Using ERA-INTERIM for regional crop yield forecasting in Europe

Allard de Wit[1,*]**, Bettina Baruth**[2]**, Hendrik Boogaard**[1]**, Kees van Diepen**[1]**,
Daniel van Kraalingen**[1]**, Fabio Micale**[2]**, Johnny te Roller**[1]**, Iwan Supit**[3]**,
Raymond van den Wijngaart**[1]

[1]**Alterra, Wageningen UR, PO Box 47, 6700 AA Wageningen, The Netherlands**
[2]**Joint Research Centre, Institute for the Protection and Security of the Citizen, Monitoring Agricultural Resources Unit (MARS),
AGRI4CAST, Italy**
[3]**Earth System Science and Climate Change Group, Wageningen UR, PO Box 47, 6700 AA Wageningen, The Netherlands**

ABSTRACT: Agrometeorological systems for regional crop yield forecasting have traditionally relied on weather data derived from weather stations for crop simulation and yield prediction. In recent years, numerical weather prediction (NWP) models have become an interesting source of weather data with the potential to replace observed weather data. This is a result of the steadily decreasing NWP grid sizes and the availability of long and consistent time-series through the so-called reanalysis projects. We evaluated the ERA-INTERIM reanalysis data set from the European Centre for Medium-range Weather Forecasting for regional crop yield forecasting. Crop simulations were carried out using 2 identical model implementations: one using interpolated observed weather, the other using weather data derived from ERA-INTERIM. Output for both sources of weather variables was generated for the EU27 and neighbouring countries and 14 crops, aggregated to national level and validated using reported crop yields from the European Statistical Office. The results indicate that the system performs very similar in terms of crop yield forecasting skill for both sources of weather variables. In 38% of the crop–country combinations, the forecasting error can be reduced by more than 10% of the baseline forecast (the trend only) for both sources of weather variables. In almost 20% of the crop–country combinations, the forecasting error can be reduced by more than 25% of the baseline forecast. The results demonstrate that the ERA-INTERIM data set is highly suitable for regional crop yield forecasting over Europe and may be used for implementing regional crop forecasting over data-sparse regions. Finally, we conclude that there is a need to improve the crop calendar and/or calibration for some of the modelled crops.

KEY WORDS: Crop simulation models · Crop yield · Regional scale · ERA-INTERIM · Europe

1. INTRODUCTION

Agrometeorological systems for regional crop monitoring and yield forecasting have traditionally relied on weather data derived from weather stations. The observations obtained from weather stations were then interpolated to obtain regular gridded data, used as input in a crop simulation model and aggregated to a regional scale for regional yield forecasting. Such an approach is necessary in order to avoid scaling issues related to non-linear responses of crop models to input data (Hansen & Jones 2000).

The set up described above has been successfully implemented in the European Crop Growth Monitoring System (CGMS), which has been operational since 1994. However, in many areas of the world, data from weather stations are difficult to come by, expensive, scarce or incomplete. Moreover, intensive error checking is often necessary in order to flag errors, substitute missing days and replace missing variables. The re-

liance on observed weather data has therefore been considered a bottleneck for implementing the CGMS in other regions of the world.

In recent years, short-term forecasts from numerical weather prediction (NWP) models have become an interesting source of weather data with the potential to replace observed weather data in the CGMS. The application of data from NWP models was initially problematic because of the large grids with sizes on the order of 2.5° × 2.5°. However, grid sizes have steadily decreased over the years and are now on the order of 0.5° × 0.5°, which is a relevant spatial scale for regional crop yield forecasting (Easterling et al. 1998, de Wit et al. 2005).

A more subtle problem of NWP model output was the inconsistency in the time-series due to incremental upgrades of the NWP model itself. Regional crop yield monitoring and yield forecasting strongly relies on the analysis of historical time-series. Therefore, any biases in the time-series caused by NWP model upgrades will distort the analysis of historic time-series of simulated and reported yields. This problem was recognized by the NWP community and has resulted in the reanalysis projects such as ERA-40, which have produced complete, coherent and consistent time-series of NWP data.

Finally, a general problem with output from NWP models is still that modelled variables often do not compare well with ground-based weather variables and complex downscaling procedures are necessary to convert NWP weather variables into realistic weather variables on the ground (Bates et al. 1998, Charles et al. 2004). For example, an often heard complaint is that precipitation predicted from NWP models does not properly reproduce the intermittency characteristics of observed precipitation records, even if it does reproduce the seasonal and monthly totals correctly. Such deficiencies in NWP-based weather data may influence the output from crop simulation models, because crop models can be sensitive to the intermittency characteristics of precipitation. However, recent research results show that the impact of poor reproduction of intermittency characteristics on crop simulation results may be limited (Robertson et al. 2007).

Recently, the European Centre for Medium-range Weather Forecasting (ECMWF) has completed the ERA-INTERIM archive which is a reanalysis of the global atmosphere since 1989 (Berrisford et al. 2009). In contrast to earlier reanalysis archives which were only available for a fixed period of time, the ERA-INTERIM archive continues in real time with a delay of around 1 mo. This particular feature makes it interesting for regional crop yield forecasting applications which need both near real-time data as well as an archive of weather data. The remaining gap of 1 mo between the end time of ERA-INTERIM and real time can be substituted with data from the operational ECMWF NWP model.

Our objective was to evaluate the ERA-INTERIM weather data reanalysis archive for regional crop yield forecasting over Europe. To achieve this objective, we generated historical time-series of crop simulation results for both the classic CGMS driven by observed weather data and an identical implementation of CGMS except that it derives its weather data from ERA-INTERIM. Regression analysis in windows of 9 yr over the period 1989–2007 was used to relate CGMS simulated results to reported crop yields at a national level obtained from the European Statistical Office (EUROSTAT). The established regression models could then be used to predict crop yield in the year following each window. The forecast error statistics over the period 1998–2007 were used to evaluate the use of ERA-INTERIM for regional yield forecasting.

2. DATA AND METHODS

2.1. Spatially distributed crop growth model

We used the World Food Studies (WOFOST) crop simulation model as a basis for our work (van Diepen et al. 1989). The WOFOST model was implemented spatially within the framework of the CGMS as part of the Monitoring Agriculture by Remote Sensing (MARS) project (Vossen & Rijks 1995). The MARS project started in 1988 and was carried out by the European Commission's Joint Research Centre with the goal of providing pre-harvest estimates of crop yield and production for a number of important crops. CGMS allows regional application of WOFOST by providing a database framework which handles model input (weather, soil, crop parameters), model output (crop simulation results such as total biomass and leaf area index), aggregation to statistical regions and yield forecasting (Genovese 1998, Boogaard et al. 2002).

The WOFOST model was selected for implementation in CGMS on both scientific and practical grounds. From a scientific perspective the model has a solid biophysical basis and included many algorithms that were developed and tested in earlier 'School of C. T. de Wit' crop models (see Bouman et al. 1996, van Ittersum et al. 2003 for an overview). The biophysical basis allowed the model to be applied over a large range of climatic and management conditions. From a practical perspective, at the time of the CGMS development in the early 1990s, WOFOST was already a fully developed simulation product with a mature, open source code base and proper version management. In addi-

tion, the model was designed to simulate a large range of crop types with a single codebase by changing crop parameters only. Moreover, calibrated crop parameters for various crops were already available based on previous work. In contrast, many contemporary models used different versions of the codebase (with varying degrees of modification) to simulate different crop types. Particularly, the single code base approach made the WOFOST model much easier to implement within CGMS.

WOFOST is a mechanistic crop growth model that describes plant growth by using light interception and CO_2 assimilation as growth driving processes and by using crop phenological development as a growth controlling process. The model can be applied in 2 different ways: (1) a potential mode, where crop growth is purely driven by temperature and solar radiation and no growth-limiting factors are taken into account; (2) a water-limited mode, where crop growth is limited by the availability of water. The difference in yield between the potential and water-limited modes can be interpreted as the effect of drought. Currently, no other yield-limiting factors (nutrients, pests, weeds, farm management) are taken into account.

The CGMS setup for Europe was implemented on a grid size of 25 × 25 km and includes the 27 member countries of the EU (EU27), Switzerland, the Balkan countries, the Maghreb, Turkey, Ukraine, Belarus, Moldavia and Russia up to the Ural Mountains. Two CGMS implementations were created which only differed in weather inputs: one system (CGMS-Classic) uses observed weather data from weather stations, and the other (CGMS-ERA) uses weather data derived from ERA-INTERIM.

2.2. Crop and soil inputs

The Soil Geographical Database of Europe (v. 4) was used as basis for implementing the soil-related components in CGMS (Lambert et al. 2003). Criteria were defined to determine which soils are suitable for agriculture. The soil hydraulic properties and maximum rooting depth were derived from the soil descriptions in the database in combination with pedotransfer rules (European Commission 2006).

Crop parameter values and crop calendars were derived from the operational CGMS; these are based on parameter sets initially constructed by Boons-Prins et al. (1993) derived from field trials executed in Belgium, the UK and the Netherlands. These crop parameter values and crop calendars were extended to other regions of Europe based on research executed in the framework of the MARS project (Bignon 1990, Falisse & Decelle 1990, Hough 1990, Russell 1990,

Carbonneau et al. 1992, Narciso et al. 1992, Russell & Wilson 1994). Since new crop varieties are constantly introduced, crop parameters that describe crop growth and development are regularly updated and calibrated as new information comes available (Willekens et al. 1998, GiSAT 2003). For all crop types the planting date of the regional crop varieties have been collected and for some crops that may not reach maturity (i.e. sugar beet, potato, and maize) the end of season has been defined as well. For each crop–region combination a fixed sowing date is assumed during the entire period.

2.3. Weather inputs

2.3.1. Observed weather data

Historical data on climate were provided by the MARS project (Micale & Genovese 2004) of the Institute for the Protection and Security of the Citizen, Joint Research Centre of the European Commission, Ispra, Italy. These data consist of daily values of maximum and minimum temperature, wind speed, global radiation, vapour pressure and rainfall, interpolated from station data to a 25 × 25 km climatic grid (Beek et al. 1992, van der Voet et al. 1994). These station data have been collected from the Global Telecommunication System of the World Meteorological Organisation as well as from national and sub-national station networks.

The data are decoded and quality checked using the AMDaC software package (MeteoConsult 1991). Checking is done by comparing each observation with the corresponding values of surrounding stations and by comparing that particular observation with observations made earlier that day. Obvious observation errors are corrected automatically; other errors are flagged for possible correction by a meteorologist. Meteorological stations with >20% missing data were excluded from the interpolation.

Only meteorological stations deemed suitable were used for interpolation. The suitability was established from a list of criteria such as station height, distance from the coast, etc. Since there were often several stations to choose from to perform the interpolation towards a particular grid cell, the number of missing data at grid level was reduced. Presently, data from nearly 7000 stations is available. Of these stations, about 3000 receive daily meteorological information. Missing global radiation values were computed as a function of either sunshine duration, a combination of cloudiness and temperature range or only the temperature range. Other missing data were replaced by long-term average values. From 1975 onward, a more or less complete coverage of Europe is available.

2.3.2. Weather data derived from ERA-INTERIM

ERA-INTERIM is a reanalysis of the global atmosphere since 1989, continuing in real time (Berrisford et al. 2009). The ERA-INTERIM atmospheric model and reanalysis system has a spatial resolution of 0.7° × 0.7° and 60 atmospheric layers. Due to an improved reanalysis system, performance of ERA-INTERIM has improved compared to previous reanalysis data sets such as ERA-40 (ECMWF 2007).

A downscaling procedure was implemented which downscales the 0.7° × 0.7° ERA-INTERIM data set to the 25 × 25 km grid of the CGMS. This downscaling procedure is executed in 3 steps. Firstly, the 3-hourly weather variables from ERA-INTERIM were converted to daily weather variables that are needed for the crop simulation model using the following approach. Daily total precipitation was derived from the 3-hourly time slots spanning 06:00 to 06:00 h UTC of the following day, in order to be compatible with precipitation estimates from weather stations that report every 12 h (usually at 06:00 and 18:00 h UTC). Daily minimum (maximum) temperature was derived as the minimum (maximum) temperature between 18:00 and 06:00 h (06:00 and 18:00 h) UTC. The remaining variables (early morning vapour pressure, total irradiation and mean wind speed) were derived over the 00:00 to 24:00 h UTC window.

Secondly, an inverse distance weighting was applied which estimates the value of each weather variable at a given CGMS grid as the weighted average of the corresponding variable at the 4 surrounding ERA-INTERIM grid nodes. Finally, variable-specific corrections were applied. The parameters in the correction were derived from regression between time-series of interpolated variables and observed variables at weather stations over Europe. This correction also takes into account the systematic offsets caused by differences in elevation between the ERA-INTERIM model grid and the CGMS model grid.

For radiation and rainfall no corrections were applied. For the radiation this was due to the limited number of observations available, while for rainfall an independent validation demonstrated that correction did not improve the quality of the rainfall estimates (JRC 2006).

2.4. Crop simulations and spatial aggregation

Both CGMS-ERA and CGMS-Classic were used to generate crop simulation results at grid level for the entire spatial domain over the period 1989–2007, which corresponds with the ERA-INTERIM archive that was available. The main outputs from these simu-

lations are the 4 biomass estimates that WOFOST provides and which are relevant for crop yield prediction: total aboveground biomass under potential production conditions, yield (harvestable product) under potential production conditions, total aboveground biomass under water-limited production conditions and yield (harvestable product) under water-limited production conditions.

The simulated crop biomass values on individual grids have to be aggregated to regions in order to establish relationships with yield statistics available from EUROSTAT. Ideally, the cultivated area of each crop should be known per CGMS grid cell in order to aggregate simulated yields to regions. However, crop area estimates are not available at the level of individual CGMS grids or, quite often, administrative regions. Moreover, the regression procedures that are used to derive relationships between simulated and reported yields are sensitive to errors in the time-series of crop area statistics.

Therefore, we decided to substitute crop area with the area of arable land within a grid cell. The area of arable land was derived from the CORINE Land Cover 2000 land cover database for the EU27 area (Nunes de Lima 2005), complemented with Global Land Cover 2000 for areas outside the EU27 (Bartholomé & Belward 2005). Aggregation of simulation results to the European national level was thus performed by weighting on the area of arable land within each CGMS grid cell within a region.

2.5. EUROSTAT reported crop yields

Official reported crop yields at national level were derived from EUROSTAT over the period 1989–2007 for 14 crops (Table 1) and all E27 countries (EUROSTAT 2005). Not all crops listed in Column 1 of Table 1 are modelled explicitly by CGMS, and in a number of cases one simulated crop is used as predictor for a family of crops.

In the present study we assume that the crop yield as reported by EUROSTAT represents the true crop yield for a country. However, uncertainties in the EUROSTAT statistics probably vary a lot between countries as a result of the differences in methodologies used to derive the yield estimates that are reported to EUROSTAT. In order to remove some obvious errors and inconsistencies, we applied the following procedure: (1) records which contained zero values or null values were deleted from the databases; (2) records where subsequent years had exactly the same numeric value were deleted from the database; and (3) a visual analysis was carried for all countries and crop types and outliers were deleted from the database. Outliers were

Table 1. Overview of the European Statistical Office (EURO-STAT) crop types and the European Crop Growth Monitoring System (CGMS) simulated crops that were used as predictors

EUROSTAT crop	CGMS crop
Wheat	Winter wheat
Soft wheat	Winter wheat
Durum wheat	Winter wheat
Barley	Winter wheat
Winter barley	Winter wheat
Spring barley	Barley
Grain maize	Grain maize
Field peas	Field beans
Field beans	Field beans
Oil seed rape	Oil seed rape
Turnips (rape)	Oil seed rape
Sunflower	Sunflower
Sugar beets	Sugar beets
Potato	Potato

typically years where yield was more than double or less than half the average yield, except for years where there were strong indications for exceptionally low or high yields (such as exceptional drought years).

2.6. Regression analysis and forecast error retrieval

Fig. 1 shows the CGMS water-limited crop yield at the end of the growing season and the EUROSTAT reported yield for wheat in Spain and sugar beet in Germany. From these figures a number of conclusions can be drawn that have an impact on the construction of a regression model to relate CGMS simulation results to EUROSTAT reported yields.

First of all, the reported yields for the 2 examples contain a trend of rising crop yields over time. For wheat in Spain, yields have increased from 1.64 t ha^{-1} (1975–1979 average) to 2.83 t ha^{-1} (2005–2009 average). For sugar beet in Germany, yields have increased from 46.6 t ha^{-1} (1988–1992 average) to 61.8 t ha^{-1} (2005–2009 average). In both countries the trend over the entire window is significant at $\alpha = 0.001$. This trend is generally called the technology trend and is caused by improved agricultural practices (with regard to yield) over time. The CGMS simulated results do not contain such a trend, as the technology level is assumed stable.

Secondly, there is a considerable mismatch in absolute terms between CGMS simulated yields and EUROSTAT reported yields. For wheat in Spain, the average simulated yield is 5.1 t ha^{-1} and the average reported yield is 2.3 t ha^{-1}. For sugar beet in Germany, the average simulated yield is 13.0 t ha^{-1} and the average reported yield is 54.4 t ha^{-1}. For wheat in Spain, the difference is caused by sub-optimal management
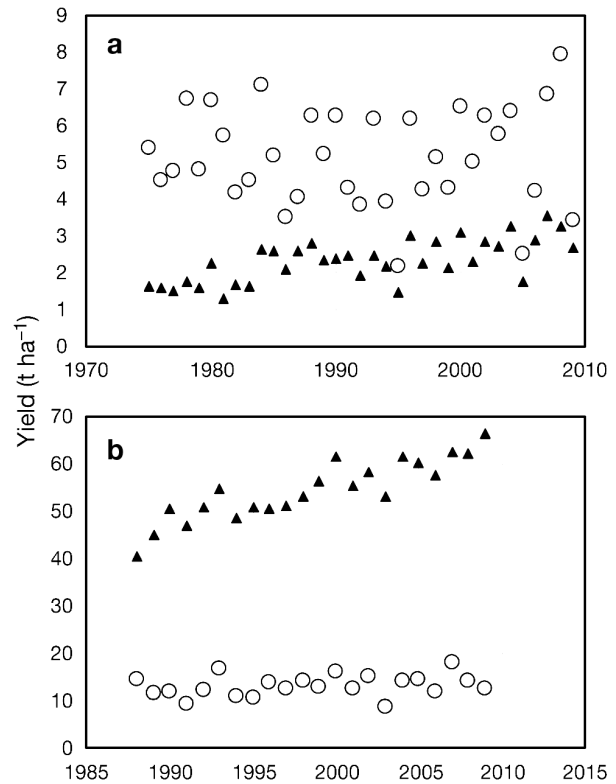

Fig. 1. Time-series of the European Crop Growth Monitoring System (CGMS) simulated yields (○) and the European Statistical Office (EUROSTAT) reported yields (▲) for (a) wheat in Spain and (b) sugar beet in Germany

by farmers, while WOFOST assumes that management is optimal in terms of nutrients, pests and diseases. For sugar beet in Germany, the difference is mainly caused by the water content of the reported yield (fresh weight), while WOFOST predicts dry weight.

Finally, the interannual variability matches quite well between the time-series of reported and simulated yields. The extreme years can be recognized in both time-series easily: in Spain the 1995 and 2005 drought years, and in Germany the 2003 and 2006 drought years and the favourable years 1993 and 2000. Nevertheless, particularly for wheat in Spain, the CGMS simulated yields show much larger fluctuations than the reported yields.

Based on the above considerations, it is clear that a regression model that predicts reported yields based on CGMS simulation results should correct for the difference in yield level, map fluctuations in simulated yield to fluctuations in reported yield and include a trend model that takes the technology trend into account.

However, the selection of a trend model is not always obvious because the trend is not necessarily stable over the entire window. Moreover, a smooth trend of any type over a large number of years assumes a continuity which might be unrealistic (Vossen 1992). For example,

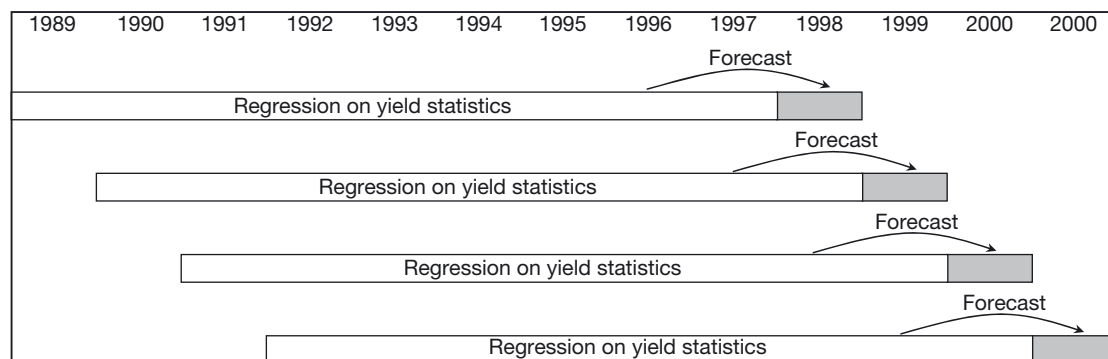| 1989 | 1990 | 1991 | 1992 | 1993 | 1994 | 1995 | 1996 | 1997 | 1998 | 1999 | 2000 | 2000 |

Fig. 2. Schematic representation of the sliding windows approach with a one-year-ahead forecast

it is well known that the 1992 reform of the European Agricultural Policy has caused a breach of trend in the reported crop yields in many countries. Therefore, the regression model should only be based on data from a limited number of years preceding the year for which prediction is to be carried out. Nevertheless, the series should be long enough to give a sufficient number of degrees of freedom in the regression analysis.

In the present study we did not analyse the length of the time-series and the type of trend that should be applied, but we conformed to the operational practice in use at the Joint Research Centre. Therefore, we used a time-series of reported crop yields for a given crop and country over a 9 yr period to determine a linear technology trend assuming a stable trend over this period. The residual variation around the trend is then to be explained by the crop yield simulation results (yield predictors) provided by CGMS-ERA and CGMS-Classic using simple linear regression. All predictors were used individually for regression analyses for both CGMS-ERA and CGMS-Classic. We did not apply multiple linear regression using multiple CGMS predictors because of the risk of high collinearity between predictors.

The following equation describes the regression model:

$$O_y = \overline{O} + b(y - \overline{y}) + c(P_Y - \overline{P}) + \varepsilon$$

where O_y is the reported yield O for year y, $\overline{O}$ is the average reported yield, $b(y - \overline{y})$ is the technology trend where b is the yearly increase/decrease in official yield, $c(P_Y - \overline{P})$ is the modification of the trend by CGMS/CGFS predictor P and regression coefficient c and ε is the residual error.

Since CGMS produces dekadal[1] crop yield predictors during the growing season, the regression procedure described above is repeated for each dekad in the growing season and a separate regression model is established at each dekad. Crop yield forecasting can be carried out by extrapolating the trend for the following year (the target year) and by feeding the WOFOST

simulation results for the target year into the regression models to predict the deviation from the trend for each dekad in the target year.

The CGMS-ERA and CGMS-Classic forecast errors were derived using the so-called one-year-ahead prediction in combination with a sliding window analysis (Fig. 2). In this approach the regression models established in the preceding time-series are used to hindcast the crop yield for the target year. The official reported EUROSTAT crop yield for the target year is then used to calculate the error. This approach has several advantages:

1. The error derived from the one-year-ahead forecast represents the most faithful error estimate as it mimics the operational conditions and takes the error from extrapolating the trend or average into account.

2. The analysis window (1989–2007) is large enough to allow 10 sliding windows of 9 yr and thus the collection of reliable error statistics (n = 10).

3. The forecast error can be easily calculated and summarized over the entire period (in contrast to, for example, *T*-values which express the significance of the regression).

4. It can be easily compared to the baseline forecast based on the error provided by using the (extrapolated) trend as one-year-ahead forecast.

Note that no tests were carried out on the significance, stability or performance of the regression model derived for each sliding window. We assumed that for cases where CGMS cannot significantly explain the interannual variability in crop yield, the resulting regression model will have poor performance when forecasting the yield for the target year as well. Cases with poor CGMS performance can then be easily recognized because the forecast error is usually larger than or close to the baseline forecast.

---

[1]The use of the term dekad refers to an FAO convention in order to distinguish 10 yr periods (decade) from 10 d periods (dekad)

2.7. Summarizing the forecast error

For analyzing the yield forecasts for each predictor, we calculated the root mean square error (RMSE) which is a widely used metric for model testing (Willmott et al. 1985). The RMSE between the forecasted yield and the official reported yield over all available years was calculated as:

$$\mathrm{RSME} = \sqrt{\frac{1}{n} \sum_{y=1998}^{y=2007} (F_y - O_y)^2}$$

where F_y is the forecasted crop yield for a particular crop yield predictor, O_y is the official reported crop yield from EUROSTAT for year y and n is the number of forecasts. We repeated this approach over all dekads in time and over all 9 cases: 8 predictors (4 predictors for CGMS-ERA and 4 for CGMS-Classic) and the control case without predictors (trend only).

We then visualized the crop yield forecasting performance by plotting the RMSE on the *y*-axis against the dekad through the year on the *x*-axis for each available combination of crop and country. Moreover, maps were constructed which show the spatial patterns in forecast accuracy and performance of CGMS-ERA, CGMS-Classic and the baseline predictor. Finally, the results were summarized by counting in how many cases the CGMS-ERA and CGMS-Classic predictors could improve the baseline predictor with a defined margin.

3. RESULTS

3.1. Selected examples of forecast error during the growing season

The total number of crop–country combinations is too large (265) to discuss each combination individually. However, we will discuss some examples that provide insight into the performance and ambiguities that arise from this analysis.

The first example is shown in Fig. 3 where the forecast error for wheat in Spain is plotted. Winter wheat in Spain is a fairly ideal case because the year-to-year weather variability is large and the crop is mainly grown under rain-fed conditions. From the example it is clear that the water-limited crop yield predictors from both CGMS-ERA and CGMS-Classic start to correlate with the official reported crop yields around Dekad 12 as the blue and green lines are clearly below the error of the trend-only case (black dotted line). From this dekad onward in time, the CGMS simulation results make a clear contribution to reducing the forecast error and by Dekad 15 the forecast error stabilizes.

On average, the forecast error could be reduced from 0.6 t ha for the trend-only case to 0.35 t ha^{-1} for the trend plus the CGMS predictor. The figure also demonstrates that the CGMS predictors related to potential production conditions are not correlated with the reported crop yield statistics, as they show no (or hardly any) improvement of forecasting error beyond the trend-only case.

A second example is provided by Fig. 4 where the forecast error for sunflower in Austria is plotted. It is clear that in this example there is no correlation between any of the CGMS predictors and the official reported yields as no predictor performs substantially better than the trend-only case (black dotted line). This example shows that unstable behaviour in the forecast error sometimes occurs at the beginning of biomass accumulation (Dekad 12 for total biomass, Dekad 17 for storage organs). This is caused by the relatively small absolute amount of biomass available at that point in
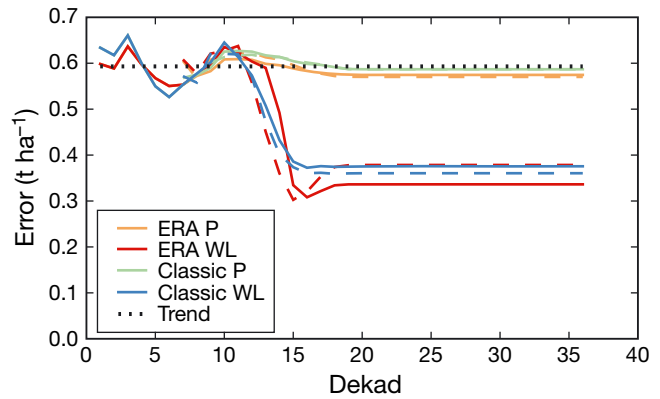
Fig. 3. Temporal evolution of the forecast error for wheat in Spain (n = 10). Solid lines: predictor 'crop total aboveground biomass' for potential (P) and water-limited (WL) production levels. Dashed lines: results for the predictor 'crop yield'

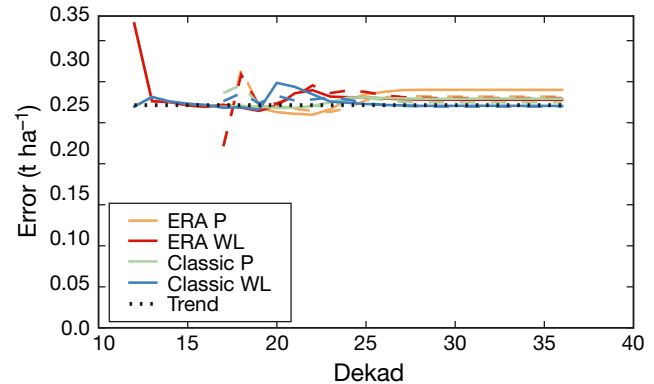Fig. 4. Temporal evolution of the forecast error for sunflower in Austria (n = 10). Solid lines: predictor 'crop total aboveground biomass' for potential (P) and water-limited (WL) production levels. Dashed lines: results for the predictor 'crop yield'
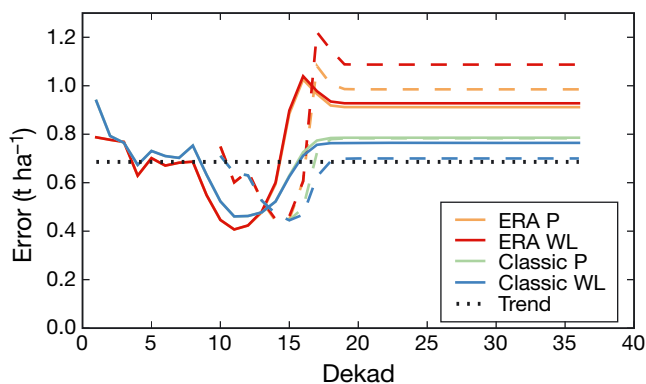
Fig. 5. Temporal evolution of the forecast error for winter barley in Bulgaria (n = 9). Solid lines: predictor 'crop total aboveground biomass' for potential (P) and water-limited (WL) production levels. Dashed lines: results for the predictor 'crop yield'

the growing season, which may lead to very large relative differences in biomass from year to year, which in turn causes large swings in the forecasted crop yield.

Finally, the course of the forecast error for winter barley in Bulgaria is shown in Fig. 5. This example is not so easy to interpret: the predictors of water-limited total biomass (solid blue and green lines) provide minimal forecast error around Dekad 11, after which the forecast error increases; and the water-limited yield predictors (dashed blue and green lines) provide minimal forecast error around Dekad 15, after which they increase again. At the end of the growing season, no predictor is substantially better than the trend-only case (black dotted line).

In the case of winter barley in Bulgaria, both systems show variable performance through the growing season, and further investigation is needed to identify the cause of forecasting error variability. Other cereals in Bulgaria (wheat, soft wheat, barley) show a similar pattern compared to Fig. 4, which is striking given that the simulation results for winter wheat are used to predict crop yield for all 4 crops (wheat, soft wheat, spring barley and winter barley). Therefore, the crop calendar and/or parameterization of the WOFOST model may need to be adapted. Another cause may be overestimation of soil rooting depth. Currently the average ratio of potential to water-limited biomass at the end of growing season for winter wheat in Bulgaria is 0.982, showing that water limitation hardly plays a role. Thus the system is rather insensitive to variability in rainfall and drought effects which often play a role at the end of the growing season.

3.2. Maps of forecast error

A straightforward way to summarize the results would be to visualize the lowest forecast error at the end of the growing season per crop, country and

system. However, the examples in the previous section have demonstrated that the results from the analysis are sometimes not straightforward to interpret. The example of winter barley in Bulgaria demonstrates that the forecast error at the end of the season is not better than the trend-only case and, therefore, we would erroneously conclude that both systems have no forecast skill at all.

To overcome this deficiency we have summarized the results in 2 different ways: (1) the predictor with the smallest forecast error at the end of the growing season (Dekad 36) was taken for CGMS-ERA and CGMS-Classic and compared with the trend; and (2) the predictor with smallest forecast error over the entire season was taken for CGMS-ERA and CGMS-Classic and compared with the trend.

Further, both analyses were not only carried out in absolute terms (t ha^{-1}), but also relative to the error of the trend. In this analysis the forecast error could be expressed as a percentage error of the trend-only case. We used this percentage error to only select those cases where the contribution of the CGMS predictors could reduce the forecast error from the trend-only case by >10%. This analysis is helpful in discriminating those cases where the CGMS predictors do have the lowest absolute forecast error, but the difference between the baseline and the CGMS forecast error is marginal.

The results are displayed as 4 maps (Fig. 6), where the upper 2 maps display the absolute forecast error at the end of the growing season and the smallest absolute forecast during the growing season. Error bars provide an indication of the forecast error for the predictor with the smallest error for CGMS-Classic (blue) and CGMS-ERA (red) as well as the forecast error of the trend only (cream). Regions are coloured according to the system with the overall lowest error. The lower 2 maps display the absolute and lowest forecast errors relative to the trend. Regions where CGMS-Classic or CGMS-ERA forecasting skill is able to reduce the error of the trend-only by >10% are coloured according to the system with lowest error.

The 4 maps summarize information about the stability and significance of the forecast. In the case of a well-behaved forecast error, as for winter wheat in Spain (Fig. 3), all 4 maps will show the same pattern since the forecast error at the end of the growing season is (nearly) equal to the smallest forecast error and all forecasts improve the trend-only forecast by >10%. In the case of awkward forecast behaviour, such as for winter barley in Bulgaria (Fig. 5), the 4 maps will show different patterns as the forecast error at the end of the growing season is much larger than the smallest forecast error during the season. We have chosen to discuss the spatial distribution of forecast error for winter
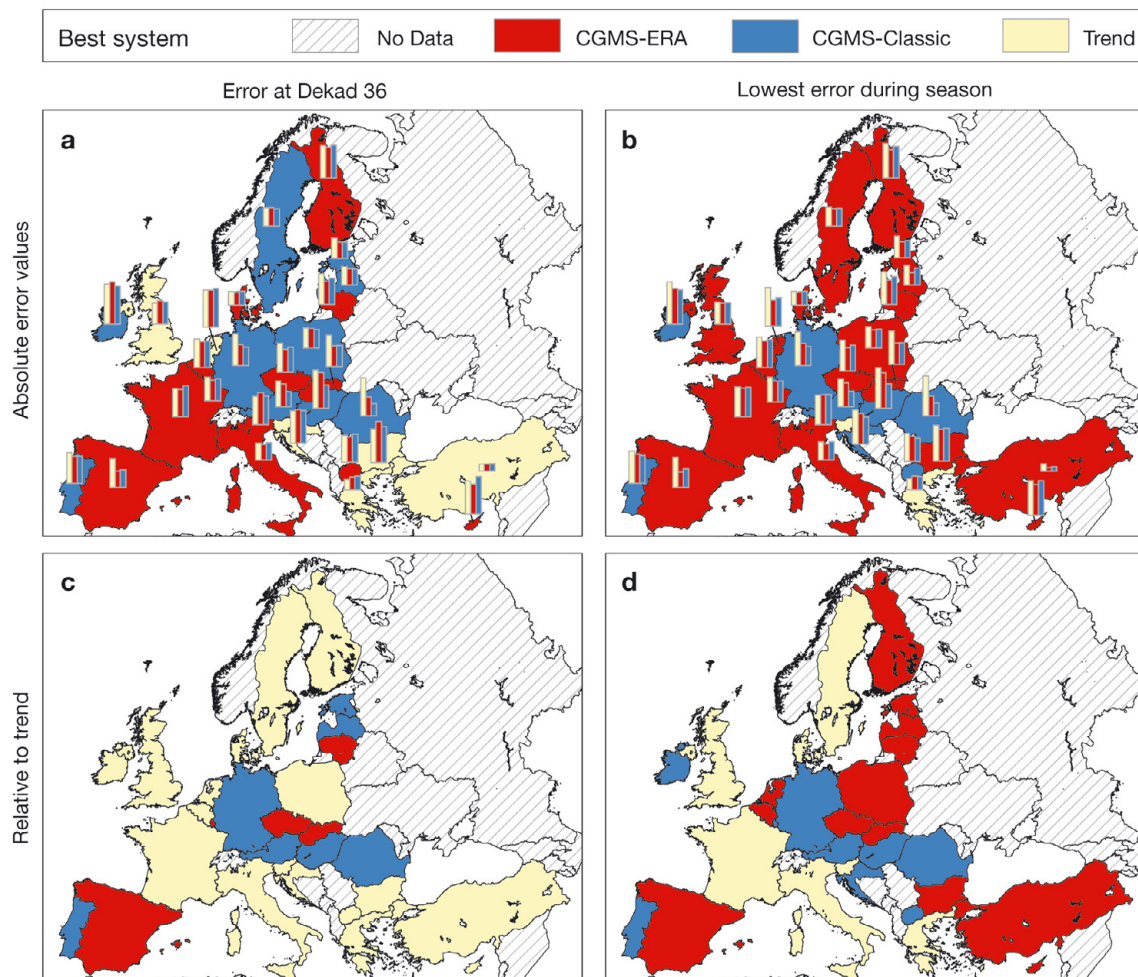
Fig. 6. Overview maps of crop yield forecast root mean square error for wheat. Regions are coloured according to the best performing system, error bars show the forecast error. (a) Absolute error at the end of the season; (b) lowest absolute error during the growing season; (c) regions where the absolute error at the end of the season is at least 10% better than the trend; and (d) regions where the lowest absolute error during the growing season is at least 10% better than the trend

wheat and grain maize because of the contrasting results for these 2 crops.

The maps of forecast error for wheat show a mixed pattern with regard to system performance (Fig. 6). With regard to the absolute error at Dekad 36, CGMS-ERA performs best in southwestern Europe (France, Italy and Spain), while CGMS-Classic performs best in central Europe (Germany, Poland, Austria, Hungary and Romania). Both CGMS-Classic and CGMS-ERA do not have any forecasting skill in the UK, Greece, Turkey, the Netherlands or Bulgaria. When looking at the map of lowest error, in the UK, Bulgaria, Turkey and the Netherlands CGMS-ERA now shows forecasting skill, while Sweden, the Baltic States and Poland flipped from CGMS-Classic to CGMS-ERA as the best performing system.

The maps with forecast error relative to the trend illustrate in which countries the forecast skill has at

least 10% lower error than the trend-only case. The results demonstrate that stable skill is reached for Spain, Portugal, Lithuania and the central European countries (Germany, Austria, Hungary, Czech Republic, Slovakia and Romania).

Analyses of the countries where the system has conflicting behaviour show that in some countries the system suffers from unstable forecast skill, which means that the forecast skill at some point during the growing season is better than the end-of-season forecast skill (Bulgaria, Cyprus, Estonia, Finland, Ireland and the Netherlands). In other countries the systems simply do not show any significant forecasting skill (Belgium, France, Greece, Croatia, Italy, Latvia, Poland, Sweden, Slovenia and the UK).

In general, forecasting skill for grain maize is high and stable for most countries (Fig. 7), and improves beyond 10% of the baseline forecast. Notable excep-
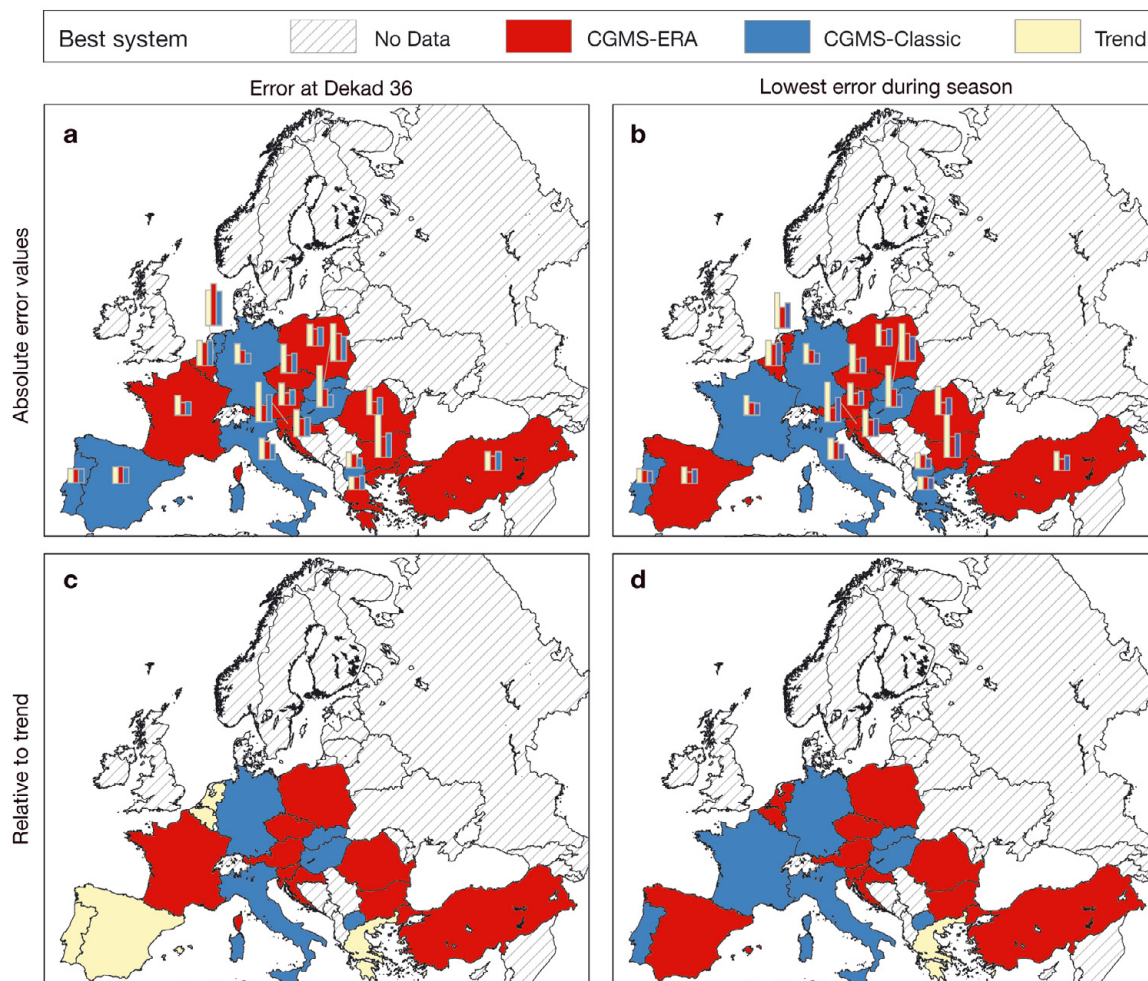
Fig. 7. As Fig. 6 but for grain maize

tions are Greece, Spain and Portugal, most likely because grain maize is irrigated in those countries. Also, in Belgium and the Netherlands the system shows either limited skill (Belgium) or instabilities in the forecast (the Netherlands). No clear pattern is visible in terms of whether CGMS-Classic or CGMS-ERA performs best. In countries with high skill, both systems generally have good performance.

3.3. Summarizing the results

The results were summarized per crop type by counting in how many countries CGMS-Classic and CGMS-ERA outperformed the trend-only case by 10 and 25% (Table 2). The table summarizes the forecasting error at the end of the growing season. Therefore, the reported statistics are a conservative estimate of the system performance because it was demonstrated that for many countries the forecasting error at the end of the growing season is not the lowest error.

With a threshold of 10%, CGMS-ERA and CGMS-Classic demonstrate yield forecasting skill in 101 and 98 out of 265 cases, respectively (38.1 and 37.0%). There are some small differences between the 2 systems, notably CGMS-ERA clearly performs better than CGMS-Classic for barley (15 vs. 11 cases). However, the overall pattern is very similar between the 2 systems.

From Table 2, we can derive the performance of CGMS-Classic and CGMS-ERA per crop type. Performance is relatively high (≥40% skill) for grain maize, sugar beets, field beans (only 13 cases), barley, sunflower and soft wheat. Performance is poor (<20% skill) for oil seed rape, turnips (rape) and field peas. The remaining crops are intermediate.

If we increase the threshold for performance to 25% improvement over the trend-only case, then the differences between CGMS-ERA and CGMS-Classic are larger and in favour of CGMS-Classic. CGMS-ERA shows yield forecasting skill in 45 out of 265 cases (17.0%) and CGMS-Classic shows skill in 52 out of

Table 2. Overall performance (% skill) of CGMS-Classic and CGMS-ERA based on forecast error at the end of the growing season, where forecast skill improves the trend by at least 10%, or at least 25%. Also shown is the number of countries (N) for which each system had the lowest error per crop type and the total number of countries where this crop is cultivated

Crop	CGMS-ERA		CGMS-Classic		Total
	N	Skill (%)	N	Skill (%)	
≥10%					
Wheat	10	34.5	11	37.9	29
Soft wheat	10	40.0	10	40.0	25
Durum wheat	1	12.5	3	37.5	8
Barley	15	51.7	11	37.9	29
Winter barley	5	27.8	5	27.8	18
Spring barley	6	33.3	6	33.3	18
Grain maize	14	73.7	13	68.4	19
Field peas	0	0.0	0	0.0	10
Field beans	7	53.8	7	53.8	13
Oil seed rape	3	15.8	3	15.8	19
Turnips (rape)	3	16.7	3	16.7	18
Sunflower	6	42.9	6	42.9	14
Sugar beets	12	54.5	12	54.5	22
Potato	9	39.1	8	34.8	23
Total	101	38.1	98	37.0	265
≥25%					
Wheat	4	13.8	6	20.7	29
Soft wheat	4	16.0	5	20.0	25
Durum wheat	1	12.5	0	0.0	8
Barley	4	13.8	6	20.7	29
Winter barley	1	5.6	2	11.1	18
Spring barley	2	11.1	2	11.1	18
Grain maize	10	52.6	12	63.2	19
Field peas	0	0.0	0	0.0	10
Field beans	1	7.7	2	15.4	13
Oil seed rape	2	10.5	2	10.5	19
Turnips (rape)	2	11.1	2	11.1	18
Sunflower	3	21.4	3	21.4	14
Sugar beets	5	22.7	4	18.2	22
Potato	6	26.1	6	26.1	23
Total	45	17.0	52	19.6	265

265 cases (19.6%). CGMS-Classic outperforms CGMS-ERA for sugar beet (3 cases), wheat (2 cases), barley (2 cases), maize (2 cases) and durum wheat (1 case). CGMS-ERA outperforms CGMS-Classic only for potato (1 case).

Finally, we aggregated all crop–country combinations and displayed the forecast error relative to the trend-only case as a cumulative frequency distribution (Fig. 8). The results confirm that CGMS-Classic slightly outperforms CGMS-ERA in the crop–country combinations where skill is high (left side of figure), but in the vast majority of crop–country combinations there is no substantial difference in forecasting skill.

4. DISCUSSION AND CONCLUSIONS

The overall conclusion is that CGMS-Classic and CGMS-ERA perform very similarly in terms of crop
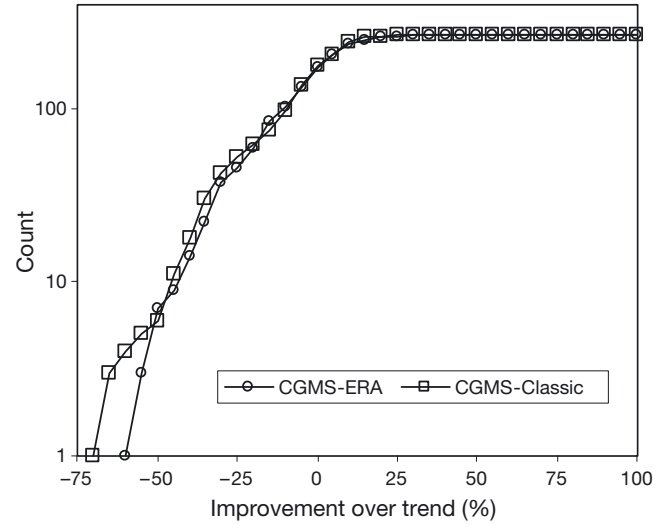

Fig. 8. Cumulative frequency distribution of forecast error over all crop–country combinations relative to the baseline forecast (trend only). Negative values imply lower forecast error than the baseline forecast, positive values imply higher forecast error

yield forecasting skill. The only difference between CGMS-ERA and CGMS-Classic is that CGMS-ERA uses weather variables derived from ERA-INTERIM as input to the WOFOST crop model, while CGMS-Classic uses interpolated observed weather variables. This implies that the deficiencies which are often observed in weather variables derived from NWP models (e.g. failure to reproduce rainfall intermittency characteristics) do not significantly affect the year-to-year variability in crop simulation results. Therefore, ERA-INTERIM is a suitable replacement for observed weather data over Europe. Moreover, the good results obtained over Europe indicate that ERA-INTERIM may be used to implement regional crop yield forecasting systems over regions with sparse coverage of weather stations, although in such regions the quality of ERA-INTERIM may be lower due to the smaller number of observations that could be taken into account in the reanalysis.

ERA-INTERIM will be maintained by ECMWF and updated weather records will become available with a delay of approximately 1 mo. This means that ERA-INTERIM does not provide weather variables in near real-time and CGMS-ERA therefore relies on weather variables from the operational ECMWF NWP model (ECMWF-OPE). However, if ECMWF-OPE has different properties (e.g. monthly total precipitation, total solar radiation) compared to ERA-INTERIM, then this implies that the WOFOST simulations in the forecast year may be biased compared to the simulations in the historic archive upon which the forecast regression equations are based.

Accordingly, future research should address if the yield forecasting performance is affected by a breach of trend in the WOFOST simulations caused by the transition of ERA-INTERIM to ECMWF-OPE in the target year. Such an analysis would require an overlapping archive of ECMWF-OPE and ERA-INTERIM data, which is not yet available. However, we do want to stress the need for an evaluation of the properties of ECMWF-OPE and ERA-INTERIM in order to evaluate the influences of possible biases in the WOFOST simulations in the target year.

In this analysis we used the database of observed weather data as a reference to evaluate the ERA-INTERIM database for crop yield prediction. Implicitly we assume that the observed data are of better quality. However, this assumption is not justified for all variables. Global incoming radiation is usually predicted fairly accurately by NWP models, while it is poorly estimated by CGMS (de Wit et al. 2005), principally using estimates of sunshine duration (20% of stations) or cloud cover and temperature (76% of stations) based on a procedure described by Supit & Van Kappel (1998). Although the limited accuracy of CGMS radiation estimates has been criticized, our results indicate that the impact of errors in radiation estimates on the CGMS yield forecasts is small.

The results of the present study provide valuable background information on the accuracy and precocity of CGMS for regional yield prediction over Europe. However, we do want to stress that the yield forecasting accuracy presented here is not representative of operational conditions. The analyses carried out for the present study consist of hundreds of thousands of regressions (per country, crop, predictor, sliding window and dekad) which prevent optimizing settings on the regression model (length of time-series, trend model, etc.) for specific cases. In this respect, it is likely that a skilled operator who makes a careful analysis of the length of time trend, the type of trend model used (none, linear or quadratic) and the impact of different predictors on the regression model will be able to obtain better forecasts than those described here.

Finally, we conclude that there is a need to improve the crop calendar and/or calibration of CGMS in general, and for some crops in particular. Barley/winter barley shows unstable forecasting skill in a number of countries where forecasting skill is high before harvest but is completely lost at the end of the season. Given that barley/winter barley is predicted with the simulation results for winter wheat, it may be worthwhile to implement winter barley as a separate crop type. Field peas show no forecasting skill at all. As field peas are currently predicted with the simulation results from field beans, this may indicate the need to implement a true field pea crop in CGMS.

LITERATURE CITED

Bartholomé E, Belward AS (2005) GLC2000: a new approach to global land cover mapping from Earth observation data. Int J Remote Sens 26:1959–1977

Bates BC, Charles SP, Hughes JP (1998) Stochastic downscaling of numerical climate model simulations. Environ Model Softw 13:325–331

Beek EG, Stein A, Jansen LLF (1992) Spatial variability and interpolation of daily precipitation amount. Stoch Hydrol Hydraul 6:304–320

Berrisford P, Dee D, Fielding K, Fuentes M, Kallberg P, Kobayashi S, Uppala S (2009) The ERA-INTERIM archive. European Centre for Medium-range Weather Forecasting, Shinfield Park, Reading

Bignon J (1990) Agrométéorologie et physiologie du maïs grain dans la communauté Européenne. Report No. EUR 13041 FR. Office for Official Publications of the EU, Luxembourg

Boogaard HL, Eerens H, Supit I, van Diepen CA, Piccard I, Kempeneers P (2002) Description of the MARS Crop Yield Forecasting System (MCYFS). Study contract no. 19226-2002-02-F1FED ISP. Joint Research Centre, Ispra

Boons-Prins ER, de Koning GHJ, van Diepen CA, Penning de Vries FWT (1993) Crop-specific parameters for yield forecasting across the European Community. Simulation Reports CABO-TT 32. Wageningen

Bouman BAM, van Keulen H, van Laar HH, Rabbinge R (1996) The 'School of de Wit' crop growth simulation models: a pedigree and historical overview. Agric Syst 52: 171–198

Carbonneau A, Riou C, Guyon D, Riom J, Schneider C (1992) Agrométéorologie de la vigne en France. Report No. EUR 13911 FR. Office for Official Publications of the EU, Luxembourg

Charles SP, Bates BC, Smith IN, Hughes JP (2004) Statistical downscaling of daily precipitation from observed and modelled atmospheric fields. Hydrol Process 18:1373–1394

de Wit AJW, Boogaard HL, van Diepen CA (2005) Spatial resolution of precipitation and radiation: the effect on regional crop yield forecasts. Agric For Meteorol 135:156–168

Easterling WE, Weiss A, Hays CJ, Mearns LO (1998) Spatial scales of climate information for simulating wheat and maize productivity: the case of the US Great Plains. Agric For Meteorol 90:51–63

ECMWF (European Centre for Medium-range Weather Forecasting) (2007) ERA-Interim: new ECMWF reanalysis products from 1989 onwards. ECMWF Newsl 110:25–35

European Commission (2006) New Soil Information for the MARS Crop Yield Forecasting System, Vol VIII. European Commission, Luxembourg

EUROSTAT (European Statistical Office) (2005) European regional and urban statistics: reference guide. EUROSTAT, Luxembourg

Falisse A, Decelle Y (1990) Froment, avoine; etude effectuée par la faculté des sciences Agronomiques de Gembloux (Belgique), Projet pilote de télédétection appliquée aux statistique agricoles. Joint Research Centre, Ispra

Genovese GP (1998) The methodology, the results and the evaluation of the MARS crop yield forecasting system. In: Rijks D, Terres JM, Vossen P (eds) Agrometeorological applications for regional crop monitoring and production assessment. Report No. EUR 17735 EN. Office for Official Publications of the EU, Luxembourg, p 67–119

GISAT (2003) Studies in support to the MARS project. Crop monographies on candidate countries. MOCA Study. Final report, Contract No. 20363-2002-12 F1ED ISP CZ, GISAT, Prague

Hansen JW, Jones JW (2000) Scaling-up crop models for climate variability applications. Agric Syst 65:43–72

Hough MN (1990) Agrometeorological aspects of crops in the United Kingdom and Ireland. A review for sugar beet, oilseed rape, peas, wheat, barley, oats, potatoes, apples and pears. Report No. EUR 13039 EN. Office for Official Publications of the EU, Luxembourg

JRC (Joint Research Centre) (2006) Verification of downscaling methods for CGMS numerical weather based. Joint Research Centre, Ispra

Lambert JJ, Daroussin J, Eimberck M, Le Bas C, Jamagne M, King D, Montanarella L (2003) Soil geographical database for Eurasia and the Mediterranean: instructions guide for elaboration at scale 1:1,000,000. Version 4.0. Report No. EUR 20422 EN. Office for Official Publications of the European Communities, Luxembourg

MeteoConsult (1991) AMDaC system manual. MeteoConsult, Wageningen

Micale F, Genovese G (eds) (2004) Meteorological data collection, processing and analysis, Vol. 1. Report No. EUR 21291 EN, Office for Official Publications of the EU, Luxembourg

Narciso G, Ragni P, Venturi A (1992) Agrometeorological aspects of crops in Italy, Spain and Greece. A summary review for common and durum wheat, barley, maize, rice, sugar beet, sunflower, soya bean, rape, potato, cotton, olive and grape crops. Report No. EUR 14124 EN, Office for Official Publications of the EU, Luxembourg

Nunes de Lima VM (ed) (2005) IMAGE2000 and CLC2000: products and methods. EUR 21757 EN, Joint Research Centre, Ispra

Robertson AW, Ines AVM, Hansen JW (2007) Downscaling of seasonal precipitation for crop simulation. J Appl Meteorol Climatol 46:677–693

Russell G (1990) Barley knowledge base. Report No. EUR 13040 EN. Office for Official Publications of the EU, Luxembourg

Russell G, Wilson GW (1994) An agro-pedo-climatological knowledge base of wheat in Europe. Report No. EUR 15789 EN. Office for Official Publications of the EU, Luxembourg

Supit I, Van Kappel RR (1998) A simple method to estimate global radiation. Sol Energy 63:147–160

van der Voet P, van Diepen CA, Oude Voshaar J (1994) Spatial interpolation of daily meteorological data: a knowledge based procedure for the regions of the European Community. SC report 53/3, DLO Winand Staring Centre, Wageningen

van Diepen CA, Wolf J, van Keulen H, Rappoldt C (1989) WOFOST: a simulation model of crop production. Soil Use Manag 5:16–24

van Ittersum MK, Leffelaar PA, van Keulen H, Kropff MJ, Bastiaans L, Goudriaan J (2003) On approaches and applications of the Wageningen crop models. Eur J Agron 18:201–234

Vossen P (1992) Forecasting national crop yields of EC countries: the approach developed by the agriculture project. In: Toselli F, Meyer-Roux J (eds) Conference on application of remote sensing to agricultural statistics. Office for Official Publications of the EU, Belgirate, p 159–176

Vossen P, Rijks D (1995) Early crop yield assessment of the EU countries: the system implemented by the Joint Research Centre. Report No. EUR 16318, Publication of the Office for Official Publications of the EC, Luxembourg

Willekens A, van Orshoven J, Feyen J (1998) Estimation of the phenological calendar, Kc-curve and temperature sums for cereals, sugar beet, potato, sunflower and rape seed across Pan Europe, Turkey and the Magreb countries by means of transfer procedures. Contract no. 13118-97-07 F1ED ISP. Joint Research Centre, Ispra

Willmott CJ, Ackleson SG, Davis RE, Feddema JJ and others (1985) Statistics for the evaluation and comparison of models. J Geophys Res 90:8995–9005