Agricultural and Forest Meteorology 206 (2015) 12–32

Contents lists available at ScienceDirect

Agricultural and Forest Meteorology

journal homepage: www.elsevier.com/locate/agrformet

Towards regional grain yield forecasting with 1 km-resolution EO biophysical products: Strengths and limitations at pan-European level

Raúl López-Lozano[a,*], Gregory Duveiller[a,b], Lorenzo Seguini[a], Michele Meroni[a], Sara García-Condado[a], Josh Hooker[a], Olivier Leo[a], Bettina Baruth[a]

[a] European Commission, Joint Research Centre (JRC), Institute for Environment and Sustainability (IES), Monitoring Agricultural Resources Unit (MARS), Via Enrico Fermi 2749, 20127 Ispra, Italy
[b] European Commission, Joint Research Centre (JRC), Institute for Environment and Sustainability (IES), Climate Risk Management Unit, Via Enrico Fermi 2749, 20127 Ispra, Italy

ARTICLE INFO

Article history:
Received 5 August 2014
Received in revised form 24 February 2015
Accepted 28 February 2015
Available online 9 March 2015

Keywords:
Remote sensing
fAPAR
Wheat
Barley
Maize

ABSTRACT

This study addresses the role of satellite Earth Observation (EO) indicators within an operational crop yield forecasting system for the European Union (EU) and neighbouring countries, by exploring the correlation between official yield statistics and indicators derived from *fAPAR* time-series at sub-national level for the period 1999–2012, and by identifying possible differences across agro-climatic conditions in Europe.

A significant correlation between *fAPAR* and official yields ($R^2 > 0.6$) was found in water-limited yield agro-climatic conditions (e.g. the Black Sea region and the Mediterranean basin) for all three crops studied. In regions where crops experience frequent water stress, most of the yield inter-annual variability is explained by substantial changes in leaf area from one year to another, and can be well captured by regional *fAPAR* time-series. By contrast, in regions characterized by high yields (e.g. northern Europe) – where water constraints are generally not frequent and, therefore, *fAPAR* inter-annual variability is low – the correlation between *fAPAR* and yield is weaker ($R^2 < 0.5$) as yield variations tend to be explained by multiple factors other than green leaf area.

These results confirm the reliability of EO time-series for operational crop yield forecasting at regional level, but also suggest that additional meteorological variables (temperature, precipitation, evapotranspiration) need to be taken into account to interpret EO products meaningfully. Moreover, specific issues related to the spatial resolution of the EO-products, and the absence of dynamic crop masks, currently impede access to crop-specific time-series in the fragmented agricultural landscapes of Europe, and restrict the use of 1-km biophysical products to major crops.

© 2015 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

1. Introduction

Recent trends in global agriculture prices have brought a new scenario for agricultural policies worldwide. Increased world demand for agricultural products and inter-annual fluctuations of global production mostly driven by climate variability have led to price volatility in the agricultural markets and social unrest in many parts of the world. In this context, operational crop yield forecasting activities play a major role in providing an objective and independent early quantitative assessment of crop yields.

Currently, several forecast systems are running at continental/global scales with the purpose of producing regular and timely assessments of the main agricultural products: UN–FAO started the Global Information and Early Warning System (GIEWS) in 1975; within the US Department of Agriculture, both the National Agricultural Statistical Service (NASS), and the Foreign Agricultural Service (FAS), have conducted in the last four decades crop monitoring and yield forecasting at country and global levels, respectively; in China the CropWatch system (Chinese Academy of Sciences) was launched in 2010; and the MARS (Monitoring Agricultural Resources) Unit of the European Commission's Joint Research Centre has been monitoring crops in Europe through the MARS Crop Yield Forecasting System (MCYFS) since 1992. All of these systems use a wide variety of data sources and methods: meteorological indicators, remotely sensed biophysical products and operational crop growth modelling (Lazar and Genovese, 2004; Van Leeuwen et al., 2011; Mueller and Seffrin, 2006; Wu et al., 2014).

Crop yield forecasting at regional, national and continental levels requires reliable and near real time indicators of vegetation status able to cover extensive areas with high temporal frequency.

* Corresponding author. Tel.: +39 332 78 65 35; fax: +39 332 78 90 29.
E-mail address: raul.lopez@jrc.ec.europa.eu (R. López-Lozano).

http://dx.doi.org/10.1016/j.agrformet.2015.02.021
0168-1923/© 2015 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

Satellite Earth Observation (EO) systems have the potential to support operational activities in crop monitoring and yield forecasting as they provide spectral vegetation indices related to plant status (e.g. the Normalized Difference Vegetation Index, NDVI, proposed by Rouse et al., 1974) or direct estimations of relevant biophysical parameters such as the leaf area index (LAI) or the fraction of absorbed photosynthetically active radiation (*fAPAR*) with global coverage on a daily basis (Baret et al., 2007; Myneni et al., 2002).

Spectral vegetation indices and biophysical products derived from EO data provide an insight on light interception and carbon assimilation of crop canopies, key variables that influence final crop yields (e.g. Doraiswamy et al., 2003; Lobell et al., 2003; Wall et al., 2008). This link is often exploited to derive site-specific (semi-) empirical regression models between spatially aggregated vegetation indices or biophysical parameters and observed yields from field measurements or official statistics (e.g. Balaghi et al., 2008; Becker-Reshef et al., 2010; Prasad et al., 2006; Kogan et al., 2013; Meroni et al., 2013; Johnson, 2014). Recently, Kowalik et al. (2014) compared the accuracy of EO-based indicators against crop model indicators implemented in the MCYFS at country level for crop yield forecasting in Europe, and showed that EO-based indicators produced more accurate yield estimates in specific countries.

When considering large geographic extents – from national to continental or global scales – agriculture may differ substantially across regions because of differences in climatic conditions and agro-management strategies. These differences may affect the relationships between biophysical indicators retrieved from EO systems, and the observed yields in the field, which in reality will be influenced by diverse limiting factors (Mueller et al., 2012).

The purpose of this study is to investigate the role of EO indicators within an operational crop yield forecasting system in the EU and neighbouring countries, and to evaluate how these differences in agriculture at a continental level may influence the relationship between biophysical indicators and crop yield. Specific objectives are: (1) to analyse the relationship between *fAPAR* time-series and yield inter-annual variability for three major European crops (wheat, barley and grain maize); and (2) to identify the agro-climatic conditions under which EO indicators better explain the observed yield inter-annual variability.

The organization of the rest of this paper is as follows. A detailed description of the methods and materials used is given in Section 2. In Section 3 the spatial distribution of the correlation coefficient between *fAPAR* and observed yields across Europe is presented for each crop, outlining the regions where *fAPAR* is highly correlated with yields, and identifying the agro-climatic variables associated with the observed differences. Some examples on specific regions are analysed to illustrate the different model performances in two different European agro-ecological settings: (1) high yielding regions in the north, characterized by water availability and inter-annual stability of *fAPAR* time-series; and (2) regions in the south where crop growth is essentially water-limited and thus subject to large fluctuations over the years. An extended analysis, identifying extreme years with unusually low or high yields across Europe in the study period, and relating model performances to specific observed weather events, is provided in Appendix A. Finally, the strengths and limitations of EO biophysical indicators for crop yield forecasting are discussed in Section 4, summarising the main challenges and limitations in the context of an operational yield forecasting system.

2. Materials and methods

2.1. Study area

This study covers European Union (EU) member states with average grain production of at least one million tons of either wheat, barley or maize in the last five years, according to official national statistics published by the European Statistical Office (Eurostat, 2013).

Additionally, given their importance to European agricultural markets, the main producers in the EU neighbourhood are also included: Morocco, Algeria and Tunisia in North Africa; Turkey and Ukraine in the Black Sea area; and Belarus and the Russian provinces (Oblasts) located west of the Ural Mountains.

The spatial scale of the analysis, and therefore, the spatial detail of the basic administrative units considered in this study (ranging from country to district/region and to province level) were set to the highest detail possible (see Section 2.2). The set of administrative units considered in this study for each crop is shown in Fig. 1.

2.2. Collection and processing of sub-national statistics

Sub-national yield, area and production statistics for wheat, barley and maize over the period 1999–2012 were provided by the National Statistical Services of the countries included in this study. Data at NUTS (Nomenclature of Territorial Units for Statistics) level 3 (province) were preferred when available for most of the EU countries and neighbouring countries. In Germany, NUTS level 2 (Governmental districts, or Regierungsbezirke) was preferred to NUTS 3 (Administrative districts, or Kreise) statistics, as the latter administrative units are considerably smaller than NUTS 3 in most other EU member states.

Also NUTS 2 level (region) data were also collected for Poland, Austria, the Netherlands, Latvia and Lithuania. Macro-regions (NUTS 1) were selected in the case of the UK. In Morocco, GAUL level 1 (Global Administrative Units Layer, from FAO) equivalent to

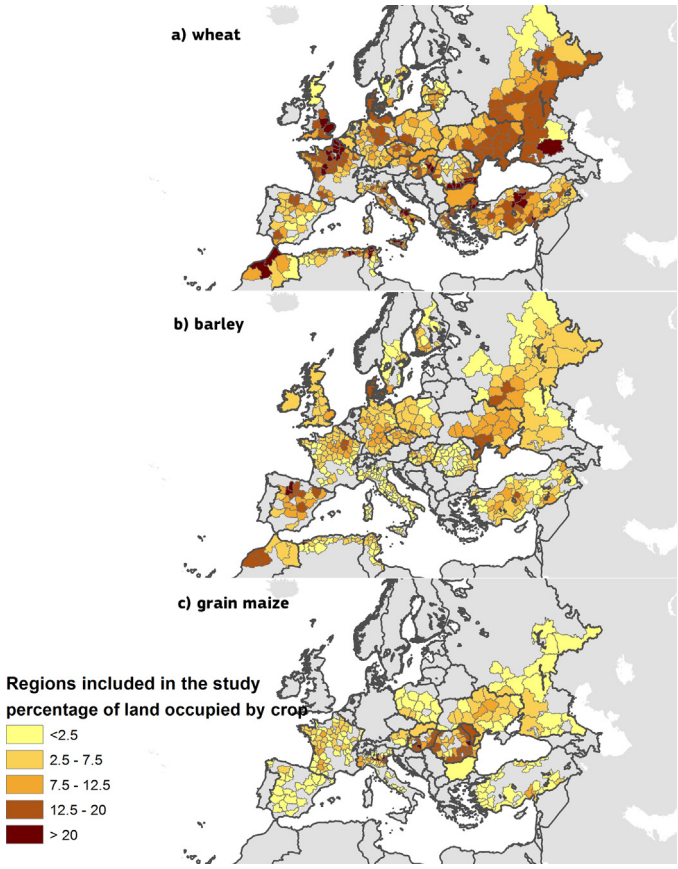

Fig. 1. Administrative units considered in the study for the different crops and percentage of land occupied by the different crops based on harvested areas statistics for 2008 (Table 1) in the different countries.

Table 1
Sources, level of disaggregation and available years of the statistics collected in the different EU member states and neighbouring countries.

Country	Regional level	Years available	Source
EU member states			
Czech Republic	NUTS3	1999–2011	Czech Statistical Office
Finland	NUTS3	1999–2011	Ministry of Agriculture and Forestry, Finland
France	NUTS3	1999–2011	Ministère de l'Agriculture, de l'Agroalimentaire et de la Forêt, France
Greece	NUTS3	1999–2008	National Statistical Service of Greece
Hungary	NUTS3	1999–2011	Hungarian Central Statistical Office
Italy	NUTS3	2000–2011	Italian National Statistical Institute
Romania	NUTS3	1999–2011	National Institute of Statistics, Romania
Spain	NUTS3	1999–2011	Ministry of Agriculture, Food and Environment
Sweden	NUTS3	1999–2011	Official Statistics of Sweden
Austria	NUTS2	1999–2011	Statistics Austria
Germany	NUTS2	1999–2010	Statistical offices of the Länder Federal Statistical Office
Latvia	NUTS2	1999–2011	Central Statistics Authority database
Lithuania	NUTS2	2000–2011	Agriculture and Environment. Statistics Lithuania
Netherlands	NUTS2	1999–2011	Statistics Netherlands
Poland	NUTS2	1999–2010	Central Statistical Office, Poland
UK	NUTS1	1999–2012	Department for Environment Food & Rural Affairs, UK
Belgium	NUTS0	1999–2012	Eurostat
Bulgaria	NUTS0	1999–2012	Eurostat
Denmark	NUTS0	1999–2012	Statistics Denmark
Ireland	NUTS0	1999–2011	Central Statistics Office, Ireland
Slovakia	NUTS0	1999–2011	Statistical Office of the Slovak Republic
EU neighbourhood			
Russia	GAUL2	1999–2012	SovEcon, Russia
Ukraine	GAUL2	1999–2011	State Statistics Service, Ukraine
Algeria	GAUL2	1999–2012	La Direction des Statistiques Agricoles et des Systémes d'Information, Algeria
Republic of Moldova	GAUL0	1999–2011	FAO
Morocco	GAUL1	1999–2011	Institut National de la Recherche Agronomique, Morocco
Tunisia	GAUL2	1999–2012	Statistical Department, Ministry of Agriculture, Tunisia
Turkey	GAUL2	1999–2011	Turkish Statistical Institute

macro-region statistics, were used. Finally country-level (NUTS 0) statistics were used in the case of Slovakia, Bulgaria, Belgium and Denmark. A summary of the sources, level of disaggregation and length of the historical archive is given in Table 1.

Statistics were processed to make them more comparable. The different crops/varieties for which the different statistical services produce statistics were grouped into three broader classes: wheat, barley and grain maize. This aggregation was necessary given the variability of the crop definitions used by the different statistical services of these countries. The class 'wheat' includes: durum wheat (*Triticum durum*) and soft wheat (*Triticum aestivum*), the latter including both winter and spring sown wheat. The same procedure was followed for barley (*Hordeum vulgare*), aggregating when necessary winter and spring barley.

Changes in the boundaries and names of administrative units that have occurred during the study period were taken into account in order to make them match the administrative units defined in the spatial framework of this study (Fig. 1). This was particularly the case for Germany where several administrative regions were grouped in 2003 and 2008. Both spatial and thematic grouping were done by aggregating production and area statistics and only then computing yields from these aggregated values.

Regions where wheat, barley and/or grain maize surface is of minor importance at country level were discarded from the analysis. This was done by ordering the regions within each country according to their relative contribution to total country surface, and then retaining only those regions that contributed most to 90% of the total national surface in any year of the study period. Regions with incomplete yield or area statistics during the period 1999–2012 were also excluded: a (subjective) threshold of 10 years of continuous data was selected. Therefore, excluded regions (those shaded grey in Fig. 1) are those where grain yield production is – according to these thresholds – either marginal or incomplete; or where there is an insufficient number of pixels of arable land (see Section 2.3.2). In the specific case of grain maize, Germany was not considered since grain yield statistics at sub-national level were not available. The final number of administrative units included was 450 for total wheat, 410 for total barley, and 276 for grain maize.

2.3. *Processing of regional* fAPAR *time-series*

Time-series of SPOT-VEGETATION *fAPAR* product for the period 1999–2012 are available in the MCYFS. These products were acquired and processed in near real time (NRT) over Europe, and then used operationally for crop monitoring by the MARS Unit (Baruth et al., 2008; Royer and Genovese, 2004).

2.3.1. *Production of* fAPAR *10-day composite products*

SPOT-VEGETATION daily top-of-atmosphere (TOA) reflectance images at 1 km spatial resolution for the period 1999–2012 were acquired from CTIV (VEGETATION Image Processing Centre) run by VITO (Flemish Technical and Research Institute, Belgium). In this product, cloud and snow pixels are masked based on the methodology from Lissens et al. (2000).

fAPAR was computed from TOA reflectances using the algorithm proposed by Gobron et al. (2000), taking the value of the Optimized vegetation normalized index (OVNI). The algorithm retrieves a polynomial function relating *fAPAR* to multispectral TOA reflectance, as modelled by a radiative transfer scheme coupling the PROSPECT leaf optical properties model (Jacquemoud and Baret, 1990), the semi-discrete canopy reflectance model (Gobron et al., 1997), and the 6S atmospheric model (Vermote et al., 1997). The specific implementation of this algorithm for SPOT-VEGETATION is described in Gobron et al. (2002).

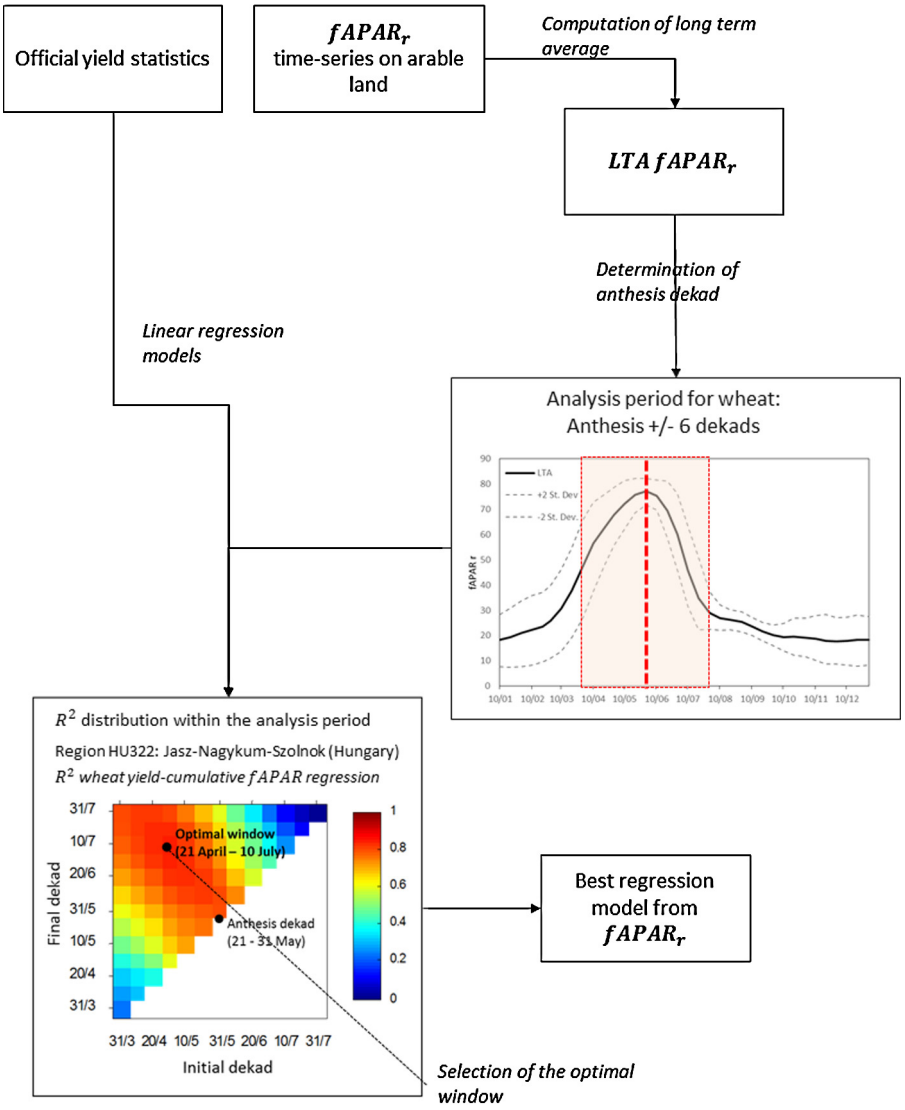

Fig. 2. Workflow followed to analyse the relationship between regional *fAPAR* time series and yield. The R^2 distribution figure is an example of wheat taken from the region Jasz-Nagykun-Szolnok in Hungary.

Dekadal (10-day) synthesis products corresponding to the period from 1 to 10th, 11 to 20th and 21st to the last day of each month, were generated from daily *fAPAR* images, following the maximum value composite method (Holben, 1986). Finally, temporal smoothing and gap-filling were applied to synthesis products, based on the method proposed by Swets et al. (1999) and modified by Klisch et al. (2007).

2.3.2. Computation of the regional fAPAR time series on arable land

The smoothed *fAPAR* time series were then aggregated from 1-km pixels to administrative units based on a static non-irrigated arable land mask. For the EU member states, this mask was derived from the land cover map of the Coordinated information of the European environment programme (CORINE, Buttner et al., 2004), taking the class 211 (Non-irrigated arable land). For areas not covered by CORINE (Ukraine, Russia, Belarus and North African countries), the mask comes from the GLC2000 Global Land Cover map (Bartholomé and Belward, 2005) class number 116 (Cultivated and managed areas).

The aggregation method follows the CORINE-NDVI (CNDVI) approach proposed by Genovese et al. (2001). An area fraction image has been specifically generated for the SPOT-VEGETATION product grid, containing the proportion of the pixel area (named area fraction, AF) belonging to the arable land class. This static image is then used to compute the aggregated *fAPAR* for region *r* as a weighted average:

$$fAPAR_r = \frac{\sum_i^n (fAPAR_i \times F_i)}{\sum_i^n (F_i)} \quad (1)$$

where $fAPAR_r$ (expressed in this paper as a percentrage) stands for the regional mean value, $fAPAR_i$ is the 10-day smoothed *fAPAR* value on pixel *i*, F_i represents the area fraction value for the same pixel, and *n* groups all the pixels included in administrative unit *r*. The adopted aggregation scheme therefore gives more weight to those pixels having a higher fraction of arable land in the computation of the regional mean value.

A threshold was established for each administrative unit to remove pixels with low F_i that may potentially contaminate the signal with information belonging to other land cover classes. This threshold was set at the value of the 90th percentile of arable land fraction values for that administrative unit. If the number of pixels above that threshold was lower than 25, the threshold value was lowered until it reached that minimum number. An administrative

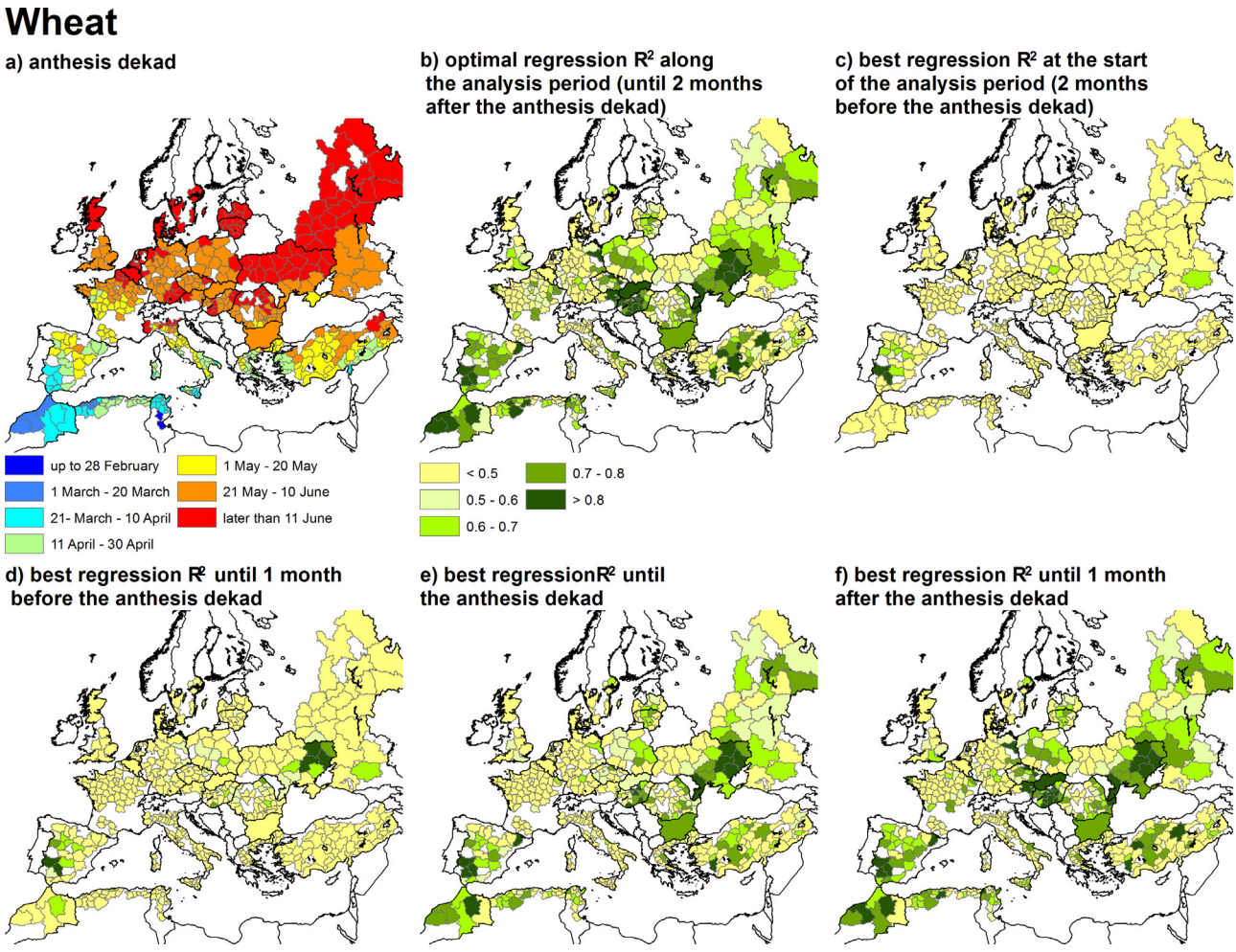

Fig. 3. (a) Dekad of anthesis computed from $fAPAR_r$ LTA (Section 2.4.1); and (b)–(f), coefficient of determination R^2 of the best regression models (Section 2.4.2) between cumulated $fAPAR_r$ and official wheat yields at administrative unit level on different moments of the analysis period.

unit is excluded from the analysis if: (i) the non-irrigated arable land area is below 2% of the total administrative unit area; or (ii) the average AF of the selected set of pixels above 90th percentile is below 0.10. In most of the regions the AF threshold was above 0.9 (pixels with more than 90% of their surface belonging to arable land).

2.4. *Analysis of the relationship between regional* fAPAR *time-series and observed yields*

The $fAPAR_r$-yield relationship is derived from the linear regression between time series of official yields and $fAPAR_r$ accumulated over an optimal temporal window selected for each individual administrative unit. The optimal temporal window is defined in two steps: first an analysis period is identified on the basis of the timing of maximum $fAPAR_r$ (when crop anthesis occur), then the optimal temporal period is searched for within this analysis period using linear regressions. A graphical representation of the methodology is given in Fig. 2.

2.4.1. *Computation of the anthesis dekad and analysis period*

Anthesis is the phenological stage of maximal flower expansion and a critical period determining yield potentials for winter cereals (Slafer and Rawson, 1994; Ugarte et al., 2007). Here, it is used as a reference date to set the analysis period to the part of the crop growing season where *fAPAR* would be relevant to estimate yield.

The average anthesis dekad is calculated for each administrative unit using the dekadal long term average (LTA) of $fAPAR_r$ over the period 1999–2012. The dekad where this average curve reaches its maximum value during the growing season is considered the anthesis dekad. The timing of anthesis was constrained to the period from 10th January–10th July for wheat and barley, and from 21st June–30th September for maize, since anthesis is unlikely outside these time intervals in all regions of Europe. Then, the analysis period is defined for every region as a 13-dekad window centered on the estimated dekad of anthesis (see the example in Fig. 2).

Here it has been assumed that anthesis dekad and, consequently, the analysis period is the same for wheat and barley. However, phenology may vary substantially between both crops and differences in the date of anthesis of wheat and barley may actually exist within an administrative unit (in some cases up to 2–3 weeks), especially where winter and spring varieties of both crops coexist. In this case the estimated anthesis dekad would represent the predominant crop/variety. The reliability of the anthesis dekad estimate is expected a priori to be low for those crops covering the lowest proportion of arable land area. What's more, the analysis period for wheat and barley, and for maize, do overlap in many regions as the arable land mask used does not separate winter from summer crops.

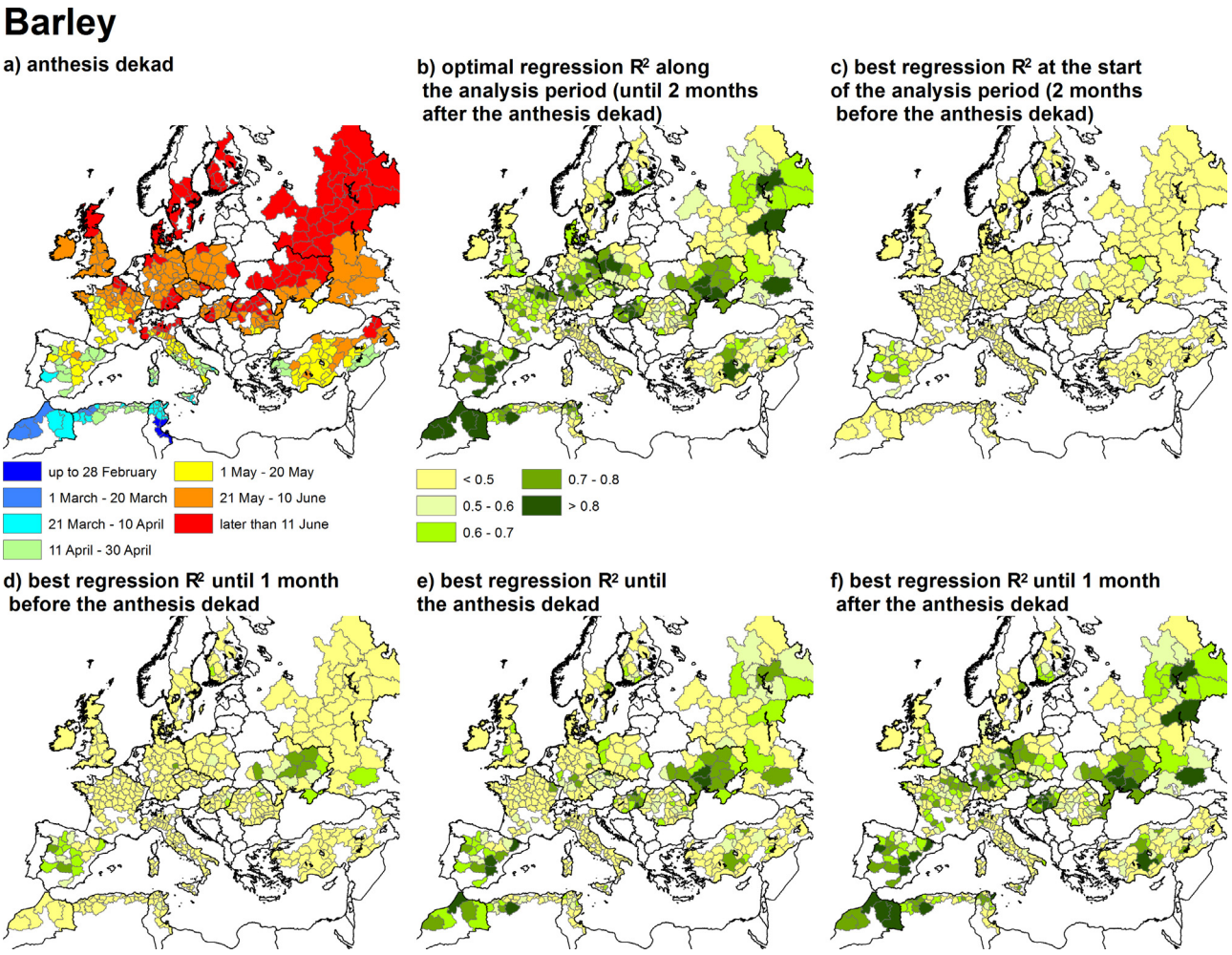

Fig. 4. (a) Dekad of anthesis computed from $fAPAR_r$ LTA (Section 2.4.1); and (b)–(f) coefficient of determination R^2 of the best regression models (Section 2.4.2) between cumulated $fAPAR_r$ and official barley yields at administrative unit level depending on the period of the analysis period.

2.4.2. Analysis of the correlation between cumulated $fAPAR_r$ *and official yield statistics*

Regression models relating cumulated $fAPAR_r$ (predictor) to official yields (predictand) over 1999–2012 were analysed for all possible temporal windows – from single-dekad windows up to the largest 13-dekad window – within the analysis period (Section 2.4.1), resulting in 91 regression models per crop for each administrative unit. The number of years included in the regression varies among countries depending on the availability of official statistics (see Table 1).

The rationale behind accumulating $fAPAR_r$ within a temporal window, as opposed to using only individual dekads, is to mitigate the influence of any possible bias on yield prediction caused by the quality of single dekad composites (e.g. a low number of valid observations within the synthesis period). Shifts in crop phenology associated with unusual weather events in some seasons (that do not necessarily affect yields) may also affect the interpretation of $fAPAR_r$ in individual dekads. Cumulative $fAPAR_r$ is therefore used to arrive at a seasonal indicator of green leaf area formation (as suitable proxy of gross primary productivity) as yield predictor.

The coefficient of determination R^2 between cumulative $fAPAR_r$ and observed yield (as a function of the initial and final dekad of all the 91 possible windows) is analysed, and the optimal temporal window for a given administrative unit is selected as the window with the highest R^2.

Inter-annual stability in land use is a priori an important assumption of this methodology. This assumption is dictated by the use of a static crop mask instead of a dynamic one. As a consequence, changes in crop areas over the years (e.g. regions re-orienting their production towards different crops/varieties) may introduce anomalies in $fAPAR_r$ time-series of arable land not necessarily related to yield variations.

2.4.3. Additional weather indicators

Additional weather indicators are analysed in an attempt to interpret the relationship between green biomass accumulation – as indicated by $fAPAR_r$ time-series – and official yields using data from 1999 to 2012 also contained in the MCYFS archive.

These include temperature, precipitation and potential evapotranspiration (ET_0) datasets produced at 25 km regular grid from interpolated daily observations coming from more than 2200 weather stations across Europe, and then spatially aggregated at administrative unit level. ET_0 in the MCYFS is derived from observed temperature, water vapour pressure and global radiation using the Penman–Monteith equation (Allen et al., 1998). Details about indicator computation, spatial interpolation and aggregation of weather observations can be found in the MCYFS documentation (Micale and Genovese, 2004; Lazar and Genovese, 2004) and the MCYFS internet site: http://marswiki.jrc.ec.europa.eu/agri4castwiki/index.php/Weather_Monitoring.

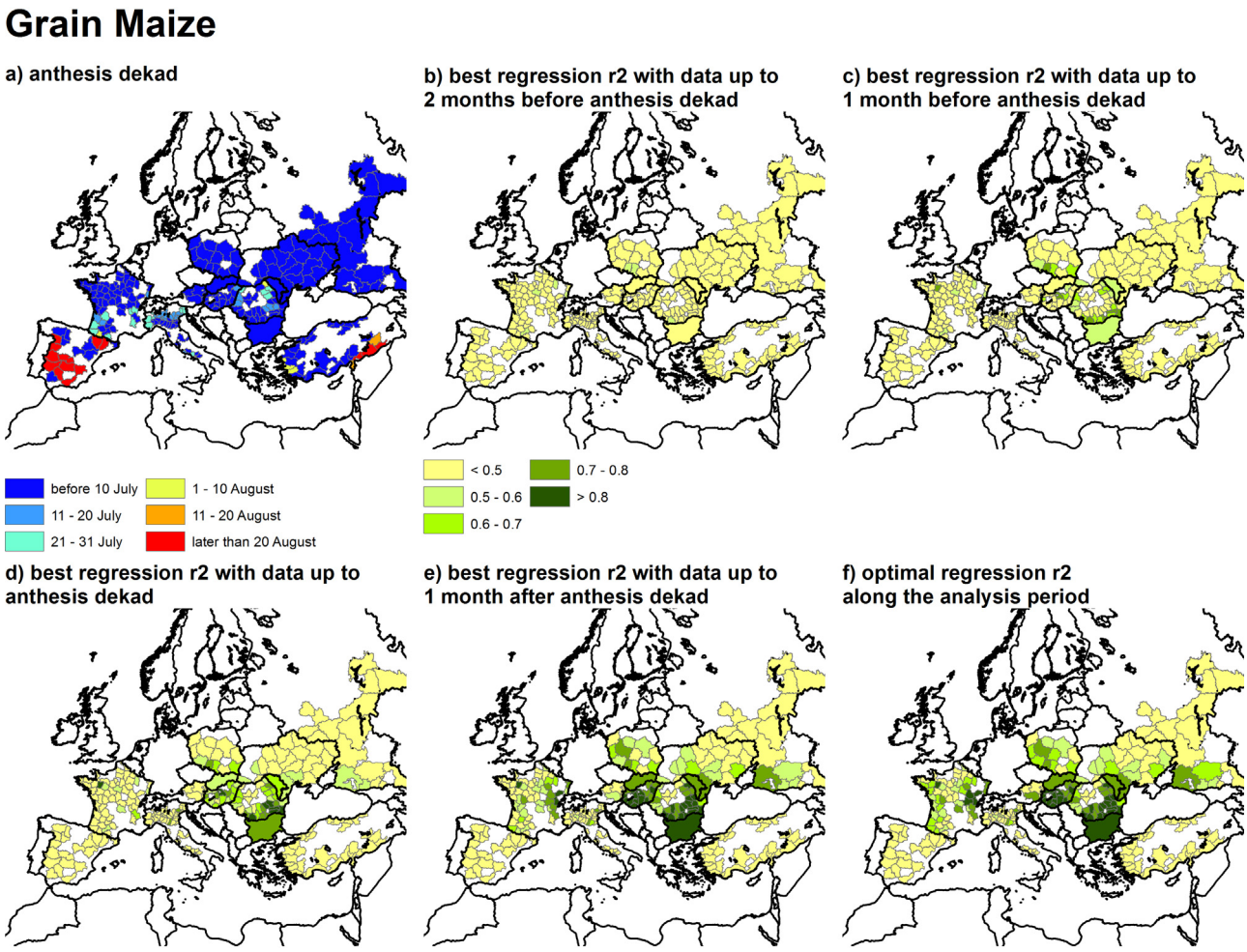

Fig. 5. (a) Dekad of anthesis computed from $fAPAR_r$ LTA (Section 2.4.1); and (b)–(f) coefficient of determination R^2 of the best regression models (Section 2.4.2) between cumulated $fAPAR_r$ and official maize yields at administrative unit level depending on the period of the growing season.

3. Results

3.1. *Correlation between regional* fAPAR *and observed yields across Europe*

Fig. 3 depicts the evolution of the coefficient of determination R^2 between cumulative $fAPAR_r$ and official wheat yields at different periods of the growing season. The anthesis dekad, extracted from the $fAPAR_r$ LTA (see Section 2.4.1) is shown in Fig. 3a and illustrates the variation across Europe, with western Mediterranean regions being the earliest ones (March–early April), then progressing towards the North following a latitudinal gradient to the latest ones in northern Europe (June).

The coefficient of determination of the optimal regression model (allowing the temporal window to come from any time within the analysis period, i.e. from two months before until two months after the anthesis dekad) is represented in Fig. 3b. In Mediterranean countries (except Italy and Greece) as well as in Hungary, Slovakia, the Black Sea area and many regions in southern Russia, $fAPAR_r$ is well correlated with yields (R^2 usually above 0.6). These results agree with previous studies applying similar methodologies in Ukraine (Kogan et al., 2013) and Morocco (Balaghi et al., 2008).

The progression of R^2 through the season (Fig. 3c–f) shows how, in the above mentioned areas, $fAPAR_r$ correlation with yields improves when the accumulation window includes dates one month after the anthesis dekad (Fig. 3f). In the post-anthesis period, when grain formation is taking place, the correlation of yield with $fAPAR_r$ is maximum in Hungary, Slovakia and Poland. In some regions of Spain, Maghreb countries and Ukraine, the information available up to the anthesis dekad (Fig. 3e) is already highly correlated with yield, indicating the existence of substantial *fAPAR* anomalies early in the season and their importance in determining the final yields.

By contrast, the strength of the correlation decreases towards the north of Europe. In most of France, Germany, the UK and Denmark the results show a much weaker relationship between yields and $fAPAR_r$ compared to the southern half of Europe. R^2 exceeds 0.6 in relatively few regions of Germany, France and the UK.

A similar spatial distribution of R^2 is seen for barley (Fig. 4b). In western Mediterranean regions (especially in Spain and North Africa) cumulative *fAPAR* exhibits a strong relationship with official yields. The correlation is especially high in Spain, one of the main EU producers with about 20% of the total barley area in the EU. The temporal progression of R^2 for barley (Fig. 4c–f) in these areas is similar to wheat. The correlation between $fAPAR_r$ and yields is already strong including data up to the anthesis dekad (second half of April, first of May Fig. 4e) and increases progressively when dekads after anthesis are included (Fig. 4f and b).

In the Black Sea area, the correlation between $fAPAR_r$ and barley yield is substantially lower than for wheat. Nevertheless, R^2 remains high in the southern half of Ukraine, where wheat and barley official yields are strongly correlated with each other, and also in some of the main winter cereal producing regions in Turkey's

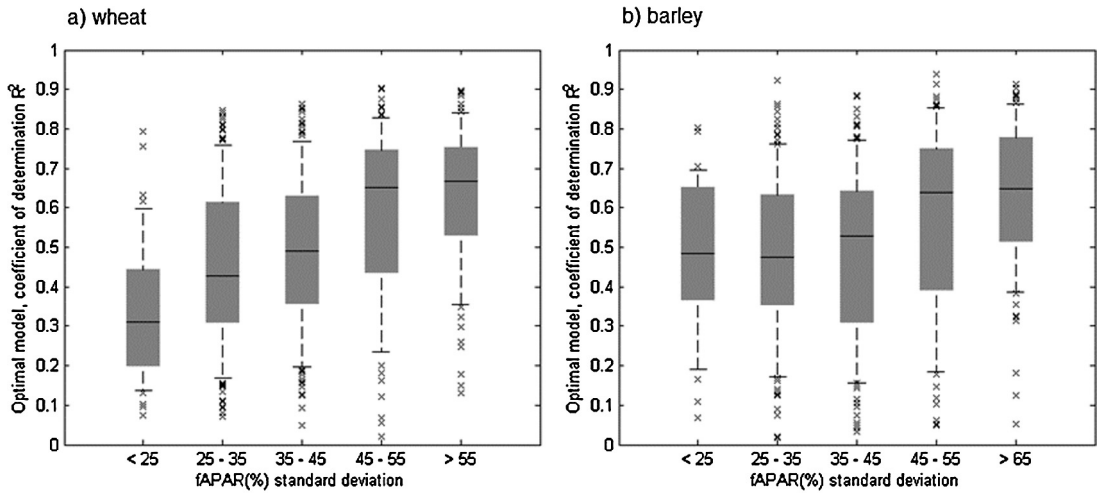

Fig. 6. Boxplots (median; boxes for 25 and 75 percentiles; whiskers for 10 and 90 percentiles; and outliers) relating the standard deviation (inter-annual variability from 1999 to 2012) of cumulated *fAPAR* (%) along the analysis period to the coefficient of determination of the optimal models for (a) total wheat (see Fig. 3b), and (b) total barley (see Fig. 4b).

central plain (e.g. Konya and Ankhara). By contrast, R^2 is less than 0.6 in most of Romania, where barley is a minor crop compared to wheat and, especially, grain maize.

In France and Germany $fAPAR_r$ is more strongly correlated with barley yields, compared to wheat. In the northeastern departments of France where barley – especially spring varieties – is cultivated over a large proportion of the arable land, R^2 reaches 0.8. Also, in the central part of the country where barley area is low compared to wheat, R^2 remains higher than 0.6. Similarly, in Denmark and in southwestern and easternmost regions of Germany, where the presence of spring barley is more prevalent than winter varieties, $fAPAR_r$ after anthesis is well correlated with observed yields (R^2 ranging from 0.6 to 0.8). These results are in agreement with Weissteiner and Kühbauch (2005), who found a significant correlation between NDVI and malting barley yields in some districts of southern Germany.

For grain maize (Fig. 5) the best results are achieved in Romania and Hungary (summing up to 40% of the total EU grain maize area), where R^2 is usually higher than 0.8. In Bulgaria, Moldova and Slovakia $fAPAR_r$ time-series are highly correlated with maize yields as well. For all these countries, a significant increase in R^2 is observed in August, during grain-filling (see Fig. 5e).

In France, the correlations between $fAPAR_r$ time-series and maize yield are moderate. In the southwest, where most of the grain maize is grown, R^2 varies between 0.5 and 0.7. Similarly, in the Alsace region and the east of France R^2 reaches higher values (>0.6).

By contrast, in other important grain maize producing countries of Europe observed yields and cumulative $fAPAR_r$ are poorly correlated. This is the case in Ukraine, Turkey and Spain, as well as northern Italy, where maize has an important share in total cereals (and R^2 is below 0.3).

3.2. *Role of agro-climatic conditions in the* fAPAR-*yield correlations*

Local agro-climatic conditions constitute an important factor explaining the spatial distribution across Europe of the correlation between $fAPAR_r$ and observed yields described in Section 3.1. Fig. 6 depicts the variation of R^2 as a function of the $fAPAR_r$ inter-annual variability (expressed as the standard deviation of regional values cumulated along all the entire analysis period) for wheat and barley between 1999 and 2012.

The R^2 is clearly influenced by the *fAPAR* inter-annual variability in the case of wheat (Fig. 6a). For regions where this variability is low (standard deviation <25%), the correlation of $fAPAR_r$ with yields tends to be low. In administrative units where *fAPAR* is more variable among the years, R^2 increases in parallel with *fAPAR* variability. Theoretically, in regions with large inter-annual variation in *fAPAR*, the correlation between *fAPAR* and yields should be high, as weather conditions are producing large variations in leaf area between years, which are in turn determining final yields. Conversely, small variation in *fAPAR* would suggest that yields are instead influenced by factors not necessarily linked to leaf area formation.

The effect of $fAPAR_r$ variability is much less evident with barley (Fig. 6b). Although R^2 remains high in regions with a high inter-annual variability (*fAPAR* standard deviation >45%), the differences are not substantial compared to the regions where *fAPAR* time-series are more stable. For instance, R^2 values observed for barley yields were moderately high (>0.6) in many regions of France and Germany (Fig. 4b) where $fAPAR_r$ variability was low and poorly correlated with wheat yields. It is noted, however, that wheat is the predominant crop in these regions, and therefore the actual *fAPAR* inter-annual variability coming from barley fields is masked, which may distort the relationship between R^2 and $fAPAR_r$ standard deviation.

Water availability is the main driving factor determining *fAPAR* variability across Europe. Fig. 7 shows the relationship between total cumulative *fAPAR* within the analysis period and its variability, with the climatic water balance (precipitation minus potential evapotranspiration) from 1st January to 30th June. Both relationships are statistically significant ($p < 0.01$).

In regions with high water deficit (e.g. climatic water balance below −100 mm) $fAPAR_r$ is low and its inter-annual variability is quite high. This is a consequence of the high variability on leaf area expansion and senescence on winter cereals between the years, both driven by seasonal rainfall regimes. By contrast, in those regions where the climatic water balance is close to zero, or even positive, cumulative $fAPAR_r$ presents little variability between the years, as water availability permits an optimal leaf area expansion most years.

Some illustrative examples from different agro-climatic conditions in Europe are given in Fig. 8. Valladolid (Spain) and Mykolayivs'ka (Ukraine) represent two regions of Europe where barley and wheat, respectively, are grown under water-limited

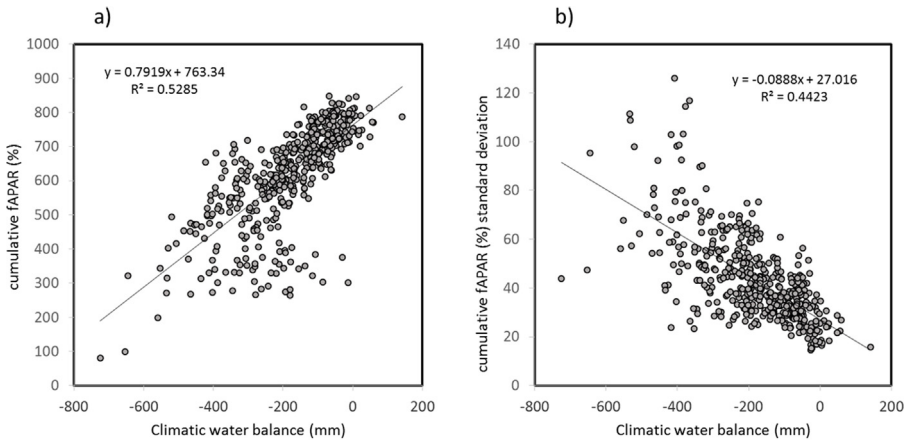

Fig. 7. Relationship between LTA of the climatic water balance (precipitation minus potential evapotranspiration, in mm) from 1st January to 30th June, and (a) average (from 1999 to 2012) cumulated *fAPAR* within the analysis period for winter wheat, and (b) its standard deviation. The whole set of 450 administrative units is included.

conditions. The $fAPAR_r$ time-series shown in Fig. 8a and b present an important inter-annual variability. The years 2005 and 2009 in Spain, and 2007 in Ukraine were characterized by severe drought conditions (cumulated precipitation from January to June of about 100 mm) determining leaf area growth and yields. The year 2003 in Ukraine – and many other regions of Europe – was characterized by unusually cold temperatures during early spring, which hampered the development of wheat and barley and negatively affected yields. By contrast, in years with abundant precipitation during the growing season (e.g. 2000 and 2008 in Spain; 2001 and 2008 in Ukraine) observed yields are quite high, and correspond to higher-than-average $fAPAR_r$ values.

Compared to the previous two examples Mecklemburg-Vorpommern (Germany) and Nord (France) represent the opposite agro-climatic conditions. These are in fact humid regions with high average wheat and barley yields, and small inter-annual variability of the $fAPAR_r$ time-series (Fig. 8c and d). Moreover, in some years with quite low yields the $fAPAR_r$ values are not particularly below the LTA, and, conversely, years with very high yields are not associated with $fAPAR_r$ values above the LTA. This is the case in 2007, when low yields observed in northern Germany and France were not the consequence of dry conditions constraining the growth of wheat and barley, but a result of the continuous rain that fell through the summer during ripening-harvest stages, which lead to a drastic yield reduction (but no appreciable influence on *fAPAR*). Wheat yields in 2011 (Nord) and 2008 (Mecklemburg-Vorpommern), for example, were among the highest of the period studied, but $fAPAR_r$ time-series were close to those of 2007.

Appendix A contains an extended analysis of the empirical $fAPAR_r$ model performance on extreme years (years with unusually low or high yields), and the influence of specific adverse weather events, both on $fAPAR_r$ time-series and yields.

In the case of maize, the influence of agro-climatic conditions on the $fAPAR_r$-yield correlations is similar to wheat and barley: R^2 is high in regions where water availability during summer is the main factor constraining yields. Nevertheless, the $fAPAR_r$ time-series computed over arable land do not allow a specific analysis of *fAPAR* for maize as it occupies a minor area compared to winter crops (except in the Black Sea area, the North of Italy and some regions of France – see Fig. 1). This can result in poor correlations between *fAPAR* and yield in many countries and complicates the analysis of the relationship between the actual $fAPAR_r$ variability of maize and the model R^2, as that presented for wheat in Fig. 6a.

Fig. 9 shows $fAPAR_r$-yield models and $fAPAR_r$ time-series for three important maize producing regions of Europe. Fig. 9a illustrates a paradigmatic case in the southeast of Europe (Romania, Hungary, Moldavia, Slovakia and Bulgaria), where maize is rarely irrigated (less than 5% of the total area of grain maize in Romania, according to Eurostat, 2010) and yields are mostly determined by precipitation during crop vegetative and reproductive phases (mainly July and August). The inter-annual variability of $fAPAR_r$ is high, reflecting the large influence of rainfall regimes during the growing season on leaf area expansion and senescence rates, which are in turn driving crop yields. The consequences of the hot and dry conditions experienced during the summer period of 2007 and 2008 (see Appendix A), leading to very low yields, can be appreciated on the $fAPAR_r$ time-series for both years (Fig. 9a). By contrast, the high $fAPAR_r$ values for 2011 reflect the favourable weather conditions observed that year with abundant rainfall during July and August, and which resulted in high yields.

In France, irrigation reduces the inter-annual variability of $fAPAR_r$ and yields (Fig. 9b). According to official statistics, about 50% of the total maize area is under irrigation, but in many regions a strong correlation between *fAPAR* and yields still holds. However, such correlation is highly influenced by the presence of two extreme observations in the regression: the year 2003 where high temperatures and the absence of rainfall in summer caused significant damage to grain maize; and 2011, characterized by abundant rainfall and high yields. Thanks to that, the $fAPAR_r$-yield correlation is strong also in some regions in the East, where the presence of maize is minor (Fig. 5).

Italy represents the opposite: the proportion of irrigated maize is higher than is the case in France (about 65%, according to Eurostat, 2010), and the inter-annual variability of (official) yield is very small (Fig. 9c), which explains the poor $fAPAR_r$-yield correlations. Moreover, in some years (e.g. 2007) when the $fAPAR_r$ time-series would suggest that grain filling during summer was constrained by water availability in many regions, the observed yields were not affected, likely due to irrigation practices.

3.3. *Effect of changes in agro-management on* fAPAR-*yield correlations*

Possible changes in agro-management during in the study period have to be accounted for when analysing the relationship between spatially aggregated EO biophysical indicators and regional yield statistics as they may produce a progressive increase of yields (a yield trend) independent of the influence of weather conditions. Yield trends are, in general, associated with one or more of the following factors: an extension of irrigated area, improvements in crop genetics, the adoption of new pest and disease controls, and the use of fertilizers. Such factors may additionally

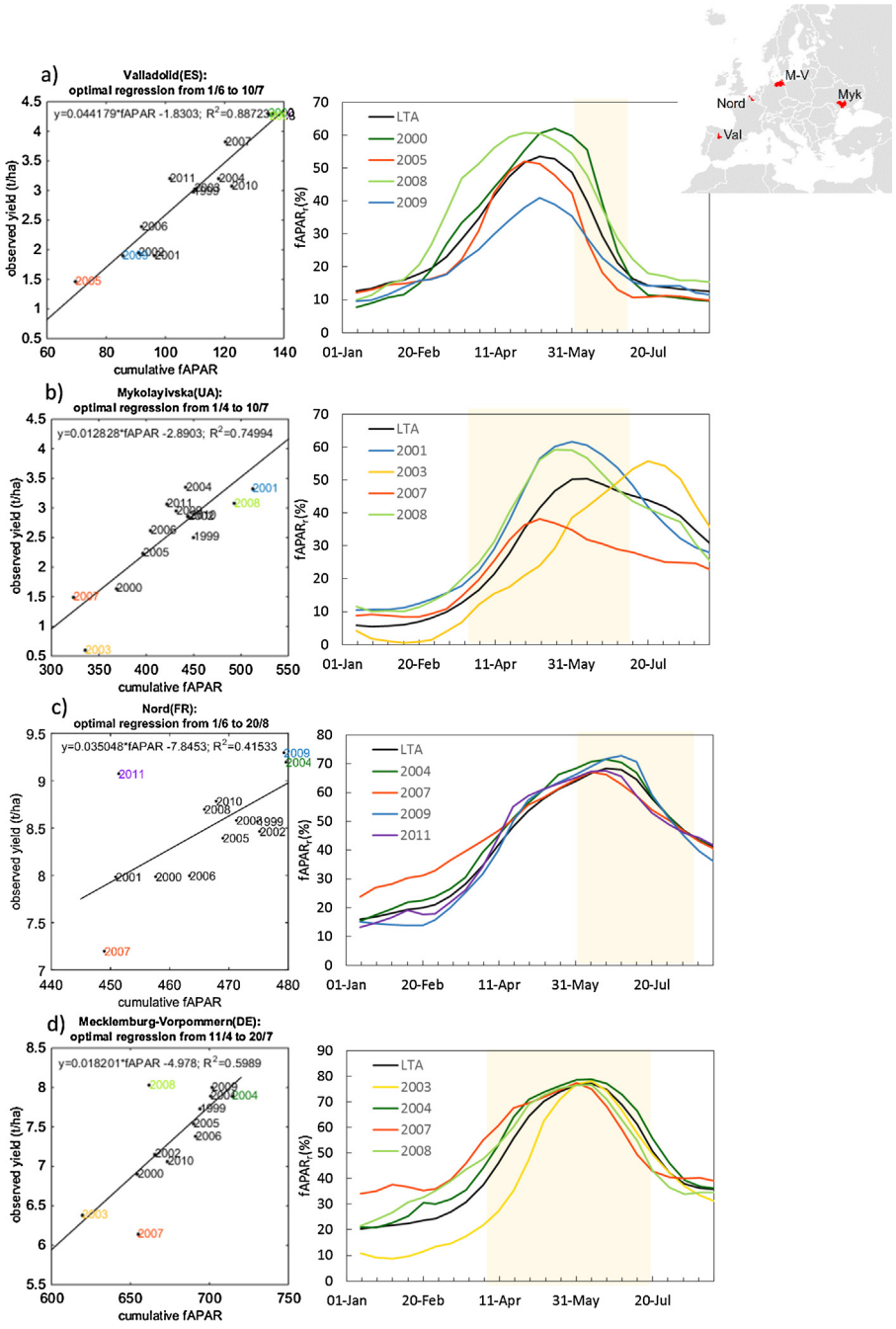

Fig. 8. Optimal $fAPAR_i$-yield regression models and their respective $fAPAR_i$ time-series in four major regions for winter cereal production across Europe. Wheat yields are predicted in all the models except on a) Valladolid, where the regression is done with barley yield.

interact with changes in climate (Brisson et al., 2010; Chloupek et al., 2004; Lobell et al., 2011).

Changes in agro-management during the period 1999–2012 can be observed in the yield series of maize for some countries. In Ukraine, Turkey and to a lesser extent Spain, the yield inter-annual variability of maize is mostly explained by a yield trend in time, rather than by weather conditions. A simple linear regression model, using year as yield predictor can help to assess how much absolute yields are explained by trends (Fig. 10 reports the R^2 values of such a model as an indicator of the strength of the trend). Yield trends are not observed for wheat and barley.

In Ukraine, grain maize production has experienced substantial change in the last decade. As reported by Lioubimtseva and Henebry (2012), land use/land cover changes are one of the main change vectors. Sown area has increased by a factor of three in the last 15 years. This increase in surface has been accompanied by a higher use of fertilizers and improvements in grain varieties, resulting in a yield trend like the one shown in Fig. 10. In Turkey the yield trend is due to a re-distribution of grain maize areas towards South-Eastern regions where an important development of irrigation infrastructures started in the early 2000s, thanks to the GAP project (acronym for South-Eastern Anatolia regional development project of the Turkish Government, see http://www.gap.gov.tr/english for further details). In Spain, 90% of the grain maize surface is irrigated, leading to high yields with a low inter-annual variability, mostly associated with a yield trend (as in the example shown in Fig. 10) which result from various technological improvements (maize varieties, fertilization, etc.) as reported by Lopez-Bellido (2009).

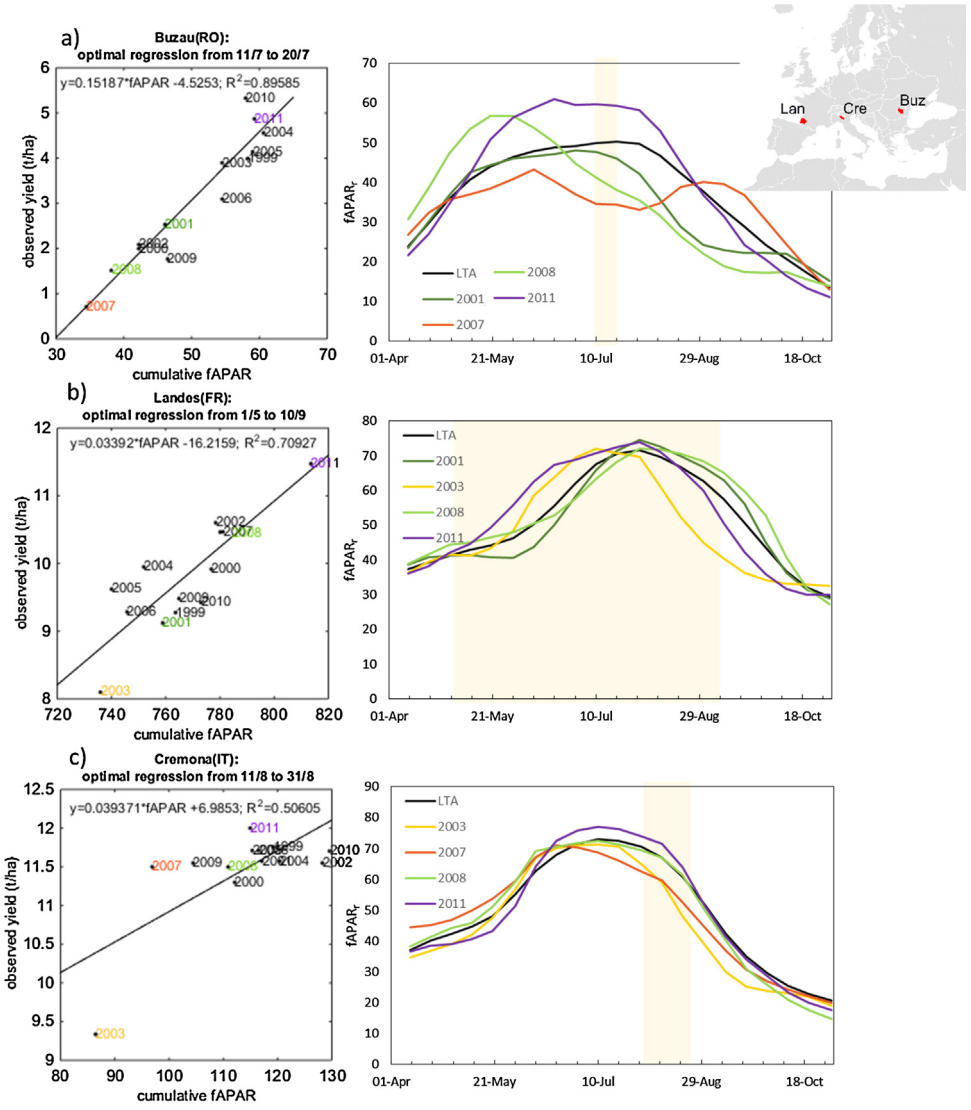

Fig. 9. Optimal $fAPAR_r$-yield regression models and their respective $fAPAR_r$ time-series in three major regions for maize production across Europe.

The relationship between *fAPAR* and yield becomes more difficult to interpret in regions where most of the yield inter-annual variability is the result of a trend, leading to poor *fAPAR*-maize yield correlations as have been observed in the above mentioned countries (as shown in Fig. 5). In the case of Turkey and Ukraine, land use/land cover changes experienced in the last few years make interpretation of the time-series even more complicated, as they are not properly captured by the static arable land mask used in this study.

An alternative approach when a yield trend exists is to use EO products to predict the yield deviations from the trend, as done by Duveiller et al. (2013) for sugarcane, where the trend represents the expected yield given average weather conditions, and based on the agro-management practices of every year. However, it is difficult to estimate the yield trend (e.g. to isolate the influence of agro-management factors on absolute yield) based only on yield time-series.

4. Discussion

In the previous section the main factors determining the correlations found between *fAPAR* time-series and wheat, barley and grain maize yields in Europe were presented. The operational use of EO biophysical products in a near-real time crop yield forecasting context requires an adequate interpretation of the information contained within the predictors, and understanding the main limitations of the methodology.

This study is based on a definite biophysical product (*fAPAR*) but, in our opinion, the main results and findings explained above could be extrapolated to other specific EO products or vegetation indices used in a similar framework. Nevertheless, the possible influence of specific features of the biophysical product used in this study, such as uncertainties in atmospheric correction or sensitivity to actual canopy conditions (e.g. saturation effects at high *fAPAR* values) will also have to be addressed by future studies using different products.

4.1. Use of EO time-series for yield assessment

In the present study *fAPAR* time-series were found reliable for quantitative yield assessment of wheat, barley and maize, where a single dominant factor – water availability – is responsible of most of the inter-annual yield variability. As discussed in Sections 3.1 and 3.2, crops growing under moderately or strongly water-limited conditions present a high inter-annual variability on leaf area development and senescence that influences the final yield, which is observable by EO systems.

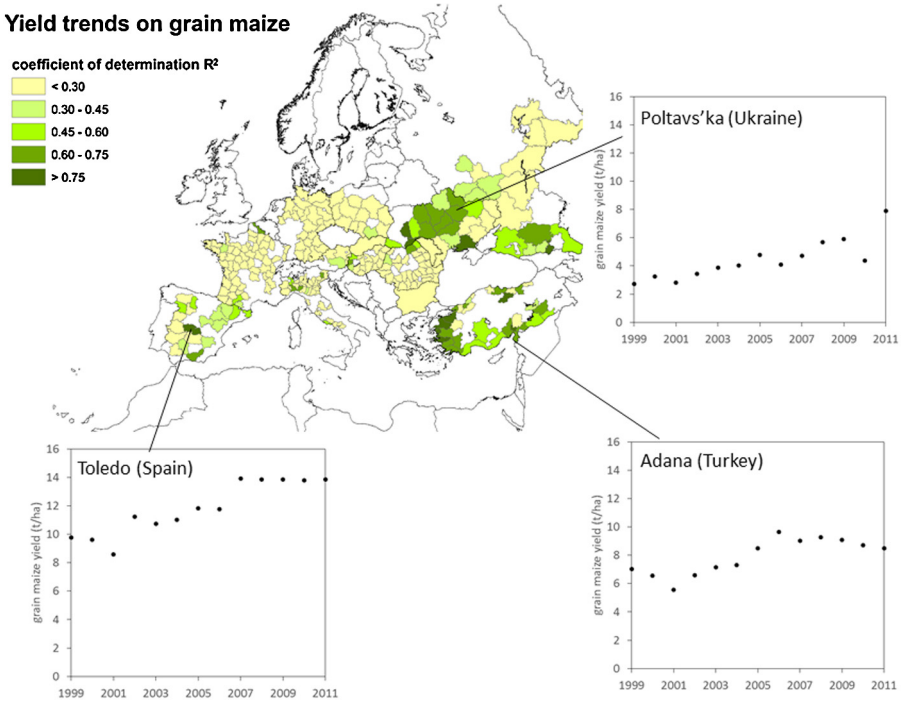

Fig. 10. Coefficient of determination R^2 for the linear yield trends in time observed on grain maize for the period 1999–2012, and yield time-series for three specific regions.

In contrast, in agro-climatic areas where crop growth is not frequently water-limited (e.g. central and northern Europe, irrigated crops) yield variations are explained by a complex interaction of different factors not necessarily related to crop leaf area dynamics. For instance in north western Europe, low wheat and barley yields in 2007 were caused by an excess of rain at harvest which cannot be detected by *fAPAR* time-series; whereas, high yields in 2008 and 2011 were not associated with positive *fAPAR* anomalies (Section 3.2). Further examples can be found in Appendix A.

In these conditions, weather indicators are needed to complement *fAPAR* time-series in order to identify the underlying factors explaining yield, as done by Newlands et al. (2014). A synergistic use of meteorological data, EO biophysical products and especially crop models, is needed to adequately interpret the interactions between weather events, leaf area formation, crop phenology, yield formation and harvest conditions (de Wit, 2007; de Wit et al., 2010).

As demonstrated by Richards and Townley-Smith (1987), the relationship between leaf area formation and yield may be complex, where the latter is not necessarily (or always) determined by the former. Therefore, *fAPAR* gives an objective assessment of canopy green biomass, but cannot fully explain actual crop yield (or yield potential). In contrast, the grain filling process is specifically simulated in crop models (Van Diepen et al., 1989; Stöckle et al., 2003) taking into account several limiting factors: leaf area, soil water content, nitrogen content, optimal growth temperatures etc. These factors can be integrated to produce a final yield assessment, thus representing a complementary view to EO-products. Kowalik et al. (2014) reported that in France and Germany crop model indicators outperform EO products for wheat yield prediction, producing lower estimation errors. In this specific context, the assimilation of remote sensing information into crop models at regional scale should be further investigated (Launay and Guerif, 2005; Dorigo et al., 2007) as a way to enhance the capabilities of crop models with remote observations of crop development and growth.

Moreover, it should be noted that the strength of the correlation between EO biophysical products and crop yield is also influenced by the reliability of the yield statistics. For this study, yield time series were required at regional scale. The only systematic observations available are the official yield statistics, and are here considered as true observed values. However, the protocols to derive these figures (and therefore their accuracy as well) may vary from one country to another.

4.2. Specific issues related to the spatial aggregation of indicators

The availability of crop-specific information plays a major role when regional time-series are used for crop yield forecasting. In this study, the use of a single static arable land mask to aggregate *fAPAR* from pixels to administrative units does not permit computing crop-specific time series. This constitutes an important limitation for those crops with a low share in the total arable land area and for those regions where land use varies substantially from one year to another. Indeed, this limitation may itself have led to some of the weak correlations observed between $fAPAR_f$ and barley and maize yields in many regions of Europe (Section 3.1). The availability of a dynamic crop classification updated yearly towards the end of the growing season would help to solve some of these problems. For example, the Cropland Data Layer (Johnson and Mueller, 2010) using MODIS and Landsat data to produce a yearly crop mask for the coterminous US represents a valuable product to derive crop-specific time-series. Unfortunately, there is currently no consistent, dynamic crop mask across EU member states, although some recent initiatives show promising results (De Bies et al., 2014).

This issue is also partially linked to the spatial resolution of the products used (Doraiswamy et al., 2005), as the full benefit of a dynamic crop mask could be achieved only with a suitable spatial resolution. Duveiller and Defourny (2010), studied the spatial resolution requirements for agricultural monitoring, concluding that in the highly fragmented agricultural landscapes of Europe the 1 km spatial resolution observations provided by the SPOT-VGT instrument are not suitable for crop-specific monitoring. Duveiller (2012) did show, however, that biophysical variables extracted from 250 m MODIS time series could characterize a crop specific signal for wheat in a complicated European landscape. The Copernicus Programme for EO recently launched by the

European Commission (http://land.copernicus.eu/global) jointly with recent projects on the exploitation of the forthcoming Sentinel multi-resolution constellation and the production of added-value datasets for agriculture such as ImagineS (funded by EC FP-7, see http://fp7-imagines.eu/) or SIGMA (see http://www.geoglam-sigma.info/Pages/default.aspx) might overcome some of these challenges in the future.

5. Conclusions

In the present study an exploratory analysis of the relationship between regional *fAPAR* time-series and yield inter-annual variability over the period 1999–2012 has been conducted, evaluating its reliability for operational crop yield forecasting activities.

In light of the results, remote sensing time-series are confirmed to be a reliable tool for regional crop yield forecasting with a strong potential to contribute effectively to operational systems such as those currently running at continental/global level (GIEWS, NASS, FAS, CropWatch or MCYFS). However, a correct interpretation of these time-series is needed under certain agro-climatic conditions.

Overall, the results indicate strong correlation between *fAPAR* and official yields in agro-climatic conditions where yield tends to be water-limited, for all three crops studied (wheat, barley and maize). Where crops are subject to frequent water stress (e.g. the Mediterranean basin, the Black Sea area) most of the yield inter-annual variability is explained by leaf area dynamics along the growing season, and can be well captured by regional *fAPAR* time-series.

In contrast, in regions of high yield where *fAPAR* inter-annual variability is generally low as consequence of rainfall regimes close to potential evapotranspiration (or intensive irrigation in the specific case of grain maize), the correlation between *fAPAR* and yield is weak. In such conditions yield variations are mainly explained by factors other than leaf area dynamics (e.g. excess of rain during harvest).

The use of remote sensing in operational crop yield forecasting at regional scales therefore requires the consideration of additional, independent indicators: weather parameters, crop model outputs, etc. EO biophysical products provide a synoptic view of the vegetation status and leaf area formation along the growing season, but the results of this study indicate that, under specific agro-climatic conditions, their interpretation may be sometimes mislead, especially when they are used alone to explain extreme low and high yields.

The relationship between spatially aggregated EO biophysical product time-series and observed yields may be highly scale-dependent. In the present work a wide variety in the size of the administrative units was considered, as a consequence of the differences in availability of regional yield data between countries. The inter-annual variability of biophysical products may be highly influenced by the size of the regions considered, especially if they include heterogeneous agro-climatic conditions. In an operational forecasting system, scale effects have to be properly addressed in order to target an adequate scale of analysis. Specific issues related to spatial aggregation to generate regional crop-specific *fAPAR* series (such as crop mask, spatial resolution) still have to be solved to improve the suitability of EO products for yield estimation.

Acknowledgments

The authors are very grateful to the National Statistical Offices and Institutions listed in Table 1 which provided the regional crop statistics (and which were essential to accomplish this study). The useful inputs of Dr. Maurits van den Berg to the present manuscript are sincerely acknowledged. The authors also thank an anonymous reviewer, whose comments have helped to improve the quality of the manuscript.

Appendix A. Analysis of extreme years in the empirical *fAPAR*-yield models across Europe

Analysis of extreme years in the empirical fAPAR-*yield models across Europe*

Extreme years are here defined as those showing unusually high or low yields compared to their historical mean. The influence of those years is typically very high in determining the coefficients of the empirical $fAPAR_r$-yield model. The purpose of this section is: (1) to identify extreme years in the empirical models between cumulative $fAPAR_r$ and observed yields across Europe, presented in Section 3.1; and (2) to analyse how $fAPAR_r$ times-series are reflecting the meteorological event causing it. This is an important task to assess the possible shortcomings in the use of EO biophysical products to predict yields when specific meteorological events occur.

The analysis presented here refers to wheat and maize only. Results for wheat are similar to those for barley and are not reported for the sake of brevity. The only differences between wheat and barley models were found in northeastern France and southern Germany where, as mentioned in Section 3.1 of the manuscript, the correlation between *fAPAR* and yield was higher for barley than for wheat.

A.1. Methods

A.1.1. Identification of the extreme years

Two different statistics are used to identify which years can be considered extreme. The first one is the leverage of the regression on each specific year. Leverage identifies those observations that are far away from the rest in the *X*,*Y* plane. It is calculated from the projection matrix (*H*), that maps the vector of observed yield values to the vector of fitted yield values:

$$H = X(X'X)^{-1}X' \tag{A.1}$$

where *X* is the vector (or matrix) of observed (cumulative) *fAPAR* values. The leverage score h_i of a model on a specific year *i* is the i^{th} diagonal of the projection matrix:

$$h_i = (H)_{ii} \tag{A.2}$$

The second one is the Cook's distance, an indicator of the influence of a specific year on the model. More specifically, Cook's distance *D*, identifies those years that have more impact on yields estimated by the model:

$$D_i = \frac{\sum_{j=1}^{n}\left(\hat{Y}_j - \hat{Y}_{j(i)}\right)^2}{p\text{MSE}} \tag{A.3}$$

where $\hat{Y}_j$ is the estimated value for year *j* using the whole set of observations, $\hat{Y}_{j(i)}$ is the estimated value for year *j* excluding year *i* from the *fAPAR*-yield correlation, MSE is the mean squared error of the model, and *p* is the number of fitted parameters.

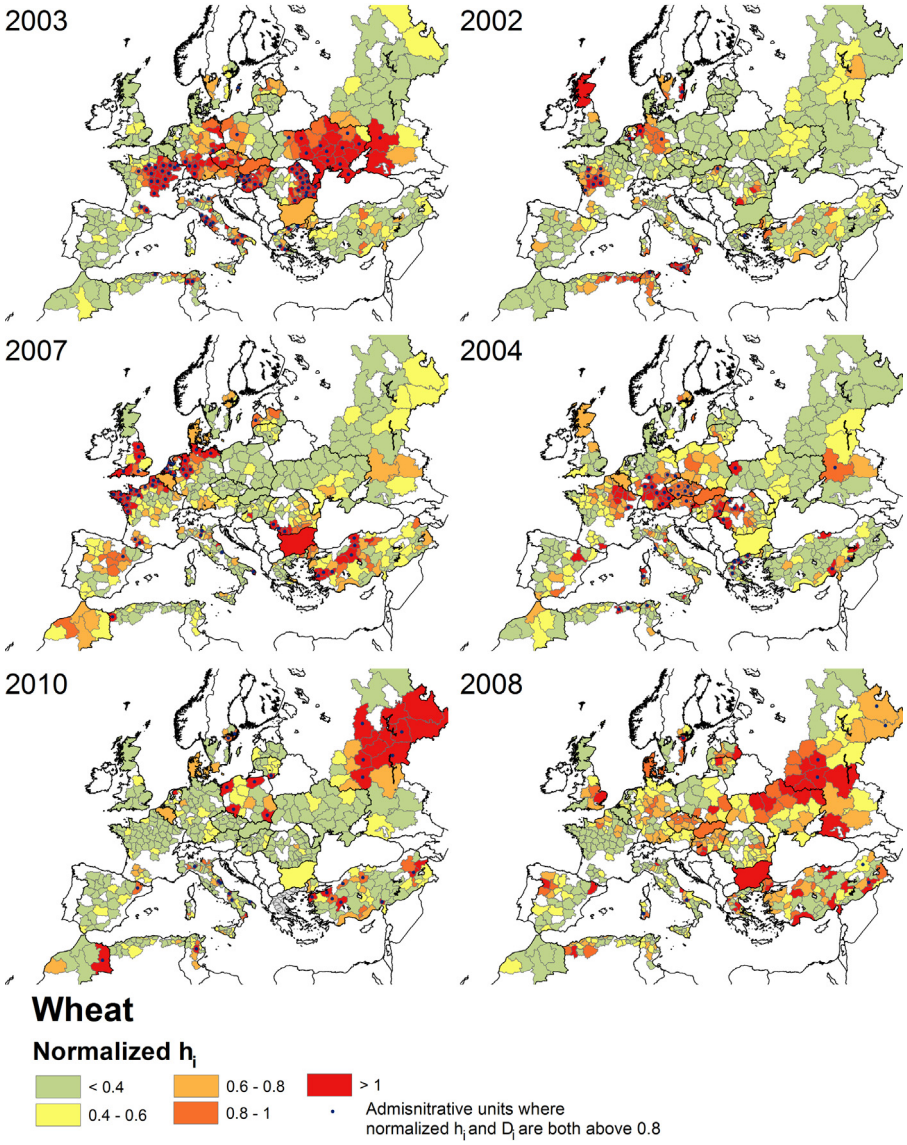

Fig. A1. Normalized leverage h_i of the optimal regression between $fAPAR_f$ and official wheat yields on years 2002–2004, 2007, 2008 and 2010. Dots in regions indicate normalized $h_i > 0.8$ and normalized $D_i > 0.8$. Maps on the left column refer to, in general, unfavourable years, whereas in the map on the right column refer to years with high yields.

Values for h_i and D_i are normalized using reference values, to make their interpretation easier. In the case of leverage, the reference threshold for normalization is $2p/n$ as proposed in Montgomery et al. (2001) where n stands for the number of observations included in the regression. Cook's distance is normalized by $4/n$ following the guidelines from Bollen and Jackman (1985).

Normalized h_i and D_i were calculated for every year on each individual *fAPAR*-yield model across Europe (Section 3.1). Looking at the spatial distribution of both indicators, a set of years where these values were close to 1 in a large number of regions of Europe was identified.

A.1.2. Analysis of meteorological data

The meteorological events leading to those extreme yields were investigated using observed weather parameters extracted from the MCYFS (see Section 2.4.3) on these specific years, including time series of temperature and precipitation, but also derived indicators such as sums of temperatures, or the fraction of days above/below a threshold temperature. Both their absolute values and their differences against the LTA were considered in the analysis.

An important factor to consider when analysing the possible link between a specific weather event and an extreme year is its spatial distribution: e.g. if the spatial context of a given meteorological event coincides with the regions where that year was considered as extreme.

A.2. Analysis of extreme years on fAPAR*-yield regressions for wheat*

The years 2002–2004, 2007, 2008 and 2010 were identified as extreme, according to the two statistics described in Section A.1, for the regions of Europe shown in Fig. A1. The regression models for nine major wheat producing regions in Europe, are shown on Fig. A2. At the continental level, 2003, 2007 and 2010 were identified as the less productive growing seasons of the study period for wheat and barley. In 2003 low yields were observed in most of the main producing regions in Europe: from France to the Black Sea area. In 2007, low yields were restricted to southern UK, northern France, Germany and some areas of the Black Sea region. In 2010 low yields were observed in southern Russian oblasts.

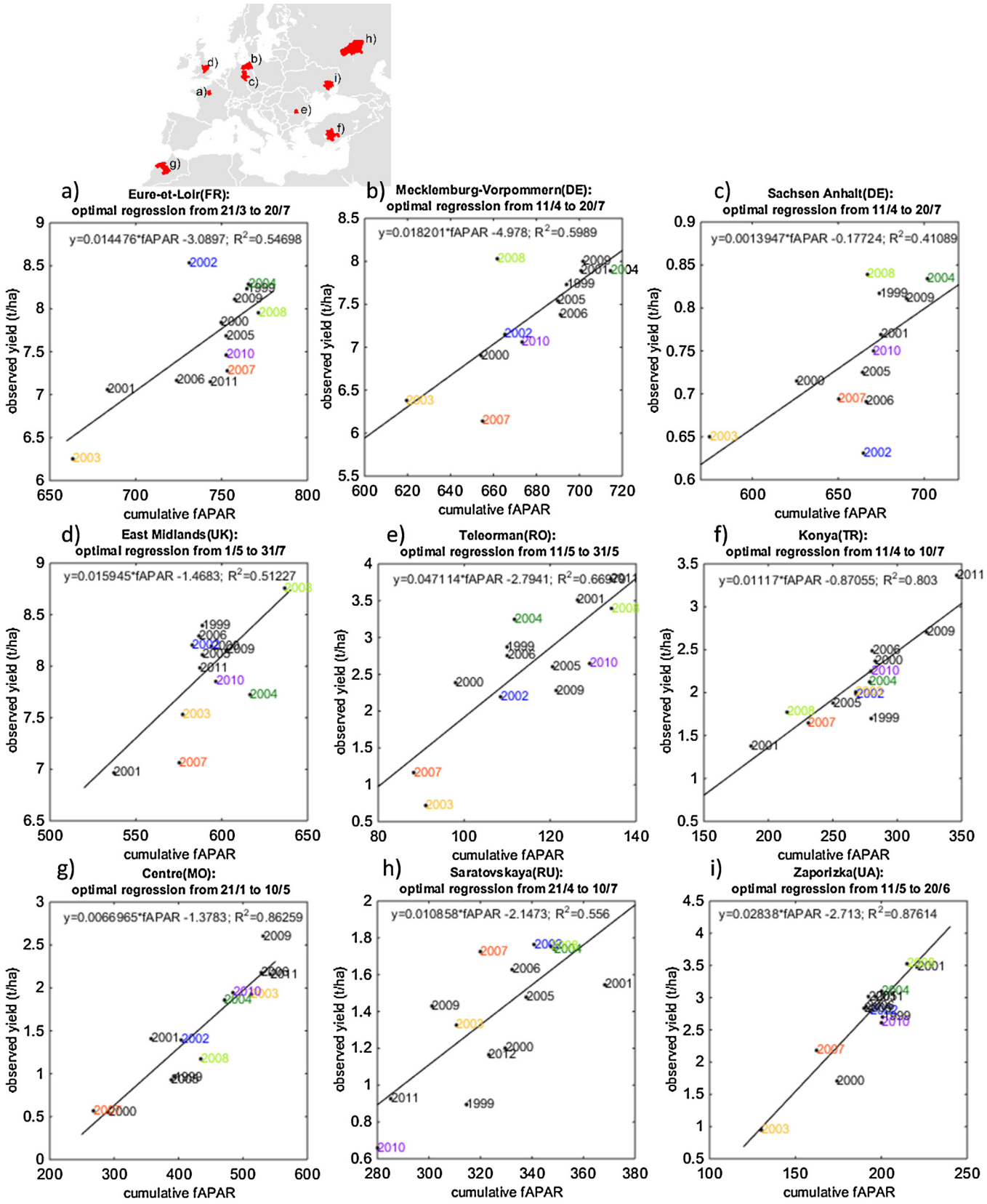

Fig. A2. Optimal $fAPAR_t$-yield regression models for wheat in nine major producing regions of Europe. Years 2002–2004, 2007, 2008 are displayed with different colours.

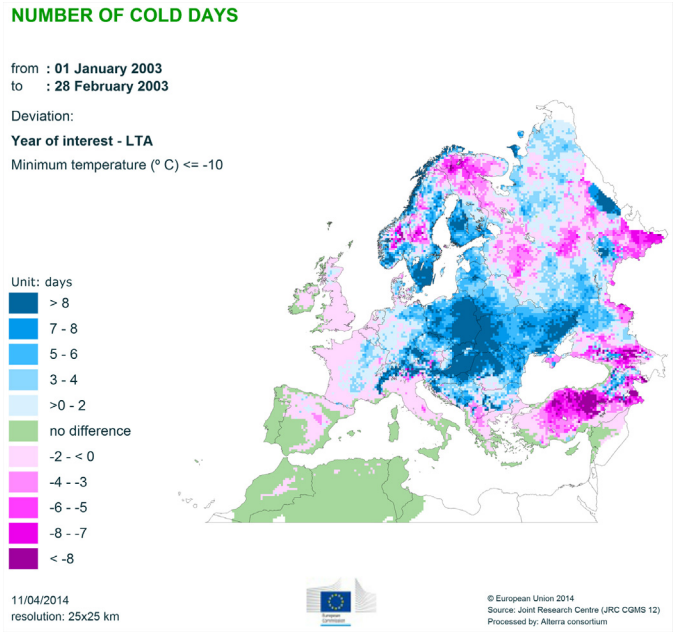

Fig. A3. Anomalies in the number of days with minimum temperature below −10 °C in January and February 2003. Extracted from the MARS crop yield forecasting system (MCYFS), see http://marswiki.jrc.ec.europa.eu for further details.

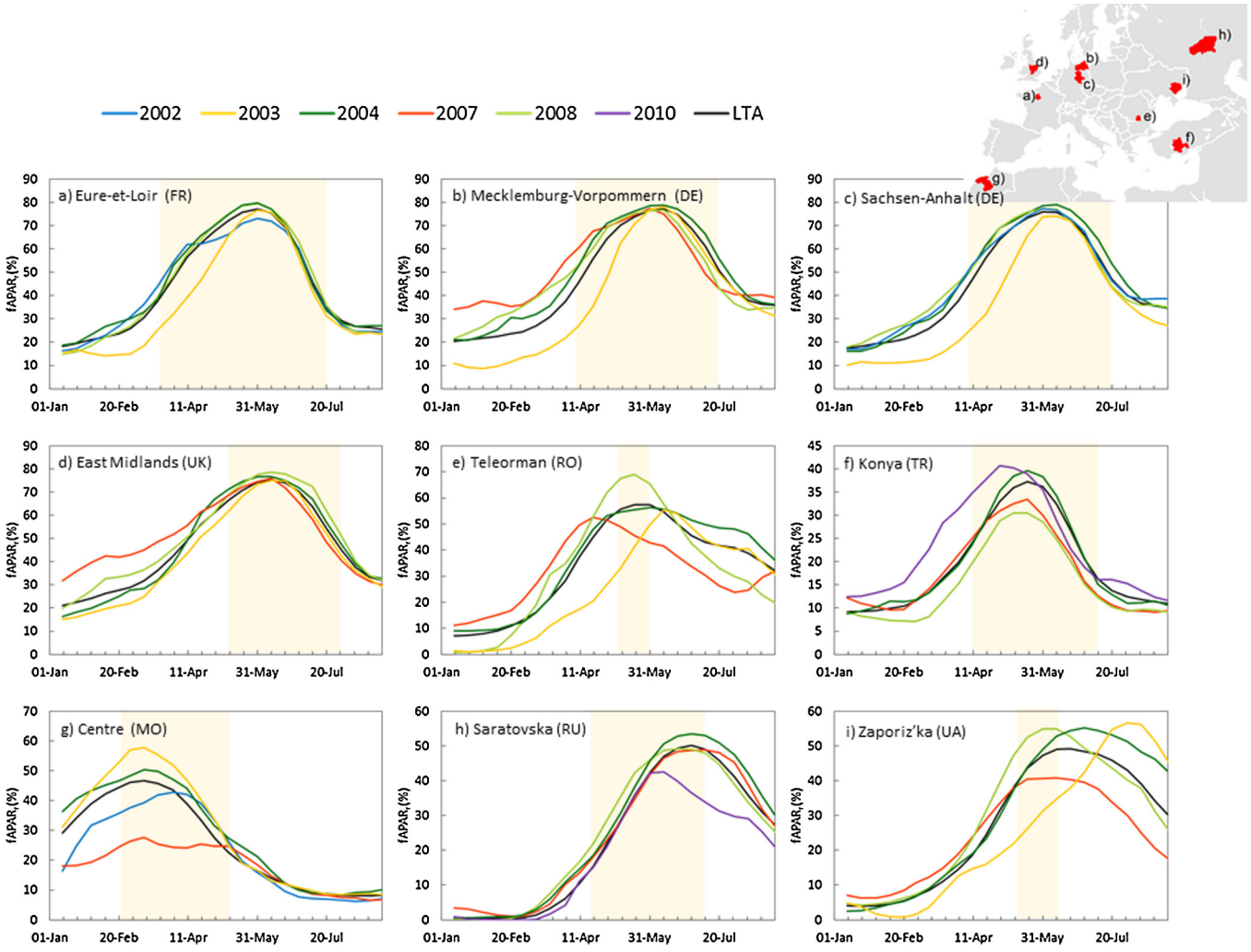

Fig. A4. *fAPAR*$_i$ time-series for the years 2002–2004, 2007, 2008, 2010, and long term average (LTA) for nine major wheat producing regions of Europe. Orange boxes indicate the temporal window used to accumulate *fAPAR* in the optimal regression for each of them. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

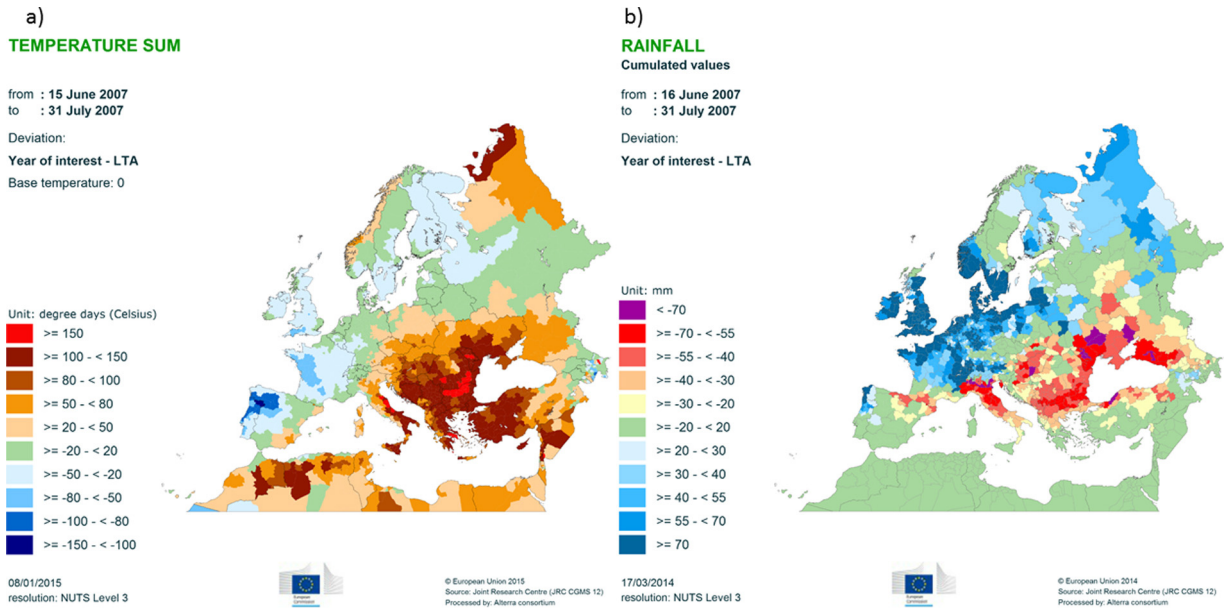

Fig. A5. (a) Anomalies in temperature sum from 15th June to 31st July 2007 in Europe; and (b) anomalies in precipitation for the same period. Values were computed at regional level from interpolated weather station observations, and extracted from the MARS crop yield forecasting system (MCYFS), see http://marswiki.jrc.ec.europa.eu for further details.

The years 2002, 2004, and 2008 were identified as those with unusually high yields in different regions of Europe. In 2002 there were exceptionally high yields in western France, in 2004 high yields were registered in most of the central European regions, and 2008 was very positive for central Europe, northern Spain (especially for barley), Ukraine and southern Russia.

A.2.1. Years with low yields: 2003, 2007 and 2010

The growing season in 2003 was affected by a very cold winter with unusually low minimum temperatures, affecting crop emergence in central, northern Europe and the Black Sea area. Fig. A3 shows the spatial distribution in Europe of this meteorological event, corresponding to the regions where 2003 was considered as extreme (Fig. A1). The ascending phase of the *fAPAR* time-series after winter shows a strong delay compared to the LTA in most of central and southeastern Europe (Fig. A4) as a consequence of these low temperatures. This was particularly the case in the Black Sea area, where the extremely low minimum daily temperatures produced the failure of winter crops during the second half of February, and fields had to be re-sown with spring varieties or even other summer crops (these effects can be observed in Fig. A4e and i). In most of the empirical models, the year 2003 is correctly estimated (Fig. A2) as the exceptionally low yields observed correspond to a negative anomaly on *fAPAR* time-series.

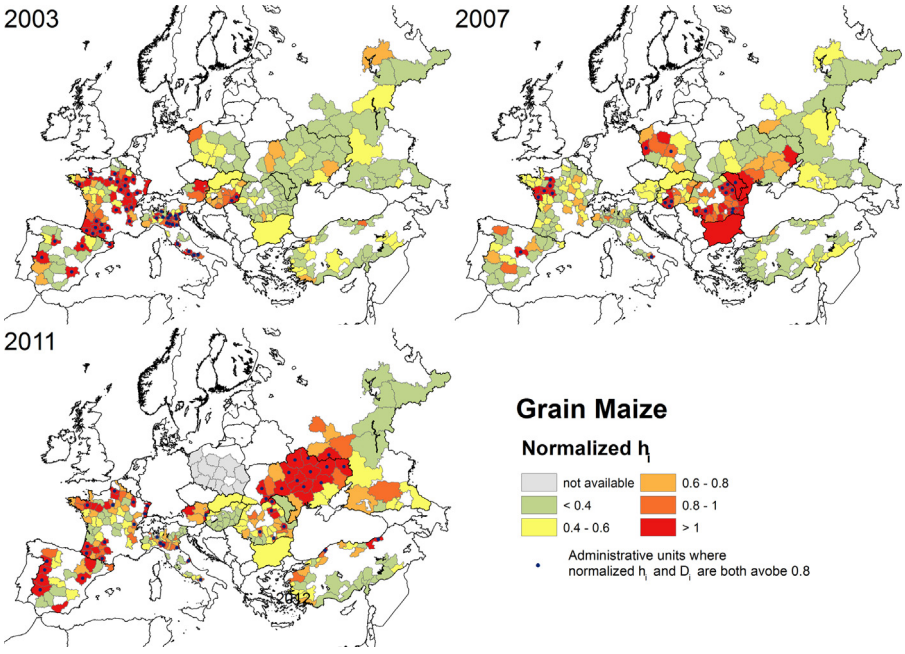

Fig. A6. Normalized leverage h_i of the optimal regression between $fAPAR_r$ and official grain maize yields on years 2003, 2007 and 2011. Dots in regions indicate normalized $h_i > 0.8$ and normalized $D_i > 0.8$.

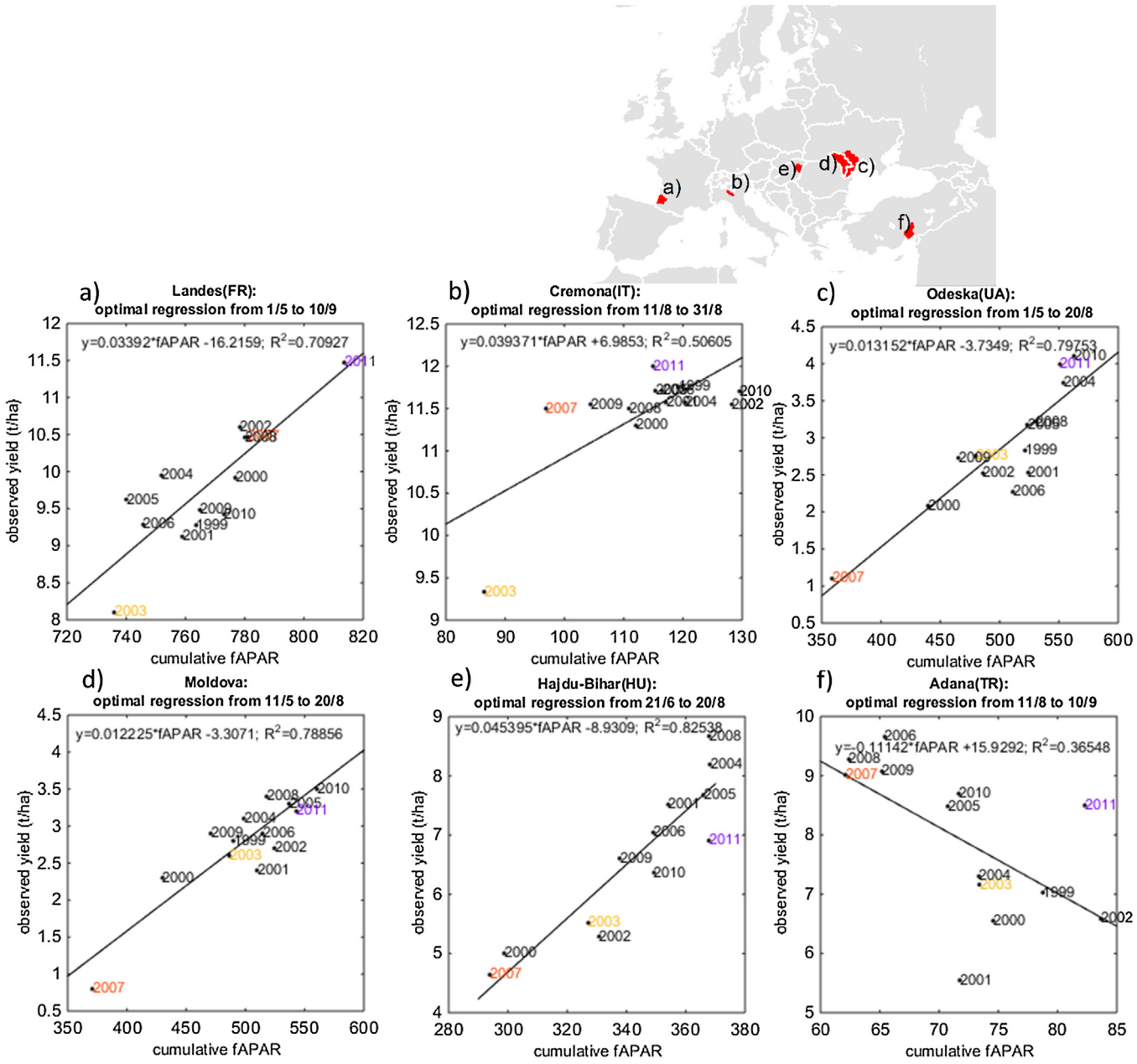

Fig. A7. Optimal $fAPAR_r$-yield regression models for maize in nine major producing regions of Europe. Years 2003, 2007 and 2011 are displayed with different colours.

In 2007 areas with low wheat and barley yields were observed in two agro-climatological groups: (1) northwestern Europe (northern coast of Germany and France, south of UK), and (2) most of the Black Sea area (Romania, Bulgaria, Turkey, Ukraine) and Morocco. The causes of low yields in these two groups of regions were completely different.

In northwestern Europe the 2007 growing season was characterized by a quick leaf area expansion at the beginning of the season with, in general, average to high *fAPAR* values (Fig. A4b and d). The low yields observed in Germany, France and the UK are due to the abundant rainfall in summer, during ripening-harvest stages (Fig. A5b). These meteorological conditions led to a drastic yield reduction (lack of sunshine, excess of moisture and specific weight reduction due to rewetting of grains), as reported by Agreste (2007) and Van der Velde et al. (2012). In all these regions, 2007 constitutes an outlier in the *fAPAR*-yield correlation (see Fig. A2b and d) that had a large influence in the empirical model (Fig. A1). This is due to the fact that the low observed yields were not a consequence of a reduction in crop leaf area, and therefore *fAPAR* time series were insufficient to predict these yields. In these agro-climatic areas the excess of rain during harvest, impacting yields negatively are not unusual. Another example of that is 2002 in some regions of central Germany (Fig. A2c), even though in most of Europe it was a favourable season.

By contrast, in the Black Sea area precipitation at the end of the spring of 2007 was scarce and temperatures unusually high (Fig. A5), leading to a strong water deficit. In Morocco the absence of precipitation from January to February produced the same adverse conditions. Water constraints negatively impacted leaf area expansion of winter crops, which is reflected in the *fAPAR* time-series (Fig. A4e–g and i). Note that normalized D_i or h_i are not as high as they were in year 2003. The main reason is that observed yields were not so low in 2007 compared to 2003, and its influence on the model is, therefore, moderate. Similarly, dry

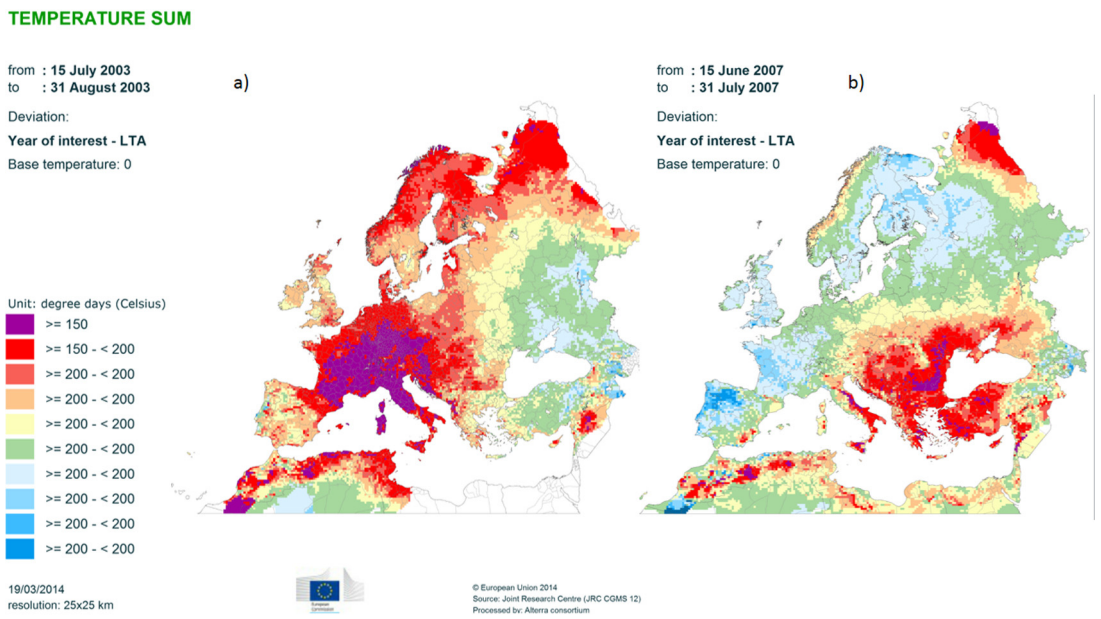

Fig. A8. Anomalies on the sum of temperatures (base temperature 0 °C) (a) from 15 July 2003 to 31 August 2003 and (b) from 15 June to 31 July 2007 in Europe, computed in 25 km grids from interpolated weather station observations. Data extracted from the MARS crop yield forecasting system (MCYFS), see http://marswiki.jrc.ec.europa.eu for further details.

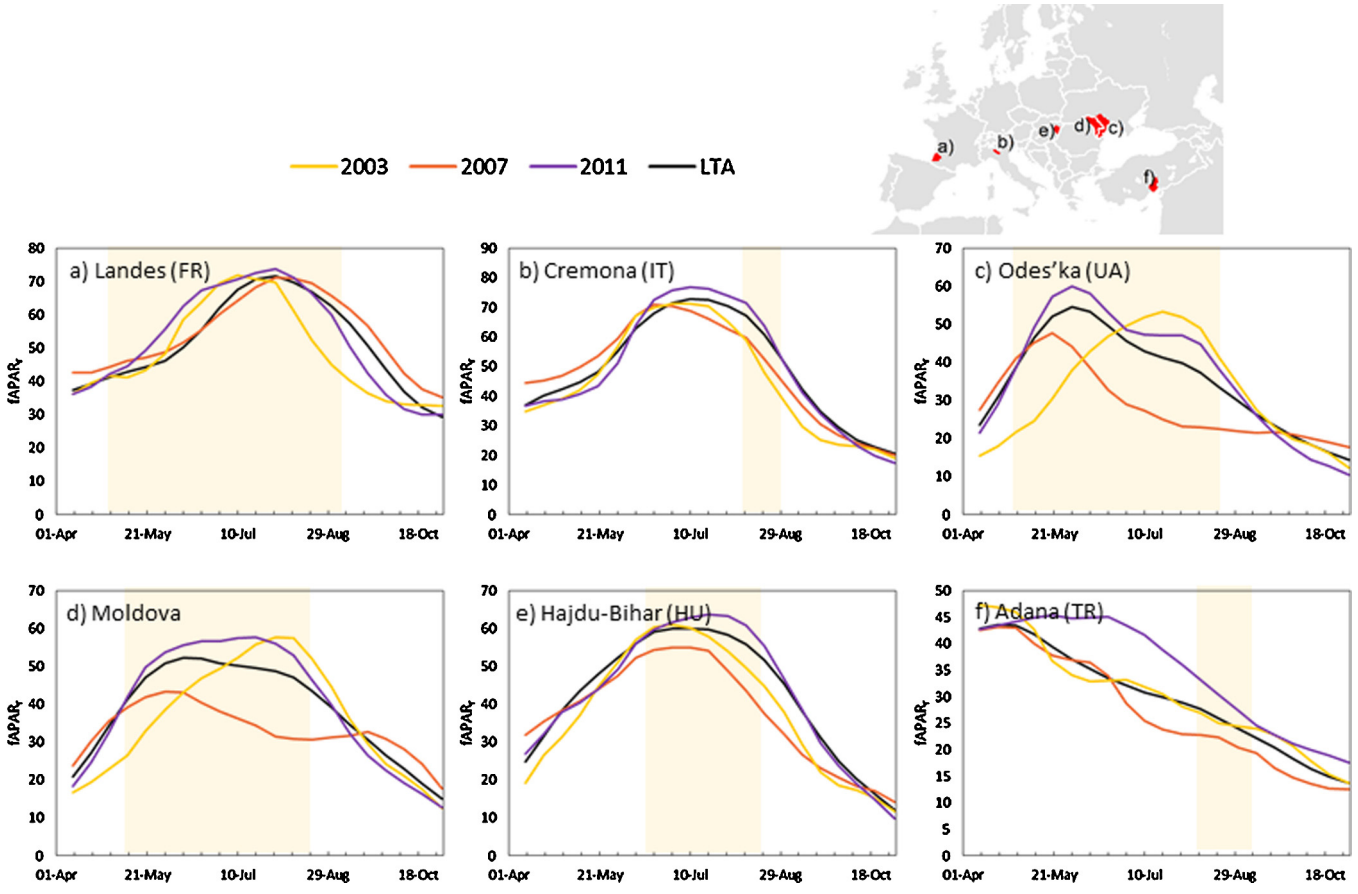

Fig. A9. *fAPAR$_f$* time-series of the years 2003, 2007, 2011, and the long term average (LTA) for six major maize producing regions of Europe. Orange boxes indicate the temporal window used to accumulate *fAPAR* in the optimal regression for each of them. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

and hot weather conditions on 2010 in southern Russia reduced drastically green leaf area during summer months as shown in Fig. A4h, which in turn explain low yields observed on 2010.

A.2.2. High yield years: 2002, 2004 and 2008

Years 2002, 2004 and 2008 were identified as the most productive ones in the studied period. Although their influence on the regression models varies spatially (Fig. A1), record yields were registered in these years across large areas of Europe. These years had favourable weather conditions during the growing season (especially rainfall), and in the critical stages of anthesis and grain filling.

The $fAPAR_r$ time-series of these years do not show a clear positive anomaly against LTA in northern European regions (Fig. A4a–e). For instance, it can be observed that in some cases $fAPAR_r$ time-series of years with exceptionally high yields are below the average (e.g. year 2002 in France, see Fig. A4a), or even comparable to the years with very low yields (Fig. A4b and c). In these humid agro-climatic areas, the inter-annual variability of $fAPAR_r$ seems quite low, and a link between high yields and high *fAPAR* values cannot be established.

By contrast, Fig. A4 shows that in the regions of southern Europe $fAPAR_r$ is usually above the LTA on years with high yields. In these agro-climatic areas, where rainfall determines yield its inter-annual variability is strongly connected to *fAPAR*.

A.3. Analysis of extreme years on the fAPAR*-yield regressions for maize*

Years 2003 and 2007 have been identified as extreme in the regressions between *fAPAR* and observed grain maize yields across Europe (Fig. A6): year 2003 was extreme for western Europe; whereas year 2007 had a strong impact on the Black Sea countries. In both years extremely low yields were observed (Fig. A7). Year 2011, with high yields registered across the main producing regions of Europe, was identified as an extreme season as well (Fig. A6).

A.3.1. Years with low yields: 2003 and 2007

In the summer of 2003 extremely high temperatures were registered in most of central Europe, including important grain maize producers such as Italy, France and, to a lesser extent, Hungary (Fig. A8a). That hot spell, in addition to the absence of rainfall, caused significant damages on grain maize yields and a drastic yield reduction (Van der Velde et al., 2010). Season 2003, indeed, highly influences the regression model between *fAPAR* and yields, especially in France and Italy (Fig. A7a and b). Fig. A9a and b shows the impact on $fAPAR_r$ time-series of this adverse weather event: a sharp decrease of green leaf area from mid-July in both countries.

Summer heat waves are quite frequent in the main EU producing regions of grain maize. Year 2007 is another example, where hot temperatures in June and July were registered in the Black Sea area (Fig. A8). These adverse weather conditions heavily constrained leaf area expansion during summer, in all of the southeast of Europe (Fig. A9), and especially in Ukraine, Romania and Moldova. In these countries observed yields were extremely low, with perhaps the exception of Hungary. The impact of 2007 on the *fAPAR*-yield regression models is high in these regions (Fig. A7), as it constitutes a high leverage point, responsible for the high R^2 that was found. Only in Turkey (Fig. A7f), where the correlation between *fAPAR* and yield is poor, has 2007 not had a significant influence in the empirical model. It should be noted that, in the case of Turkey, the existence of yield trends (see Section 3.3) and recent changes in crop distribution make it difficult to establish proper relationships between *fAPAR* and yield, which may result in artefacts (e.g. negative correlations) as in Fig. A7f. In northern Italy, although the hot and dry conditions experienced in 2007 seem to have had an impact on *fAPAR* (Fig. A9b) the observed yield was not reduced compared to the average.

A.3.2. High yield season of 2011

The summer of 2011 was particularly favourable for grain maize yield, as abundant precipitation were registered in most of the main producing regions, leading to high yields (Fig. A7), especially in France. The effects of these favourable weather conditions are appreciable in the high values of *fAPAR* in most of these regions (Fig. A9). However, the link between high *fAPAR* time-series and high yield is not evident, for example, in Turkey and the north of Italy, where maize is frequently irrigated.

References

Agreste, 2007. Bilan conjoncturel. Agreste Conjoncture 10–11.

Allen, R.G., Pereira, L.S., Raes, D., Smith, M., 1998. Crop Evapotranspiration Guidelines for Computing Crop Water Requirements: Guidelines for Computing Crop Water Requirements. FAO, Rome.

Balaghi, R., Tychon, B., Eerens, H., Jlibene, M., 2008. Empirical regression models using NDVI, rainfall and temperature data for the early prediction of wheat grain yields in Morocco. Int. J. Appl. Earth Obs. Geoinf. 10, 438–452, http://dx.doi.org/10.1016/j.jag.2006.12.001.

Baret, F., Hagolle, O., Geiger, B., Bicheron, P., Miras, B., Huc, M., Berthelot, B., Nino, F., Weiss, M., Samain, O., Roujean, J.L., Leroy, M., 2007. LAI, fAPAR and fCover CYCLOPES global products derived from VEGETATION: part 1: principles of the algorithm. Remote Sens. Environ. 110, 275–286.

Bartholomé, E., Belward, A.S., 2005. GLC: a new approach to global land cover mapping from Earth observation data. Int. J. Remote Sens. 26, 1959–1977, http://dx.doi.org/10.1080/01431160412331291297.

Baruth, B., Royer, A., Klisch, A., Genovese, G., 2008. The use of remote sensing within the MARS crop yield monitoring system of the european commission. ISPRS Arch. 37, 935–940.

Becker-Reshef, I., Vermote, E., Lindeman, M., Justice, C., 2010. A generalized regression-based model for forecasting winter wheat yields in Kansas and Ukraine using MODIS data. Remote Sens. Environ. 114, 1312–1323.

Bollen, K., Jackman, R., 1985. Regression diagnostics: an expository treatment of outliers and influential cases. Sociol. Methods Res. 13, 510–542.

Brisson, N., Gate, P., Gouache, D., Charmet, G., Oury, F.-X., Huard, F., 2010. Why are wheat yields stagnating in Europe? A comprehensive data analysis for France. Field Crops Res. 119, 201–212, http://dx.doi.org/10.1016/j.fcr.2010.07.012.

Buttner, G., Feranec, J., Jaffrain, G., Mari, L., Maucha, G., Soukup, T., 2004. The CORINE land cover project. EARSel E Proc. 3, 331–346.

Chloupek, O., Hrstkova, P., Schweigert, P., 2004. Yield and its stability, crop diversity, adaptability and response to climate change, weather and fertilisation over 75 years in the Czech Republic in comparison to some European countries. Field Crops Res. 85, 167–190, http://dx.doi.org/10.1016/S0378-4290(03)162-X.

De Bies, C.A.J., Gallego, F.J., Vrieling, A., Veldkamp, A., 2014. Integrating LUCAS, CORINE and SPOT NDVI-data to prepare detailed EU-Wide crop intensity maps Presented at the GEO-UA 2014, Kiev, Ukraine, May 2014.

de Wit, A., 2007. Regional Crop Yield Forecasting Using Probabilistic Crop Growth Modelling and Remote Sensing Data Assimilation, Environmental Sciences Group. Wageningen University, Wageningen, The Netherlands.

de Wit, A., Baruth, B., Boogaard, H., Van Diepen, K., Van Kraalingen, D., Micale, F., Te Roller, J., Supit, I., Van Den Wijngaart, R., 2010. Using ERA-INTERIM for regional crop yield forecasting in Europe. Clim. Res. 44, 41–53.

Doraiswamy, P.C., Moulin, S., Cook, P.W., Stern, A., 2003. Crop yield assessment from remote sensing. Photogramm. Eng. Remote Sens. 69, 665–674.

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

Dorigo, W.A., Zurita-Milla, R., de Wit, A.J.W., Brazile, J., Singh, R., Schaepman, M.E., 2007. A review on reflective remote sensing and data assimilation techniques for enhanced agroecosystem modeling. Int. J. Appl. Earth Obs. Geoinf. 9, 165–193.

Duveiller, G., 2012. Caveats in calculating crop specific pixel purity for agricultural monitoring using MODIS time series. Presented at the Proceedings of SPIE – The International Society for Optical Engineering, http://dx.doi.org/10.1117/12.974625.

Duveiller, G., Defourny, P., 2010. A conceptual framework to define the spatial resolution requirements for agricultural monitoring using remote sensing. Remote Sens. Environ. 114, 2637–2650.

Duveiller, G., López-Lozano, R., Baruth, B., 2013. Enhanced processing of 1-km spatial resolution fAPAR time series for sugarcane yield forecasting and monitoring. Remote Sens. 5, 1091–1116, http://dx.doi.org/10.3390/rs5031091.

Eurostat, 2010. Irrigation: Number of Farms, Areas and Equipment by Size of Irrigated Area and NUTS 2 Regions. Eurostat.

Eurostat, 2013. Crops Products – Annual Data. Eurostat.

Genovese, G., Vignolles, C., Nègre, T., Passera, G., 2001. A methodology for a combined use of normalised difference vegetation index and CORINE land

cover data for crop yield monitoring and forecasting. A case study on Spain. Agronomie 21, 91–111.
Gobron, N., Pinty, B., Verstraete, M.M., Govaerts, Y., 1997. A semidiscrete model for the scattering of light by vegetation. J. Geophys. Res. Atmos. 102, 9431–9446.
Gobron, N., Pinty, B., Verstraete, M.M., Taberner, M., 2002. VEGETATION – An optimized FAPAR algorithm – Theoretical Basis Document. EUR Report No. 20146 EN, Institute for Environment and Sustainability.
Gobron, N., Pinty, B., Verstraete, M.M., Widlowski, J.-L., 2000. Advanced vegetation indices optimized for up-coming sensors: design, performance, and applications. Remote Sens. IEEE Trans. Geosci. 38, 2489–2505, http://dx.doi.org/10.1109/36.885197.
Holben, B.N., 1986. Characteristics of maximum-value composite images from temporal AVHRR data. Int. J. Remote Sens. 7, 1417–1434, http://dx.doi.org/10.1080/01431168608948945.
Jacquemoud, S., Baret, F., 1990. PROSPECT: a model of leaf optical properties spectra. Remote Sens. Environ., http://dx.doi.org/10.1016/0034-4257(90)90100-Z.
Johnson, D.M., 2014. An assessment of pre- and within-season remotely sensed variables for forecasting corn and soybean yields in the United States. Remote Sens. Environ. 141, 116–128, http://dx.doi.org/10.1016/j.rse.2013.10.027.
Johnson, D.M., Mueller, R., 2010. The 2009 cropland data layer. Photogramm. Eng. Remote Sens. 76, 1201–1205.
Klisch, A., Royer, A., Lazar, C., Baruth, B., Genovese, G., 2007. Extraction of phenological parameters from temporally smoothed vegetation indices ISPRS Archives XXXVI-8/W48 Workshop Proceedings: Remote Sensing Support to Crop Yield Forecast and Area Estimates Stresa, Italy.
Kogan, F., Kussul, N., Adamenko, T., Skakun, S., Kravchenko, O., Kryvobok, O., Shelestov, A., Kolotii, A., Kussul, O., Lavrenyuk, A., 2013. Winter wheat yield forecasting in Ukraine based on earth observation, meteorologicaldata and biophysical models. Int. J. Appl. Earth Obs. Geoinf. 23, 192–203, http://dx.doi.org/10.1016/j.jag.2013.01.002.
Kowalik, W., Dabrowska-Zielinska, K., Meroni, M., Raczka, T.U., de Wit, A., 2014. Yield estimation using SPOT-VEGETATION products: a case study of wheat in European countries. Int. J. Appl. Earth Obs. Geoinf., http://dx.doi.org/10.1016/j.jag.2014.03.011.
Launay, M., Guerif, M., 2005. Assimilating remote sensing data into a crop model to improve predictive performance for spatial applications. Agric. Ecosyst. Environ. 111, 321–339.
Lazar, C., Genovese, G., 2004. Methodology of the MARS crop yield forecasting system. Agro-meteorological Modelling, Processing and Analysis, 2. Office for the Official Publications of the European Communities, Luxembourg.
Lioubimtseva, E., Henebry, G.M., 2012. Grain production trends in Russia, Ukraine and Kazakhstan: new opportunities in an increasingly unstable world? Front. Earth Sci. 6, 157–166, http://dx.doi.org/10.1007/s11707-012-0318-y.
Lissens, G., Kempeneers, P., Fierens, F., 2000. Development of a Cloud, Snow and Cloud Shadow Mask for VEGETATION Imagery. EC-JRC, Belgirate, Italy, pp. 303–306.
Lobell, D.B., Asner, G.P., Ortiz-Monasterio, J.I., Benning, T.L., 2003. Remote sensing of regional crop production in the Yaqui Valley, Mexico: estimates and uncertainties. Agric. Ecosyst. Environ. 94, 205–220, http://dx.doi.org/10.1016/S0167-8809(02) 21-X.
Lobell, D.B., Schlenker, W., Costa-Roberts, J., 2011. Climate trends and global crop production since 1980. Science 333, 616–620, http://dx.doi.org/10.1126/science.1204531.
Lopez-Bellido, L., 2009. Cuestiones Referentes Al Sector De Herbáceos, Más Relevantes Para la Definición de la Política de Seguros Agrarios: Situación Actual y Tendencias a Corto y Medio Plazo. Universidad de Cordoba.
Meroni, M., Marinho, E., Sghaier, N., Verstrate, M.M., Leo, O., 2013. Remote sensing based yield estimation in a stochastic framework – case study of durum wheat in tunisia. Remote Sens. 5, 539–557, http://dx.doi.org/10.3390/rs5020539.
Micale, F., Genovese, G., 2004. Methodology of the MARS crop yield forecasting system. Meteorological Data Collection, Processing and Analysis, 1. Office for the Official Publications of the European Communities, Luxembourg.
Montgomery, D.C., Peck, E.A., Vining, G.G., 2001. Introduction to Linear Regression Analysis. Wiley.
Mueller, N.D., Gerber, J.S., Johnston, M., Ray, D.K., Ramankutty, N., Foley, J.A., 2012. Closing yield gaps through nutrient and water management. Nature 490, 254–257, http://dx.doi.org/10.1038/nature11420.
Mueller, R., Seffrin, R., 2006. New methods and satellites: a program update on the NASS cropland data layer acreage program. Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. 36, 97–102.
Myneni, R.B., Hoffman, S., Knyazikhin, Y., Privette, J.L., Glassy, J., Tian, Y., Wang, Y., Song, X., Zhang, Y., Smith, G.R., Lotsch, A., Friedl, M., Morisette, J.T., Votava, P., Nemani, R.R., Running, S.W., 2002. Global products of vegetation leaf area and fraction absorbed PAR from year one of MODIS data. Remote Sens. Environ. 83, 214–231.
Newlands, N.K., Zamar, D.S., Kouadio, L.A., Zhang, Y., Chipanshi, A., Potgieter, A., Toure, S., Hill, H.S.J., 2014. An integrated, probabilistic model for improved seasonal forecasting of agricultural crop yield under environmental uncertainty. Front. Environ. Sci. 2, 17, http://dx.doi.org/10.3389/fenvs.2014.00017.
Prasad, A.K., Chai, L., Singh, R.P., Kafatos, M., 2006. Crop yield estimation model for Iowa using remote sensing and surface parameters. Int. J. Appl. Earth Obs. Geoinf. 8, 26–33, http://dx.doi.org/10.1016/j.jag.2005.06.002.
Richards, R., Townley-Smith, T., 1987. Variation in leaf area development and its effect on water use: yield and harvest index of droughted wheat. Aust. J. Agric. Res. 38, 983–992.
Rouse, J., Haas, R., Schell, J., 1974. Monitoring the Vernal Advancement and Retrogradation (Greenwave Effect) of Natural Vegetation. Texas A and M University, College Station.
Royer, A., Genovese, G., 2004. Methodology of the MARS crop yield forecasting system. Remote Sensing Information, Data processing and Analysis, 3. Office for the Official Publications of the European Communities, Luxembourg.
Slafer, G.A., Rawson, H.M., 1994. Sensitivity of wheat phasic development to major environmental factors: a re-examination of some assumptions made by physiologists and modellers. Aust. J. Plant Physiol. 21, 393–426.
Stöckle, C.O., Donatelli, M., Nelson, R., 2003. CropSyst: a cropping systems simulation model. Eur. J. Agron. 18, 289–307.
Swets, D.L., Reed, B.C., Rowland, J.D., Marko, S.E., 1999. A Weighted Least-squares Approach to Temporal NDVI Smoothing. ASPRS, Portland, Oregon.
Ugarte, C., Calderini, D.F., Slafer, G.A., 2007. Grain weight and grain number responsiveness to pre-anthesis temperature in wheat, barley and triticale. Field Crops Res. 100, 240–248, http://dx.doi.org/10.1016/j.fcr.2006.07.010.
Van der Velde, M., Tubiello, F.N., Vrieling, A., Bouraoui, F., 2012. Impacts of extreme weather on wheat and maize in France: evaluating regional crop simulations against observed data. Clim. Change 113, 751–765, http://dx.doi.org/10.1007/s10584-011-0368-2.
Van der Velde, M., Wriedt, G., Bouraoui, F., 2010. Estimating irrigation use and effects on maize yield during the heatwave in France. Agric. Ecosyst. Environ. 135, 90–97, http://dx.doi.org/10.1016/j.agee.2009.08.017.
Van Diepen, C.A., Wolf, J., Van Keulen, H., Rappoldt, C., 1989. WOFOST: a simulation model of crop production. Soil Use Manage. 5, 16–24.
Van Leeuwen, W., Hutchinson, C., Drake, S., Doorn, B., Kaupp, V., Haithcoat, T., Likholetov, V., Sheffner, E., Tralli, D., 2011. Benchmarking enhancements to a decision support system for global crop production assessments. Exp. Syst. Appl. 38, 8054–8065, http://dx.doi.org/10.1016/j.eswa.2010.12.145.
Vermote, E.F., Tanré, D., Deuzé, J.L., Herman, M., Morcrette, J.-J., 1997. Second simulation of the satellite signal in the solar spectrum, 6s: an overview. IEEE Trans. Geosci. Remote Sens. 35, 675–686.
Wall, L., Larocque, D., Léger, P.-M., 2008. The early explanatory power of NDVI in crop yield modelling. Int. J. Remote Sens. 29, 2211–2225.
Weissteiner, C.J., Kühbauch, W., 2005. Regional yield forecasts of malting barley (*Hordeum vulgare* L.) by NOAA-AVHRR remote sensing data and ancillary data. J. Agron. Crop Sci. 191, 308–320, http://dx.doi.org/10.1111/j.1439-037X2005.00154.x.
Wu, B., Meng, J., Li, Q., Yan, N., Du, X., Zhang, M., 2014. Remote sensing-based global crop monitoring: experiences with China's cropwatch system. Int. J. Digital Earth 7, 113–137, http://dx.doi.org/10.1080/17538947.2013.821185.