

Yield estimation and forecasting for winter wheat in Hungary using time series of MODIS data

Péter Bognár^a, Anikó Kern^a, Szilárd Pásztor^a, János Lichtenberger^b, Dávid Koroncay^a and Csaba Ferencz^a

^aDepartment of Geophysics and Space Sciences, Eötvös Loránd University, Budapest, Hungary; ^bGeodetic and Geophysical Institute, Research Centre for Astronomy and Earth Sciences, Hungarian Academy of Sciences, Sopron, Hungary

ABSTRACT

Wheat is one of the most important crops in Hungary, which represents approximately 20% of the entire agricultural area of the country, and about 40% of cereals. A robust yield method has been improved for estimating and forecasting wheat yield in Hungary in the period of 2003–2015 using normalized difference vegetation index (NDVI) derived from the data of the Moderate Resolution Imaging Spectroradiometer. Estimation was made at the end of June – it is generally the beginning of harvest of winter wheat in Hungary – while the forecasts were performed 1–7 weeks earlier. General yield unified robust reference index (GYURRI) vegetation index was calculated each year using different curve-fitting methods to the NDVI time series. The correlation between GYURRI and country level yield data gave correlation coefficient (r) of 0.985 for the examined 13 years in the case of estimation. Simulating a quasi-operative yield estimation process, 10 years' (2006–2015) yield data was estimated. The differences between the estimated and actual yield data provided by the Hungarian Central Statistical Office were less than 5%, the average difference was 2.5%. In the case of forecasting, these average differences calculated approximately 2 and 4 weeks before the beginning of harvest season were 4.5% and 6.8%, respectively. We also tested the yield estimation procedure for smaller areas, for the 19 counties (Nomenclature of Territorial Units for Statistics-3 level) of Hungary. We found that, the relationship between GYURRI and the county level yield data had r of 0.894 for the years 2003–2014, and by simulating the quasi-operative forecast for 2015, the resulting 19 county average yield values differed from the actual yield as much as 8.7% in average.

1. Introduction

One prominent application of remote sensing is monitoring the surfaces covered by vegetation, the biosphere itself, including both the actual state and any changes on various timescales. Within the scope of this, both the estimation of the growing area of the main

agricultural species and the acquisition of basic data for yield estimation has considerable economic importance (Doraiswamy et al. 2003). The connection between the spectral properties of certain plant species and their yield has been well known for some time (e.g. Knipling 1970; Tucker et al. 1980), and the usefulness of spectral data from satellite imagery was studied extensively since the launch of the first civil Earth observation satellite (Landsat-1) in 1972 (Rembold et al. 2013). These studies became even more intensive with the appearance of the Advanced Very High Resolution Radiometer (AVHRR) instrument on board the National Oceanic and Atmospheric Administration (NOAA) satellites (e.g. Hochheim and Barber 1998; Doraiswamy et al. 2003; Vicente-Serrano, Cuadrat-Prats, and Romo 2006; Bognár et al. 2011). The methodology of the yield estimation procedures has a wide range (Atzberger 2013). Certain procedures incorporate the satellite data into existing agrometeorological or plant-physiological models (e.g. Wiegand, Richardson, and Kanemasu 1979; Dorigo et al. 2007; Reynolds et al. 2010; Curnel et al. 2011; Huang et al. 2015). Others work with direct mathematical relation between the satellite data and the yield of the given plant, where some of them also include some basic meteorological or agronomical data into the relationship (e.g. Maselli et al. 1992; Hamar et al. 1996; Schut et al. 2009; López-Lozano et al. 2015). The procedures are applied for various species, for example, for corn (e.g. Hayes and Decker 1996; Bolton and Friedl 2013; Shao et al. 2015), for spring wheat (e.g. Hochheim and Barber 1998; Labus et al. 2002; Kouadio et al. 2014), for barley (e.g. Kastens et al. 2005), for rice (e.g. Prasad et al. 2007), for millet (e.g. Rasmussen 1992), for soybean (e.g. Esquerdo, Zullo, and Antunes 2011; Bolton and Friedl 2013), for sorghum (e.g. Kogan, Salazar, and Roytman 2012), for sugar beet (e.g. Clevers 1997; Ferencz et al. 2004), for canola (e.g. Mkhabela et al. 2011), for cotton (e.g. Anderson et al. 2016), for potato (e.g. Rahman et al. 2012), or even for opium (Waine et al. 2014). Among all these, wheat is the most widely-grown food crop type in the world and the most important cereal crop traded on international markets (Becker-Reshef et al. 2010). In the past years, several yield estimation methods for winter wheat were presented. In Kansas, the AVHRR-based vegetation health index was used for the period of 1982–2004. The differences between official and estimated yield data were less than 8% (Salazar, Kogan, and Roytman 2007). The seasonal Moderate Resolution Imaging Spectroradiometer (MODIS)-based maximum normalized difference vegetation index (NDVI) data were used to develop a regression-based model for forecasting winter wheat yield in Kansas and in Ukraine. Six weeks prior to harvest the difference between forecasted and reported wheat yield were 7% (root mean square error (RMSE) = 0.18 t ha⁻¹) in Kansas and 15% (RMSE = 0.44 t ha⁻¹) in Ukraine for the period of 2001–2008 (Becker-Reshef et al. 2010). The same model was applied to the USA, Ukraine, and China, extended with the Growing Degree Day parameter. The accuracy of forecasted total production of winter wheat was about 10% roughly 2 months prior to harvest (Franch et al. 2015). For the 24 regions of Ukraine, three methods were applied for forecasting winter wheat yield: an empirical regression-based model using MODIS-derived NDVI, an empirical model using meteorological data and the adaptation for Ukraine the Crop Growth Monitoring System. The best results for 2010 and 2011 in terms of RMSE value several weeks prior to harvest were 0.3 and 0.5 t ha⁻¹, respectively (Kogan et al. 2013; Kussul et al. 2013). In Punjab, Pakistan, the linear regression of peak-season MODIS-derived vegetation indices were used, and the forecasted yields between 2008 and 2013 were within 11.5% of final reported values (Dempewolf et al. 2014).

Our aim was to develop a yield estimation and forecast method for winter wheat over the country based only on remotely sensed daily data of the MODIS sensor on board the Terra and Aqua satellites. The advantage of the daily data in phenology related studies is clear, since it is able to preserve the natural changes in the phenology in contrast to the temporal composite datasets covering longer periods, such as the 16-day temporal composite MODIS data included in the official MOD13 and MYD13 products (Hmimina et al. 2013). Forecasts are performed before the entire crop has been harvested, whereas estimates are made shortly after the crop has been harvested (Basso, Cammarano, and Carfagna 2013). It is important to underline that our robust method does not require either meteorological, biophysical, agronomical, etc., parameters, or high-resolution satellite data about the location of the cultivated fields of the given agricultural plants. Moreover, similarly to other studies, due to the simplicity of our method, it also does not take into account other information such as the effect of the bidirectional reflectance distribution function (BRDF) (e.g. Soudani et al. 2008; Hmimina et al. 2013; Balzarolo et al. 2016). Naturally, these supplementary parameters can help to improve the accuracy of the estimation of yield and can produce better results (Koloti et al. 2015). Nevertheless, our aim was to develop a simple and easily adaptive method and test its accuracy simulating a quasi-operative yield estimation procedure for a longer period (more than 10 years) without any of the above-mentioned ancillary information.

In this article, we are focusing on the estimation and forecast of yield for wheat in Hungary. First, we attempted to improve our estimation procedure developed earlier (Bognár et al. 2011). Using the data of the period 2003–2015, we tested a number of new curve-fitting techniques and developed a forecast method which enables yield forecast to be given at different dates preceding harvest. Finally, we investigated whether the above-mentioned procedure can also be applied to smaller regions, notably the 19 counties in Hungary, with adequate accuracy.

2. Study area and databases used

2.1. Study area

The study area was the whole area of Hungary (located in Central Europe), consisting of 19 counties (Nomenclature of Territorial Units for Statistics (NUTS)-3 level). Hungary has a moderate climate with a mean annual precipitation of 500–700 mm and an annual mean temperature of about 10°C. The two main agricultural crops for Hungary are corn and winter wheat. In 2014, winter wheat represented approximately 27% of the entire agricultural area, where the sowing area of winter wheat was 1.1 Mha, which is almost 12% of the entire area of Hungary (HCSO, 2016).

2.2. Remote-sensing database

In order to perform crop yield estimations, a 10 year long MODIS direct broadcast (DB) dataset received by the Eötvös Loránd University (ELU) receiving station were used. Preprocessing of the MODIS data was based on freely available software packages, such as the MODIS Level 1DB (v1.8) SeaWiFS Data Analysis System (SeaDAS) software from the Ocean Biology Processing Group (OBPG, SeaDAS (SeaWiFS Data Analysis System, NASA

2016) and the MODIS Level 2 (v3.0) as part of the International MODIS/AIRS Processing Package (IMAPP, Huang et al. 2004; CIMSS (Cooperative Institute for Meteorological Satellite Studies) 2016), supporting DB ground stations in meteorological and environmental research. The generated so-called Level 1b calibrated and geolocated radiance values, and the determined cloud masks are fully identical with the official Collection 6 MOD02, MOD03, and MOD35 MODIS products, respectively. Atmospheric correction of the derived surface reflectances was based on the IMAPP MODIS Land Surface Reflectance algorithm software, which eliminates the effect of atmospheric scattering and absorption caused by atmospheric gases and aerosols. Although this DB version of the official MOD09 algorithm is the same as the one used for global MODIS Collection 5 processing at National Aeronautics and Space Administration Goddard Space Flight Center (NASA GSFC), differences can be present due to the simplified requirements of the ancillary data.

After preprocessing the DB MODIS overpasses, we calculated the country- and county-averaged (NUTS-3 level) reflectances for Hungary both from the raw and the atmospherically corrected reflectances of MODIS band 1 and 2 with $250 \text{ m} \times 250 \text{ m}$ horizontal resolution, where the information of the derived cloud masks was also taken into account. The area averaged reflectances were derived from every overpass based only on the cloud-free pixels, where the cloud mask was shifted by one pixel in every direction to avoid any side effect. Area-averaged raw and atmospherically corrected reflectances were also determined for the main land-cover types in Hungary, such as croplands, grasslands and forests. The distinctions between the main land-cover types were principally based on the database of International Geosphere Biosphere Programme (IGBP) included in the official MODIS Land-Cover product (MCD12Q1). However, we improved this land-cover dataset for Hungary with the CORINE Land-Cover (CLC-2000) database (Büttner, Feranec, and Jaffrain 2002) to improve the representation of grasslands due to the well-known grassland related problems in MCD12Q1 (Jung et al. 2006). In case of missing data (due to the malfunction of the receiving station or DB turn-offs by the Deep Space Network), the Level 1b data were downloaded from NASA/Reverb Echo Site (EOSDIS-ECHO (Earth Observing System Data and Information System Earth Observing System ClearingHouse) 2016) and processed in the same way as described earlier. From the results of the presented pre- and post-processed raw MODIS data, we used two time series covering the period 2005–2015: the cropland specific area-averaged reflectances for the whole of Hungary based on both the raw and the atmospherically corrected reflectances, separately. Information about the geometry of the observation and illumination (such as sensor and solar zenith and azimuth angles) was also derived and their cropland specific mean values for the cloud free pixels were calculated as well. Considering the differences between the Terra/MODIS and Aqua/MODIS sensors and their different specific data processing, the resulting datasets are distinguished and are handled separately. From the derived reflectances the normalized difference vegetation index (NDVI) was calculated and used as the basis for the further studies. The usual definition of NDVI is:

$$\text{NDVI} = \frac{\rho_2 - \rho_1}{\rho_2 + \rho_1}, \quad (1)$$

where ρ_1 and ρ_2 are surface reflectance estimates in the MODIS visible red and near-infrared channels, respectively.

Due to the fact that in the meanwhile a new algorithm (called Collection 6) was introduced into the processing of the raw Level-1A Direct Broadcast MODIS data and that the new official MOD02 products also became available (produced by NASA), we investigated the effects of these on the calculated country averaged and county averaged NDVI values. The results showed that there are no remarkable differences between the NDVI values calculated from the two datasets, therefore we can assume that no error in the accuracy was introduced by using only the Collection 5 data.

To improve the accuracy of the calculations within the method, it would be ideal to have NDVI datasets which fully represent all the agricultural lands in Hungary during the vegetation period. However, this kind of data occurs only when negligible cloud cover is present. Such days are few, and thus the resulting NDVI data series are always scanty for some periods, therefore they cannot be used for more detailed calculations. On the other hand, if the cloud cover has a higher percentage value, the average NDVI of the clear areas usually do not represent accurately the average NDVI of the whole country. We tested the robust method (described later) for various limits in the cloud cover values, and the results confirm our earlier experience that the optimum solution is to use the 50% cloud coverage limit in this procedure. In this case, there is enough number of images to get adequate amount of data for the process, while the clear parts represent the whole areas with an appropriate level. Henceforth, we used the images where cloud coverage was in the range of 0–50%.

2.3. Yield data

The verification of the yield estimation and forecasting method requires accurate official yield data. In Hungary, this data is provided by HCSO. For winter wheat, preliminary yield data is made available several weeks after harvesting, the final data is published generally at the end of the year. The yield data is available for the whole country and for the 19 counties (NUTS-3 level) of Hungary.

In Europe, the Monitoring Agriculture with Remote Sensing (MARS) project was launched in 1988 with the main objective of providing early yield and area estimates all over Europe based on remote-sensing techniques (Rembold et al. 2013). The predicted data are published continuously in MARS bulletins which are available online too. We used these bulletins to compare our results with the estimation of the MARS programme (MARS, 2016).

3. Methodology

3.1. The robust method

The two main crop types in Hungary are winter wheat and corn. Our aim was to construct a yield estimating procedure for winter wheat using MODIS NDVI data based on our formerly developed 'robust yield estimating method'. This method was originally developed for NOAA AVHRR Greenness (GN) data (Ferencz et al. 2004) – where GN is the difference of reflectance values of near-infrared and visible red channels (e.g. Jackson

1983; Hamar et al. 1996) – and has been successfully applied for the 1996–2000 interval both for corn and wheat (Bognár et al. 2011). In this work, we adopted a similar method for the MODIS NDVI data.

The knowledge of the crop area is a key element of the total crop production of a country (Gallego et al. 2014), and for crop yield estimation, the ideal approach would be to use crop-specific masks (Atzberger 2013). However, if the number of the small fields (which are smaller than the MODIS pixel with nominal pixel size) is high, like it is in Hungary, then the classification method requires high-resolution satellite data and even ground-truth data. Moreover, because of the crop rotation, the classification must be made in every year. The essential point of our robust method is that there is no need for a detailed classification to identify the acreage of different crops since the general conditions of the vegetation-covered surfaces also represent the individual crop conditions in the case of the species with similar vegetation cycles (Dabrowska-Zielenska et al. 2002; Ferencz et al. 2004; Bognár et al. 2011), and thus we use the general (non-crop specific) cropland mask described in the previous section.

As described in the previous section, we prepared NDVI temporal profiles over the whole agricultural areas in Hungary for the period of 2003–2015.

The beginning of harvesting the winter wheat occurs in Hungary around the end of June, therefore, we considered the data only until this time of the year. We also fixed the starting point of the season to middle of February, since before this date the bare or snow covered soil dominates the measured signal. Therefore we defined the vegetation cycle as the period between day of the year (DOY) 55 and 181. (Remark: the global changing could modify the limits of this period in the future.) Naturally, the NDVI time series between these two dates contain information of the whole vegetation area, but in this period the dominant crop is the winter wheat in Hungary, therefore, we can use these time series hereafter. The steps of the robust method are shown on Figure 1.

In some of the years, the number of useful images was small, especially for satellite Aqua, therefore, we created the ‘mixed’ NDVI temporal series for every year using the combination of Terra and Aqua data. Despite this, it was obvious that the temporal distribution of the points is not always uniform, and that there still are some years where longer gaps exist in the series, which lead to unrealistic features in the interpolated curves. To overcome this, we defined an interval length threshold of 20 days. Where the gap was longer than this threshold, the datasets were complemented depending on the length of the gap according to the following algorithm: one or two extra points were added where the length of the gap was between 21 and 30 days or between 31 and 40 days, respectively. The values of the extra points were determined by the average NDVI value of all the images with 0–80% cloud coverage in the given interval. In the case of the ‘mixed’ series, there were only two occasions when these points were needed. A gap occurred in 2006 between days 133 and 166 among the 0–50% cloud coverage images, so here we added two points at days 144 and 155, and there was a gap in 2010 between days 122 and 146, so there was one extra point added at day 134 from the 0% to 80% cloud coverage images. Hereafter, we refer to the latter data series as the ‘complemented’ data series or point sets to distinguish them from the ‘original’ data series and point sets.

The next step is to fit a mathematically simple function on the measured NDVI points to eliminate the stochastic fluctuation of the individual NDVI data, which can be caused

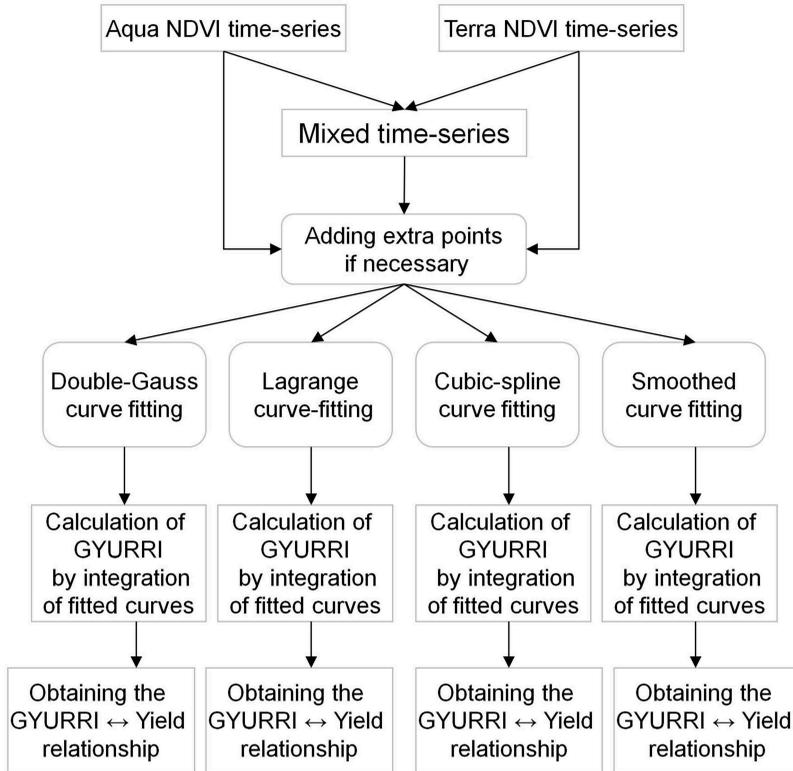


Figure 1. Flow chart of the applied robust method.

by the occasionally inadequate atmospheric correction, the mistakes of cloud detection, the combined effect of the varying viewing and illumination geometry (expressed as bidirectional distribution function (BRDF)), or the different occurrence of the clear pixels (resulting different area to be averaged). Based on our earlier experiences the double-Gaussian function is an adequate approximate of a given NDVI temporal series (Hamar et al. 1996):

$$\text{NDVI}_t = A_1 e^{-((t-t_{01})/\Delta t_1)^2} + A_2 e^{-((t-t_{02})/\Delta t_2)^2} + S, \quad (2)$$

where NDVI_t is the fitted double-Gaussian curve, S is the value of the so-called soil-line, A_1 and A_2 are the amplitudes, t_{01} and t_{02} are the location of the maxima of the Gaussian curves in time and Δt_1 and Δt_2 are the half-widths, respectively. Adding a soil-line is necessary because both before and after the vegetation cycle the NDVI values do not settle to zero but to a certain value which corresponds to soil spectral reflectance. Based on our experience with the existing data, the soil line is fixed to a constant 0.3 NDVI value.

To ensure the starting and ending of this double-Gaussian function on this soil-line – by simple mathematical concept – we added artificial points to the NDVI time series. These points are at the DOY 10, 20, 30 and 250, 270, 290, where we set NDVI to be equal to 0.3 (Ferencz et al. 2004). In accordance with earlier experience regarding mathematical considerations (Ferencz et al. 2004), we fixed the half-width of the first Gaussian

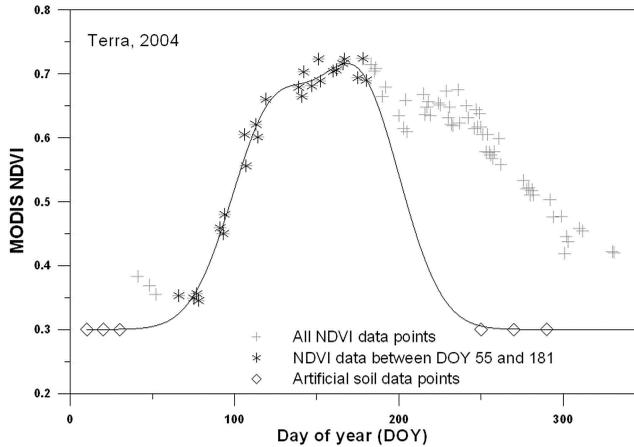


Figure 2. Double-Gaussian curve fit on Terra data in 2004 between days 55 and 181 using artificial soil data points.

function at $\Delta t_1 = 35$ days, thus reducing the number of parameters to be fitted to five (to determine the numbers and values of these artificial soil-points, the value of soil-line and the value of Δt_1 , several combinations of these values were tested, and the minimum standard deviation of curve fitting was found using the above-mentioned parameters). Then, we carried out the double-Gaussian fitting for every single year. An example of the fitted curve for 2004 based on Terra/MODIS data is shown in [Figure 2](#).

Besides the double-Gaussian curve, three other curve-fitting methods were established:

- (1) Lagrange interpolation. Since the DOY–NDVI point sets, on which the approximation curves must sit in a way to show a realistic NDVI variation through the years, often contains random variations within shorter periods, here we first applied a certain averaging in the point set. This averaging is characterized by an interval length in days. The algorithm groups the original data points with growing intervals based on their date differences. At the end in every group the date-span of the points is smaller than the characterizing interval length. Then to every group an averaged data point (DOY, NDVI) is assigned, resulting the formation of the averaged point to be set. On the averaged point set we use a small order Lagrange interpolation, leading to a continuous curve for the approximation.
- (2) Cubic spline. Here, we use the same averaged point set as described in the Lagrange interpolation paragraph. This kind of interpolation gives nicer curves to the eye than the Lagrangian, but the results of the two are very close to each other.
- (3) Smoothing. Here, there is one parameter, an interval length. A smoothed NDVI is a weighted average of all the points in the point set (the original one). The weight function is a Gaussian decreasing with the given interval length.

[Figure 3](#) shows the different interpolation curves for a given point set derived from a Terra image in 2008.

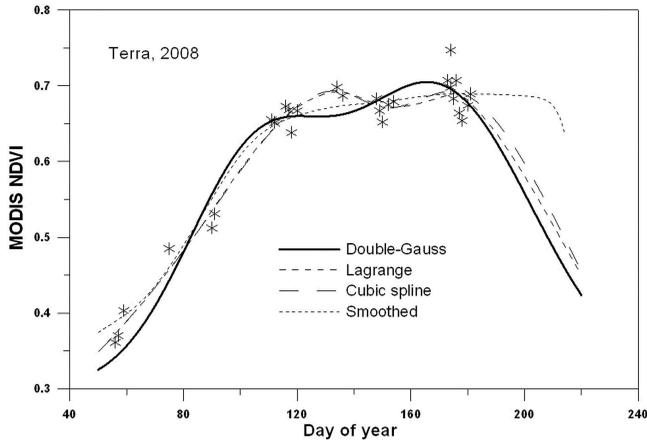


Figure 3. Different curve fits on 2008 Terra data (a double-Gaussian curve, a second-order Lagrange interpolation curve on 20 day averaged point, a cubic spline on a 20 day averaged point set and a smoothed curve with 20 day interval parameter).

In the next step, the General Yield Unified Robust Reference Indices (GYURRI) were calculated for each year individually from the fitted curves (Ferencz et al. 2004):

$$\text{GYURRI} = \int_{d_1}^{d_2} (\text{NDVI})_t dt, \quad (3)$$

where t is measured in DOY units, NDVI_t is the fitted curve (double-Gaussian, Lagrange interpolation, cubic spline, or smoothed curve) and d_1 and d_2 are the first and last days of the integral. At first glance, it seems reasonable to settle the integration limits to the same value as the limits of the whole vegetation cycle, that is, taking 55 and 181 days for d_1 and d_2 , respectively. However, after experimenting with the data we concluded that the results are better if we select a certain standardized subinterval for the integration. The subinterval we fixed here (in the country-wide estimation) is between $d_1 = 113$ and $d_2 = 166$ (see the remarks on global changes discussed earlier).

3.2 Yield forecasting

The above-mentioned method is based on NDVI data measured until the approximate date of harvest (end of June, DOY 181), but it is interesting to examine the possibility of creating forecasts for the crop yield few weeks before harvest time with an acceptable accuracy. Since our previous experience with NOAA data was positive in this respect (Bognár et al. 2011), we improved the forecasting method for the current MODIS NDVI data.

The essence of the method is that we create extrapolated points for the period after the prediction date based on points before this date in the NDVI series, and carry out the process described in the previous section on this extrapolated series, and so we obtain the predicted yield. The extrapolation principle is that we deduce the relationship between the average NDVI in the 20 days preceding prediction date and the NDVI values on certain

post-prediction dates, using the NDVI data in all the available data from previous years. Prediction coefficients are calculated, which are multiplied with the pre-prediction 20 day average to give an estimate for the NDVI values on the individual days. We carry out this calculation for 10 different prediction dates, starting with day 175 and counting down by steps of 5 days until day 130. At prediction day 175, there is only one extrapolated point, it is for day 178. When predicting on day 170, there are two extrapolated points on day 173 and 178, and so on until we get to prediction day 130, where there are 10 extrapolated points, the first starts on day 133 followed by every fifth day until day 178. Let prediction day be day k , denote the year with y , Let the NDVI average between days $k-20$ and k in year y be denoted as $A_{k,y}$. Denote the 5 day averages as $N_{\text{middleday},y}$ (e.g. $N_{133,y}$ is the mean of the NDVI values between DOY 131 and 135 in year y , $N_{136,y}$ is the average NDVI value for DOY 136 and 140 and so on). The prediction coefficients at day k for day d ($C_{k,d}$, $d > k$) are:

$$C_{k,d} = \frac{\sum_{y=y_b}^{y_e} \frac{N_{d,y}}{A_{k,y}}}{y_e - y_b + 1}, \quad (4)$$

where y_b and y_e are the first and last year in the calculation. Thus, the extrapolated NDVI value for day d in year y when predicting on day k is

$$\text{NDVI}_{d,k,y} = A_{k,y} C_{k,d}. \quad (5)$$

Comparing the double-Gaussian functions on the original and the extrapolated datasets, the accuracy of the extrapolation is visible. In a given year, if there is a sufficient amount of data within the 20 day time window prior to the forecast date, and no extreme meteorological event occurs after this date, the extrapolated data will differ only slightly from the real measured data, as it is seen for example in year 2005 (Figure 4). Otherwise, the difference will be higher.

After the calculation of extrapolated data, the robust method can be applied using the original NDVI data until the day of forecasting and the extrapolated NDVI data between the day of forecasting and DOY 181.

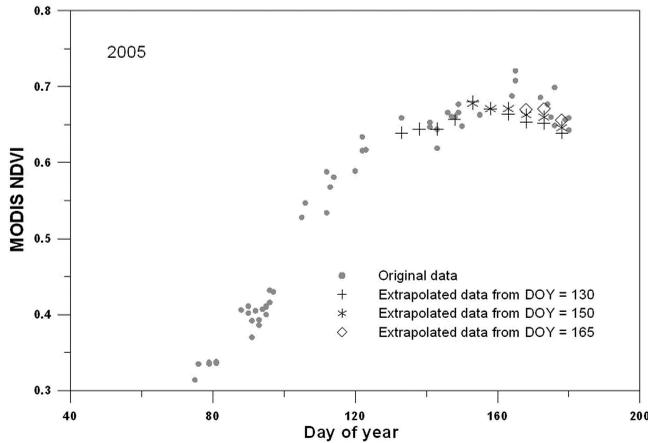


Figure 4. Extrapolated NDVI values in 2005 on three different days (day of the year (DOY) = 130, 150, and 165).

4. Results and discussion

4.1. Results of yield estimation

Initially we used all three (Aqua, Terra and mixed) original NDVI temporal series, fitting curves to them in each year – first we applied only the double-Gaussian fitting (see an example for year 2003 in [Figure 5](#)).

Following the methodology, annual GYURRI values were obtained, that can be directly compared to the country average yield data of the same year provided by HCSO. Fitting a line to the obtained data pairs ([Figure 6](#)), the highest correlation coefficient ($r = 0.975$) was obtained in the case of the mixed dataset, while it was lower for the Terra ($r = 0.946$) and Aqua data-series ($r = 0.797$). Based on this result, in the rest of the investigation we used exclusively the mixed dataset.

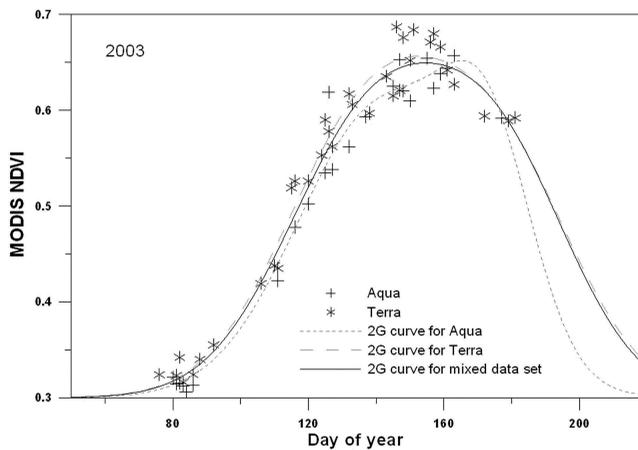


Figure 5. Double-Gaussian curve fit on the Terra, Aqua, and mixed NDVI data sets in 2003.

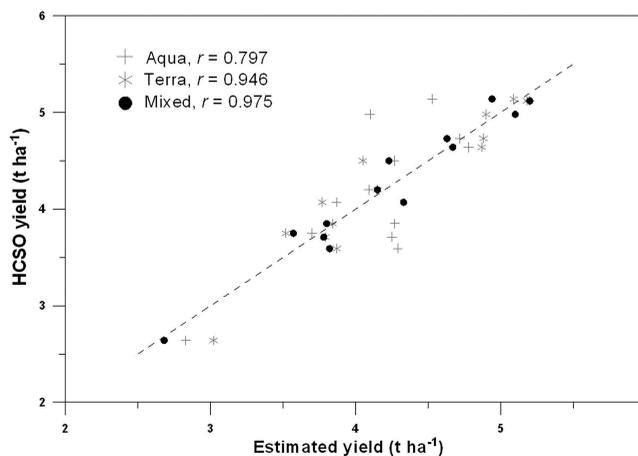


Figure 6. Results of the robust yield forecasting method based on the original Aqua, Terra, and mixed data series in interval 2003–2015, without complementing the series (dashed line represents the $y = x$ relation).

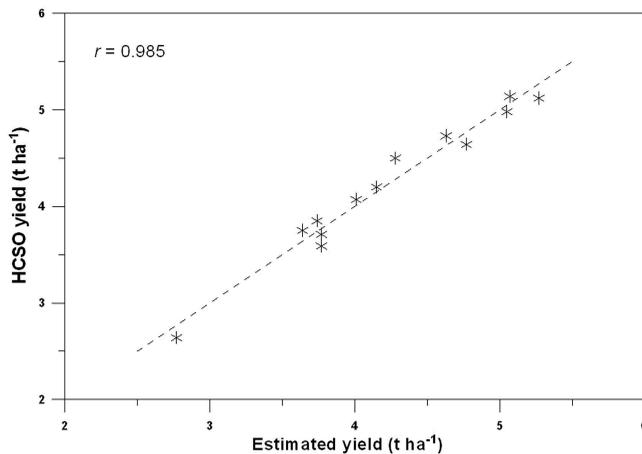


Figure 7. Result of the robust yield forecasting method based on the complemented mixed data series in interval 2003–2015 (dashed line represents the $y = x$ relation).

The derived yield values gave better results (when compared to the official HCSO crop yields) using the complemented time series, where the correlation coefficients between the complemented and official datasets was $r = 0.985$ and the RMSE (Root Mean Squared Error) was 0.12 t ha^{-1} (Figure 7). The deviation of the estimation from the HCSO data can be calculated using the parameters of the linear relationship (Table 1). These results indicate a less than (or equal to) 5% deviation of the calculated crop yields during the investigated 13 years, where the mean absolute difference is 2.7%.

In the next step, we repeated the same procedure with different curve fitting methods and variable parameters. The strongest correlation was observed with the third order Lagrange interpolation of using a 20-day averaging interval ($r = 0.947$). The cubic spline approximation and the smoothed curve gave their best correlation also with 20 day averaging interval ($r = 0.944$) and at 20 day characteristic smoothing interval ($r = 0.936$), respectively. All these correlations are slightly weaker than that of the double-Gaussian

Table 1. Differences between estimated (EST) and Hungarian Central Statistical Office (HCSO) wheat yield (EST = 0.5912 general yield unified robust reference index (GYURRI) – 16.379).

| Year | Calculated GYURRI | EST yield (t ha^{-1}) | HCSO yield (t ha^{-1}) | Difference EST – HCSO (t ha^{-1}) | Difference EST – HCSO (%) |
|------|-------------------|----------------------------------|-----------------------------------|--|---------------------------|
| 2003 | 32.39 | 2.77 | 2.64 | 0.13 | 4.9 |
| 2004 | 36.61 | 5.27 | 5.12 | 0.15 | 2.9 |
| 2005 | 34.94 | 4.28 | 4.50 | -0.22 | -4.9 |
| 2006 | 34.48 | 4.01 | 4.07 | -0.06 | -1.5 |
| 2007 | 34.08 | 3.77 | 3.59 | 0.18 | 5.0 |
| 2008 | 36.25 | 5.05 | 4.98 | 0.07 | 1.4 |
| 2009 | 34.04 | 3.74 | 3.85 | -0.11 | -2.8 |
| 2010 | 34.09 | 3.77 | 3.71 | 0.06 | 1.7 |
| 2011 | 34.72 | 4.15 | 4.20 | -0.05 | -1.3 |
| 2012 | 33.87 | 3.64 | 3.75 | -0.11 | -2.8 |
| 2013 | 35.77 | 4.77 | 4.64 | 0.13 | 2.8 |
| 2014 | 35.54 | 4.63 | 4.71 | -0.10 | -2.1 |
| 2015 | 36.28 | 5.07 | 5.14 | -0.07 | -1.4 |

Table 2. Differences between EST and HCSO wheat yield in case of quasi-operative wheat yield estimation.

| Year of estimation | Years using for linear fitting | Fitted parameters | | GYURRI | EST yield (t ha ⁻¹) | HCSO yield (t ha ⁻¹) | Difference EST–HCSO (t ha ⁻¹) | Difference EST–HCSO (%) |
|--------------------|--------------------------------|-------------------|----------|--------|---------------------------------|----------------------------------|---|-------------------------|
| | | <i>a</i> | <i>b</i> | | | | | |
| 2006 | 2003–2005 | 0.5994 | -16.6833 | 34.48 | 3.98 | 4.07 | -0.09 | -2.2 |
| 2007 | 2003–2006 | 0.5983 | -16.6228 | 34.08 | 3.77 | 3.59 | 0.18 | 5.0 |
| 2008 | 2003–2007 | 0.6063 | -16.9350 | 36.25 | 5.04 | 4.98 | 0.06 | 1.2 |
| 2009 | 2003–2008 | 0.5985 | -16.6737 | 34.04 | 3.70 | 3.85 | -0.15 | -3.9 |
| 2010 | 2003–2009 | 0.5904 | -16.3713 | 34.09 | 3.76 | 3.71 | 0.05 | 1.3 |
| 2011 | 2003–2010 | 0.5922 | -16.4391 | 34.72 | 4.12 | 4.20 | -0.08 | -1.9 |
| 2012 | 2003–2011 | 0.5928 | -16.4509 | 33.87 | 3.63 | 3.75 | -0.12 | -3.2 |
| 2013 | 2003–2012 | 0.5865 | -16.2208 | 35.77 | 4.76 | 4.64 | 0.12 | 2.6 |
| 2014 | 2003–2013 | 0.5772 | -15.9084 | 35.54 | 4.61 | 4.71 | -0.10 | -2.1 |
| 2015 | 2003–2014 | 0.5841 | -16.1376 | 36.28 | 5.05 | 5.14 | -0.09 | -1.8 |

EST = *a* GYURRI + *b*.

curves, which confirms that the application of the double-Gaussian curve fitting technique in the most robust method.

These results were obtained by fitting a line for all of the data of all 13 years (2003–2015), giving only some general information on the deviation of the method. To simulate a real operative yield estimating process, we used the linear relationship between the GYURRI indices and yield data of *n* years and made a prediction for the yield in the (*n* + 1) th year from the actual GYURRI index. First we took four GYURRI–yield points from years 2002–2005. With the line parameters fitted, we estimated the yield of the following year (2006) using the actual GYURRI index. Then, for the next year we used five (2002–2006) GYURRI–yield regression points and with the obtained line we estimated the 2006 yield, and so on. The result of this procedure is presented in Table 2. It is seen that the deviations of the estimated data from official HCSO data never exceed the 5%.

The results are encouraging, especially if we compare them with the estimations of the MARS programme for the years between 2006 and 2015 (Figure 8). As mentioned

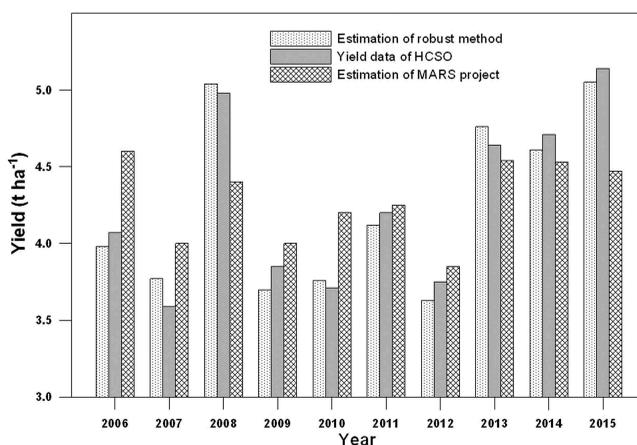


Figure 8. Comparison of the end of June estimations of the robust yield estimation method with the estimations of the Monitoring Agriculture with Remote Sensing (MARS) programme.

Table 3. Differences of the estimated yields of the quasi-operative robust yield forecasting method, the Monitoring Agriculture with Remote Sensing (MARS) programme and the published yields of the HCSO.

| Year of estimation | HCSO yield (t ha ⁻¹) | EST yield at 30 June (t ha ⁻¹) | MARS project | | | | |
|--------------------|----------------------------------|--|-----------------------------|--------------------|----------------------------------|---|--|
| | | | Yield (t ha ⁻¹) | Date of estimation | Source: MARS Bulletin (no./vol.) | Difference EST – HCSO (t ha ⁻¹) | Difference MARS – HCSO (t ha ⁻¹) |
| 2006 | 4.07 | 3.98 | 4.60 | 30/06 | 14/4 | -0.09 | 0.53 |
| 2007 | 3.59 | 3.77 | 4.00 | 17/07 | 15/4 | 0.18 | 0.41 |
| 2008 | 4.98 | 5.04 | 4.40 | 8/07 | 16/4 | 0.06 | -0.58 |
| 2009 | 3.85 | 3.70 | 4.00 | 14/07 | 17/4 | -0.15 | 0.15 |
| 2010 | 3.71 | 3.76 | 4.20 | 20/07 | 18/4 | 0.05 | 0.49 |
| 2011 | 4.20 | 4.12 | 4.25 | 26/07 | 19/11 | -0.08 | 0.05 |
| 2012 | 3.75 | 3.63 | 3.85 | 26/07 | 20/7 | -0.12 | 0.10 |
| 2013 | 4.64 | 4.76 | 4.54 | 20/07 | 21/7 | 0.12 | -0.10 |
| 2014 | 4.71 | 4.61 | 4.53 | 15/07 | 22/7 | -0.10 | -0.18 |
| 2015 | 5.14 | 5.05 | 4.47 | 24/07 | 23/7 | -0.09 | -0.67 |

earlier, our yield estimations were performed on 30 June (or in a leap year on 29 June), thus we used the estimations of the MARS programme, which were published after this date. For these 10 years, our estimations deviated from HCSO data by an average value of 0.104 t ha⁻¹, while the ESA MARS programme gave 0.326 t ha⁻¹ difference (Table 3). The results of the robust method corroborate its applicability, where the estimation of wheat yield at the time of the harvest shows adequate accuracy.

4.2. Results of yield forecasting

We performed forecasting on 10 dates between day 130 and 175 in 5 day steps (Table 4). At day 130, the correlation of the GYURRI-yield values for the investigated 13 years (2003–2015) were low ($r = 0.737$), the average relative difference between forecasted and HCSO data was 10.2%, the RMSE was 0.47 t ha⁻¹. At middle of May (DOY = 135), one and a half months before harvest, the average difference of forecasting and HCSO data was less than 10%, and the biggest deviation was 18.1% in 2007.

Table 4. The results of the forecasting method in the 2003–2015 period.

| DOY | r | RMSE (t ha ⁻¹) | Difference (%) |
|-----|-------|----------------------------|----------------|
| 130 | 0.737 | 0.47 | 10.2 |
| 135 | 0.832 | 0.39 | 8.5 |
| 140 | 0.853 | 0.36 | 7.7 |
| 145 | 0.873 | 0.34 | 6.6 |
| 150 | 0.855 | 0.36 | 6.8 |
| 155 | 0.843 | 0.37 | 6.7 |
| 160 | 0.926 | 0.26 | 5.4 |
| 165 | 0.955 | 0.21 | 4.5 |
| 170 | 0.964 | 0.18 | 4.0 |
| 175 | 0.970 | 0.17 | 3.6 |

DOY is the prediction date, r is the GYURRI-yield correlation coefficient for the 13 years. 'Difference' is the average of the relative differences of the predicted and real yield values in percentage.

Obviously the results of forecasting became better as the date of forecasting approached the date of harvest. Three weeks before harvest (on DOY = 160), the correlation was 0.857 and the RMSE was 0.26 t ha⁻¹. Five days later (DOY = 165) the absolute average difference was less than 5%, the differences of forecasted and HCSO data were within 10% in every year, while the biggest difference was 9.6%, in year 2012.

It is important to note that in 2010 the forecasts before DOY = 160 significantly overestimated the real yield (for example on DOY = 150 the difference between forecasted and HCSO data was 26.7%) due to the longer data gap mentioned earlier. This overestimation increased the average deviation significantly (without year 2010 the average deviations of the predictions on day 150 are only 4.9% in average). These results are encouraging compared to the results of other analyses. For example, the RMSE between the official and predicted wheat yields 6 weeks prior to harvest was 0.44 t ha⁻¹ in Ukraine, which is equivalent to a 15% error (Becker-Reshef et al. 2010). In Argentina, an *r* value of 0.87 was found between the predicted yields of the calibrated empirical model and the official yields 30 days before the harvest, where the RMSE value was 0.58 t ha⁻¹ representing 14% of the mean observed yields (Lopresti, di Bella, and Degioanni. 2015).

Using the quasi-operative method described in the previous section, we calculated forecasts with the above-described quasi-operative method for years 2011–2015. For each year, we calculated the prediction coefficients, for example, in 2011 the calculation is extended to 8 years' data, 2003–2010. (As we go back in years, previous years' data becomes fewer and fewer, and thus the prediction coefficients fluctuate more and more, and the process becomes unstable. Therefore, it is not worth using the quasi-operative process if there is not enough available data on the previous years.) Our results then were compared to the June predictions of the MARS programme. MARS publishes its results monthly in MARS bulletins, but the forecast dates change year by year. For the comparison we always selected an earlier date in our predictions than the prediction date MARS was. The comparison results are shown in Table 5. In four of the five years in the comparison, our predictions were closer to the actual yield. In 2015, the MARS programme gave a prediction that was 1 t ha⁻¹ lower than the actual yield, while the error of our prediction was only 0.24 t ha⁻¹. The only case when MARS gave better prediction was 2012, where our prediction was 0.12 t ha⁻¹ later, while that of the MARS programme was 0.1 t ha⁻¹ above the real yield value. It is notable that the error of our prediction never exceeded 5%.

Table 5. Deviations of the predictions of the quasi-operative robust yield forecasting method (EST) and the MARS programme from the published yields of the HCSO.

| Year | Forecast | | | | | | Difference | | | |
|----------------------------|----------|------|-------|------|--------------------------|-----------------------|------------|-----------------------|-------------|-------|
| | EST | | MARS | | Source: | | EST – HCSO | | MARS – HCSO | |
| HCSO (t ha ⁻¹) | Yield | Date | Yield | Date | MARS Bulletin (no./vol.) | (t ha ⁻¹) | (%) | (t ha ⁻¹) | (%) | |
| 2011 | 4.20 | 4.03 | 14/6 | 3.79 | 17/6 | 19/7 | -0.17 | -4.0 | -0.41 | -9.8 |
| 2012 | 3.75 | 3.63 | 18/6 | 3.85 | 20/6 | 20/6 | -0.12 | -3.2 | +0.10 | +2.7 |
| 2013 | 4.64 | 4.86 | 9/6 | 4.33 | 10/6 | 21/6 | +0.22 | +4.7 | -0.31 | -6.7 |
| 2014 | 4.71 | 4.76 | 9/6 | 4.53 | 10/6 | 22/6 | +0.05 | +1.1 | -0.18 | -3.8 |
| 2015 | 5.14 | 4.90 | 19/6 | 4.19 | 20/6 | 23/6 | -0.24 | -4.7 | -0.95 | -18.5 |

4.3. Results of the county level yield estimation

Administratively Hungary is divided into 19 counties (NUTS-3 level). These are the smallest regional entities of which the HCSO officially publishes annual yield data. We investigated how the accuracy of prediction of the robust method can be kept in smaller areas, on the county level. To attempt this, we assembled the county level NDVI temporal series of the cultivated areas. Here, from the start we only used the mixed MODIS series derived combining both Terra and Aqua data.

First we investigated the method for individual counties searching for the best correlation as a function of curve fitting methods (double-Gaussian, Lagrange, cubic spline and 'smoothed') and GYURRI calculation limits (d_1 , d_2). From the 19 counties, in 6 cases the double-Gaussian, in 6 cases the Lagrange (order 2, 25 day averaging interval) and in 7 cases the cubic spline (also with 25 day averaging interval) proved to be better. The limits of integration also varied, the lower limit (d_1) between DOY = 88 and 112, the upper limit (d_2) between DOY = 151 and 193. This, and the fact that in more than half of the cases (10 out of 19) the correlation (r) was below 0.9 (the average was 0.883) lead us to conclude that the stability of the process is reduced on these smaller areas.

Next, we examined what do we get if we handle all the 19 counties together, that is applying the same curve fitting method with standard integration limits to every county in every year. To perform this, a so-called 'relief-stratum' correction was applied for each county, in order to take into account the different relief and other local properties (Ferencz et al. 2004). For the 2003–2014 data, we got the best correlation ($r = 0.896$, $RMSE = 0.39 \text{ t ha}^{-1}$) with the Lagrange interpolation (second order, 25 day averaging interval, $d_1 = 102$, $d_2 = 178$) curves (Figure 9). With cubic spline curves, using the same interval parameter lead to $r = 0.889$, the double-Gaussians lead to $r = 0.88$, and the smoothed curves resulted in $r = 0.867$. Therefore, we took the curves as Lagrangians with the above-mentioned parameters and modelled the quasi-operative procedure, and made predictions for the 19 counties in 2015. The predicted yields and the real yield values deviated by more than 12% in average, the RMSE was 0.72 t ha^{-1} . After these investigations, we concluded that the

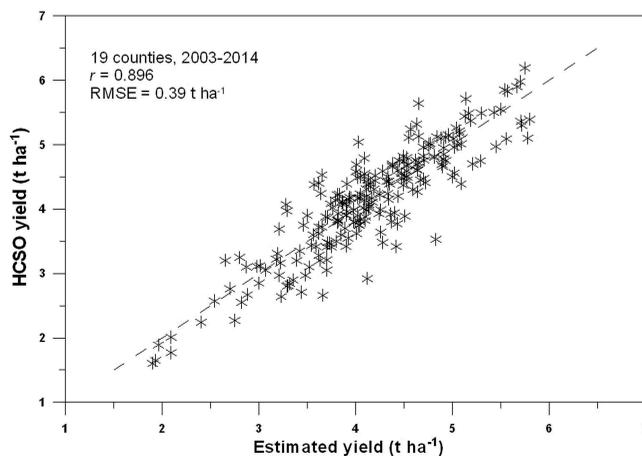


Figure 9. Results of the robust yield forecasting for 19 counties of Hungary in 2003–2014 (dashed line represents the $y = x$ relation).

robust yield forecasting method becomes less reliable for smaller areas, and we cannot predict yield values reliably with high accuracy. This lower accuracy can be explained by the less number of cloud-free measured data used for the yield estimation of a given county.

5. Conclusions

Our robust yield estimating method, which was originally developed for NOAA AVHRR data using GN vegetation index, has been successfully adapted for Terra and Aqua daily MODIS NDVI data. Using this method, it is possible to estimate the yield of winter wheat for Hungary at the end of June – which is generally the beginning of the harvest – at better than 5% accuracy. The wheat yield can also be forecasted with sufficient accuracy several weeks before harvest: at the middle of June the difference between forecasted and reported yield data for wheat is less than 5% in five investigated years, while at the end of May the average difference was 6.8% for years between 2003 and 2015. The method proved to be reliable in years with different weather conditions. However, if a sudden stress event occurs between the day of forecasting and harvest (sudden drought, diseases, hailstorm, fire, etc.) then the forecasted yield value will be inevitably higher than the harvested yield data. The method was based on Terra and Aqua MODIS data. The reception of these data is a low-budget solution, and therefore it is suitable for use in developing countries as well. At the same time, the county-level yield forecast produced less accurate results. Based on the estimated winter wheat crop yields using the country-averaged NDVI values, the calculated Nash–Sutcliffe model efficiency coefficient (Nash and Sutcliffe 1970) is 0.9758, which means that the presented crop yield estimation method and the created model corresponds to a very good model efficiency.

Acknowledgements

The research has been supported in part by the Hungarian Scientific Research Fund (OTKA PD–111920). The authors wish to thank the IMAPP team at the Space Science and Engineering Center, University of Wisconsin-Madison and the SeaDAS group for developing and kindly providing us the applied software. The authors wish to thank NASA for producing and distributing the MOD13 NDVI data. Earth Observing System Data and Information System (EOSDIS) 2009. Earth Observing System Clearing House (ECHO)/Reverb Version 10.91.5 [online application]. Greenbelt, MD: EOSDIS, Goddard Space Flight Center (GSFC) National Aeronautics and Space Administration (NASA). URL: <https://wist.echo.nasa.gov/api/>.

References

- Anderson, M. C., A. Cornelio, P. C. Zolin, C. R. Sentelhas, K. Hain, M. T. Semmens, F. Yilmaz, J. Gao, A. Otkin, and R. Tetrault. 2016. "The Evaporative Stress Index as an Indicator of Agricultural Drought in Brazil: An Assessment Based on Crop Yield Impacts." *Remote Sensing of Environment* 174: 82–99. doi:10.1016/j.rse.2015.11.034.
- Atzberger, C. 2013. "Advances in Remote Sensing of Agriculture: Context Description, Existing Operational Monitoring Systems and Major Information Needs." *Remote Sensing* 5: 948–981. doi:10.3390/rs5020949.
- Balzarolo, M., S. Vicca, A. L. Nguy-Robertson, D. Bonal, J. A. Elbers, Y. H. Fu, T. Grünwald, et al. 2016. "Matching the Phenology of Net Ecosystem Exchange and Vegetation Indices Estimated with MODIS and FLUXNET In-Situ Observations." *Remote Sensing of Environment* 174: 290–300. doi:10.1016/j.rse.2015.12.017.
- Basso, B., D. Cammarano, and E. Carfagna. 2013. "Review of Crop Yield Forecasting Methods and Early Warning Systems." Report Presented to First Meeting of the Scientific Advisory Committee of the Global Strategy to Improve Agricultural and Rural Statistics, Rome, July 18–19.
- Becker-Reshef, I., E. Vermote, M. Lindeman, and C. Justice. 2010. "A Generalized Regression-Based Model for Forecasting Winter Wheat Yields in Kansas and Ukraine Using MODIS Data." *Remote Sensing of Environment* 114: 1312–1323. doi:10.1016/j.rse.2010.01.010.
- Bognár, P., C. Ferencz, S. Pásztor, G. Molnár, G. Timár, D. Hamar, J. Lichtenberger, B. Székely, P. Steinbach, and O. Ferencz. 2011. "Yield Forecasting for Wheat and Corn in Hungary by Satellite Remote Sensing." *International Journal of Remote Sensing* 32: 4759–4767. doi:10.1080/01431161.2010.493566.
- Bolton, D. K., and M. A. Friedl. 2013. "Forecasting Crop Yield Using Remotely Sensed Vegetation Indices and Crop Phenology Metrics." *Agricultural and Forest Meteorology* 173: 74–84. doi:10.1016/j.agrformet.2013.01.007.
- Büttner, G., J. Feranec, and G. Jaffrain. 2002. "Corine Land Cover Update 2000." Technical Report 89, European Environment Agency. http://www.eea.europa.eu/publications/technical_report_2002_89
- CIMSS (Cooperative Institute for Meteorological Satellite Studies). Accessed May 25, 2016. <http://cimss.ssec.wisc.edu/imapp>
- Clevers, J. G. P. W. 1997. "A Simplified Approach for Yield Prediction of Sugar Beet Based on Optical Remote Sensing Data." *Remote Sensing of Environment* 61: 221–228. doi:10.1016/S0034-4257(97)00004-7.
- Curnel, Y., A. J. W. de Wit, G. Duvellier, and P. Defourny. 2011. "Potential Performances of Remotely Sensed LAI Assimilation in WOFOST Model Based on an OSS Experiment." *Agricultural and Forest Meteorology* 151: 1843–1855. doi:10.1016/j.agrformet.2011.08.002.
- Dabrowska-Zielinska, K., F. Kogan, A. Ciolkosz, M. Gruszczynska, and W. Kowalik. 2002. "Modelling of Crop Growth Conditions and Crop Yield in Poland Using AVHRR-Based Indices." *International Journal of Remote Sensing* 23: 1109–1123. doi:10.1080/01431160110070744.
- Dempewolf, J., B. Adusei, I. Becker-Reshef, M. Hansen, P. Potapov, A. Khan, and B. Barker. 2014. "Wheat Yield Forecasting for Punjab Province from Vegetation Index Time Series and Historic Crop Statistics." *Remote Sensing* 6: 9653–9675. doi:10.3390/rs6109653.
- Doraiswamy, P. C., S. Moulin, P. W. Cook, and A. Stern. 2003. "Crop Yield Assessment from Remote Sensing." *Photogrammetric Engineering and Remote Sensing* 69: 665–674. doi:10.14358/PERS.69.6.665.
- Dorigo, W. A., R. Zurita-Milla, A. J. W. de Wit, J. Brazile, R. Singh, and M. E. Schaepman. 2007. "A Review on Reflective Remote Sensing and Data Assimilation Techniques for Enhanced Agroecosystem Modeling." *International Journal of Applied Earth Observation and Geoinformation* 9: 165–193. doi:10.1016/j.jag.2006.05.003.
- EOSDIS-ECHO (Earth Observing System Data and Information System Earth Observing System ClearingHouse) Accessed May 25, 2016. <https://wist.echo.nasa.gov/api/>

- Esquerdo, J., J. Zullo, and J. F. G. Antunes. 2011. "Use of NDVI/AVHRR Time-Series Profiles for Soybean Crop Monitoring in Brazil." *International Journal of Remote Sensing* 32: 3711–3727. doi:10.1080/01431161003764112.
- Ferencz, C., P. Bognár, J. Lichtenberger, D. Hamar, G. Tarcsai, G. Timár, G. Molnár, et al. 2004. "Crop Yield Estimation by Satellite Remote Sensing." *International Journal of Remote Sensing* 25 :4113–4149. doi:10.1080/01431160410001698870.
- Franch, B., E. F. Vermote, I. Becker-Reshef, M. Claverie, J. Huang, J. Zhang, C. Justice, and J. A. Sobrino. 2015. "Improving the Timeliness of Winter Wheat Production Forecast in the United States of America, Ukraine and China Using MODIS Data and NCAR Growing Degree Day Information." *Remote Sensing of Environment* 161: 131–148. doi:10.1016/j.rse.2015.02.014.
- Gallego, F. J., N. Kussul, S. Skakun, O. Kravchenko, A. Shelestov, and O. Kussul. 2014. "Efficiency Assessment of Using Satellite Data for Crop Area Estimation in Ukraine." *International Journal of Applied Earth Observation and Geoinformation* 29: 22–30. doi:10.1016/j.jag.2013.12.013.
- Hamar, D., C. Ferencz, J. Lichtenberger, G. Tarcsai, and I. Ferencz-Árkos. 1996. "Yield Estimation for Corn and Wheat in the Hungarian Great Plain Using Landsat MSS Data." *International Journal of Remote Sensing* 17: 1689–1699. doi:10.1080/01431169608948732.
- Hayes, M. J., and W. L. Decker. 1996. "Using NOAA AVHRR Data to Estimate Maize Production in the United States Corn Belt." *International Journal of Remote Sensing* 17: 3189–3200. doi:10.1080/01431169608949138.
- HCSO (Hungarian Central Statistic Office) Accessed May 25, 2016. <http://www.ksh.hu/docs/hun/xftp/idoszaki/mezo/mezoszerepe14.pdf>
- Hmimina, G., E. Dufrene, J. Y. Pontailier, N. Delpierre, M. Aubinet, B. Caquet, A. Grandcourt, et al. 2013. "Evaluation of the Potential of MODIS Satellite Data to Predict Vegetation Phenology in Different Biomes: An Investigation Using Ground-Based NDVI Measurements." *Remote Sensing of Environment* 132: 145–158. doi:10.1016/j.rse.2013.01.010.
- Hochheim, K. P., and D. G. Barber. 1998. "Spring Wheat Yield Estimation for Western Canada Using NOAA NDVI Data." *Canadian Journal of Remote Sensing* 24: 17–27. doi:10.1080/07038992.1998.10874687.
- Huang, H., L. E. Gumley, K. Strabala, J. Li, E. Weisz, T. Rink, K. C. Baggett, J. E. Davies, W. L. Smith, and J. C. Dodge. 2004. "International MODIS and AIRS Processing Package (IMAPP): A Direct Broadcast Software Package for the NASA Earth Observing System." *Bulletin of the American Meteorological Society* 85: 159–161. doi:10.1175/BAMS-85-2-159.
- Huang, J., L. Tian, S. Liang, H. Ma, I. Becker-Reshef, Y. Huang, W. Su, X. Zhang, D. Zhu, and W. Wu. 2015. "Improving Winter Wheat Yield Estimation by Assimilation of the Leaf Area Index from Landsat TM and MODIS Data into the WOFOST Model." *Agricultural and Forest Meteorology* 204: 106–121. doi:10.1016/j.agrformet.2015.02.001.
- Jackson, R. D. 1983. "Spectral Indices in N-Space." *Remote Sensing of Environment* 13: 409–421. doi:10.1016/0034-4257(83)90010-X.
- Jung, M., K. Henkel, M. Herold, and G. Churkina. 2006. "Exploiting Synergies of Global Land Cover Products for Carbon Cycle Modeling." *Remote Sensing of Environment* 101: 534–553. doi:10.1016/j.rse.2006.01.020.
- Kastens, J. H., T. L. Kastens, D. L. A. Kastens, K. P. Price, E. A. Martinko, and R. Lee. 2005. "Image Masking for Crop Yield Forecasting Using AVHRR NDVI Time Series Imagery." *Remote Sensing of Environment* 99: 341–356. doi:10.1016/j.rse.2005.09.010.
- Knipling, E. B. 1970. "Physical and Physiological Basis for the Reflectance of Visible and Near-Infrared Radiation from Vegetation." *Remote Sensing of Environment* 1: 155–159. doi:10.1016/S0034-4257(70)80021-9.
- Kogan, F., N. Kussul, T. Adamenko, S. Skakun, O. Kravchenko, O. Kryvobok, O. Kryvobok, et al. 2013. "Winter Wheat Yield Forecasting in Ukraine Based on Earth Observation, Meteorological Data and Biophysical Models." *International Journal of Applied Earth Observation and Geoinformation* 23: 192–203. doi:10.1016/j.jag.2013.01.002.
- Kogan, F. N., L. Salazar, and L. Roytman. 2012. "Forecasting Crop Production Using Satellite-Based Vegetation Health Indices in Kansas, USA." *International Journal of Remote Sensing* 33: 2798–2814. doi:10.1080/01431161.2011.621464.

- Koloti, A., N. Kussul, A. Shelestov, S. Skakun, B. Yailymov, R. Basarab, M. Lavreniuk, T. Oliinyk, and V. Ostapenko. 2015. "Comparison of Biophysical and Satellite Predictors for Wheat Yield Forecasting in Ukraine." In: *The 36th International Symposium on Remote Sensing Environment*, ISRS36-185-4. Berlin, Germany. doi: [10.5194/isprsarchives-XL-7-W3-39-2015](https://doi.org/10.5194/isprsarchives-XL-7-W3-39-2015)
- Kouadio, L., N. K. Newlands, A. Davidson, Y. Zhang, and A. Chipanshi. 2014. "Assessing the Performance of MODIS NDVI and EVI for Seasonal Crop Yield Forecasting at the Ecodistrict Scale." *Remote Sensing* 6: 10193–10214. doi:[10.3390/rs61010193](https://doi.org/10.3390/rs61010193).
- Kussul, O., N. Kussul, S. Skakun, O. Kravchenko, A. Shelestov, and A. Kolotii. 2013. "Assessment of Relative Efficiency of Using MODIS Data to Winter Wheat Yield Forecasting in Ukraine." In *2013 IEEE International Geoscience and Remote Sensing Symposium-IGARSS*, 3235–3238. Melbourne, Australia. doi: [10.1109/IGRASS.2013.6723516](https://doi.org/10.1109/IGRASS.2013.6723516)
- Labus, M. P., G. A. Nielsen, R. L. Lawrence, R. Engel, and D. S. Long. 2002. "Wheat Yield Estimates Using Multi-Temporal NDVI Satellite Imagery." *International Journal of Remote Sensing* 23: 4169–4180. doi:[10.1080/01431160110107653](https://doi.org/10.1080/01431160110107653).
- López-Lozano, R., G. Duveiller, L. Seguini, M. Meroni, S. García-Condado, J. Hooker, O. Leo, and B. Baruth. 2015. "Towards Regional Grain Yield Forecasting with 1km-Resolution EO Biophysical Products: Strengths and Limitations at Pan-European Level." *Agricultural and Forest Meteorology* 206: 12–32. doi:[10.1016/j.agrformet.2015.02.021](https://doi.org/10.1016/j.agrformet.2015.02.021).
- Lopresti, M. F., C. M. Di Bella, and A. J. Degioanni. 2015. "Relationship between MODIS–NDVI Data and Wheat Yield: A Case Study in Northern Buenos Aires Province, Argentina." *Information Processing in Agriculture* 2: 73–84. doi:[10.1016/j.inpa.2015.06.001](https://doi.org/10.1016/j.inpa.2015.06.001).
- MARS (Monitoring Agriculture with Remote Sensing). Accessed May 25, 2016. <https://ec.europa.eu/jrc/en/mars/bulletins>
- Maselli, F., C. Conese, L. Petkov, and M. A. Gilabert. 1992. "Use of NOAA–AVHRR NDVI Data for Environmental Monitoring of Crop Forecasting in the Sahel. Preliminary Results." *International Journal of Remote Sensing* 13: 2743–2749. doi:[10.1080/01431169208904076](https://doi.org/10.1080/01431169208904076).
- Mkhabela, M. S., P. Bullock, S. Raj, S. Wang, and Y. Yang. 2011. "Crop Yield Forecasting on the Canadian Prairies Using MODIS NDVI Data." *Agricultural and Forest Meteorology* 151: 385–393. doi:[10.1016/j.agrformet.2010.11.012](https://doi.org/10.1016/j.agrformet.2010.11.012).
- Nash, J. E., and J. V. Sutcliffe. 1970. "River Flow Forecasting through Conceptual Models Part I — A Discussion of Principles." *Journal of Hydrology* 10: 282–290. doi:[10.1016/0022-1694\(70\)90255-6](https://doi.org/10.1016/0022-1694(70)90255-6).
- Prasad, A. K., R. P. Singh, V. Tare, and M. Kafatos. 2007. "Use of Vegetation Index and Meteorological Parameters for the Prediction of Crop Yield in India." *International Journal of Remote Sensing* 28: 5207–5235. doi:[10.1080/01431160601105843](https://doi.org/10.1080/01431160601105843).
- Rahman, A., K. Khan, N. Y. Krakauer, L. Roytman, and F. Kogan. 2012. "Using AVHRR-Based Vegetation Health Indices For Estimation Of Potato Yield In Bangladesh." *Journal of Civil and Environmental Engineering* 2: 111. doi:[10.4172/2165-784X.1000111](https://doi.org/10.4172/2165-784X.1000111).
- Rasmussen, M. S. 1992. "Assessment of Millet Yields and Production in Northern Burkina Faso Using Integrated NDVI from the AVHRR." *International Journal of Remote Sensing* 13: 3431–3442. doi:[10.1080/01431169208904132](https://doi.org/10.1080/01431169208904132).
- Rembold, F., C. Atzberger, I. Savin, and O. Rojas. 2013. "Using Low Resolution Satellite Imagery for Yield Prediction and Yield Anomaly Detection." *Remote Sensing* 5: 1704–1733. doi:[10.3390/rs5041704](https://doi.org/10.3390/rs5041704).
- Reynolds, C. A., M. Yitayew, D. C. Slack, C. F. Hutchinson, A. Huete, and M. S. Petersen. 2010. "Estimating Crop Yields and Production by Integrating the FAO Crop Specific Water Data and Ground-Based Ancillary Data." *International Journal of Remote Sensing* 21: 3487–3508. doi:[10.1080/014311600750037516](https://doi.org/10.1080/014311600750037516).
- Salazar, L., F. Kogan, and L. Roytman. 2007. "Use of Remote Sensing Data for Estimation of Winter Wheat Yield in the United States." *International Journal of Remote Sensing* 28: 3795–3811. doi:[10.1080/01431160601050395](https://doi.org/10.1080/01431160601050395).
- Schut, A. G. T., D. J. Stephens, R. G. H. Stovold, M. Adams, and R. L. Craig. 2009. "Improved Wheat Yield and Production Forecasting with a Moisture Stress Index, AVHRR and MODIS Data." *Crop and Pasture Science* 60: 60–70. doi:[10.1071/CP08182](https://doi.org/10.1071/CP08182).

- SeaDAS (SeaWiFS Data Analysis System, NASA); Accessed May 25, 2016). <http://seadas.gsfc.nasa.gov/modis1db/>
- Shao, Y., J. B. Campbell, G. N. Taff, and B. Zheng. 2015. "An Analysis of Cropland Mask Choice and Ancillary Data for Annual Corn Yield Forecasting Using MODIS Data." *International Journal of Applied Earth Observation and Geoinformation* 38: 78–87. doi:10.1016/j.jag.2014.12.017.
- Soudani, K., G. Le Maire, E. Dufrêne, C. François, N. Delpierre, E. Ulrich, and S. Cecchini. 2008. "Evaluation of the Onset of Green-Up in Temperate Deciduous Broadleaf Forests Derived from Moderate Resolution Imaging Spectroradiometer (MODIS) Data." *Remote Sensing of Environment* 112: 2643–2655. doi:10.1016/j.rse.2007.12.004.
- Tucker, C. J., B. N. Holben, J. H. Elgin Jr., and J. E. McMurtrey. 1980. "Relationship of Spectral Data to Grain Yield Variation." *Photogrammetric Engineering and Remote Sensing* 45: 657–666.
- Vicente-Serrano, S. M., J. M. Cuadrat-Prats, and A. Romo. 2006. "Early Prediction of Crop Production Using Drought Indices at Different Time-Scales and Remote Sensing Data: Application in the Ebro Valley (North-East Spain)." *International Journal of Remote Sensing* 27: 511–518. doi:10.1080/01431160500296032.
- Waine, T. W., D. M. Simms, J. C. Taylor, and G. R. Juniper. 2014. "Towards Improving the Accuracy of Opium Yield Estimates with Remote Sensing." *International Journal of Remote Sensing* 35: 6292–6309. doi:10.1080/01431161.2014.951743.
- Wiegand, C. L., A. J. Richardson, and E. T. Kanemasu. 1979. "Leaf Area Index Estimates for Wheat from Landsat and Their Implications for Evapotranspiration and Crop Modeling." *Agronomy Journal* 71: 336–342. doi:10.2134/agronj1979.00021962007100020027x.