Multi-model seasonal forecasting for the North Atlantic and Europe

Francisco J. Doblas-Reyes, Renate Hagedorn, Magdalena Alonso Balmaseda, and Tim N. Palmer

European Centre for Medium-Range Weather Forecasts (ECMWF)
f.doblas-reyes@ecmwf.int

ABSTRACT

The forecast quality of the DEMETER seasonal multi-model ensemble predictions has been assessed over the North Atlantic/Europe region. The analysis shows that single-model ensembles have positive though low skill in almost every season. The skill decreases with lead time, although the agreement in skill between the models increases. This may be due to a more relevant role of the decadal variability versus the initial conditions at longer lead times. The simple multi-model constructed from equally weighted single-model ensembles proves to have an average forecast quality superior to any of the single models and to persistence predictions, in particular in a probabilistic framework. In spite of the relatively low skill, a perfect model approach indicates that potential skill is much higher than actual skill. However, potential skill estimates depend strongly on the model used. It is suggested that a multi-model system may provide more realistic predictability estimates. The relevance of these results for operational seasonal forecasts and its users is discussed.

Keywords: Seasonal ensemble forecasting, multi-model, European climate, predictability.

Predicción estacional multi-modelo para el Atlántico Norte y Europa

RESUMEN

En este trabajo se evalúa la calidad de las predicciones estacionales por conjuntos multimodelo creadas en el proyecto DEMETER. La evaluación se realiza para la región del Atlantico Norte y Europa. Los resultados muestran que los modelos individuales tienen una calidad predictiva baja aunque positiva durante casi todas las estaciones del año. La calidad decrece con el tiempo de predicción, observándose al mismo tiempo un aumento en el grado de acuerdo en la calidad entre los modelos. Ello puede deberse al hecho de que la variabilidad decadal juegue un papel más importante que las condiciones iniciales cuando el tiempo de predicción aumenta. El multimodelo simple construido a partir de los conjuntos de los modelos individuales sin ponderación demuestra tener, en promedio, una calidad superior a la de cualquier modelo individual así como a la de predicciones basadas en la persistencia de las anomalías iniciales. Esto es más evidente en el caso de predicciones probabilistas. A pesar de la calidad relativamente baja de las predicciones, una evaluación de la predecibilidad usando la aproximación de modelo perfecto indica que la calidad potencial es superior a la calidad observada. Sin embargo, las estimaciones de la calidad potencial dependen del modelo que se use. Se sugiere en el texto que el sistema multimodelo puede proporcionar una estimación más realista de la predecibilidad. También se discute la repercusión de estos resultados para la predicción operativa y los usuarios.

Palabras clave: Predicción estacional por conjuntos, multimodelo, clima europeo, predecibilidad.

1. INTRODUCTION

The model-based evidence for seasonal forecasting in the North Atlantic/Europe (NAE, 35°N-75°N, 60°W-45°E) region is discussed in this paper and the advantages of the multi-model briefly described. The analysis is carried out on the basis of the forecast quality properties of the DEMETER multi-model experiment (Palmer et al. 2004) to take into account the inherent probabilistic nature of long-range predictions. The DEMETER project[1] has been funded under the European Union V[th] Framework Environment Programme to assess the skill and economic value of multi-model ensemble seasonal forecasts. One of the principal aims of DEMETER has been to advance the concept of multi-model ensemble prediction by installing a number of state-of-the-art global coupled ocean-atmosphere models on a single supercomputer and producing a series of multi-model ensemble hindcasts with common archiving and common diagnostic software.

Ensemble weather and climate forecasting (Epstein 1969; Molteni et al. 1996; Palmer 2002) has been principally developed so as to estimate forecast uncertainty due to unavoidable errors in the initial conditions. However, initial conditions are not the sole source of uncertainty in weather and climate forecasting. Additional contributions to forecast uncertainty come from model error, although its estimation is still an open problem that suffers from a lack of analytical solutions (Palmer 2001). A pragmatic method to deal with this problem is based on the multi-model approach (Tracton and Kalnay 1993; Vislocky and Fritsch 1995; Fritsch et al. 2000; Palmer et al. 2000). The multi-model approach seeks to combine predictions issued with different forecast systems to produce a more skilful and reliable estimate of the forecast probability density function (PDF). In the case of dynamical forecasting, the use of several models as members of an ensemble is a way of taking into account the model uncertainty in the simulation of the climate system. If, in addition, each model can produce an ensemble of simulations with slightly different initial conditions, the multi-model ensemble system will sample the two sources of forecast uncertainty mentioned above, i.e. model error and uncertainty linked to initial conditions.

As described in Barnston et al. (2003) and Hagedorn et al. (2004), a simple way of issuing predictions with a multi-model system consists in pooling together all the ensemble members from several single-model ensembles and assigning an equal weight to each model. The underlying hypothesis in this simple multi-model construction is that each model is independent and equally skilful. There is substantial evidence of the fact that different single models have very different performances, the ranking of model skills changing as a function of lead time, variable, region, etc. (e.g. Shukla et al. 2000). Therefore, the simple multi-model system has shown an overall clear superiority in terms of forecast quality over any single-model ensemble (Hagedorn et al. 2004). In particular, the superiority of the multi-model concept is more clearly evidenced in a probabilistic framework (Palmer et al. 2000; Doblas-Reyes et al. 2000;

---

[1] A complete description of the project and its main results can be found on the DEMETER website: http://www.ecmwf.int/research/demeter.

Table 1. Institutions whose global coupled ocean-atmosphere
models comprise the DEMETER multi-model prediction system

Institution	Acronym
European Centre for Research and Advanced Training in Scientific Computation (France)	CERFACS
European Centre for Medium-range Weather Forecasts	ECMWF
Istituto Nazionale de Geofisica e Vulcanología (Italy)	INGV
Laboratoire d'Océanographie Dynamique et de Climatologie (France)	LODYC
Météo-France (France)	MetFr
Met Office (UK)	UKMO
Max-Planck Institut für Meteorologie (Germany)	MPI

Palmer et al. 2004). A substantial component of this superiority is due to the higher reliability of the predictions issued with the simple multi-model.

The paper is organised as follows. The DEMETER experiment and the data used are described in Section 2. Section 3 summarizes the forecast quality of seasonal hindcasts over the North Atlantic/Europe region. An outline of the results and short discussion of the results can be found in Section 4 along with a short description of a multi-model operational strategy.

2. DATA

The data have been generated as part of the DEMETER experiment (Palmer et al. 2004). The DEMETER multi-model prediction system comprises the global coupled ocean-atmosphere models of the seven institutions, indicated in the Table 1.

To assess seasonal dependence on forecast skill, the DEMETER hindcasts have been started four times a year from 1st February, 1st May, 1st August, and 1st November at 00 GMT. The atmospheric and land-surface initial conditions are taken from the ECMWF Re-Analysis[2] (ERA-40) dataset. The ocean initial conditions are obtained from ocean-only runs forced by ERA-40 fluxes, except in the case of MPI that used a coupled initialization method. Each hindcast has been integrated for 6 months and comprises an ensemble of 9 members. Hindcasts have been produced over the period 1958-2001, although the common period to the 7 models is 1980-2001. A significant part of the DEMETER dataset is freely available for research purposes through a public online data retrieval system installed at ECMWF[3].

The performance of the DEMETER system has been evaluated from this comprehensive set of hindcasts (Palmer et al. 2004; Hagedorn et al. 2004). In addition,

---

[2] http://www.ecmwf.int/research/era.
[3] Data can be retrieved from the ECMWF public data server in GRIB and NetCDF from http://www.ecmwf.int/research/demeter/data.

a complete set of diagnostics can be accessed on-line[4]. The predictions correspond to the seasonal averages for the simulation period going from months 2 to 4 and 4 to 6, i.e., results are described for the following seasons: December to February (DJF), March to May (MAM), June to August (JJA) and September to November (SON) in the case of the 2-4 month predictions and for February to April (FMA), May to July (MJJ), August to October (ASO) and November to January (NDJ) for 4-6 month predictions. A key result is that probability scores for the simple multi-model ensemble are, on average, better than those from any of the single-model ensembles, this improvement not being only a consequence of the larger ensemble size of the multi-model. Overall, the results indicate that the multi-model ensemble is a viable pragmatic approach to the problem of representing model uncertainty in seasonal-to-interannual prediction, leading to a more reliable and skilful forecasting system than that based on any one single model.

Due to the model biases, generally the raw model output in long-range dynamical forecasting is not useful (Déqué 1991), so that anomalies have to be computed. Anomalies are expressed as departures from the corresponding long-term climatology (1980-2001). Calculation of model and reference anomalies has been carried out in one-year out cross-validation mode (Michaelsen 1987) to obtain unbiased estimators. This implies eliminating the target year from the computation.

ERA-40 data have been used as verification reference except for the precipitation, for which the GPCP (*Global Precipitation Climatology Project*) data set has been employed.

3. SEASONAL FORECAST QUALITY

The skill of seasonal averages of climate variables in the NAE region is low, when compared with the tropical regions, albeit positive[5]. Fig. 1a-c shows the grid-point correlation for wintertime (DJF) 2-metre temperature from ensemble-mean 2-4 month hindcasts initiated on 1st November over the period 1980-2001 for three single models. While the models tend to agree on the levels of predictive skill, there is some disagreement over Europe, where the correlation ranges from less than 0.2 to above 0.4. LODYC and ECMWF have a similar skill distribution. These models share the atmospheric component of the coupled model, which suggests that the skill is mostly sensitive to the atmospheric component of the coupled model. The significance of these differences is unclear due to the relatively small sample used and the existence of spatial autocorrelation of the fields. In addition, part of the differences in skill between the different models may be associated with relatively small spatial shifts in the predicted signals, which would appear as a low skill. Similar conclusions can be applied to precipitation, although with a lower level of skill.

[4] http://www.ecmwf.int/research/demeter/verification.
[5] A comprehensive set of diagnostics including the tropical regions is available in the DEMETER online verification site.
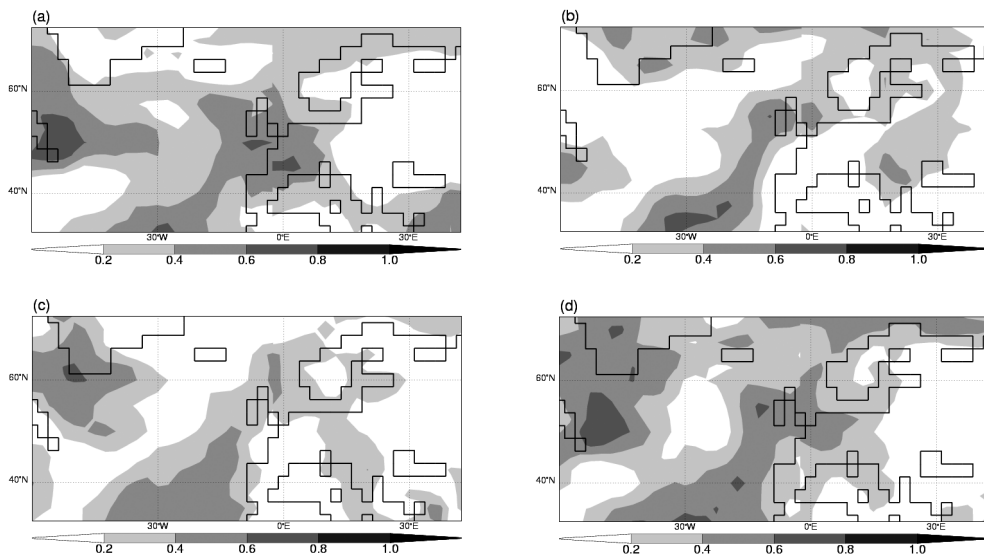
Figure 1. Grid-point correlation for wintertime (DJF) 2-metre temperature from 2-4 month ensemble-mean hindcasts initiated on 1st November over the period 1980-2001 for a) CERFACS, b) ECMWF, c) LODYC and d) simple multi-model ensembles.

In spite of the low positive scores depicted by each individual model, the simple multi-model ensemble mean (Fig 1d) displays a slightly more optimistic picture. Positive correlations are found over most of western Europe and relatively high values appear in many areas. This is the first indication of the superiority of the simple multi-model versus single-model ensembles and extends to every variable that has been verified.

Fig. 2 shows the seasonal evolution of the anomaly correlation coefficient (*ACC*) for ensemble-mean 2-4 and 4-6 month seasonal hindcasts of several variables from each of the seven DEMETER single models (open bars) over Europe (35°N-75°N, 12.5°W-42.5°E). The maximum skill is found in different seasons for different variables: spring for 2-metre temperature, summer for precipitation, and summer and winter for geopotential height. The seasonality of the skill is not clear in the case of the 4-6 month hindcasts, the correlation being much lower at this longer lead time especially for geopotential height and precipitation. The 4-6 month 2-metre temperature hindcast skill is positive for most of the models and shows a good agreement between the different models. The reasons of the higher skill with increasing lead time for temperature should be found in the stronger persistence of this variable. When most of the models agree in having some skill, the simple multi-model (solid bars) does as well, sometimes even better, than the best model. This suggests that although locally or for some period a single-model ensemble may have more skill than the simple multi-model ensemble, overall the simple multi-model is on average superior to any single model (Hagedorn et al. 2004). It
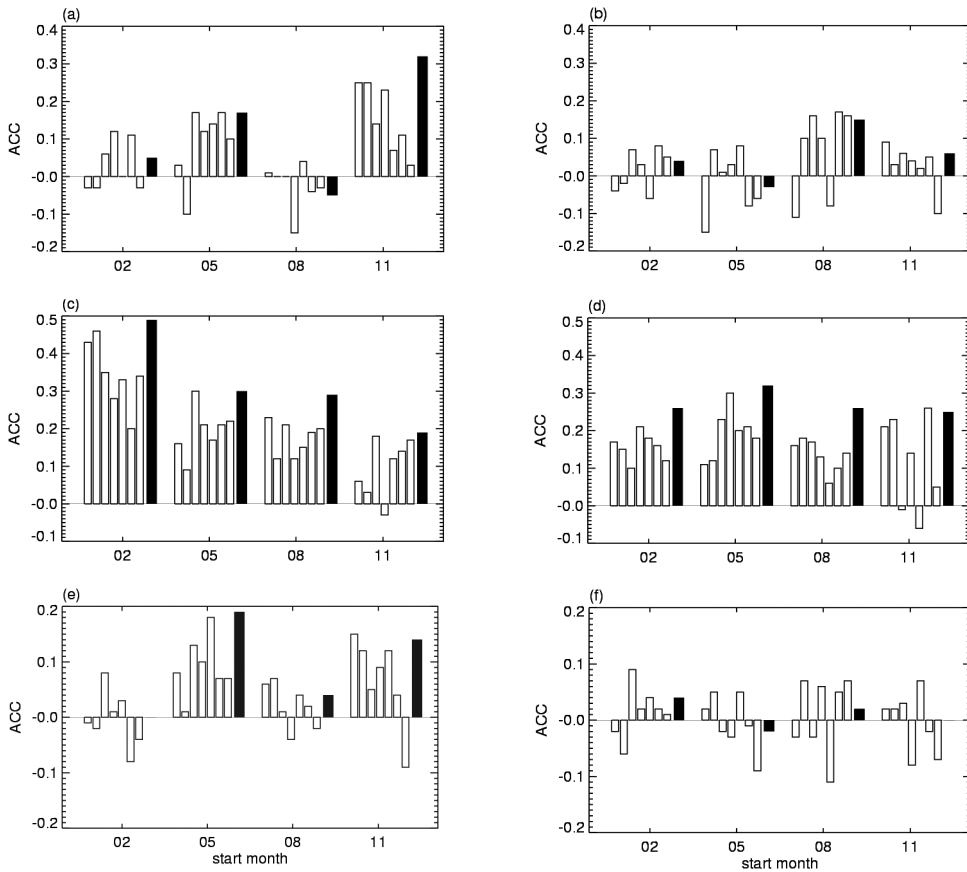
Figure 2. Seasonal evolution of the ensemble-mean anomaly correlation coefficient for 500 hPa geopotential height (top row), 2-metre temperature (middle row) and precipitation (bottom row) from each of the seven DEMETER single models (open bars) and the simple multi-model (solid bars) over Europe (35°N-75°N, 12.5°W-42.5°E). The left (right) column corresponds to 2-4 (4-6) month seasonal hindcasts. The hindcast sample covers the period 1980-2001. The horizontal axis indicates the start date of the hindcasts: 02, 05, 08 and 11 for 1st of February, May, August and November, respectively.

will be shown that the skill improvement of the simple multi-model with regard to single-model ensembles is even more obvious in a probabilistic setting. However, the simple multi-model may not be the optimal forecasting system. Some models have negative correlations in Fig. 2. Including the predictions from these models in the multi-model ensemble may reduce the skill. An alternative way of constructing multi-model predictions consists in calibrating and combining the single models by giving different weights to each model. Ongoing research on this topic (Doblas-Reyes et al. 2004; Stephenson et al. 2004) has produced positive results, especially for the tropics.
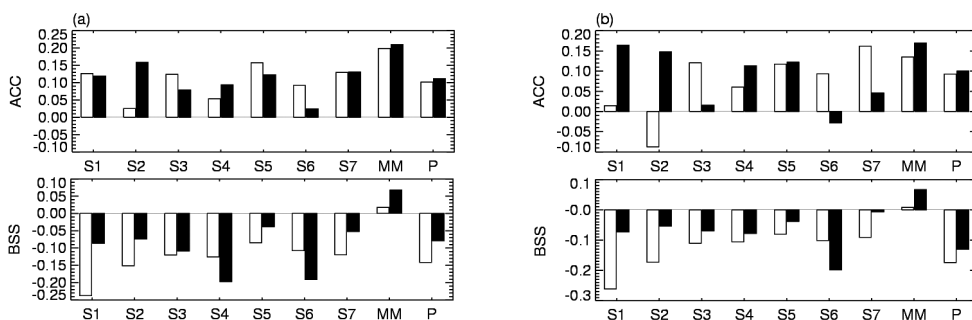
Figure 3. Ensemble-mean anomaly correlation coefficient (*ACC*, top) and Brier skill score (*BSS*, bottom; event «anomalies above normal») for precipitation (open bars) and 2-metre temperature (solid bars) over the a) North Atlantic/Europe (35°N-75°N, 60°W-45°E) and b) Europe (35°N-75°N, 12.5°W-42.5°E) regions. The pairs of bars correspond, from left to right, to the seven single-model ensembles, the simple multi-model ensemble and the persistence forecast. The hindcast sample covers the period 1980-2001.

Fig. 3a shows the *ACC* (top) and Brier skill score (bottom) for 2-metre temperature (solid bars) and precipitation (open bars) 2-4 month winter (DJF) hindcasts over the NAE region for each of the seven DEMETER models (S1 to S7), the simple multi-model (MM) and a persistence forecast (P). The persistence forecast has been obtained from the seasonal anomaly of the three months previous to the start date of the hindcasts (August, September, October). The event probability for the persistence forecast is initially 1 if the event occurs in the three months prior to the start date of the dynamical hindcasts and 0 otherwise. The probabilistic persistence prediction is then linearly relaxed towards a climatological probability forecast with a time scale of 8 months. The Brier score, *b*, is a measure of probabilistic forecast quality for dichotomous events (the event verified in Fig. 3 is «anomalies above normal»). The score is

defined as $b = \dfrac{1}{N}\sum_{i=1}^{N}(p_i - v_i)^2$ where *N* is the number of predictions (22 in the case

of the period 1980-2001), p_i the forecast probability and $v_i = 1$ or 0, depending on whether the event actually occurs or not. The Brier score is similar to the conventional root-mean-square score, although in probability space. It is positive, and equals zero only for a correct certain forecast. The Brier score from a climatological forecast, b_{clim}, can be similarly calculated by using $p_i = p_{clim}$, the climatological frequency that the event occurs. The Brier skill score, *BSS*, is then a measure of how good the ensemble

forecast is relative to a climatological forecast: $BSS = 1 - \dfrac{b}{b_{clim}}$. If $BSS \leq 0$ then

the ensemble forecast is no better, or even worse, than the climatological forecast, while *BSS* = 1 for a perfect deterministic forecast. Fig. 3a shows that both the ensemble-mean *ACC* and *BSS* of the single models is similar to that of persistence. In particular,

the coupled models and the persistence forecast have a negative *BSS*, implying that their skill is lower than the reference climatological forecast. On the contrary, this figure confirms that the simple multi-model ensemble not only is superior to the single-model ensembles but also has higher skill than persistence, especially in the probabilistic framework for which the *BSS* is positive. The improvement of the simple multi-model is due mainly to an increase in the «reliability» of the predictions (a measure of how well the forecast probability matches the verification frequency, Atger 1999), associated with an increase in spread that samples more efficiently the forecast uncertainty. The comparison between *ACC* and *BSS* illustrates the importance of formulating predictions that include a measure of the forecast uncertainty, such as probabilistic forecasts, instead of ensemble mean or deterministic predictions. The simple multi-model skill superiority is not only due to the higher skill over the ocean observed in Fig. 1. Fig. 3b shows results similar to those in Fig. 3a for the smaller region covering the European continent (35°N-75°N, 12.5°W-42.5°E) used in Fig. 2.

Part of the increased skill of the simple multi-model over the single-model ensembles can be attributed to its larger ensemble size. However, tests also show that a 54-member simple multi-model ensemble (taken from the seven-model multi-model) shows higher *BSS* than a 54-member ensemble of just the best-performing individual model.

The low, though positive, skill described over the region may convey a slightly pessimistic message. However, estimates of predictability or potential skill introduce a different picture. Fig. 4 shows the potential skill of the coupled models. The potential skill is assessed under the assumption that the model is perfect (perfect model approach) and that the model spread bears a direct relationship with the forecast skill. Although it can not be considered as a measure of the maximum skill that can be achieved by improving the model, it is an indicator of predictability. In this example, the *ACC* is used as a measure of potential skill to estimate predictability as in Déqué (1997). The potential skill is computed by taking as verification each ensemble member in turn, assuming it represents the truth, and using the other ensemble members to compute the ensemble mean prediction. The results are given for the same DEMETER prediction systems as in Fig. 1. A comparison of both figures shows that the single-model potential skill is higher than the actual skill. The potential skill is larger over the ocean than over the continent, where it barely presents values above 0.4. Results for other variables indicate that it is not necessarily the case that the model with highest potential skill (CERFACS in the case of Fig. 4) has the highest actual skill. The models sharing the same ocean module (LODYC and CERFACS) display a similar geographical distribution of the potential skill. This suggests that the differences in potential skill are truly sensitive to the atmospheric component and not simply due to sampling uncertainties.

Fig. 4 indicates that the differences in potential skill between the models are high. This implies that no model can be considered perfect when estimating predictability and that these estimates are model dependent. In addition, it is well known that most of the models underestimate the ensemble spread. In those cases predictability may be overestimated. Given that the spread of the simple multi-model seems to be a more realistic measure, the simple multi-model can also be used to obtain a predictability estimate. The ensemble mean of each single model has been taken as truth each time
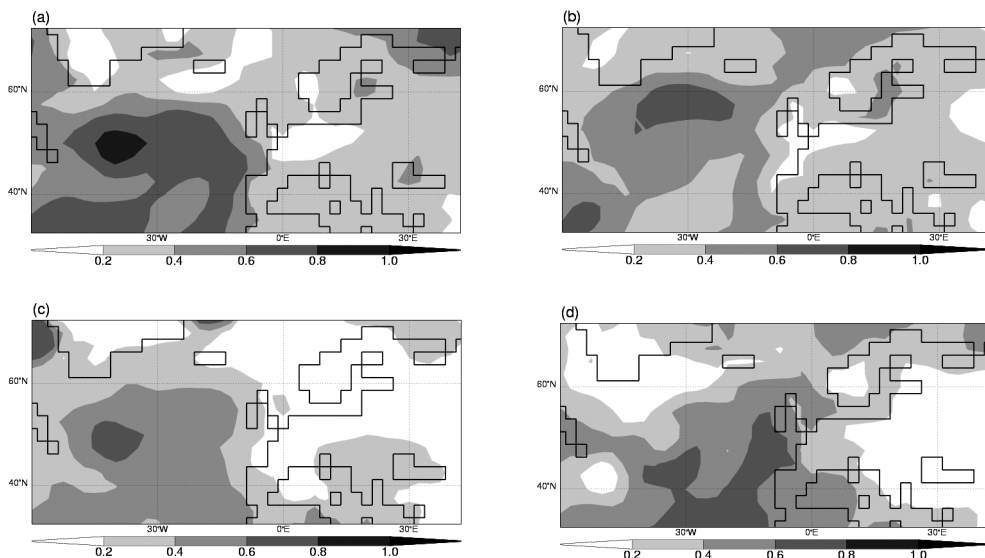
Figure 4. Grid-point potential skill for wintertime (DJF) 2-metre temperature from 2-4 month ensemble-mean hindcasts initiated on 1st November over the period 1980-2001 for a) CERFACS, b) ECMWF, c) LODYC and d) simple multi-model ensembles. The potential skill of the single models is computed by taking as verification each ensemble member in turn, assuming it represents the truth, and using the other ensemble members to compute the ensemble-mean prediction. In the case of the simple multi-model, the ensemble mean of each single model is taken as truth each time.

and the *ACC* has been computed. The results are displayed in Fig. 4d, where the geographical distribution and range of values similar to that of the actual skill (Fig. 1d). In other words, the gap between actual and potential skill is narrower in this case because the potential skill has decreased and the actual skill is higher. In addition, the similar geographical distribution implies that there may be a link between multi-model spread and forecast skill. Therefore, the simple multi-model potential skill may be regarded as a more meaningful estimate of predictability as it samples both initial conditions and model uncertainty.

4. CONCLUSIONS

An assessment of the forecast quality of the seasonal multi-model ensemble predictions produced within the DEMETER project has been carried out for the NAE region. There is positive seasonal climate skill, although it is low in particular for precipitation. The skill has a strong seasonality that depends on the variable. The equally weighted multi-model (or simple multi-model) system provides higher skill than single-model systems. In addition, it exceeds the skill of persistence. Although

the simple multi-model skill is not always higher than the skill of the best single model, it has been found that there is no single model that is consistently better for all variables and lead times. This implies that the simple multi-model performs, on average, better than any single model selected at random. The simple multi-model superiority over the single-model ensembles is more clearly evidenced in the case of probabilistic predictions. However, the possibility of additionally increasing the multi-model skill by using ways of combining the models is under investigation. It is expected that a scheme assigning different coefficients to the single models depending on their past performance may further improve the multi-model skill.

A predictability assessment using the perfect model approach indicates that potential skill is much higher than actual skill. However, predictability estimates appear to depend strongly on the model used, suggesting that they are not reliable. It has been suggested that the simple multi-model may provide a more meaningful measure of predictability as it takes into account not only the impact of the uncertainty in the initial conditions, but also the model error.

These results imply that a thorough assessment of every multi-model forecast system needs to be carried out. This requires the availability of long samples. The outcome of the assessment needs to be considered in the context of end-user models (Morse et al. 2004) and users' decision-making process (Archer 2003). This is the framework where the actual predictive potential of a multi-model seasonal forecast system can be determined.

The success of the multi-model approach in DEMETER has motivated the creation of an operational multi-model system at ECMWF. As a consequence, multi-model initiatives similar to DEMETER are also envisaged by other organizations, like the Asian Pacific Climate Network (APCN) and the International Research Institute (IRI).

5. ACKNOWLEDGMENTS

6. REFERENCES

ARCHER, E. R. M. (2003). Identifying underserved end-user groups in the provision of climate information. *Bull. Amer. Meteor. Soc.,* 84, 1525-1532.

ATGER, F. (1999). The skill of ensemble prediction system. *Mon. Wea. Rev.,* 127, 1947-1953.

BARNSTON, A. G.; S. J. MASON, L. GODDARD, D. G. DEWITT & S. E. ZEBIAK (2003). Multimodel ensembling in seasonal climate forecasting at IRI. *Bull. Amer. Meteor. Soc.,* 84, 1783-1796.

DÉQUÉ, M. (1991). Removing the model systematic error in extended range forecasting. *Ann. Geophys.,* 9, 242-251.

DÉQUÉ, M. (1997). Ensemble size for numerical seasonal forecasts. *Tellus A*, 49, 74-86.

DOBLAS-REYES, F. J.; M. DÉQUÉ & J. P. PIEDELIÈVRE (2000). Multi-model spread and probabilistic seasonal forecasts in PROVOST. *Quart. J. Roy. Meteor. Soc.,* 126, 2069-2088.

DOBLAS-REYES, F. J.; R. HAGEDORN & T. N. PALMER (2004). The rationale behind the success of multi-model ensembles in seasonal forecasting. Part II: Calibration and combination. *Tellus A*, submitted.

EPSTEIN, E. S. (1969). A scoring system for probability forecasts of ranked categories. *J. Appl. Meteor.,* 8, 985-987.

FRITSCH, J. M.; J. HILLIKER, J. ROSS & R. L. VISLOCKY (2000). Model consensus. *Wea. Forecasting*, 15, 571-582.

HAGEDORN, R.; F. J. DOBLAS-REYES & T. N. PALMER (2004). The rationale behind the success of multi-model ensembles in seasonal forecasting. Part I: Basic concept. *Tellus A*. submitted.

MICHAELSEN, J (1987). Cross-validation in statistical climate forecast models. *J. Climate Appl. Meteor.,* 26, 1589-1600.

MOLTENI, F.; R. BUIZZA, T. N. PALMER & T. PETROLIAGIS (1996). The ECMWF ensemble prediction system: Methodology and validation. *Quart. J. Roy. Meteor. Soc.,* 122, 73-119.

MORSE, A. P.; F. J. DOBLAS-REYES, M. B. HOSHEN, R. HAGEDORN, M. C. THOMSON & T. N. PALMER (2004). First steps towards the integration of a dynamic malaria model within a probabilistic multi-model forecast system. *Tellus A*, submitted.

PALMER, T. N. (2001). A nonlinear dynamical perspective on model error: A proposal for non-local stochastic-dynamic parametrization in weather and climate prediction models. *Quart. J. Roy. Meteor. Soc.,* 127, 279-304.

PALMER, T. N. (2002). The economic value of ensemble forecasts as a tool for risk assessment: From days to decades. *Quart. J. Roy. Meteor. Soc.,* 128, 747-774.

PALMER, T. N.; C. BRANKOVIC & D. S. RICHARDSON (2000). A probability and decision-model analysis of PROVOST seasonal multi-model ensemble integrations. *Quart. J. Roy. Meteor. Soc.,* 126, 2013-2034.

PALMER, T. N.; A. ALESSANDRI, U. ANDERSEN, P. CANTELAUBE, M. DAVEY, P. DÉLÉCLUSE, M. DÉQUÉ, E. DÍEZ, F. J. DOBLAS-REYES, H. FEDDERSEN, R. GRAHAM, S. GUALDI, J. F. GUÉRÉMY, R. HAGEDORN, M. HOSHEN, N. KEENLYSIDE, M. LATIF, A. LAZAR, E. MAISONNAVE, V. MARLETTO, A. P. MORSE, B. ORFILA, P. ROGEL, J. M. TERRES & M. C. THOMSON (2004). Development of a European multimodel ensemble system for seasonal-to-interannual prediction (DEMETER). *Bull. Amer. Meteor. Soc.,* 85, 853-872.

SHUKLA, J.; L. MARX, D. PAOLINO, D. STRAUS, J. ANDERSON, J. PLOSHAY, D. BAUMHEFNER, J. TRIBBIA, C. BRANKOVIC, T. N. PALMER, Y. CHANG, S. SCHUBERT, M. SUÁREZ & E. KALNAY (2000). Dynamical seasonal prediction. *Bull. Amer. Meteor. Soc.,* 81, 2593-2606.

STEPHENSON, D. B.; C. A. D. S. COELHO, F. J. DOBLAS-REYES & M. BALMASEDA (2004). A unified framework for the combination of multi-model climate predictions: The role of forecast assimilation. *Tellus A*, submitted.

TRACTON, M. S. & E. KALNAY (1993). Operational ensemble prediction at the National Meteorological Center. Practical aspects. *Wea. Forecasting*, 8, 379-398.

VISLOCKY, R. L. & J. M. FRITSCH (1995). Improved model output statistics forecast through model consensus. *Bull. Amer. Meteor. Soc.,* 76, 1157-1164.