of the year when each step of the analysis can be implemented

The comparison analysis of local extremes can be conducted for Morocco at the 13th dekad of the year, for Tunisia at the 4th dekad of the year, and for Algeria at the 1st dekad of the year. For example, in Tunisia there is a temporal resemblance between wheat yield and DMP extremes (fig.5). The local extremes of DMP in this country become evident starting from the 3d dekad of the year, after which the comparison analysis with historical yield data can be done. These findings are still based on a short time series and have to be used with caution. They are updated every year as soon as new data become available.

Regression analysis can be done at the 7th dekad of the year for Algeria and Tunisia, and at the 10th dekad for Morocco, when all necessary data for the calculation for the current season can be obtained.

Concerning the similarity analysis, it can be run every 2 months, but the most reliable results for these countries can be expected only after the last dekad of April.

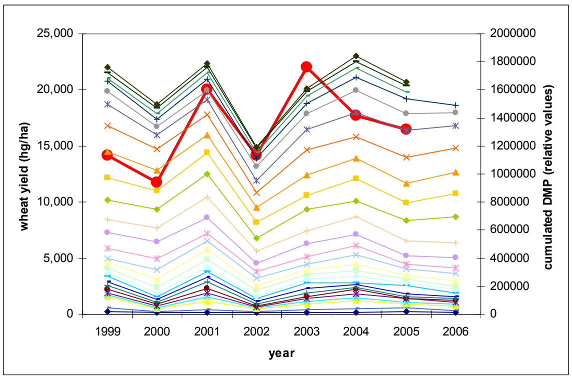

Figure 5. Co-variance between extremes of wheat yield in Tunisia (red line) and extremes of DMP indicator (each line corresponds to the dekad of the growing season)

Thus, preliminary yield prediction (mainly qualitative) can be implemented in different dekads of the year, while the final prediction based on the results of all steps can be done only when all of them are finalized. Thus, while the first results on wheat yield prediction in the Maghreb countries can be released already in December, the final prediction can be obtained only at the end of April.

4. CONCLUSIONS

Simple remote sensing indicators derived from SPOT-VEGETATION instrument can be successfully used as a source of crop yield predictors for the Mediterranean and Central Asian countries. In spite of low resolution of input data, the quality of crop yield forecasting is high enough. The quality varies from country to country, and from crop to crop. Potentially, the quality can be improved by adding into analysis other crop growth indicators (meteorological and statistical).

The elaborated methodology opens opportunity for simple quantitative crop yield forecasting for areas with relatively low data availability. It was developed for immediate operational use and each single phase can separately be improved by following the evidence of new studies in the field. Moreover, it can be successfully applied for other regions, and using remote sensing data of other spatial resolution. In the frame of this methodology the preliminary crop yield figures in many countries can be received during the crop flowering, and in some cases – before crop flowering.

REFERENCES

Global Land Cover 2000 database, JRC EC, http://www-gvm.jrc.it/glc2000/ProductGLC2000.htm (accessed 01 Feb. 2007)

Gommes, R. 2001. An Introduction to the Art of Agrometeorological Crop Yield Forecasting using multiple Regression. Report. Crop Monitoring and Forecasting Group Crop Yield Forecasting and Agrometeorology Sub-Project, UTF/BGD/029, ASIRP/DAE, Dhaka.

Lewis, J.E., Rowland, J., Nadeau, A., 1998. Estimating Maize production in Kenya using NDVI: some statistical considerations. *International Journal of Remote Sensing,* 19, 13, 2609-2617.

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

Pasmussen M.S., 1997. Operational yield forecasting using AVHRR NDVI data: prediction of environmental and inter-annual variability, *International Journal of Remote Sensing*, 18**,** pp. 1059-1077.

Rembold F., Savin I., Nègre T., 2005. Developing a simple operational multistep procedure for quantitative yield/production estimation. *Proceedings of the AfricaGIS2005 Conference*, 31 October to 4 November 2005, The Geo-Information Society of South Africa Tshwane (Pretoria), South Africa ISBN 1-920-01710-0, pp. 257-269.

Reynolds, C., Yitayew, D., Slack, D., Hutchinson, C., Huetes, A., Petersen, M., 2000. Estimating crop yields and production by integrating the FAO Crop Specific Water Balance model with real-time satellite data and ground-based ancillary data.

International Journal of Remote Sensing, Vol 21, No. 18, 3487-3508.

Savin I., Nègre T., 2003. Relative time NDVI mosaics as an indicator of crop growth. *Proceedings of the "Remote Sensing for Agriculture, Ecosystems and Hydrology IV"*. 22 – 25 September 2002, Agia Pelagia, Crete, Greece, pp. 100-108.

ACKNOWLEDGMENTS

The authors wish to thank the Flemish institute for technological research VITO (Belgium) for pre-processing the SPOT-VEGETATION data and for calculation of NDVI and DMP values, which were used in our investigation.