Universitat Jaume I of Castellón, Spain


http://geotec.uji.es


@geotecUJI


https://www.linkedin.com/company/geotec-geospatial-technologies-research-group---universitat-jaume-i

https://www.youtube.com/user/geotecUJI

A high resolution, integrated system for rice yield forecast at district level

Valentina Pagani[1,*], Tommaso Guarneri[2], Lorenzo Busetto[3], Luigi Ranghetti[3], Mirco Boschetti[3], Ermes Movedi[2], Manuel Campos-Taberner[4], Francisco Javier Garcia-Haro[4], Dimitrios Katsantonis[5], Dimitris Stavrakoudis[6], Elisabetta Ricciardelli[7], Filomena Romano[7], Francesco Holecz[8], Francesco Collivignarelli[8], Carlos Granell[9], Sven Casteleyn[9], Roberto Confalonieri[1,*]

[1]Università degli Studi di Milano, DEMM, Cassandra lab, via Celoria 2, 20133 Milan, Italy

[2]Università degli Studi di Milano, DISAA, Cassandra lab, via Celoria 2, 20133 Milan, Italy

[3]Institute for Electromagnetic Sensing of the Environment, Italian National Research Council, Via Bassini 15, 20133 Milan, Italy

[4]Department of Earth Physics and Thermodynamics, Faculty of Physics, Universitat de València, Dr. Moliner, Burjassot 46100, València, Spain

[5]Hellenic Agricultural Organization, Plant Breeding and Genetic Resources Institute, Thermi-Thessalonikis, Ellinikis Georgikis Scholis, Greece

[6]Laboratory of Forest Management and Remote Sensing, School of Forestry and Natural Environment, Aristotle University of Thessaloniki, Thessaloniki 54124, Greece

[7]Institute of Methodologies for Environmental Analysis, Italian National Research Council, C.da S. Loja, Zona Industriale, 85050 Tito Scalo (PZ), Italy

[8]SARMAP, Cascine di Barico 10, 6989 Purasca, Switzerland

[9]Geospatial Technologies Research Group, Institute of New Imaging Technologies, Universitat Jaume I, Avda. Sos Baynat, s/n, 12071 Castellón, Spain

* Corresponding Author, Tel. +39 02 50316515, E-mail: valentina.pagani@unimi.it, roberto.confalonieri@unimi.it Fax: +39 02 50316578

Abstract

To meet the growing demand from public and private stakeholders for early yield estimates, a high-resolution (2 km × 2 km) rice yield forecasting system based on the integration between the WARM model and remote sensing (RS) technologies was developed. RS was used to identify rice-cropped area and to derive spatially distributed sowing dates, as well as for the dynamic assimilation of RS-derived leaf area index (LAI) data within the crop model. The system – tested for the main European rice production districts in Italy, Greece and Spain – performed satisfactorily: more than 66% of inter-annual yield variability was explained in six out of eight combinations of ecotype × district, with a maximum of 89% of variability explained for Tropical Japonica cultivars in the Vercelli district (Italy). In seven out of eight cases, the assimilation of RS-derived LAI allowed improving the forecasting capability, with minor differences due to the assimilation technology used (updating or recalibration). In particular, RS allowed reducing the uncertainty by capturing factors not properly reproduced by the simulation model (given the uncertainty due to large-area simulations). As an example, the season 2003 in the Serres (Greece) district was characterized by severe blast epidemics, whose effect on canopy vigor was captured by RS-derived LAI products. The system – extending the one used for rice within the EC-JRC-MARS forecasting system – was pre-operationally used in 2015 and 2016 to provide early yield estimates to private companies and institutional stakeholders within the EU-FP7 ERMES project.

Keywords

Assimilation; blast disease; *Oryza sativa* L.; remote sensing; WARM model.

1. Introduction

There is an increasing demand for systems able to provide timely yield forecasts, given the potential interest for a variety of actors within the agricultural sector, including private companies and institutional stakeholders (e.g., Supit, 1997; Bannayan and Crout, 1999; Wang et al., 2010; Fang et al., 2011). While industries and private companies are interested in crop yield forecasts for reasons such

as the need of defining selling strategies or planning milling operations (Everingham et al., 2002), the interest of public institutions deals with the need of regulating agricultural markets and mitigating volatility of prices in case of speculative actions on food commodities (e.g., OECD and FAO, 2011; Kogan et al., 2013). Simple forecasting systems – based, e.g., on agroclimatic indicators (Balaghi et al., 2012) – demonstrated their usefulness under conditions characterized by large year-to-year fluctuations in yields and when those fluctuations are driven by one or two key factors. Other approaches are more complex, relying on remote sensing (Mkhabela et al., 2005; Wang et al., 2010; Duveiller et al., 2013; Son et al., 2014) or crop simulation models (Vossen and Rijks, 1995; Supit, 1997; Bezuidenhout and Singels, 2007a/b; de Wit et al., 2010; Kogan et al., 2013; Pagani et al., 2017). Crop models are indeed able to interpret reality, e.g., by capturing the effects of weather anomalies or other factors affecting crop yields better than simpler systems. As an example, they are able to simulate the effect of thermal shocks-induced spikelet sterility (Shimono et al., 2005). A forecasting system solely based on remote sensing would fail in contexts where sterility is an issue, since sterility can severely affect yields even without any damage to the canopy. Forecasting systems solely based on remote sensing are unsuitable also in contexts characterized by a good yield potential because of signal saturation (Sader et al., 1989; Dobson et al., 1995; Zhao et al., 2016). However, crop models are demanding in terms of data needs and, when applied on large areas, they can be affected by many sources of uncertainty, due to the poor quality of input data (weather, soil), to the lack of information on management (e.g., sowing dates, irrigation practices, cultivars/hybrids grown), as well as to the model structure (Sándor et al., 2016; Confalonieri et al., 2016a), to the experience of users (Diekkrüger et al., 1995; Confalonieri et al., 2016b), and to the uncertainty in the data used for their calibration (Kersebaum et al., 2015; Confalonieri et al., 2016c).

The availability of powerful platforms for gridded model runs and for automatic calibration of parameters, as well as the availability of consistent archives (e.g., leaf area index estimates for the period 2000 - 2016 from ESA - Copernicus or NASA – MODIS) and new generations (e.g., Sentinel satellites from the Copernicus program) of remote sensing products (Lefebvre et al., 2016), is increasing the potential of forecasting systems integrating crop models and remote sensing

technologies. The combined use of these two kinds of technology can indeed markedly reduce their intrinsic limits, because the potentialities of both technologies can be increased by their integration (Fang et al., 2011; Ines et al., 2013; Ma et al., 2013). As an example, the uncertainty in the sowing dates provided to crop models can be reduced through the analysis of temporal profile or remote sensing products (e.g., Boschetti et al., 2009). Moreover, in case of differences in vigour among varieties or of factors not accounted for by simulation models (e.g., insects, Wu and Wilson, 1997; weeds, Kropff et al., 1992), they can be implicitly included in the simulation via the assimilation of remote sensing-derived leaf area index data varying in time and space (Launay and Guèrif, 2005; Dorigo et al., 2007).

Two main strategies are available to integrate remote sensing information into crop simulators, each presenting pros and cons for different species and agroclimatic/operational contexts (Dorigo et al., 2007): recalibration and updating. The recalibration method is based on the automatic adjustment of model parameters targeting the minimization of the error between model outputs and remote sensing-derived state variables (e.g., Bouman et al., 1995). The updating method is instead based on the update of model state variables when the remote sensing data are available, using algorithms to convert them into simulated variables and to redefine all model outputs accordingly (McLaughlin, 2002). The latter is easy to implement and does not increase the computational time but penalizes the internal coherence of the simulation because it generate discontinuities in state variables (Bouman et al., 1995; Zhao et al., 2013; Jin et al., 2018). Recalibration instead – besides the higher complexity and requirements in terms of computational time/power – exposes to the risk of degrading the quality of parameter sets because at each assimilation event and for each spatial simulation unit, few RS-derived LAI values are used to change parameter values that were developed using many data (usually for different state variables) from dedicated multi-site and multi-year experiments (Dorigo et al., 2007).

The aim of this study was to develop and test a high-resolution rice yield forecasting system targeting the main European rice districts based on different techniques (recalibration, updating) to assimilate remote sensing information in crops models. This was done in the framework of the EU-FP7 ERMES project, whose aim was developing services and disseminating added-value information for the rice

sector (www.ermes-fp7space.eu). The system is an evolution of the one used for rice within the European Commission Joint Research Centre MARS (EC-JRC-MARS) activities.

2. Materials and methods

2.1. Study areas

The system was developed targeting Italy, Spain and Greece, which are responsible of 52% (1386100 t, 219500 ha), 25% (861103 t, 110419 ha) and 7% (229900 t, 30720 ha) of the total European rice production, respectively (FAOSTAT, 2014). The rice production districts selected for each of the three countries are shown in Fig. 1. For the Italian district "Lombardo-Piemontese", we considered the province of Vercelli and the area of Lomellina (located in the Pavia province), respectively including 31% and 27% of the Italian rice area (National Rice Authority [Ente Nazionale Risi]; www.enterisi.it). For Spain, the system included the areas of the Ebro delta and of the "Parc natural de l'Albufera", with the two rice districts located in the provinces of Tarragona and Valencia representing about 30% of the Spanish rice production. In Greece, the Central Macedonia region was selected, with rice districts located around Thessaloniki (the main Greek producing site), and Serres. According to the Koppen climate classification, Spanish and Greek areas are characterized by a Mediterranean climate, with hot and dry summers, whereas the climate in the Italian district is temperate with warm and humid conditions during the summer months. In general, the rice season in the three countries starts in March/April and ends in September/October, even though the length of the cycle strictly depends on the cultivated variety and on the seasonal weather conditions. The most common water management adopted in the three countries is based on continuous flooding; however, dry sowing (usually coupled with delayed flooding at the 3^{rd}-5^{th} leaf stage) is increasingly used in Italy, with 38.2% of the total rice surface in 2015 (last year for which official data are available) and an average value of 27.7% for the previous five years. According to data made available by Italian National Rice Authority, Japonica varieties belonging to the market category "Lungo A" are the most cultivated in Italy, followed by Japonica varieties "Tondo" and the Tropical Japonica varieties "Lungo B", with the latter – especially grown in the province of Vercelli – representing about 15% of the national rice production. In Spain

and Greece, the most cultivated varieties belong to Japonica and Tropical Japonica groups, respectively, with the exception of the Greek district of Serres, where the colder climate is more suitable for the cultivation of Japonica varieties.

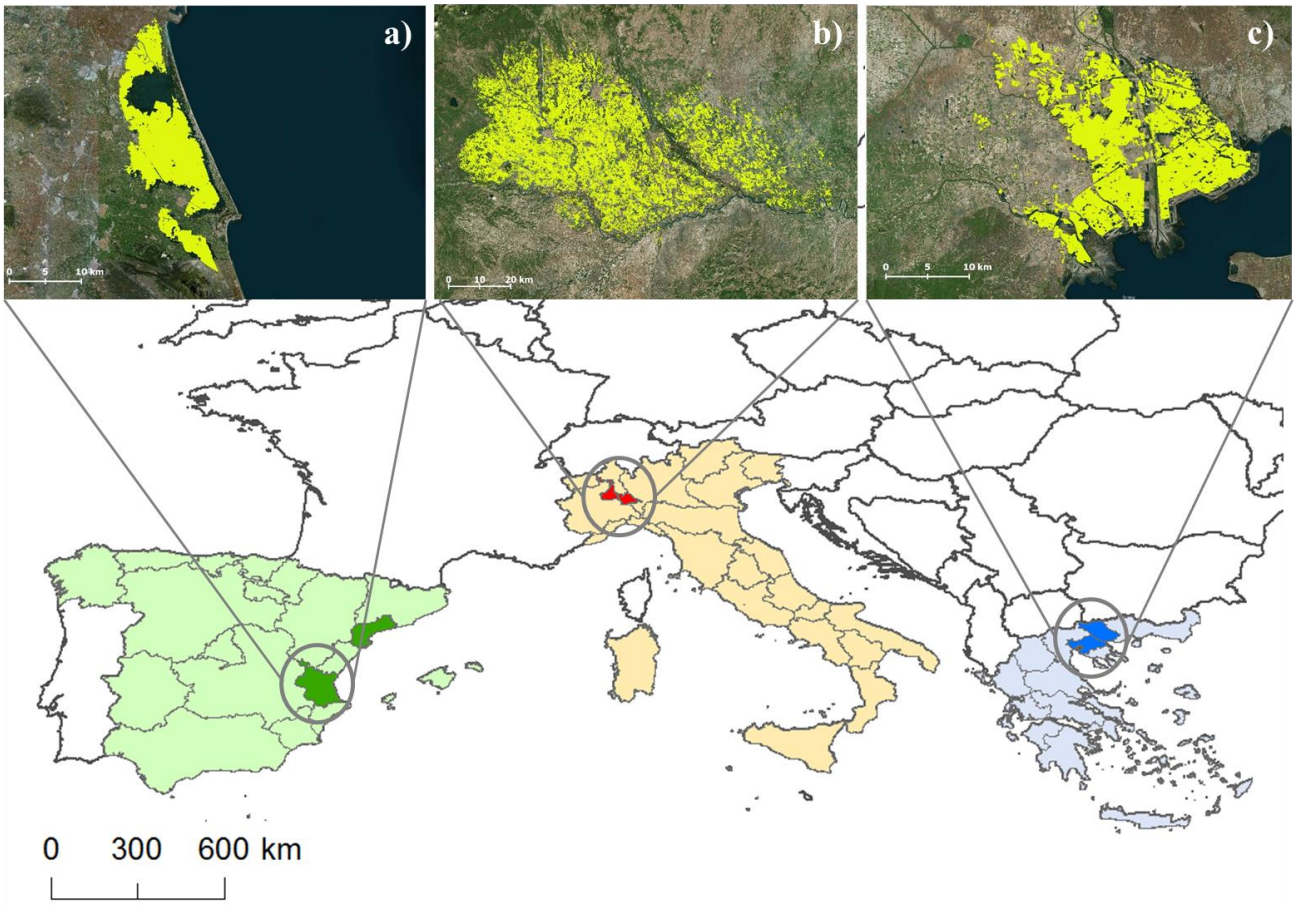


Figure 1*: rice production districts and related 2016 rice distribution maps for the a) Spanish (Valencia district), b) Italian (Lomellina and Vercelli districts) and c) Greek (Thessaloniki district) study areas. Source: EU-FP7 ERMES; Campos-Taberner et al. (2017).*

2.2. Crop model and assimilation tool

This study was carried out using the rice-specific model WARM (e.g., Confalonieri et al., 2009a; Pagani et al., 2014). The model has been used in both research (e.g., Paleari et al., 2017) and operational contexts, e.g., it is one of the rice models used within the AgMIP project (Agricultural Model Intercomparison and Improvement Project; Rosenzweig et al., 2013) and it is the model used by the EC-JRC for rice yield forecasts in Europe; http://ies-webarchive-ext.jrc.it/mars/mars/About-us/AGRI4CAST/Models-Software-Tools/Crop-Growth-Modelling-System-CGMS.html. WARM was selected among other rice models for both scientific and practical reasons. Given it was originally developed for being used within operational contexts at regional scale, it is parsimonious in terms of

data needs compared to other approaches (Confalonieri et al., 2009b). Nevertheless, it reproduces the effect of key factors affecting yield fluctuations in Europe, such as blast disease and spikelet sterility due to pre-flowering thermal shocks. Moreover, it demonstrated its accuracy in reproducing rice growth and development under the conditions explored by the crop in the study areas (Confalonieri et al., 2009a) and parameter sets were available for different groups of popular European varieties because of previous studies. Another reason was the availability of an integrated automatic optimization tool that allowed assimilating remote sensing-derived leaf area index (LAI) data based on recalibration techniques.

WARM estimates biomass accumulation based on radiation use efficiency, which is modulated according to temperature limitation, saturation of enzymatic chains, senescence, sterility, and diseases. Daily accumulated aboveground biomass (estimated with daily or hourly time step) is partitioned to plant organs using a set of beta (panicles) and parabolic (leaves) functions of a SUCROS-type development stage code driven by a single parameter (partitioning to leaves at emergence). Partitioning to stems is then equal to 1 minus the coefficients for leaves and panicles. Green leaf area index increase is derived from daily increase in leaf biomass and a development stage-dependent specific leaf area. Leaf senescence is simulated when daily-emitted leaf area units reach a thermal time threshold. Spikelet sterility due to cold shocks around young microspore stage and at flowering, as well as due to heat stress at flowering, is calculated by weighing hourly stresses by development-dependent bell-shape functions to reproduce the between- and within- plant heterogeneity in development. Concerning leaf and neck blast, disease onset is estimated based on hydrothermal time (Arai and Yoshino, 1987; Kim, 2000), whereas the daily infection efficiency is computed according to Magarey et al. (2005). Duration of the phases of latency, incubation and infectious is based on hourly air temperature. Leaf area affected by blast lesions is estimated using a compartmental susceptible-infected-removed model (Bregaglio et al., 2016). Effects of neck blast are reproduced by reducing the fraction of assimilates partitioned to panicles (Bregaglio et al., 2016). For both spikelet sterility and diseases, hourly weather data are generated from daily inputs.

176 The assimilation of remote sensing information was carried out using recalibration and updating
177 techniques. For the former, the optimization method used was a multi-start and bounded (for
178 parameter ranges) version of the downhill simplex (Nelder and Mead, 1965). The simplex has N+1
179 vertices interconnected by line segments and polygon faces in an N-dimensional parameter
180 hyperspace, and it moves through this space according to three basic rules: reflection, contraction, and
181 expansion. Although other optimization methods not using derivatives are available (e.g., Kirkpatrick
182 et al., 1983; Glover, 1986), the simplex guarantees a very favourable ratio between performance and
183 complexity (Matsumoto et al., 2002; Press et al., 2007). The updating method was instead
184 implemented (i) by deriving leaf biomass from RS-derived LAI and simulated specific leaf area and
185 (ii) by using the relationships among the relative weights of different plant organs at the previous time
186 step to update all plant-related state variables.
187 For both the assimilation methods, the procedure was triggered only (i) before flowering (to avoid
188 uncertainty due to green or senescent leaf area), (ii) in case at least three exogenous data were
189 available for each elementary simulation unit, and (iii) for leaf area index data within a biophysical
190 range for rice in each specific phenological stage. Concerning the recalibration method, parameters
191 whose values were optimized were specific leaf area at emergence and at mid-tillering and radiation
192 use efficiency.
193

194 ## 2.3. Input data

195 Dedicated processing chains were developed to produce the near real-time weather data and RS-
196 derived information used to feed the WARM simulation model. The adopted spatial resolution
197 (defining the size of the elementary simulation unit) was 2 km × 2 km, considered as a good
198 compromise between the high resolution of remote sensing data and the lower resolution of the
199 weather database (which was downscaled; details in the following sections). Despite it would have
200 been possible to further downscale weather data, the uncertainty in the information on management
201 practices and the need of developing an operational system potentially extendible to other rice districts
202 suggested to avoid further increases in the spatial resolution. Indeed, given other sources of

uncertainty (including those related with an extreme downscaling of weather data), further reductions in the size of the elementary simulation unit would have only increased the computational power needed to run the system, without increasing the quality of the information provided.

2.3.1. Weather data

An archive (continuously updated for near real-time simulations) of weather data at 2 km × 2 km spatial resolution was created starting from 1st of January 2003 to provide daily input to the WARM model. Source data were derived from the European Centre for Medium-Range Weather Forecast ERA-Interim (for the historical series) and TIGGE (for near real-time) databases (ECMWF; www.ecmwf.int; de Wit et al., 2010) for the following daily variables: maximum and minimum air temperatures, maximum and minimum air relative humidity, rainfall, average wind speed, global solar radiation. Leaf wetness duration, needed for the simulation of blast infections, was estimated according to Sentelhas et al. (2008). The spatial resolution of the ECMWF database used in this study was 0.125° (about 17 km). Data were downscaled to a regular 2 km × 2 km grid based on kriging methodology (Cressie, 1993). To allow correcting biases detected for some of the variables, dedicated calibration procedures were developed by targeting the EC-JRC MARS weather database as a reference (http://agri4cast.jrc.ec.europa.eu/DataPortal/Index.aspx).

2.3.2. Satellite remote sensing data

Satellite data were exploited to retrieve information on (i) current rice cultivated area (for identifying the area covered by rice in each 2 × 2 km elementary unit to perform the upscaling of model outputs), (ii) spatially distributed, season-specific sowing dates, and (iii) LAI for the updating/recalibration of the model. In order to guarantee continuity and redundancy, earth observation thematic products were derived either using existing operational mapping services (e.g., for LAI) or dedicated processing chains able to exploit operational, free of charge data provided by the ESA and NASA space agencies. See next sections for details.

2.3.2.1. Rice distribution maps

Spatially explicit information on rice cultivated areas for the three study sites was derived at 20 m resolution from Sentinel-1 SAR data processing (Fig.1.a/b/c). The map was produced using the processing chain module within MAPscape-RICE (Sarmap®). The mapping method involved two main steps according to Nelson et al. (2014): automatic pre-processing of SAR data and rule-based classification. Firstly, the multi-temporal spaceborne SAR Single Look Complex data were converted into terrain geocoded backscattering coefficient (σ^0) following strip mosaicking, co-registration, time-series speckle filtering, terrain geocoding, radiometric calibration and normalization, anisotropic non-linear diffusion (ANLD) Filtering, Removal of atmospheric attenuation. Secondly, a SAR specific multi-temporal σ^o rule-based rice detection algorithm (MSRD) was then applied. In this work, for Greece and Spain exclusively Sentinel-1A 12 days VV/VH Ascending (12 in Greece, 13 in Spain) and Descending (12 in Greece, 13 in Spain) data have been used (hence enabling an almost weekly monitoring) to produce the maps. For Italy, due to the complex agricultural system, 4 Landsat-8 images acquired between March 2016 and June 2016 were additionally used. In this second case, maps of Enhanced Vegetation Index (EVI; Huete et al. 2002) – a spectral index sensitive to crop LAI – and Normalized Difference Flooding Index (NDFI; Boschetti et al., 2014) – sensitive to surface standing water – were calculated from Landsat images and used as additional info to improve the differentiation between rice and other summer crops (maize, sunflower, soybean). MAPscape-RICE is therefore based on operational, free of change SAR and Optical data provided by ESA and NASA that guarantees consistency of data provision and offer the data redundancy needed to support the production of crop masks.

2.3.2.2. Sowing date maps

Spatially distributed estimates of sowing date at district level were based on the use of the PhenoRice algorithm (Boschetti et al., 2009, 2017). Currently the method works on integrated time series of TERRA and AQUA 250 m 16-days composite MODIS vegetation indices products (MOD13Q1 and MYD13Q1, respectively). The algorithm identifies a MODIS pixel as a rice crop when (i) a clear and unambiguous flood condition is detected using NDFI, and (ii) a consistent rapid crop growth is

recognized analyzing EVI. In synthesis, sowing date was estimated in correspondence with agronomic flooding and a minimum in the EVI curve, more details can be found in Boschetti et al. (2009). The PhenoRice approach was applied in three study areas (Italy, Greece and Spain) to estimate dates of crop establishment to be used primarily as direct input to crop modeling solutions, providing spatially and temporally dynamic crop calendars. MODIS imageries acquired from 2003 to 2014 were used to analyze the inter-annual and spatial variability of rice growth dynamics. Fig. 2 provides the statistics derived for sowing dates in Italy (a), Spain (b) and Greece (c). It is interesting to notice that the method was able to detect the anomalous condition that occurred in Italy in 2013 (Fig. 2.a), when crop establishment was 1-month delayed compared to the 10-year average. This observation was confirmed by the National Rice Authority in the 2013 Rice Season Report, that described how an extremely rainy and cold spring forced farmers to delay rice sowing up to mid-June (ENR, 2013). The new Sentinel-3 data represents an operational data source that can guarantee a backup solution for MODIS data.

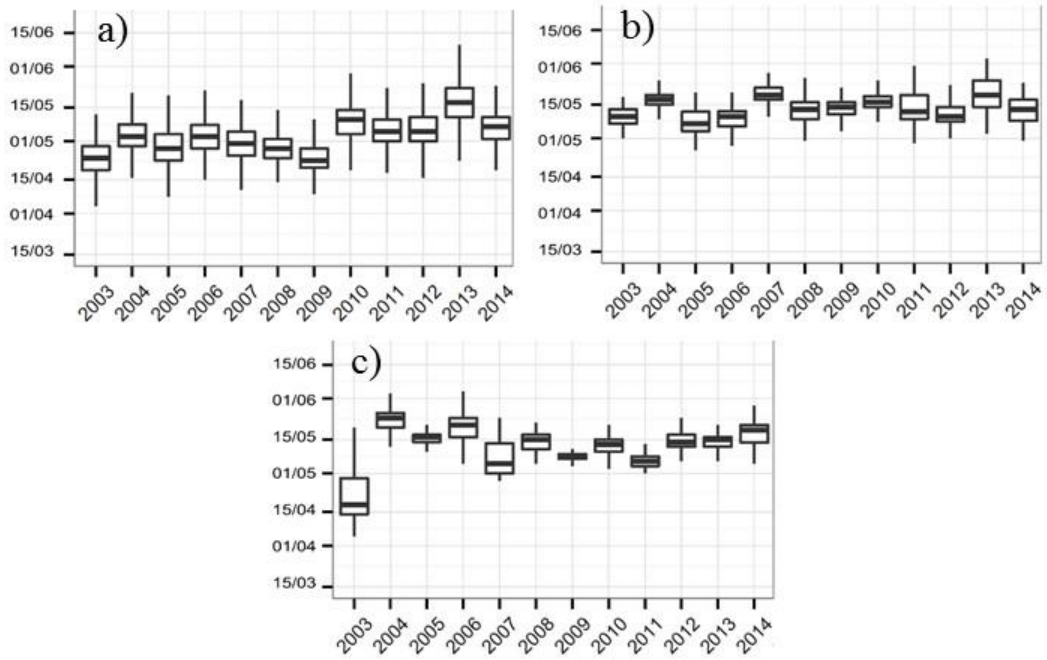


Figure 2*: Inter-annual variability of rice sowing dates from 2003 to 2014 in a) Italy, b) Spain and c) Greece.*

2.3.2.3. Leaf Area Index maps

LAI values (Fig. 3) – used for being assimilated into the WARM model – were derived from operational multi-temporal biophysical products derived from SPOT/VEGETATION and PROBA-V

in the framework of the Copernicus Global Land Services. The GEOV1 dataset is a multisensory product, developed to guarantee temporal continuity of biophysical variables over the globe on the near real-time basis. The GEOV1 LAI retrieval processing chain relies on neural networks trained using MODIS and CYCLOPES products (Baret et al., 2013) to generate remote sensing LAI estimates from SPOT/VEGETATION (1999 to May 2014) and PROBA-V (June 2014 up to date) sensors at 1/112° spatial resolution in a Plate Carrée projection (regular latitude/longitude grid) every 10 days. LAI maps were downloaded from Copernicus servers and downscaled at the 2 km × 2 km regular grid used as elementary simulation unit. The downscaling was performed using only values with quality flags indicating the best confidence estimates and using dedicated crop masks for the study areas.

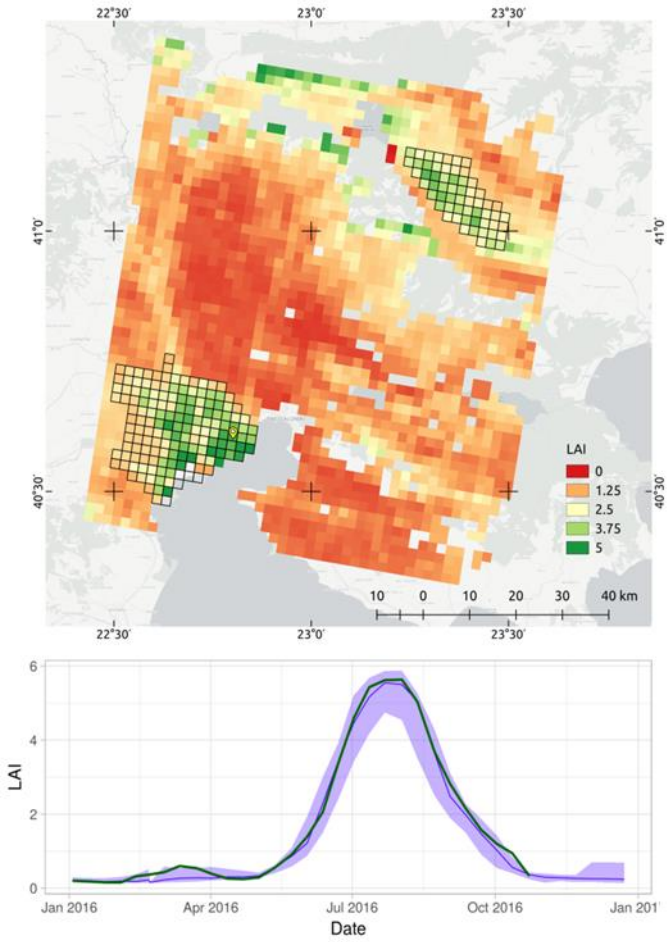


***Figure 3**: LAI map over the Greek rice area derived from Proba-V on 2 July 2016 and the time-trend of LAI in 2016 (from the ERMES geoportal, http://ermes.dlsi.uji.es/).*

LAI values used within this study were validated during the EU-FP7 ERMES project (www.ermes-fp7space.eu) following international CEOS (Committee on Earth Observation Satellites) protocols

(Fernandes et al., 2014). The procedure consisted in the intercomparison with high resolution LAI maps derived from decametric satellite imagery (Campos-Taberner et al., 2017). Results showed that GEOV1-derived LAI maps allow to properly describe the temporal and spatial variability of rice crops over the study areas. The possibility of using MODIS LAI product (MOD15A2) as a backup solution in case of missing GEOV1 data was also successfully assessed.

2.4. Spatially distributed simulations and forecasting methodology

WARM model simulations were run on 2 km × 2 km elementary units, according to the spatial resolution of the weather database (see section 2.3.1). The model parameters for Japonica varieties (Confalonieri et al., 2009a) were used for the simulations carried out in Spain (both districts) and in the Greek district of Serres, whereas the WARM parameters for Tropical Japonica varieties were used for Thessaloniki. Concerning the Italian districts, simulations were run for both the market categories "Lungo B" (belonging to the Tropical Japonica group) and "Tondo" (belonging to the Japonica group). For the latter, the available set of parameters was refined using data on phenological development, LAI and final yield provided by "Ente Nazionale Sementi Elette" (ENSE) between 2006 and 2012, and collected during dedicated experiments carried out within the ERMES project (www.ermes-fp7space.eu) in 2014-2015 in the Pavia Province. The market category "Lungo A" (belonging to the Japonica group) was excluded from the analysis since specific datasets were not available for the calibration of WARM parameters.

According to the agro-management practices in the study areas, simulations were carried out under potential conditions for water and nutrients, whereas the effects of blast disease and cold shocks around flowering (inducing spikelet sterility) were taken into account by means of dedicated WARM modules. For each elementary simulation unit (regular grid of 2 km × 2 km) and for each 10-day period, a variety of information was aggregated at district level based on the percentage of rice cover, as derived by rice maps, for each elementary simulation unit. This information included potential, blast- and cold shock-limited crop model state variables, the same state variables for model runs

including the assimilation of remote sensing information (for both the updating and recalibration assimilation methods), and key agro-climatic indicators (Table 1). According to the forecasting methodology developed and used within the MARS forecasting system of the European Commission (Vossen and Rijks, 1995; https://ec.europa.eu/jrc/en/scientific-tool/agri4cast-mars-crop-yield-forecasting-system-wiki) and in related yield forecasting systems (de Wit et al., 2010; Kogan et al., 2013), model outputs and agro-climatic indicators were then related to official yield statistics for the time series 2003-2014 using multiple linear step-wise regressions ("statistical post-processing", hereafter). The methodology is fully described by Pagani et al. (2017). To avoid losing robustness because of overfitting, the maximum allowed number of regressor was four. The forecasting event was triggered at the10-day period corresponding to the physiological maturity. Official yield statistics for the rice ecotypes considered (with the exception of Tarragona, for which only global rice statistics were available) were supplied by the Spanish and Greek Ministries of agriculture and by the Italian National Rice Authority. Before the analysis, yield statistics were examined to identify and possibly remove the presence of significant technological trends due to, e.g., improved machineries or genotypes. Indeed, if information on improved technological solutions (which affect historical yield statistics) are not provided to the model and the same management and *genotype* are used for the whole time series because of lack of information given the scale considered, the model will be able only to reproduce the effect of the year-to-year variability in weather. In case of significant technological trends, their elimination from the historical series of yield statistics prevents the step-wise regression analysis being performed on series of data (model outputs and historical yields) that are not coherent, being historical yields affected by a factor (technological improvement) not accounted for by the model. In particular, the presence of linear trends was first tested and the significance of quadratic ones was then verified. The predictive ability of each regression model was tested by performing a leave-one-out cross-validation on the available time series of historical yields and by calculating the following performance metrics: mean absolute error (MAE; minimum and optimum = 0 t ha^{-1}), relative root mean square error (RRMSE; minimum and optimum = 0%; normalized for the mean of reference values; Jørgensen et al., 1986), modelling efficiency (EF; from -

340 ∞ to +1, optimum +1; Nash and Sutcliffe, 1970), and metrics from the linear regression between
341 official and predicted yields (slope, intercept, R^2).
342
343 ***Table 1****: List of crop model outputs and agro-climatic indicators used as independent variables within the*
344 *forecasting system (P=potential conditions; B-l=blast-limited; C-l=cold shock-limited; U=updated;*
345 *R=recalibrated).*

Indicator name	Unit	Description	Model configuration
Model outputs			
DVS	-	Development stage code	P
AGB	t	Aboveground biomass	P, B-l, U, R
SB	t	Stem biomass	P, B-l, U, R
YIELD	t ha^{-1}	Storage organs biomass	P, B-l, U, R
LAI	m^2 m^{-2}	Leaf area index	P, B-l, U, R
GLAI	m^2 m^{-2}	Green leaf area index	P, B-l, U, R
BlastInf	-	Cumulated efficiency percentage of potential blast infections	B-l
Coldster	-	Cumulated efficiency percentage of potential cold-induced spikelets sterility	C-l
Agro-climatic indicators			
T_{MAX}	°C	Cumulated daily maximum temperature	-
T_{MIN}	°C	Cumulated daily minimum temperature	-
Rain	mm	Cumulated rainfall	-

346

347 ## 3. Results and discussion

348 The cross validation allowed identifying the best statistical models for each combination ecotype ×
349 production district identified among those proposed by the step-wise regression analysis (Table 2).
350 The significance level for all regression models – sorted on the basis of the beta coefficients – was
351 lower than 0.05.
352 The forecasting system achieved satisfactory performances in six out of eight cases. Unsatisfying
353 forecasting reliability was indeed obtained only for Tropical Japonica varieties in Thessaloniki and
354 Japonica varieties in Lomellina. Without considering these two cases, average RRMSE and R^2 of the
355 most reliable regression model were equal to 2.9% and 0.78, respectively. The best results were
356 obtained for the Japonica ecotype in Valencia and the Tropical Japonica ecotype (market category
357 "Lungo B") in Vercelli, for which the amount of inter-annual yield variability explained was,
358 respectively, 89% (33% of which explained by a technological trend) and 83% (no statistical
359 technological trend).

Despite the low value of R^2 for Tropical Japonica in Thessaloniki and Japonica in Lomellina (Fig. 4), the values of EF were positive (0.24 and 0.16, respectively) indicating that the forecasting system – even in these cases, is a better predictor than the mean of official yield statistics (Table 2). The marked over- and underestimations showed in some years (i.e., 2012-2014 for Thessaloniki, and 2009-2010 for Lomellina; Fig. 4.a) were likely caused by factors not accounted for by the simulation model. This hypothesis is supported by the analysis of the seasons 2012, 2013/2014 and 2014 in Thessaloniki. The low yields recorded in the district in 2012 were caused by an extreme heat wave during the reproductive period (process not simulated by the current modelling solution), whereas the high yields in 2013 and 2014 were due to the introduction of the successful high yielding variety Ronaldo (information provided by the Cereal Institute of the Hellenic Agricultural Organization (DEMETER); www.nagref-her.gr/en).

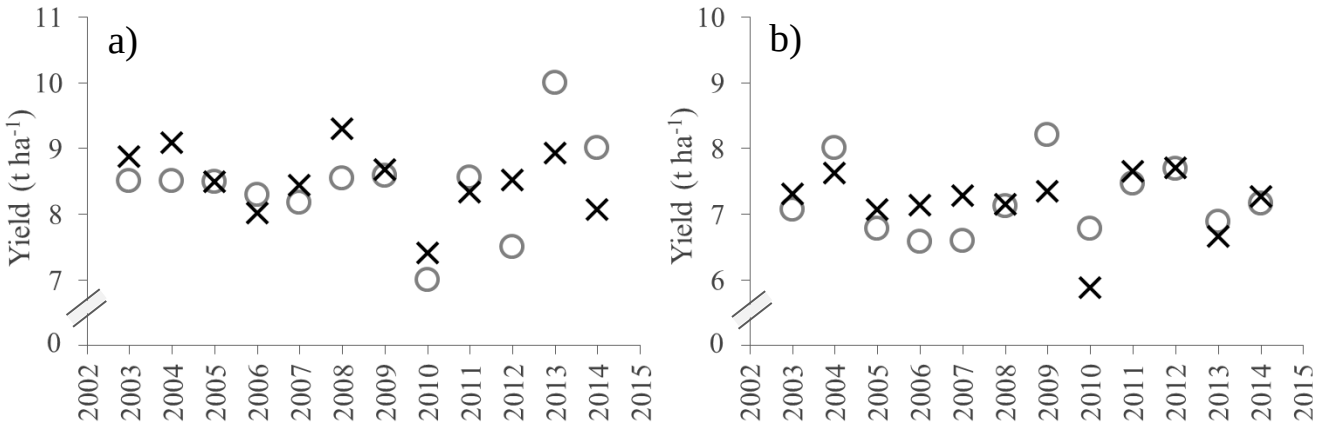


Figure 4*: Comparison between official (grey circles) and forecasted (black crosses) yields for the cross-validation for a) Tropical Japonica cultivars in Thessaloniki and b) Japonica cultivars (market category "Tondo") in Lomellina.*

In seven out of eight cases, the assimilation of remote sensing-derived LAI allowed improving the forecasting capability; in particular, the updating and recalibration assimilation methods led to improving forecasts in two and five cases, respectively. This is in agreement with results obtained by other authors: e.g., Ma et al. (2013) almost halved the error on yield estimates by assimilating MODIS-derived LAI into a forecasting system based on the WOFOST crop model. Similar results were obtained by Ines et al. (2013) for maize yield forecast using a DSSAT-based system.

For the rice district of Tarragona, simulated state variables (including blast-limited conditions) allowed achieving satisfactory predictions even without the assimilation of remote sensing information (Table 2), with EF and R^2 for cross-validation equal to 0.60 and 0.71, respectively, although 37% of the inter-annual yield variability was explained by a significant technological trend. For other combinations ecotype × district, the assimilation of remote sensing information allowed to increase the predicting capability, although the forecasting system solely based on the crop model explained a relevant part of yield fluctuations in the series of historical yields series. As an example, assimilation led to increasing R^2 from 0.70 to 0.83 for Tropical Japonica in the province of Vercelli (Table 2). In other cases, the reduction of uncertainty due to the assimilation of remote sensing-derived LAI led to a substantial improvement in yield forecasts. As an example, the model showed marked under- and overestimations in most of the years for Japonica varieties in Serres, even under blast-limiting conditions (negative EF, R^2=0.09; Table 2 and Fig. 5.a), for which some of the model outputs achieved significant values when used as regressors. In particular, the yield forecasted for 2003 was in line with the average values for the time series, whereas official yield statistics for the same year were severely affected by blast disease (information provided by the Cereal Institute of the Hellenic Agricultural Organization (DEMETER); www.nagref-her.gr/en). Among the reasons for explaining the crop model failure in identifying 2003 as a year particularly affected by blast disease, a key role is likely played by the lower resistance of varieties grown at the beginning of the 2000s and by the general uncertainty due to the lack of information on fungicides distribution for large-area simulations. Concerning the former, indeed, given the pathogen pressure change greatly between years, the effect of improved varieties is hardly detectable by medium-term technological trends. The assimilation of remote sensing LAI allowed reducing the uncertainty in simulations by detecting the overall lower vigor of rice canopies in the district due to the disease. Indeed, assimilation led to markedly increase the system capability to reproduce the inter-annual yield fluctuations (a value of 0.80 was achieved for both EF and R^2), including the correct estimate of the poor yields recorded for 2003 (Fig.5.b).

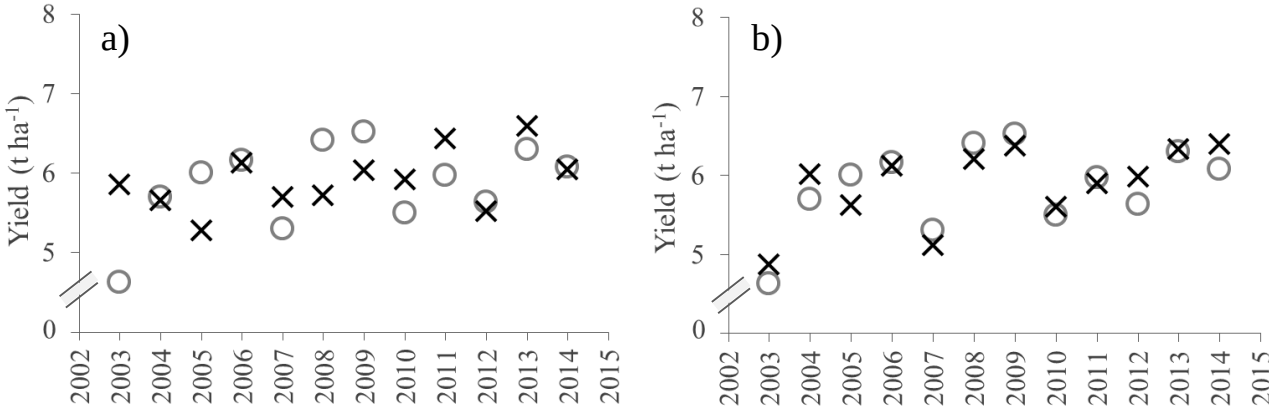


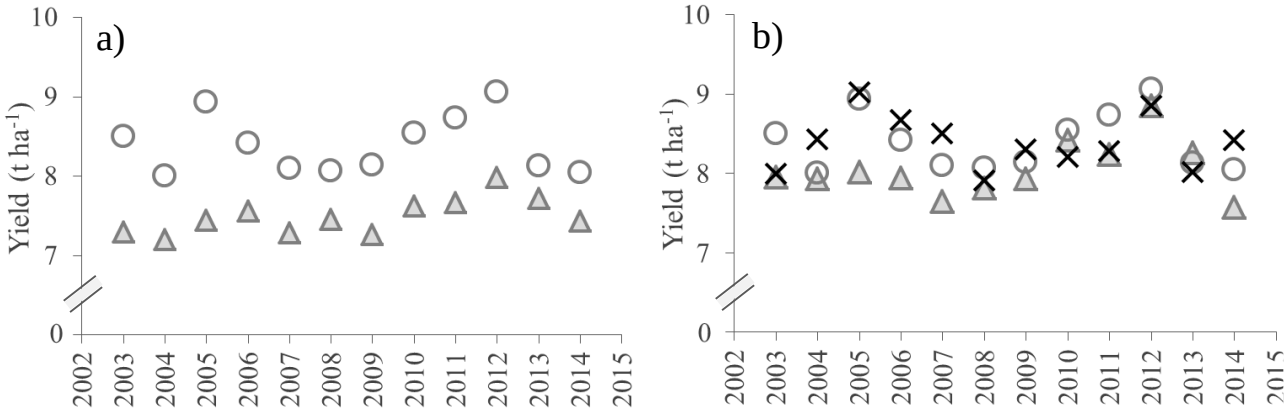
Figure 5*: Comparison between official (grey circles) and forecasted (black crosses) yields for the cross-validation for Japonica varieties in Serres. Forecasting system was based on a) crop model outputs and agro-climatic indicators, and on b) model outputs updated using remote sensing-derived LAI and agro-climatic indicators.*

In most of the cases, the statistical post-processing of simulated outputs was important for reducing the different sources of uncertainty (in model structure, parameterizations, management, upscaling procedure, etc.) affecting large area simulations and, thus, to allow the system to correctly reproducing the fluctuations along the historical series of yield statistics. However, when the model was run in contexts not severely affected by extreme weather conditions or by unreproducible (given the scale) season- or site-specific effects involved with the application of management practices, good results were obtained from the combined use of crop modelling and remote sensing technologies, even without post-processing results. As an example, in the rice district of Valencia, yields simulated by WARM under potential conditions for Japonica cultivars were sufficiently coherent – in terms of overall time trend – with official yield statistics, although they were characterized by a general underestimation during the whole time series (Fig. 6.a). The assimilation of remote sensing-derived LAI values (via recalibration of model parameters) allowed to improve the forecasting capability of the crop model even in absolute term, especially in the second part of the series (Fig. 6.b). However, the statistical post-processing of simulated results led to further improve forecasts (Fig. 6.b) by including in the regression model – besides simulated yield – a second variable, i.e. the cumulated rainfall, that presented a negative correlation with official yields. The reason for the importance of cumulated rainfall is related with its role in affecting rice productivity because of less radiation available for photosynthesis (in turn due to more cloudy days) and because of higher humidity, that

431 favored blast infections. A further negative effect of abundant rainfall in the last part of the crop cycle

432 is the possible problems during harvesting procedures.

433



434

435 ***Figure 6****: Comparison between official (grey circles) and yields forecasted for Japonica cultivars in Valencia*
436 *using a) only the crop model (grey triangles), b) the crop model with assimilation of remote sensing LAI (grey*
437 *triangles) and the statistical post-processing of simulated results (including LAI assimilation) (black crosses).*

***Table 2**: Comparison between official and forecasted yields for the 2003-2014 time series. Forecasted yields derive from the statistical post-processing of simulated model outputs (M in the column "Regressors") and agro-climatic indicators (A-c) (Table 1). Best regression models (maximum allowed regressor = 4) were identified by cross-validation. U and R (column "Regressors" and subscripts in column "Regression model") refer to the assimilation of remote sensing LAI in the crop model using updating and recalibration techniques, respectively. Empty cells indicate the absence of significant regression models. Cells with dashes indicate that no improvement was obtained from assimilation. MAE: mean absolute error (minimum and optimum = 0 t ha^{-1}); RRMSE: relative root mean square error (minimum and optimum = 0%; Jørgensen et al., 1986), EF: modelling efficiency (from - ∞ to +1, optimum +1; Nash and Sutcliffe, 1970); Slope, Int and R^2 refer to the linear regression between official and predicted yields.*

Country	Ecotype	District	Regressors	Regression model	MAE	RRMSE (%)	EF	Slope	Int	R^2
Spain	Japonica	Valencia	M, A-c	Rain**, Yield*	0.23	3.01	0.66	1.07	-0.60	0.66
			M, A-c, **U**	-	-	-	-	-	-	-
			M, A-c, **R**	Rain***, **Yield$_R$****	0.11	1.72	0.89	1.04	-0.31	0.89
		Tarragona	M, A-c	GLAI$_B$*, LAI**, Yield**, AGB$_B$**	0.22	4.61	0.60	0.72	1.71	0.71
			M, A-c, **U**	-	-	-	-	-	-	-
			M, A-c, **R**	-	-	-	-	-	-	-
Greece	Tropical Japonica	Thessaloniki	M, A-c							
			M, A-c, **U**	**GLAI$_U$**, LAI$_B$**, **LAI$_U$***	0.50	8.00	0.07	0.55	3.80	0.21
			M, A-c, **R**	GLAI**, LAI*, **AGB$_R$***	0.50	7.24	0.24	0.74	2.10	0.28
	Japonica	Serres	M, A-c	Yield$_B$*, AGB$_B$*	0.41	9.08	-0.08	0.43	3.30	0.09
			M, A-c, **U**	GLAI***, **Yield$_U$***, Tmax***	0.20	3.94	0.80	0.97	0.13	0.80
			M, A-c, **R**	LAI**, LAI$_R$*, Tmax*, Yield$_B$**	0.33	7.05	0.35	0.79	1.17	0.39
Italy	Tropical Japonica *(market category "Lungo B")*	Lomellina	M, A-c	GLAI**, DVS**, Tmax**, Tmin*	0.23	3.88	0.11	0.55	3.16	0.40
			M, A-c, **U**	DVS**, **LAI$_U$**, LAI$_B$**, **AGB$_U$****	0.16	2.78	0.54	0.75	1.78	0.62
			M, A-c, **R**	DVS**, **LAI$_R$***, Yield**, **Yield$_R$*****	0.14	2.31	0.69	0.91	0.67	0.70
		Vercelli	M, A-c	ColdSter**, LAI***	0.21	3.34	0.70	0.96	0.29	0.70
			M, A-c, **U**	-	-	-	-	-	-	-
			M, A-c, **R**	**GLAI$_R$***, LAI***, **LAI$_R$****, **Yield$_R$****	0.16	2.50	0.83	1.00	-0.03	0.83
	Japonica *(market category "Tondo")*	Lomellina	M, A-c							
			M, A-c, **U**							
			M, A-c, **R**	**SB$_R$****, **AGB$_R$***	0.37	6.58	0.16	0.60	2.90	0.30
		Vercelli	M, A-c	LAI**	0.19	3.16	0.48	0.90	0.72	0.49
			M, A-c, **U**	**LAI$_U$****, DVS*, Rain*, AGB$_B$***	0.16	2.57	0.66	0.78	1.55	0.72
			M, A-c, **R**	-	-	-	-	-	-	-

* $p \leq 0.05$; ** $p \leq 0.01$; *** $p \leq 0.001$

The forecasting reliability obtained for Italy is in line with (or better than) the best results obtained by de Wit et al. (2010) using the CGMS-WOFOST model for grain yield estimates in Europe, whereas the results we obtained for Japonica cultivars in Valencia (56% of the variance explained without considering the trend) are slightly better than those achieved by de Wit et al. (2010) for barley, field beans and sugar beets. Results achieved for other combinations rice ecotype × district are less reliable, although they can be considered as in agreement with most of the results normally obtained using generic and crop-specific (e.g., Kogan et al., 2013; Ines et al., 2013) yield forecasting systems. Comparing our system with other rice specific ones, the values of R^2 obtained by Son et al. (2014) with an approach based on MODIS-derived vegetation indices for the Mekong River Delta (Vietnam) ranged from 0.40 to 0.71. These values are similar to the ones we found, although the inter-annual yield fluctuations in Vietnam are larger than those characterizing rice cultivation in Europe, and the mean error obtained by Son et al. (2014) was often higher than those achieved in this study. However, the approach proposed by Son et al. (2014) is simpler and easier to set-up and maintain.

4. Conclusions

A high-resolution rice yield forecasting system based on the WARM model was run on 2 km × 2 km elementary simulation units covering the main European rice districts in Italy, Spain and Greece. The system integrated remote sensing information to define rice-cropped area and to derive sowing dates varying in time and space, as well as for assimilating exogenous LAI information into the simulation (using both updating and recalibration techniques). The system demonstrated its reliability in forecasting yields at district level even under conditions characterized by small year-to-year fluctuations in rice productivity. However, despite WARM includes algorithms for the simulation of processes with a relevant impact on yields in the study areas (e.g., blast disease and spikelet sterility due to thermal shocks during young microspore stage and flowering), the forecasting system in a few cases was not able to reproduce the effects of unfavorable seasons. This highlights the need for further improvements, given the importance of predicting yields especially in case of unfavorable conditions.

Although further studies are needed to increase the predicting capability in some of the districts (i.e., Thessaloniki and Lomellina), the system demonstrated its usefulness during 2015 and 2016, when it was used under pre-operational conditions during the activities performed within the EU-FP7 ERMES project (http://www.ermes-fp7space.eu/). In this context, yield forecast bulletins were regularly issued to public authorities and private companies in Italy, Greece and Spain. Feedbacks received from public and private stakeholders encouraged the continuation of the forecasting system operationally in the next seasons.

5. Acknowledgements

This study has been partially funded under the EU FP7 collaborative project, grant agreement no 606983, ERMES: An Earth obseRvation Model based RicE information Service (ERMES). Carlos Granell and Sven Casteleyn were partly funded by the Ramón y Cajal Programme of the Spanish government (grant numbers RYC-2014-16913 and RYC-2014-16606 respectively).

6. References

Arai, N., Yoshino, R., 1987. Studies on the sporulation of rice blast fungus: (1) relation between sporulation and temperature. Ann. Phytopathol. Soc. Jap. 53, 371-372.

Balaghi, R., Jlibene, M., Tychon, B., Eerens, H., 2012. Agrometeorological Cereal Yield Forecasting in Morocco. INRA, Rabat, Maroc, 149 p.

Bannayan, M., Crout, N.M.J., 1999. A stochastic modelling approach for real-time forecasting of winter wheat yield. Field Crop. Res. 62, 85-95. DOI: 10.1016/S0378-4290(99)00008-8

Baret, F., Weiss, M., Lacaze, R., Camacho, F., Makhmara, H., Pacholcyzk, P., Smets, B., 2013. Geov1: LAI and FAPAR essential climate variables and FCOVER global time series capitalizing over existing products. Part1: Principles of development and production. Remote Sens. Environ. 137, 299-309. DOI: 10.1016/j.rse.2012.12.027

Bezuidenhout, C.N., Singels, A., 2007a. Operational forecasting of South African sugarcane production: Part 1 – System description. Agric. Syst. 92, 23-38. DOI: 10.1016/j.agsy.2006.02.001

Bezuidenhout, C.N., Singels, A., 2007b. Operational forecasting of South African sugarcane production: Part 2 – System evaluation. Agric. Syst. 92, 39-51. DOI: 10.1016/j.agsy.2006.03.002

Boschetti, M., Stroppiana, D., Brivio, PA., Bocchi, S., 2009. Multi-year monitoring of rice crop phenology through time series analysis of MODIS images. Int. J. Remote Sens. 30, 4643-4662. DOI: 10.1080/01431160802632249

Boschetti, M., Nutini, F., Manfron, G., Brivio, P.A., Nelson, A., 2014. Comparative Analysis of Normalised Difference Spectral Indices Derived from MODIS for Detecting Surface Water in Flooded Rice Cropping Systems. PLoS ONE 9, e88741. DOI:10.1371/journal.pone.0088741

Boschetti, M., Busetto, L., Manfron, G., Laborte, A., Asilo, S., Pazhanivelan, S., Nelson, A., 2017. PhenoRice: A method for automatic extraction of spatio-temporal information on rice crops using satellite data time series. Remote Sens. Environ. 194, 347-365. DOI: 10.1016/j.rse.2017.03.029

Bouman, B.A.M., 1995. Crop modeling and remote-sensing for yield prediction. Nether. J. Agric. Sci. 43, 143-161.

Bregaglio, S., Titone, P., Cappelli, G., Tamborini, L., Mongiano, G., Confalonieri, R., 2016. Coupling a generic disease model to the WARM rice simulator to assess leaf and panicle blast impacts in temperate climate. Eur. J. Agron. 76, 107-117. DOI: 10.1016/j.eja.2016.02.009

Campos-Taberner, M., García-Haro, F.J., Camps-Valls, G., Grau-Muedra, G., Nutini, F., Busetto, L., Katsantonis, D., Stavrakoudis, D., Minakou, C., Gatti, L., Barbieri, M., Holecz, F., Stroppiana, D., Boschetti, M., 2017. Exploitation of SAR and optical sentinel data to detect rice crop and estimate seasonal dynamics of leaf area index. Remote Sens. 9, 248. DOI: 10.3390/rs9030248

Confalonieri, R., Rosenmund, A.S., Baruth, B., 2009a. An improved model to simulate rice yield. Agron. Sustain. Dev. 29, 463-474. DOI: 10.1051/agro/2009005

Confalonieri, R., Acutis, M., Bellocchi, G., Donatelli, M., 2009b. Multi-metric evaluation of the models WARM, CropSyst, and WOFOST for rice. Ecol. Model. 220, 1395-1410. DOI: 10.1016/j.ecolmodel.2009.02.017

Confalonieri, R., Bregaglio, S., Adam, M., Ruget, F., Li, T., Hasegawa, T., Yin, X., Zhu, Y., Boote, K., Buis, S., Fumoto, T., Gaydon, D., Lafarge, T., Marcaida, M., Nakagawa, H., Ruane, A.C., Singh, B., Singh, U., Tang, L., Tao, F., Fugice, J., Yoshida, H., Zhang, Z., Wilson, L.T., Baker, J., Yang, Y., Masutomi, Y., Wallach, D., Acutis, M., Bouman, B., 2016a. A taxonomy-based approach to shed light on the babel of mathematical models for rice simulations. Envir. Modell. Softw. 85, 332-341. DOI: 10.1016/j.envsoft.2016.09.007

Confalonieri, R., Orlando, F., Paleari, L., Stella, T., Gilardelli, C., Movedi, E., Pagani, V., Cappelli, G., Vertemara, A., Alberti, L., Alberti, P., Atanassiu, S., Bonaiti, M., Cappelletti, G., Ceruti, M., Confalonieri, A., Corgatelli, G., Corti, P., Dell'Oro, M., Ghidoni, A., Lamarta, A., Maghini, A., Mambretti, M., Manchia, A., Massoni, G., Mutti, P., Pariani, S., Pasini, D., Pesenti, A., Pizzamiglio, G., Ravasio, A., Rea, A., Santorsola, D., Serafini, G., Slavazza, M., Acutis, M., 2016b. Uncertainty in crop model predictions: What is the role of users? Environ. Modell. Softw. 81, 165-173. DOI: 10.1016/j.envsoft.2016.04.009

Confalonieri, R., Bregaglio, S., Acutis, M., 2016c. Quantifying uncertainty in crop model predictions due to the uncertainty in the observations used for calibration. Ecol. Model. 328, 72-77. DOI: 10.1016/j.ecolmodel.2016.02.013

Cressie, N.A.C, 1993. Statistics for spatial data. Wiley, New York. 900 pp.

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. DOI: 10.3354/cr00872

Diekkrüger, B., Söndgerath, D., Kersebaum, K.C., McVoy, C.W., 1995. Validity of agroecosystem models. A comparison of results of different models applied to the same data set. Ecol. Model. 81, 3-29. DOI: 10.1016/0304-3800(94)00157-d

547 Dobson, M.C., Ulaby, F.T., Pierce, L.E., Sharik, T.L., Bergen, K.M., Kellndorfer, J., Kendra, J.R., Li,
548 E., Lin, Y.C., Nashashibi, A., Sarabandi, K., Siqueira, P., 1995. Estimation of forest biophysical
549 characteristics in northern Michigan with SIR-C/X-SAR. IEEE Trans. Geosci. Remote Sens. 33,
550 877-895. DOI: 10.1109/36.406674
551 Dorigo, W.A., Zurita-Milla, R., de Wit, A.J.W., Brazile, J., Singh, R., Schaepman, M.E., 2007. A
552 review on reflective remote sensing and data assimilation techniques for enhanced
553 agroecosystem modeling. Int. J. Appl. Earth Obs. 9, 165-193. DOI: 10.1016/j.jag.2006.05.003
554 Duveiller, G., López-Lozano, R., Baruth, B., 2013. Enhanced processing of 1-km spatial resolution
555 fAPAR time series for sugarcane yield forecasting and monitoring. Remote Sens. 5, 1091-1116.
556 DOI: 10.3390/rs5031091
557 Everingham, Y.L., Muchow, R.C., Stone, R.C., Inman-Bamber, N.C., Singels, A., Bezuidenhout,
558 C.N., 2002. Enhanced risk management and decision-making capability across the sugarcane
559 industry value chain based on seasonal climate forecasts. Agric. Syst. 74, 459-477. DOI**:**
560 10.1016/S0308-521X(02)00050-1
561 Fang, H., Liang, S., Hoogenboom, G., 2011. Integration of MODIS LAI and vegetation index
562 products with the CSM–CERES–Maize model for corn yield estimation. Int. J. Remote Sens.
563 32, 1039–1065. DOI:10.1080/01431160903505310
564 Fernandes, R., Plummer, S., Nightingale, J., 2014. Global Leaf Area Index Product Validation Good
565 Practices. CEOS Working Group on Calibration and Validation – Land Product Validation Sub-
566 Group. Version 2.0. Public version available at http://lpvs.gsfc.nasa.gov/.
567 Glover, F., 1986. Future paths for integer programming and links to artificial intelligence. Comput.
568 Oper. Res. 5, 533-549. DOI: 10.1016/0305-0548(86)90048-1
569 Huete, A., Didan, K., Miura, T., Rodriguez, E.P., Gao, X., Ferreira, L.G., 2002. Overview of the
570 radiometric and biophysical performance of the MODIS vegetation indices. Remote Sens.
571 Environ. 83, 195–213. DOI:10.1016/s0034-4257(02)00096-2

572 Ines, A.V.M., Das, N.N., Hansen, J.W., Njoku, E.G., 2013. Assimilation of remotely sensed soil
573 moisture and vegetation with a crop simulation model for maize yield prediction. Remote Sens.
574 Environ. 138, 149-164. DOI: 10.1016/j.rse.2013.07.018
575 Jin, X., Kumar, L., Li, Z., Feng, H., Xu, X., Yang, G., Wang, J., 2018. A review of data assimilation
576 of remote sensing and crop models. Eur. J. Agron. 92, 141-152. DOI: 10.1016/j.eja.2017.11.002
577 Jørgensen, S.E., Kamp-Nielsen, L., Christensen, T., Windolf-Nielsen, J., Westergaard, N., 1986.
578 Validation of a prognosis based upon a eutrophication model. Ecol. Model. 32, 165-182. DOI:
579 10.1016/0304-3800(86)90024-4
580 Kersebaum, K.C., Boote, K.J., Jorgenson, J.S., Nendel, C., Bindi, M., Frühauf, C., Gaiser, T.,
581 Hoogenboom, G., Kollas, C., Olesen, J.E., Rötter, R.P., Ruget, F., Thorburn, P.J., Trnka, M.,
582 Wegehenkel, M., 2015. Analysis and classification of data sets for calibration and validation of
583 agro-ecosystem models. Environ. Modell. Softw. 72, 402-417. DOI:
584 10.1016/j.envsoft.2015.05.009
585 Kim, K.R., 2000. Weather-driven models for rice leaf blast and their implementation to forecast
586 disease development on the near real-time basis. PhD Thesis. Seoul National University, Suwon,
587 Korea.
588 Kirkpatrick, S., Gelatt, C.D., Vecchi, M.P., 1983. Optimization by simulated annealing. Science 220,
589 671-680.
590 Kogan, F., Kussul, N., Adamenko, T., Skakun, S., Kravchenko, O., Kryvobok, A., Kolotii, A., Kussul,
591 O., Lavrenyuk, A., 2013. Winter wheat yield forecasting in Ukraine based on Earth
592 observation, meteorological data and biophysical models. Int. J. Appl. Earth Obs. 23, 192-203.
593 DOI: 10.1016/j.jag.2013.01.002
594 Kropff, M.J, Spitters, C.J.T., Schnieders, B.J., Joenje, W., De Groot, W., 1992. An eco-physiological
595 model for interspecific competition, applied to the influence of Chenopodium album L. on sugar
596 beet. II. Model evaluation. Weed Res. 32, 451-463. DOI: 10.1111/j.1365-3180.1992.tb01906.x

597 Launay, M., Guèrif, M., 2005. Assimilating remote sensing data into a crop model to improve
598 predictive performance for spatial applications. Agr. Ecosys. Environ. 111, 321-339. DOI:
599 10.1016/j.agee.2005.06.005
600 Lefebvre, A., Sannier, C., Corpetti, T., 2016. Monitoring urban areas with Sentinel-2A data:
601 application of the update of the Copernicus High Resolution Layer imperviousness degree.
602 Remote Sens. 8, 606. DOI: 10.3390/rs8070606
603 Ma, G., Huang, J., Wu, W., Fan, J., Zou, J., Wu, S., 2013. Assimilation of MODIS-LAI into the
604 WOFOST model for forecasting regional winter wheat yield. Math. Comput. Model. 58, 634-
605 643. DOI:10.1016/j.mcm.2011.10.038
606 Magarey, R.D., Sutton, T.B., Thayer, C.L., 2005. A simple generic infection model for foliar fungal
607 plant pathogens. Phytopathol. 95, 92-100. DOI: 10.1094/phyto-95-0092
608 Matsumoto, T., Du, H., Lindsey, J.S., 2002. A parallel simplex search method for use with an
609 automated chemistry workstation. Chemometr. Intell. Lab. 62, 129-147. DOI:10.1016/s0169-
610 7439(02)00010-2
611 McLaughlin, D., 2002. An integrated approach to hydrologic data assimilation: interpolation,
612 smoothing, and filtering. Advan. Water Resourc. 25, 1275-1286. DOI: 10.1016/s0309-
613 1708(02)00055-6
614 Mkhabela, M.S., Mkhabela, M.S., Mashinini, N.N., 2005. Early maize yield forecasting in the four
615 agro-ecological regions of Swaziland using NDVI data derived from NOAAs-AVHRR. Agric.
616 For. Meteorol. 129, 1-9. DOI: 10.1016/j.agrformet.2004.12.006
617 Nash, J.E., Sutcliffe, J.V., 1970. River flow forecasting through conceptual models. Part I. A
618 discussion of principles. J. Hydrol. 10, 282-290. DOI:10.1016/0022-1694(70)90255-6
619 Nelder, J.A., Mead, R., 1965. A simplex method for function minimization. Comp. J. 7, 308-313.
620 DOI: 10.1093/comjnl/7.4.308
621 Nelson, A., Setiyono, T., Rala, A.B., Quicho, E.D., Raviz, J.V., Abonete, P.J., Maunahan, A.A.,
622 Garcia, C.A., Bhatti, H.Z.M., Villano, L.S., Thongbai, P., Holecz, F., Barbieri, M.,
623 Collivignarelli, F., Gatti, L., Quilang, E.J.P., Mabalay, M.R.O., Mabalot, P.E., Barroga, M.I.,

Bacong, A.P., Detoito, N.T., Berja, G.B., Varquez, F., Wahyunto, Kuntjoro, D., Murdiyati, S.R., Pazhanivelan, S., Kannan, P., Nirmala Mary, P.C., Subramanian, E., Rakwatin, P., Intrman, A., Setapayak, T., Lertna, S., Minh, V.Q., Tuan, V.Q., Duong, T.H., Quyen, N.H., Kham, D.V., Hin, S., Veasna, T., Yadav, M., Chin, C., Ninh, N.H., 2014. Towards an operational SAR-based rice monitoring system in Asia: examples from 13 demonstration sites across Asia in the RIICE project. Remote Sens. 6, 10773-10812. DOI: 10.3390/rs61110773

OECD/FAO, 2011. OECD-FAO Agricultural Outlook 2011-2020. OECD Publishing, Paris. ISBN 978-92-64-106758

Pagani, V., Francone, C., Wang, Z., Qiu, L., Bregaglio, S., Acutis, M., Confalonieri, R., 2014. Evaluation of WARM for different establishment techniques in Jiangsu (China). Eur. J. of Agron. 59, 78-85. DOI: 10.1016/j.eja.2014.05.010

Pagani, V., Stella, T., Guarneri, T., Finotto, G., van den Berg, M., Marin, F.R., Acutis, M., Confalonieri, R., 2017. Forecasting sugarcane yields using agro-climatic indicators and Canegro model: a case study in the main production region in Brazil. Agricultural Systems, 154, 45-52. DOI: 10.1016/j.agsy.2017.0302

Paleari, L., Movedi, E., Cappelli, G., Wilson, L.T., Confalonieri, R., 2017. Surfing parameter hyperspaces under climate change scenarios to design future rice ideotypes. Glob. Change Biol. DOI: 10.1111/gcb.13682

Press, W.H., Teukolsky, S.A., Vetterling, W.T., Flannery, B.P., 2007. Numerical Recipes 3rd Edition: the art of scientific computing, 2nd edition, Cambridge University Press, Cambridge, UK.

Rosenzweig, C., Jones, J.W., Hatfield, J.L., Ruane, A.C., Boote, K.J., Thorburn, P., Antle, J.M., Nelson, G.C., Porter, C., Janssen, S., Asseng, S., Basso, B., Ewert, F., Wallach, D., Baigorria, G., Winter, J.M., 2013. The Agricultural Model Intercomparison and Improvement Project (AgMIP): Protocols and pilot studies. Agric. Forest Meteorol. 170, 166-182. DOI: 10.1016/j.agrformet.2012.09.011

649 Sader, S.A., Waide, R.B., Lawrence, W.T., Joyce, A.T., 1989. Tropical forest biomass and
650 successional age class relationships to a vegetation index derived from Landsat TM data.
651 Remote Sens. Environ. 28, 143-156. DOI: 10.1016/0034-4257(89)90112-0
652 Sándor, R., Barcza, Z., Hidy, D., Lellei-Kovács, E., Bellocchi, G., 2016. Modelling of grassland
653 fluxes in Europe: evaluation of two biogeochemical models. Agric. Ecosyst. Environ. 215, 1-19.
654 DOI: 10.1016/j.agee.2015.09.00
655 Sentelhas, P.C., Dalla Marta, A., Orlandini, S., Santos, E.A., Gillespie, T.J., Gleason, M. L.,2008.
656 Suitability of relative humidity as an estimator of leaf wetness duration. Agric. For. Meteorol.
657 148, 392–400. DOI: 10.1016/j.agrformet.2007.09.011
658 Shimono, H., Hasegawa, T., Moriyama, M., Fujimura, S., Nagata, T., 2005. Modeling spikelet sterility
659 induced by low temperature in rice. Agron. J. 97, 1524-1536. DOI: 10.2134/agronj2005.0043
660 Son, N.T., Chen, C.F., Chen, C.R., Minh, V.Q., Trung, N.H., 2014. A comparative analysis of
661 multitemporal MODIS EVI and NDVI data for large-scale rice yield estimation. Agr. Forest
662 Meteorol. 197, 52-64. DOI: 10.1016/j.agrformet.2014.06.007
663 Supit, I., 1997. Predicting national wheat yields using a crop simulation and trend models. Agric. For.
664 Meteorol. 88, 199-214. DOI: 10.1016/s0168-1923(97)00037-3
665 Vossen, P., Rijks, D.A., 1995. Early crop yield assessment of the E.U. countries: The system
666 implemented by the Joint Research Centre. European Commission (Ed.), EUR 16318 EN,
667 Luxembourg.
668 Wang, Y.P., Chang, K.W., Chen, R.K., Lo, J.C., Shen, Y., 2010. Large-area rice yield forecasting
669 using satellite imageries. Int. J. Appl. Earth Obs. 12, 27-35. DOI: 10.1016/j.jag.2009.09.009
670 Wu, G.W., Wilson, L.T., 1997. Growth and yield response of rice to rice water weevil injury. Environ.
671 Entomol. 26, 1191-1201. DOI: 10.1093/ee/26.6.1191
672 Zhao, Y., Chen, S., Shen, S., 2013. Assimilating remote sensing information with crop model using
673 Ensemble Kalman Filter for improving LAI monitoring and yield. Ecol. Model. 270, 30-42. DOI:
674 10.1016/j.ecolmodel.2013.08.016

675 Zhao, P., Lu, D., Wang, G., Wu, C., Huang, Y., Yu, S., 2016. Examining spectral reflectance
676 saturation in Landsat imagery and corresponding solutions to improve forest aboveground
677 biomass estimation. Remote Sens. 8, 469. DOI: 10.3390/rs8060469
678