Tellus (2005), 57A, 476–487
Printed in UK. All rights reserved

TELLUS

Seasonal weather forecasts for crop yield modelling in Europe

By PIERRE CANTELAUBE and JEAN-MICHEL TERRES*, European Commission Joint Research Centre/Institute for Environment and Sustainability, TP 262, I-21020 Ispra (VA), Italy

(Manuscript received 30 March 2004; in final form 2 December 2004)

ABSTRACT

Within the European DEMETER project, ensembles of global coupled climate models have shown some skill for seasonal climate prediction. Meteorological outputs of the seasonal prediction system were used in a crop yield model to assess the performance and usefulness of such a system for crop yield forecasting.

An innovative method for supplying seasonal forecast information to crop simulation models was developed. It consisted in running a crop model from each individual downscaled member output of climate models. An ensemble of crop yield was obtained and a probability distribution function (PDF) was derived. Preliminary results of wheat yield simulations in Europe using downscaled DEMETER seasonal weather forecasts suggest that reliable crop yield predictions can be obtained using an ensemble multi-model approach. When compared to the operational system, for the same level of accuracy, earlier crop forecasts are obtained with the DEMETER system. Furthermore, PDFs of wheat yield provide information on both the yield anomaly and the uncertainty of the forecast. Based on the spread of the PDF, the user can directly quantify the benefits and risk of taking weather-sensitive decisions.

It is shown that the use of ensembles of seasonal weather forecast brings additional information for the crop yield forecasts and therefore has valuable benefit for decision-making in the management of European Union agricultural production.

1. Introduction

Agriculture is highly dependent on climate and, as such, crop yield variability is affected by year-to-year climatic variability, with regards to both extreme events and changes in historical patterns of regional climate (Hoogenboom, 2000; Ogallo et al., 2000). This vulnerability could have significant effects on crop production because of the many uncertainties for the growing season (Jones et al., 2000), and may result in economic and food security risks in some parts of the world.

During the last decades, however, scientists have shown that, at the global scale, some climate variability is related to large-scale interactions between the oceans and the atmosphere (Wallace and Gutzler, 1981). As a consequence, improvements in the understanding of the role of climatic phenomena, such as the El Niño Southern Oscillation (ENSO), have increased the predictability of climate fluctuations in several parts of the world (Neelin et al., 1998). This predictive capability has improved to such an extent that seasonal time-scale predictions are now made routinely at a number of operational meteorological centres around the world (Palmer et al., 2004).

*Corresponding author.
e-mail: jean-michel.terres@jrc.it

The usefulness for such information (climate knowledge) for many sectors is evident, in particular for agricultural systems.

For Hammer et al. (2001), climate prediction offers considerable opportunities for agricultural decision makers via the possible improvement of the management system (i.e. increased food production/profit, reduced risks, improved food security policy). For Jones et al. (2000), climate predictions of a certain level of accuracy, three to six months ahead of time, would enable managers to take decisions to decrease unwanted impacts and to take advantage of expected favourable conditions.

Sivakumar et al. (2000) stressed that ‘agrometeorologists should make efforts to promote more active use of seasonal to interannual climate forecasts in agricultural planning and operations’ (Agenda 21, global plan of action agreed at the United Nations Conference on Environment and Development). Hansen (2002) recognizes the considerable progress that several groups throughout the world have made in applying seasonal climate prediction in agricultural purpose (notably in food production improvement), and discusses key issues and forthcoming challenges.

These studies show also that agricultural analysts and agrometeorologists have investigated the potential economic value of such information. Solow et al. (1998) found that an increase in forecast accuracy (based on ENSO) has substantial economic value to United States (US) agriculture. Jones et al. (2000) also estimated the potential benefits of climate forecasting to agriculture in the US as very high. Petersen and Fraser (2001) estimated that a seasonal forecasting technology, which would provide a 30% decrease in seasonal uncertainty, would increase annual profits by approximately 5% in Western Australia. Meza and Wilks (2003) investigated the potential economic value of sea surface temperature anomalies forecasts for farmers and agricultural decision-making purposes in Chile, assuming that in the future the likelihood of the prediction of the conditions of the ocean will increase to (sufficient) accurate level. For the African continent, Phillips et al. (1998) note that improvement in climate forecasts (and also in crop simulation models) would potentially reduce agricultural risk associated with climate in Zimbabwe. Nnaji (2001) found that there is a potential for rainfall forecasting in northern Nigeria, to aid decision-making in planting crops.

The seasonal forecasts used in the above studies are often based on ENSO categories. Cataloguing experiences in the application of climate prediction in agriculture, Hammer et al. (2001) note that almost all studies use statistical forecasts based on the identification of analogues, from ENSO categories, as a means to derive seasonal forecasts.

This might be a limitation for the application of climate prediction in agriculture. For instance, in their case study (maize in Zimbabwe), Phillips et al. (1998) found that forecasts based only on ENSO categories did not provide enough quality information for management decision-making. Hammer et al. (2001) remark that the use of ENSO categories forecasts impacts not only the current level of forecasting skill but also the method of forecast delivery. Finally, even perfect ENSO prediction is far from perfect climate prediction (Solow et al., 1998).

Another observation coming from the literature is the very low number of such studies over Europe. This could be explained by the fact that seasonal forecasts are often based directly on ENSO categories and that ENSO impacts mainly the tropics and countries bordering the Pacific Ocean. Thus, the ENSO signal in Europe is either too weak or too difficult to link with agricultural production. North Atlantic and European regional climate variability are associated with other modes of variability, the most important of these being the North Atlantic Oscillation (NAO; Marshall et al., 2001). Pavan et al. (2000) found that the NAO explains approximately 30% of the climate variability over the Euro-Atlantic region. Furthermore, the link between European agricultural production variability and some modes of climate variability (among which is the NAO) was established by Cantelaube et al. (2004). Since European agriculture is mostly intensive as opposed to low-input farming systems, weather remains the main source of uncertainty for crop yield assessment and crop management (Vossen, 1995).

The DEMETER (development of a European multi-model ensemble system for seasonal to interannual prediction) project (http://www.ecmwf.int/research/demeter), coordinated by the European Centre for Medium-Range Weather Forecasts (ECMWF) aims primarily to construct and to evaluate a reliable ensemble-based system for predicting fluctuations in climate on seasonal time-scales (Palmer et al., 2004). The DEMETER approach combines the use of several quasi-independent climatic models to delimit the uncertainties due to model formulation (multi-model) together with the use of perturbed initial parametrizations to account for the uncertainties in initial conditions (ensemble). Predicted day-to-day evolution of the weather is sensitive to initial conditions (Palmer, 1993; Shukla, 1998). Multi-model ensembles are better to produce reliable probability forecasts of seasonal climates than any single model (Doblas-Reyes et al., 2000; Palmer et al., 2000).

In the framework of DEMETER, seven global coupled atmosphere–ocean climate models from different organizations in Europe were incorporated in a unique multi-model ensemble system (six of them installed on a single supercomputer at the ECMWF). An ensemble is composed of 63 members: seven models run with nine different initializations. A set of bias-corrected data (hindcasts) for standard meteorological parameters has been produced and evaluated for the period 1958–2001 (Doblas-Reyes et al., 2005; Hagedorn et al., 2005).

An innovative aspect of DEMETER was the coupling of application models to each individual member of the ensemble. As an application partner, the Joint Research Centre (JRC) uses a crop model for estimating crop yield in Europe, providing the Directorate General for Agriculture (DG AGRI) of the European Commission (EC) with real-time yield estimates during the growing season for the European Union (EU) Member States (see http://mars.jrc.it/marsstat/ Crop_Yield_Forecasting/ crop_yield_forecasting_system.htm; Meyer-Roux and Vossen, 1994; Rabbinge and van Diepen, 2000). Hence, by coupling the JRC crop model directly to the output of climate models, each DEMETER member would provide a yield prediction, which together would result in an ensemble of crop yield predictions; this, in turn, would give a seasonal forecast probability distribution of crop yield in Europe.

With a wheat production around 100 million tons (European Commission, 1999), ranking the EU as the second largest producer in the world after Asia (FAOSTAT 2001), DG AGRI requires timely forecasting of the production as information for management of the Common Agricultural Policy (CAP; European Commission, 1997), for setting quotas of marketable good, for managing level of supply and stocks and for fixing agricultural prices.

Within DEMETER, weather seasonal hindcasts were used as input for the crop model to simulate wheat yield in Europe. The main tasks consisted of assessing reliability of the DEMETER-based crop yield forecasts and, taking into account the probabilistic nature of these forecasts, evaluating the probabilities of yield anomalies. If any improvement in the existing system is found, the possibility of including multi-model seasonal forecasts with routine forecasting system in the field of EU agricultural production could be considered.

2. Materials and methods

2.1. *Joint Research Centre crop model*

The JRC model – the crop growth modelling system (CGMS) – consists of two modules. The first module is based on the crop growth simulation model, WOFOST (Van Diepen et al., 1989), driven by meteorological conditions, crop parameters and environmental factors such as soil characteristics (to simulate the soil water balance). It describes the crop cycle from sowing to maturity on a daily time-scale; crop growth is simulated in combination with phenological development. In the second module, a regression analysis from historical statistical yield data and WOFOST simulated crop growth indicators (biomass, storage organs) is performed to forecast the crop yield of the current year at the level of administrative region (Supit, 1997). The regression analysis takes into account a linear time trend[1] and simulated crop growth as influenced by meteorological conditions.

Olesen and Bindi (2002) list the climatic constraints to European agriculture. Biological systems are based primarily on photosynthesis, and are thus dependent on incoming radiation; however, the potential for production set by the radiation is greatly modified by temperature and rainfall. Temperature mainly affects the duration of the growing period and rainfall (through soil water availability) may affect the duration of growth and the production of the plant (leaf area and the photosynthetic efficiency). Meteorological parameters required as input for the JRC crop model are precipitation, maximum and minimum temperature, global radiation and evapotranspiration (computed here from wind speed, dew point temperature and surface net longwave radiation, using the Penman formula; Penman 1948). These meteorological parameters should be available on a daily basis.

2.2. *Use of seasonal weather forecast to reduce weather uncertainty*

The JRC crop model can make a yield forecast at any time during the growing season. However, if the forecast is issued early in the season, the weather conditions leading up to harvest time are unknown and are therefore a major source of uncertainty. Furthermore, early in the season, simulated crop indicators are still at a very early stage (e.g. tillering for winter wheat) and the regression built on these crop indicators is not robust. In operational mode, JRC yield forecasts generally improve as the growing season progress (Genovese, 1998). Information on main patterns of meteorological conditions several months in advance and their use as input in the crop model could provide better crop growth indicators, which in turn could improve the statistical regression for the yield prediction (strategy illustrated in Fig. 1). If skilful, DEMETER data could provide more reliable crop yield forecasts early in the season, giving valuable information to agricultural production managers.

Weather data from DEMETER climate models being of probabilistic nature, their input in the crop model results in a forecast probability distribution of crop yield. The potential usefulness of the DEMETER system can be evaluated by comparing the hindcasted crop yield probability distribution with official yield figures.

This study was carried out for 4 yr (1995–1998) for winter wheat (*Triticum aestivum* L.). Hindcasted yield was estimated at national level for each of the 15 EU Member States and compared to both real-time yield forecasts issued with the operational JRC system and official yields from the EUROSTAT data base (http://europa.eu.int/comm/eurostat/).

2.3. *DEMETER ensembles and downscaling*

DEMETER data used by the crop model consist of six-month ahead hindcasts (63 members of daily data) covering the period February–July for each growing season of the test period (1995–1998).

Ensembles from DEMETER climate models are issued at a low spatial resolution ($1.5^\circ \times 1.5^\circ$). Representation of orography being very coarse at such a scale, in particular for European conditions, large-scale weather systems do not yield a reliable representation of local weather conditions, notably for precipitation. It was therefore necessary to downscale meteorological data to a finer surface resolution, which provides more reliable regional details. The applied method consists of statistical downscaling of monthly mean values (temperatures, precipitations, radiation and evapotranspiration) using singular value decomposition analysis and model output statistics (method fully described by Feddersen and Andersen, 2005), then redistributed into daily series (for use in the JRC crop model) with a weather-generator based on the Richardson WGEN model (Richardson, 1981). The downscaling method was trained using the ECMWF 40-yr reanalysis data set, ERA40 (http://www.ecmwf.int/research/era/), against the JRC $50 \times 50\ \mathrm{km}^2$ gridded weather data (for the period 1987–1998). All meteorological parameters used as input for the crop model were downscaled by the Danish Meteorological Institute (Feddersen and Andersen, 2005).

The downscaling increased the correlation between ERA40 and JRC sets of daily data, for all four input parameters (Cantelaube and Terres, 2003). As an example, for precipitation in 1988, downscaling reduced the root mean square error

[1] A trend is observed on the national yields time series, due notably to improvement in farm managing practice, technological progress, improved seeds, etc. The yield time series are thus separated into two components: a time trend and an annual variation due to weather conditions. This allows us to define the annual anomalies as the difference between the actual yield value and the trend. The trend is also sometimes used to propose a forecast only based on the past years, independent of climatic conditions.

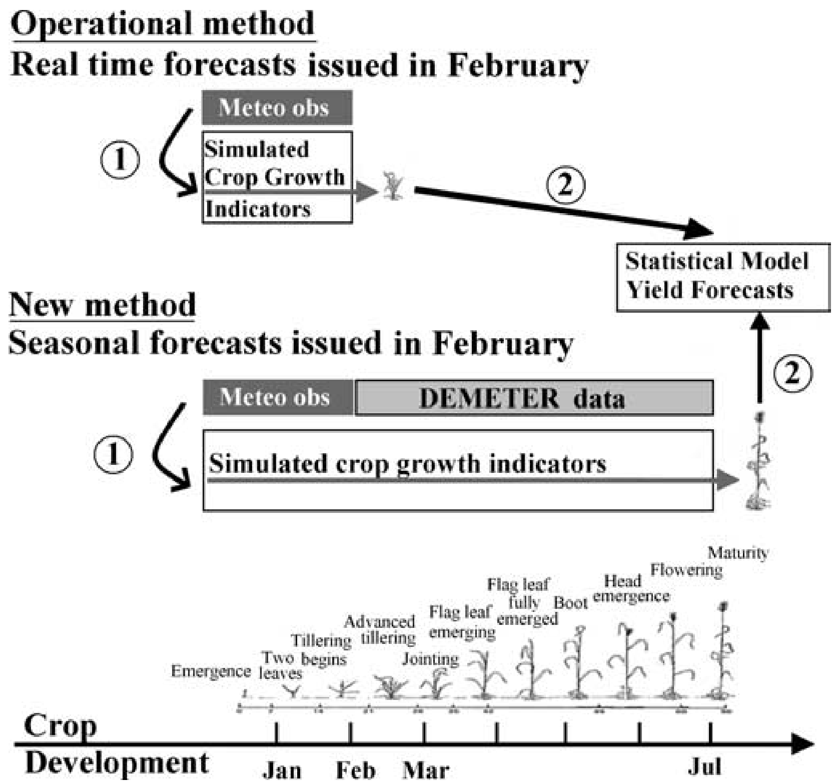

Fig 1. Crop growth monitoring system and yield forecasting system at the JRC: current operational system and new strategy for seasonal forecasting.

(RMSE) by 34% and increased the R^2 value from 0.29 to 0.7 (over Europe; Fig. 2, top). An interesting result was also found in the ability of downscaled data to reproduce realistic sequences of wet and dry days from the monthly values (through the weather generator), which matters for the crop model since it calculates a daily soil water balance (Rijks et al., 1998).

Downscaled weather inputs also improved crop modelling outputs. Discrepancies between simulated crop indicators, using JRC meteorological inputs or ERA40 reanalysis data, were less when the latter were downscaled. Figure 2 (bottom) shows an example of simulated biomass in 1988: the R^2 value increases from 0.62 to 0.69 and the RMSE is reduced by 4% for the whole of Europe (more than 10% in some parts).

For crop yield forecasting in Europe, GCM outputs need to be downscaled to a smaller spatial resolution to take into account local weather conditions. In the following sections, DEMETER ensembles will refer to downscaled data.

3. Results: crop yield seasonal forecasts assessment

Results of the yield simulations were compared to both official yield figures (EUROSTAT data, used as 'reference' for wheat yield of EU Member States) and to simulations made with the JRC operational system (used as 'control', labelled JRC in the rest of the text and figures). The comparison to the EUROSTAT reference provides an assessment of the skill of seasonal yield forecasts, and the comparison to 'JRC control' provides an indication of the possible improvement of the operational crop yield forecasting system, both in terms of accuracy and in terms of precocity of the given information for a similar level of accuracy.

Furthermore, the probabilistic nature of the DEMETER seasonal forecasts allows us to quantify the accuracy of the forecasted yield by using the probability distribution of crop yield anomalies (crop yield anomaly is defined as the yield deviance from the trend).

Probability density functions (PDFs) were used to fully exploit the information provided by the ensemble of seasonal crop yield forecasts. A probability of occurrence is associated with an event through PDF function. Building the PDF with the ensemble of simulated yields, probability $\Pr[x > x_0]$ of yield anomaly x to be greater than x_0 is computed (area under the PDF curve and right of x_0 on the x-axis and similarly probability $\Pr[x < x_1]$ of yield anomaly x to be less than x_1 is the area under the PDF curve and left of x_1 on the x-axis). This method provides, for each simulation, a probability of yield anomaly (deviation from the trend), with this probability of the anomaly being either negative ($\Pr[x < 0]$) or positive ($\Pr[0 < x] = 1 - \Pr[x < 0]$).

3.1. Prediction skill and earliness

When the crop model is run with DEMETER hindcasts, results vary from country to country, for both the ensemble mean deviation from EUROSTAT official figures and the yield ensemble

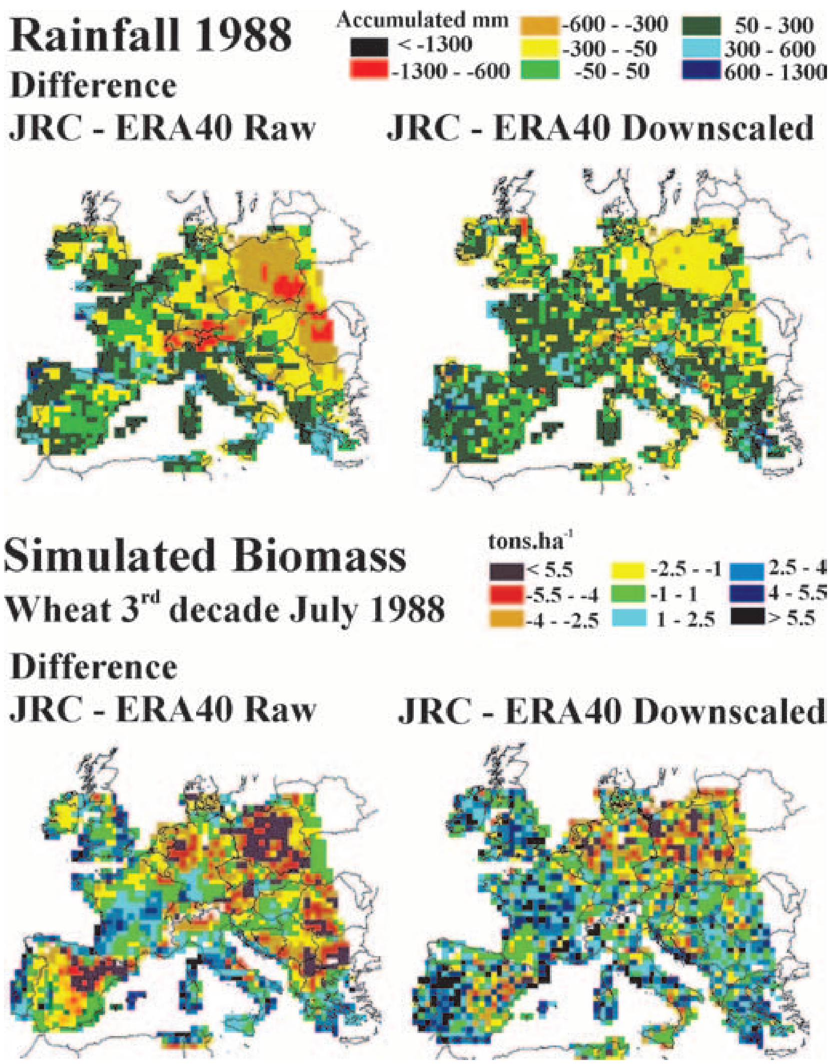

Fig 2. Top panel: maps of difference JRC minus reanalysis data ERA40 raw (left) and JRC minus downscaled (right) for rainfall in 1988 (accumulated January to end of July). Bottom panel: maps of difference between biomass simulated using JRC data and raw ERA40 data (left) or JRC data and downscaled ERA40 data (right), at the third decade of July 1988 in Europe.

variability (country results are presented in Table 1). The average percentage error (as a ratio to the official yield[2]) of the DEMETER-based yield simulations over the EU for the period 1995–1998 is 5.9% (weighted average of national values, the weight being the contribution of each member state to the total European production). Note that Portugal, which showed the largest discrepancies (see below and Table 1) was excluded from determining this average accuracy. The correlation between EUROSTAT yields and the DEMETER ensemble means yields is 0.65 for all the simulations made over the EU Member States. This correlation reaches 0.73 for the three main wheat producers: France, Germany and the United Kingdom (UK).

Variability of the ensemble of 63 simulated yields is low for Germany [the four-yearly standard deviation (s.d.) computed from the ensemble of 63 yields is lower than 0.12 tons ha^{-1}, i.e. 2% of the average simulated yield] and for France (s.d. ranges between 2% and 3%). On the other hand, the spread of the yield forecasts ensemble is larger in southern countries (s.d. greater than 10% in Greece, Portugal and Spain).

In terms of accuracy, DEMETER-based yield simulations perform better in EU northern countries (Table 1). The relative percentage of error ranges from 1.7% in Germany to 31.3% in Portugal. Good results are also obtained in the UK (average error for the 4 yr is 3.8%, with a poor result in 1997; percentage error is 9.6%) and in France (average error 4.7%). This result is particularly interesting since France, Germany and the UK are the main contributors to the European wheat production (France produces around 38% of the EU wheat production, Germany 21% and UK around 15%).

In Nordic countries such as Sweden and Denmark, DEMETER-based yield forecasts are also accurate (the average percentage errors for the 4 yr are 3.6% and 5.1%, respectively).

[2]The percentage error e is given by $e = 100\,x(\hat{y} - \mu)/\mu$ where y is the forecasted yield and μ is the official yield (EUROSTAT).

Table 1. Relative percentage error (ratio to the EUROSTAT value) of the DEMETER ensemble mean based yield forecasts, country by country, for each year of the study. The average column is the averaged error in absolute value of the 4 yr. Last column shows the RMSE computed on the 4 yr

% error	1995	1996	1997	1998	Average	RMSE
Denmark	−16.9	0.4	−0.6	2.4	**5.1**	**8.5**
France	3.5	−4.2	−3.6	−7.5	**4.7**	**5**
Germany	−0.7	−3.6	−1.3	1.1	**1.7**	**2**
Greece	10.5	12.7	−6.7	9.3	**9.8**	**10**
Ireland	0.17	11.4	−6.7	2.7	**5.2**	**6.7**
Italy	15.9	8.8	23.8	−5.3	**13.4**	**15.2**
Luxembourg	3.2	−8.2	3.8	−0.7	**4**	**4.8**
Portugal	35.8	4.1	34.9	50.2	**31.3**	**35.5**
Spain	51.9	−39.6	−5.6	−9.4	**26.6**	**33.1**
Sweden	−4.5	−2.7	−0.4	−6.9	**3.6**	**4.9**
UK	−0.6	3.3	9.6	1.8	**3.8**	**5.2**

Table 2. Weighted average percentage error (from actual values, in absolute value) for JRC and DEMETER (ensemble mean) wheat yield forecasts for Europe (national level) excluding Portugal. Weights correspond to the contribution of each member states on the EU wheat production. DEMETER forecasts are issued at the end of February

% error		1995	1996	1997	1998	Average
Average (abs. value)	JRC Feb	6.9	8.5	4.6	8.5	*7.5*
	JRC April	*8.3*	*8.1*	*6.3*	*7.9*	*7.7*
	JRC June	7.9	6.9	5.3	7.8	*7.0*
	JRC Aug	*6.7*	*5.8*	*4.9*	*4.3*	*5.4*
	DEMETER	6.9	6.3	5.9	4.8	*5.9*

On the other hand, accurate results were not found in southern countries. In particular, in Portugal, where for the 4 yr (1995–1998) the error is high (35%, systematic positive bias). Also, the low yield in Spain in 1995 due to a severe drought was not captured either by the DEMETER hindcasts or by the JRC operational system.

When analysing the result accuracy as a function of earliness, temporal analysis of yield errors (from EUROSTAT) of DEMETER-based data on one side and JRC operational yield forecasts on the other side shows that the percentage error obtained with DEMETER at the end of February (5.9%, excluding Portugal and weighting by the contribution of each member states in the EU wheat production) lies between the average error found at the end of June (7%) and the error found at the end of August (5.4%) using the JRC operational system (Table 2). Therefore, in terms of ensemble mean, for the period 1995–1998, DEMETER-based wheat yield forecasts issued in February are more accurate than JRC operational forecasts until July; that is to say, until the late stages of the wheat growing season.

3.2. *Probabilistic yield anomaly forecasts*

In order to assess the DEMETER-based yield forecasts taking into account the probabilistic nature of these forecasts (and not only the ensemble mean as in the previous section), one should analyse the probability distribution of yield anomaly forecasts.

Results for the UK and Germany are shown in Figs. 3 and 4. The plot at the top shows the yield time series. As an example, the expected value based on the time trend is shown for one of the 4 yr (1996 for Fig. 3 and 1995 for Fig. 4); from the time trend computed on the eight past years (1988–1995 for Fig. 3 and 1987–1994 for Fig. 4, solid line) a time trend based prediction is extrapolated (dashed line). The four anomalies are highlighted in the bottom-right corner of the figure. Plots at the bottom part of the figures show for each year (1995–1998) the PDF of yield anomalies obtained from DEMETER-based crop yield simulations (the hatched area corresponds to the probability of negative anomaly).

For the UK (Fig. 3), the 1996 positive probability and 1998 negative probability were accurately estimated. That is, in February, multi-model ensemble simulations indicate (correctly) that there is 87% (96%) chance to have a positive (negative) anomaly. In both cases, the main PDF peak is near the official yield anomaly (EUROSTAT value, solid vertical line). One should emphasize that already several months before harvest the final yield deviation from the trend was forecasted accurately.

For 1997, the forecasted yield significantly overestimated the observed yield. That year, the official yield presented a very large negative anomaly in the UK (almost 0.8 tons ha^{-1}) and this was not clearly depicted from DEMETER-based crop yield simulations, but the negative sign of the anomaly was correctly predicted (with a probability of 70%). In 1995, probabilities were more balanced (53–47%) and the main peak foresaw a normal year (no anomaly, the observed yield is close to the trend), which was confirmed at harvest by official figure.

In Germany (Fig. 4), DEMETER-based yield simulations correctly predicted anomalies, which were assigned by a probability higher than 80% for 3 yr out of four. However, in 1997, the difference from the anomaly is very low (−0.05 tons ha^{-1}) and DEMETER-based simulations overestimate it (the value of the ensemble mean was around −0.15 tons ha^{-1}), but at least exclude a positive anomaly. The secondary peak of the bimodal 1998 PDF fits perfectly with the actual values (probability for the correct sign was 82%), indicating that a substantial number of ensemble members permitted a very accurate simulation of the yield. Similar results are found in 1995, but slightly less accurate (lower probability for the correct sign, 62%). In 1996, simulations predicted the sign even if the yield itself was underestimated. Nevertheless, in 1996, already in February, DEMETER-based yield forecasts permitted an anticipation of a high yield anomaly with a 95% probability.

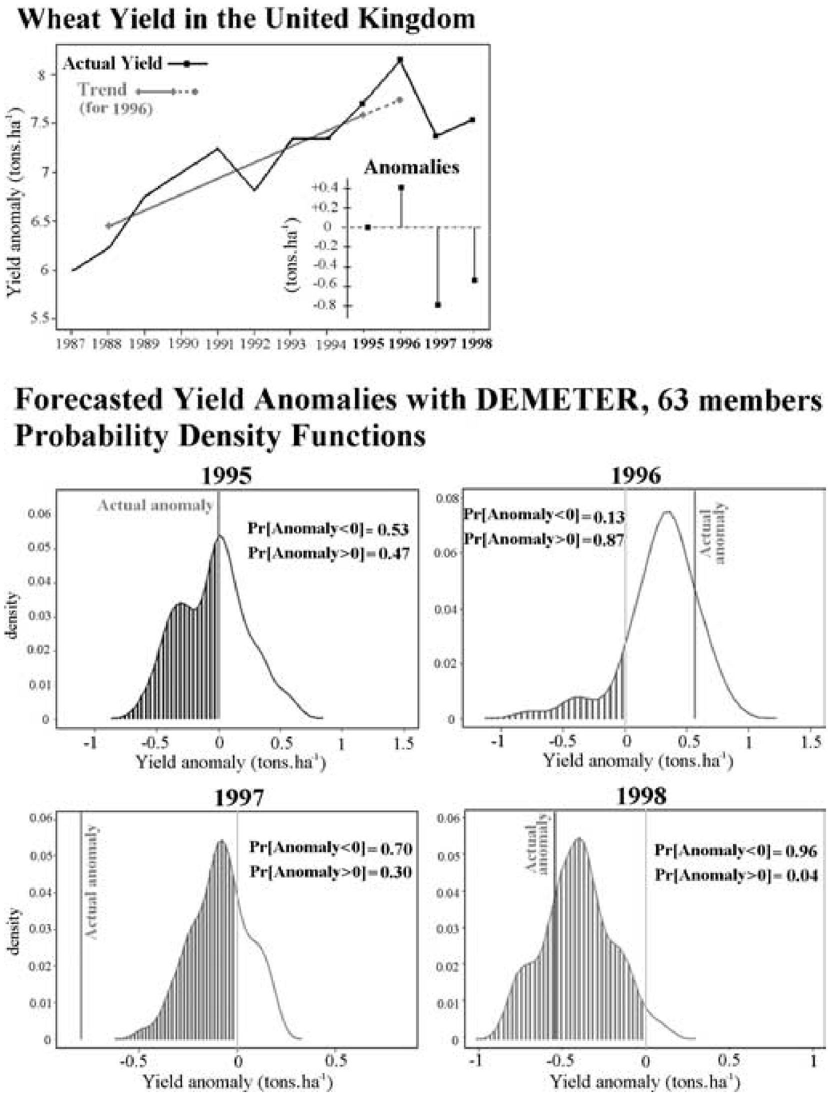

Fig 3. Top: wheat yield time series (1987–1998) with anomalies associated for the last 4 yr. As an example, for 1 yr (1996) the time trend computed over the eight past years is plotted (solid line, 1988–1995), from which is extrapolated the time trend based forecast for the current year (dashed line, 1996). The yearly anomaly is then the deviance between this time trend expected value and the actual yield. Bottom: PDFs of forecast anomalies with the DEMETER ensemble for the UK 1995–1998.

From the PDFs, the yield anomaly forecasts might be issued in the form of an interval (this is done under the assumption that the 63 members have the same probability), allowing us to give anomaly forecasts as an interval with a specified probability. This method also permits us to assess the accuracy of the forecast by testing whether or not the actual anomalies lie in these intervals. The 90% probability intervals for the simulations made in the UK and in Germany are shown in Fig. 5. For both countries, there are three hits within the 90% interval out of four. When considering all Member States, the official yield lies in the 90% interval for 80% of the cases (Portugal excluded).

4. Discussion

4.1. DEMETER hindcasts assessment

The evaluation of DEMETER hindcasts based yield forecasts in Section 3 was conducted by comparison against official figures and against operational JRC forecasts. However, other sources of errors, independent from input meteorological data, could contribute to the observed deviations. Therefore, to assess the performance of the seasonal hindcasts independently of other sources of errors, downscaled ERA40 reanalysis data were used as a 'perfect forecast' to provide a reference set. The differences obtained between DEMETER- and ERA40-based yield simulations could then be considered as a measure of the bias of DEMETER seasonal hindcasts versus climatology.

Focus is given to southern countries where the largest discrepancies were found (Table 1). In Portugal, ERA40 and DEMETER accumulated precipitation amount (January–August) are very different. Correlation coefficients between reanalysis data and DEMETER ensembles range between −0.3 and 0.8 (0.5 with DEMETER ensemble mean). Still, in Portugal, the percentage error of ERA40-based yield simulations is 1.6 times lower than

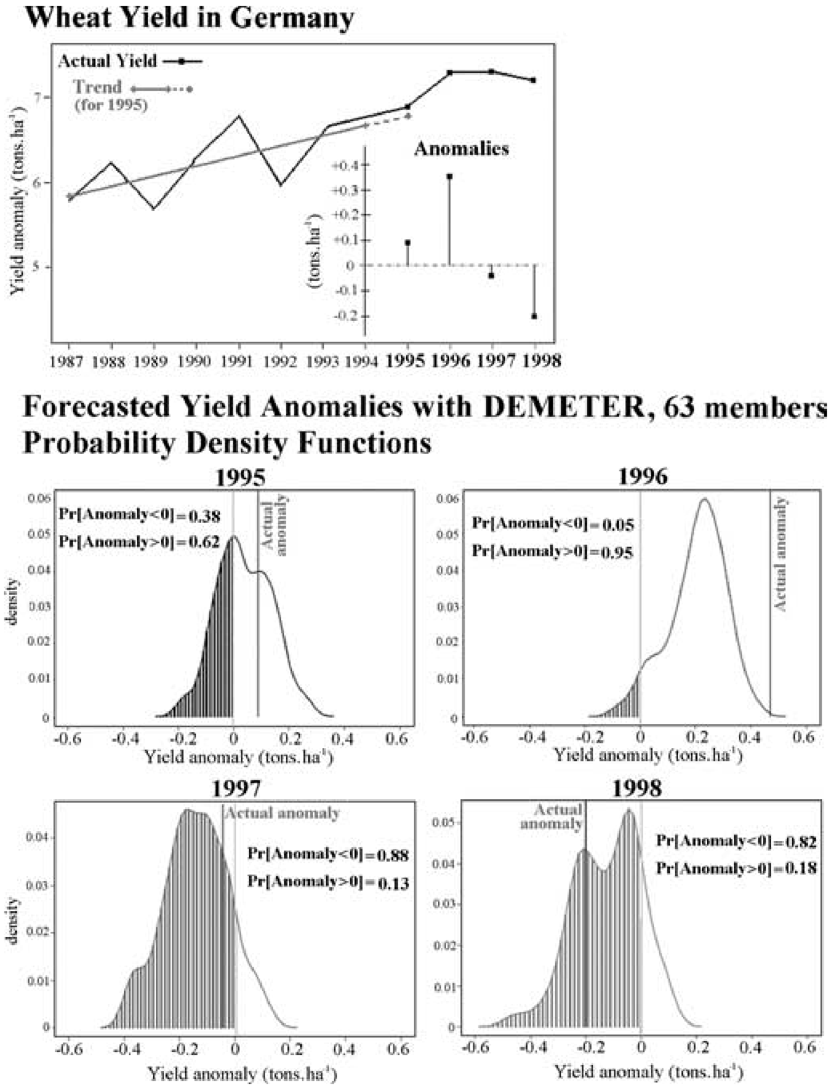

Fig 4. Same as Fig. 4, but for Germany. The time trend plotted in example in this plot refers to the 1987–1994 period and is extrapolated to provide a time trend based forecast for 1995.

the average error observed for DEMETER-based yield simulations, highlighting the bias of DEMETER.

This result was unexpected since it was found (Cantelaube et al., 2004) that, in Portugal, wheat yield shows some correlation ($R = 0.47$) with one of the main patterns of European climate variability (similar to the Eurasian type 1 pattern of Barnston and Livezey 1987), which in turn is strongly correlated with El Niño sea surface temperature anomalies and which is highly predictable (Pavan et al., 2000). Consequently, one could expect better performances of seasonal yield forecasts in Portugal.

On the other hand, also in Cantelaube et al. (2004), main modes of climate variability and winter wheat yield time series were not found well correlated in Greece and Italy. In Greece, the crop model driven by ERA40 data performs quite well (the ERA40 error is in the range of errors observed for most European countries, i.e. 6%), demonstrating the crop model ability to provide accurate results. Thus, DEMETER hindcasts contribute significantly to the error observed for the simulated yield. In Italy, ERA40 and DEMETER yield simulations are in agreement but far from official EUROSTAT yield (percentage errors are 13.5% and 15.2%). Downscaling or crop simulation could be additional sources of error contributing to the poor performances in this country. As a matter of fact, downscaling is delicate due to the complex Italian orography and climate patterns are contrasted between the influence of the Alps in the north and the hot and dry Mediterranean climate in the south. An analysis at finer regional level would be necessary (see Marletto et al., 2005).

Finally, problems in Spain for 1995 and 1996 might bias conclusions coming from this study. In fact, 1995 was characterized by rather favourable weather over the whole of Europe except for Spain and Portugal, which experienced an exceptional drought (lasting until the winter 1995/1996) and this was not reproduced by DEMETER. DEMETER is not very accurate for these 2 yr and, since 2 yr represents half of the tested period, DEMETER

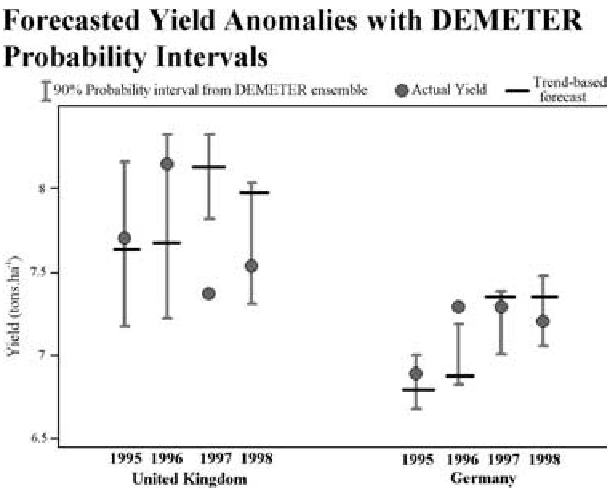

Fig 5. Probability interval (90%) built from the PDFs (DEMETER ensemble, 63 members) for the UK and Germany. The dots indicate the observed yields.

performances cannot be assessed for Spain. (Note that some missing meteorological data in the training set might have also affected the downscaling.)

4.2. *Probability distribution of yield anomaly forecasts*

Contrary to other domains using probabilistic forecasts, which can be simplified to a yes/no statement (e.g. 'rain/no rain', 'flood/no flood'), crop yield anomaly forecasts cannot be verified by defining indices and scores based on statements such as 'hits', 'false alarm', 'correct' or 'missed'. Crop yield anomalies forecasts are not categorical, and cannot be simplified to an 'observed/not observed' event.

This is primarily because of the non-binary and non-symmetric aspects of a yield anomaly (The inverse event of 'above normal' is not equivalent to the event 'below normal', since the yield could be 'normal' close to the trend). Secondly, 'no-anomaly/normal yield' means that the observed yield is 'close' to expected values, suggested by the trend; it is supposed that thresholds need to be fixed to define this 'normal yield' zone (for instance, the yield could be defined abnormal if it differs by more than ±2% from the expected trend). Finally, if the positive and negative anomalies forecasts obtained from the simulation are balanced, it is difficult to take a decision. The probability to have a negative anomaly around 50% could be either 'the ensemble-based yield simulation forecast a normal year' (Fig. 6, right; simulations for Greece 1997, PDF with a single peak centred on the null value – all ensembles agree on a no-anomaly situation) or 'PDF of ensemble-based yield simulations has a bimodal shape with one peak negative and one positive' (Fig. 6, left; simulations for Greece, 1996), when no decision is possible.

A plot summarizing the results is presented in Fig. 7; observed anomalies (EUROSTAT, in percentage of expected value, i.e. the trend) are plotted against the forecasted probabilities of occurrence of positive anomalies. Three observations marked by diamonds in the figure correspond to important negative anomalies (large errors in Portugal for 1995, 1997 and 1998), which were not depicted by DEMETER-based crop yield simulations (probability to have a negative anomaly is around 50%). Spain (square) and Denmark (triangle), both for 1995, are pointed out because the forecasted sign of the anomaly is false and associated with a high probability.

Figure 7 shows also that some of the biggest negative anomalies (in percentage) are associated with probability close to 50% (central part of the plot), indicating that either DEMETER has underestimated the intensity of the anomaly (see Fig. 3, UK 1997) or the simulations from the ensembles were scattered (see Fig. 6; Greece, 1996). However, the biggest positive anomalies are depicted with more accuracy (right-hand side of Fig. 7).

Furthermore, the figure shows that above a certain level of probability (for instance, focusing on the probability to have an anomaly, either negative or positive, higher than 60%, which includes 70.5% of the cases) the sign of the observed anomaly was correctly depicted, with the only two exceptions being Spain and

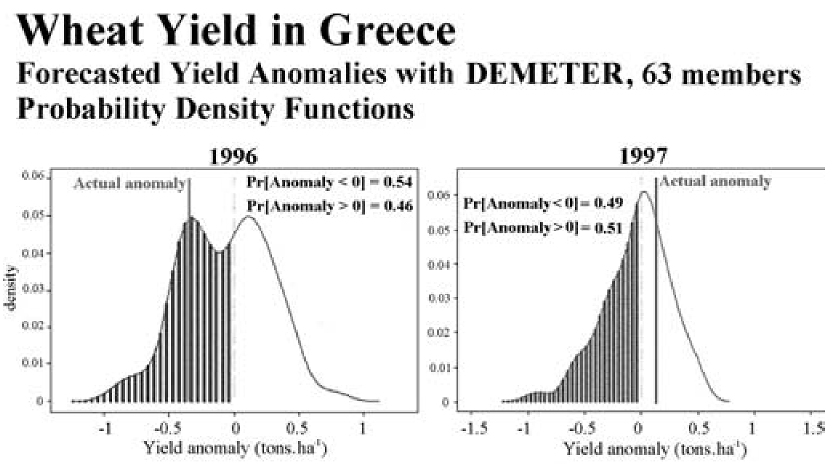

Fig 6. PDF of forecasted anomalies with the DEMETER ensemble for Greece, 1996 and 1997.

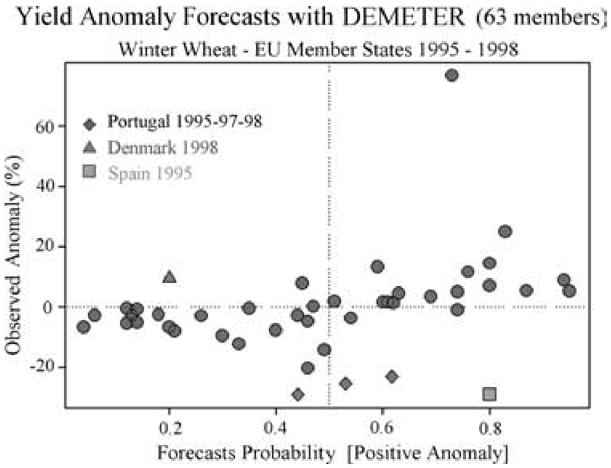

Fig 7. Reliability plot: observed anomalies plotted versus probability of yield anomaly forecasts from downscaled DEMETER ensembles for the 44 simulations made in Europe for the period 1995–1998 (39 dots plus five identified outliers).

Denmark in 1995 (7% of false sign predicted). This means that forecasts with high probability are reliable enough to indicate the future anomaly sign or at least to exclude correctly either a positive or a negative sign. As mentioned earlier, the difficulty lies in assessing probabilities around 50%; however, more years of data would be needed to assess whether robust decision criteria can be determined.

Anticipating the topic discussed in the next subsection, it must be kept in mind that, besides the seasonal forecasts themselves and their skills, an important issue is the way they could be used and integrated by end-users/decision makers. In this contest, the PDFs of yield anomalies are very useful because they could help experts from the EC DG AGRI to have several months in advance the general trend of the harvest to come, associated with a confidence level.

4.3. *Use of seasonal yield forecasts for European wheat production managers*

As pointed out by several research papers, incorporating climate forecasts in agricultural decision-making is the next challenge to tackle. 'Any benefit from climate prediction depends on the existence and understanding of decision options that are sensitive to the incremental information provided by forecasts, and are compatible with decision-makers goals, resources and constraints' (Hansen, 2002). Hammer et al. (2001) pointed out the role of linkages between scientist/analyst and decision maker in seeking relevant decisions for introduction of this new technology. Furthermore, the challenge that will determine the ultimate success of the advance in climate knowledge (i.e. improvement in seasonal forecasts) is that 'policymakers have to learn how to use this information' (Lemos et al., 2002). In Europe, the decision makers from the EC DG AGRI can certainly take advantage of the information provided by such multi-model ensemble based yield forecasts.

Coupling DEMETER ensembles to the JRC crop forecasting system is of interest to DG AGRI decision makers (as the main users of the JRC operational real-time forecasting system). Obviously, verification on longer time series will be necessary for implementation and fine-tuning.

5. Conclusion

Results from the integration of climatic seasonal forecasts in crop yield modelling in Europe suggest that reliable crop yield predictions can be obtained using an ensemble multi-model system.

The introduction of DEMETER data into the JRC crop yield forecasting system has shown that, for 1995–1998, wheat yield predictions issued in February are more accurate than those made with the operational system later in the season (July).

Furthermore, the use of ensemble seasonal forecast for crop yield demonstrates that useful information can be obtained through the probability associated with the yield forecast and suggests that yield anomalies could be anticipated. Tools such as PDFs allow us to associate a probability to the forecast and thus provide an indication of its reliability. In years and regions where there is little predictability, the resulting probability distribution will be broad, warning against specific precautionary actions. On the other hand, for years and regions where there is strong predictability, the probability distribution will be sharp, and reliable decisions for precautionary action can be made (management of agricultural markets, fixing set-aside programmes, etc.). Gain in time, coupled with (at least) comparable accuracy, constitutes a potential for money saving for the agricultural policy planners.

Further efforts in the field of crop yield seasonal forecasting are envisaged through the JRC participation in the FP6 EU-funded project ENSEMBLES, the successor of the DEMETER project.

6. Acknowledgment

The DEMETER project has been funded by the EU under the contract EVK2-1999-00197. The authors would like to thank Uffe Andersen from DMI for the work he achieved for the downscaling. The work would not have been possible without the support of Alfred de Jager (EC-JRC, IES) for the management and adaptation of the crop model input database. Thanks also to F. J. Doblas-Reyes and the DEMETER team at the ECMWF for their continuous help and availability.

References

Barnston, A. G. and Livezey, R. E. 1987. Classification, seasonality and persistence of low-frequency atmospheric circulation patterns. *Mon. Weather Rev.* **115**, 1083–1126.

Cantelaube, P. and Terres, J. M. 2003. *Use of Seasonal Forecasts on Crop Yield Modelling*. Publication of the European Commission, EUR 20996 EN.

Cantelaube, P., Terres, J. M. and Doblas-Reyes, F. J. 2004. Influence of climate variability on European agriculture—an analysis of winter wheat production. *Clim. Res.* **27**, 135–144.

Doblas-Reyes, F. J., Déqué, M. and Piedelièvre J.-Ph. 2000. Multi-model spread and probabilistic seasonal forecasts in PROVOST. *Q. J. R. Meteorol. Soc.* **126**, 2069–2088.

Doblas-Reyes, F. J., Hagedorn, R. and Palmer, T. N. 2005. The rationale behind the success of multi-model ensembles in seasonal forecasting – II. Calibration and combination. *Tellus* **57A**, 234–252.

European Commission. 1997. Towards a Common Agricultural and Rural Policy for Europe. Convened by Commission of European Communities. Directorate General for Agriculture (http://europa.eu.int/comm/agriculture/publi/buck_en/cover.htm).

European Commission. 1999. La politique Agricole commune, synthèse 1999. Directorate General for Agriculture. Office des publications officielles des Communautés européennes, Luxembourg, 2000, 30 pp.

FAOSTAT. 2001. *FAO Statistical Databases*. Food and Agriculture Organization of the United Nation. Rome, Italy.

Feddersen, H. and Andersen, U. 2005. A method for statistical downscaling of seasonal ensemble predictions. *Tellus* **57A**, this issue.

Genovese, G. 1998. The methodology, the results and the evaluation of the MARS crop yield forecasting system, in Agrometeorological Application for Regional Crop Monitoring and Production Assessment, Account of the EU Support group on Agrometeorology (SUGRAM), Office for Official Publications of the European Communities, Luxembourg. Report No. EUR 17735 EN, 67–119.

Hagedorn, R., Doblas-Reyes, F. J. and Palmer, T. N. 2005. The rationale behind the success of multi-model ensembles in seasonal forecasting – I. Basic concept. *Tellus* **57A**, 219–233.

Hammer, G. L., Hansen, J. W., Phillips, J. G., Mjelde J. W., Hill, H. and co-authors. 2001. Advances in application of climate prediction in agriculture. *Agricultural Systems* **74**, 515–553.

Hansen, J. W. 2002. Applying seasonal climate prediction to agriculture production. *Agricultural Systems* **74**, 305–307.

Hoogenboom, G. 2000. Contribution of agrometeorology to the simulation of crop production and its application. *Agric. For. Meteorol.* **103**, 137–157.

Jones, J. W., Hansem, J. W., Royce F. S. and Messina, C. D. 2000. Potential benefits of climate forecasting in Agriculture. *Agric. Ecosyst. Environ.* **82**, 169–184.

Lemos, M. C., Finan, T. J., Fox, R. W., Nelson D. R. and Tucker J. 2002. The use of seasonal forecasting in policymaking: lessons from the north-east Brazil. *Clim. Change* **55**, 479–507.

Marletto, V., Zinoni, F., Criscuolo, L., Fontana, G., Marchesi, S. and co-authors. 2005. Evaluation of downscaled DEMETER multi-model ensemble seasonal hindcasts in a northern Italy location by means of a model of wheat growth and soil water balance. *Tellus* **57A**, 488–497.

Marshall, J., Kushnir, Y., Battisti, D., Chang, P., Czaja, A. and co-authors. 2001. North Atlantic climate variability: phenomena, impact and mechanisms. *Int. J. Climatol.* **21**, 1863–1898.

Meyer-Roux, J. and Vossen, P. 1994. The first phase of the MARS project, 1988–1993: overview, methods and results. In: *Proceedings of the Conference on the MARS Project: Overview and Perspectives, Belgirate, 17–18 November 1993*. Publication EUR 15599 EN of the Office for Official Publications of the EU, Luxembourg.

Meza, F. J. and Wilks, D. S. 2003. Value of operational forecasts of seasonal average sea surface temperature anomalies for selected rainfed agricultural locations of Chile. *Agric. For. Meteorol*. **116**, 137–158.

Neelin, J. D., Battisti, D. S., Hirst, A. C., Jin, F., Wakata, Y. and co-authors. 1998. ENSO theory. *J. Geophys. Res.* **103**, 14 261–14 290.

Nnaji, A. O. 2001. Forecasting seasonal rainfall for agricultural decision-making in northern Nigeria. *Agric. For. Meteorol*. **107**, 193–205.

Ogallo, L. A., Boulahya, M. S. and Keane, T. 2000. Applications of seasonal to interannual climate predictions in agricultural planning and operations. *Agric. For. Meteorol*. **103**, 159–166.

Olesen, J. E. and Bindi, M. 2002. Consequences of climate change for European agricultural productivity, land use and policy. *Eur. J. Agronomy* **16**, 239–262.

Palmer, T. N. 1993. Extended-range atmospheric prediction and the Lorentz model. *Bull. Am. Meteorol. Soc.* **74**, 49–65.

Palmer, T. N., Brankovic, Č. and Richardson, D. S. 2000. A probability and decision-model analysis of PROVOST seasonal multi-model ensemble integrations. *Q. J. R. Meteorol. Soc.* **126**, 2013–2034.

Palmer, T. N., Alessandri, A., Andersen, U., Cantelaube, P., Davey, M. and co-authors. 2004. Development of a European multi-model ensemble system for seasonal-to-interannual prediction (DEMETER). *Bull. Am. Meteorol. Soc.* **85**, 853–872.

Pavan, V., Molteni, F. and Brankovic, Č. 2000. Wintertime variability in Euro-Atlantic region in observations and in ECMWF seasonal ensemble experiments. *Q. J. R. Meteorol. Soc.* **126**, 2143–2173.

Penman, H. L. 1948. Natural evaporation from open water, bare soil and grass. *Proceedings of the Royal Society, Series A* **193**, 120–146.

Petersen, E. and Fraser, R. W. 2001. An assessment of the value of seasonal forecasting technology for Western Australian farmers. *Agricultural Systems* **70**, 259–274.

Phillips, J. G., Cane, M. A. and Rosenzweig, C. 1998. ENSO, seasonal rainfall patterns and simulated maize yield variability in Zimbabwe. *Agric. For. Meteorol* **90**, 39–50.

Rabbinge, R. and van Diepen, C. A. 2000. Change in agriculture and land use in Europe. *Eur. J. Agronomy* **13**, 85–100.

Richardson, C. W. 1981. Stochastic simulation of daily precipitation, temperature and solar radiation. *J. Appl. Meteorol*. **17**, 182–190.

Rijks, D., Terres, J. M. and Vossen, P. 1998. Agrometeorological Applications for Regional Crop Monitoring and Production Assessment, Account of the EU Support Group on Agrometeorology (SUGRAM). Publication EUR 17735 EN of the Office for Official Publications of the European Communities, Luxembourg.

Shukla, J. 1998. Predictability in the midst of chaos: a scientific basis for climate forecasting. *Science* **282**, 728–731.

Sivakumar, M. V. K., Gommes, R. and Baier, W. 2000. Agrometeorology and sustainable agriculture. *Agric. For. Meteorol* **103**, 11–26.

Solow, A., Adams, R. M., Bryant, K. J., Legler, O. M., O'Brien, J. J. and co-authors. 1998. The value of improved ENSO prediction to US agriculture. *Clim. Change* **39**, 47–60.

Supit, I. 1997. Predicting national wheat yields using a crop simulation and trend models. *Agric. For. Meteorol.* **88**, 199–214.

Van Diepen, C. A., Wolf, J. and van Keulen, H. 1989. WOFOST: a simulation model of crop production. *Soil Use and Management* **5**, 16–24.

Vossen, P. 1995. Early crop production assessment of the European Union, the system implemented by the MARS–STAT Project. Agrometeorological models: theory and applications in the MARS project. Publication EUR 16008 EN of the Office for Official Publications of the European Communities, Luxembourg.

Wallace, J. M. and Gutzler, D. S. 1981. Teleconnections in the geopotential height field during the Northern Hemisphere winter. *Mon. Wea. Rev*. **109**, 784–812.