Development of a European Multi-Model Ensemble System for Seasonal to Inter-Annual Prediction (DEMETER)

T. N. Palmer, F. J. Doblas-Reyes, R. Hagedorn
t.palmer@ecmwf.int
European Centre for Medium-Range Weather Forecasts, Reading, UK

1. Introduction

Seasonal forecasts are clearly of value to a wide cross section of society, for personal, commercial and humanitarian reasons. Dynamical seasonal forecasts have been made using ensemble systems with slightly different initial conditions. However, if uncertainties in initial conditions are the only perturbations represented in a seasonal-forecast ensemble, then the resulting measures of predictability will not be reliable; the reason being that the model equations are also uncertain. One approach to solve this problem relies on the fact that global climate models have been developed somewhat independently at different climate institutes. An ensemble comprising such quasi-independent models is referred to as a multi-model ensemble. The ability of multi-model ensembles to produce more skilful probability seasonal forecasts will be presented in this contribution.

The DEMETER project[1] (Development of a European Multi-model Ensemble System for Seasonal to Interannual Prediction) has been funded under the European Union Vth Framework Environment Programme to assess the skill and potential economic value of multi-model ensemble seasonal forecasts. The principal aim of DEMETER was 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 to produce a series of six-month multi-model ensemble hindcasts with common archiving and common diagnostic software.

2. Description of the experiment

The DEMETER system comprises 7 global coupled ocean-atmosphere models: CERFACS (European Centre for Research and Advanced Training in Scientific Computation, France), ECMWF (European Centre for Medium-range Weather Forecasts), INGV (Istituto Nazionale de Geofisica e Vulcanologia, Italy), LODYC (Laboratoire d'Océanographie Dynamique et de Climatologie, France), Météo-France (France), Met Office (UK) and MPI (Max-Planck Institut für Meteorologie, Germany). In order to assess seasonal dependence on forecast skill, the DEMETER hindcasts have been started from 1st February, 1st May, 1st August, and 1st November initial conditions. Each hindcast has been integrated for 6 months and comprises an ensemble of 9 members. In its simplest form, the multi-model ensemble is obtained by merging the ensemble hindcasts of the seven models, thus comprising 7x9 ensemble members. The performance of the DEMETER system has been evaluated from a comprehensive set of hindcasts over a substantial amount of years, with the main focus in 1987-1999 (Palmer et al., 2003).

To enable a fast and efficient post-processing and analysis of this complex data set, much attention was given to the definition of a common archiving strategy for all models. A large subset of atmosphere and ocean variables, both daily data and monthly means has been stored into the ECMWF's Meteorological Archival and Retrieval System (MARS). A significant part of the DEMETER data set (monthly averages of a large subset of surface and upper-air fields) is freely available for research purposes through an online data retrieval system installed at ECMWF[2].

A comprehensive verification system to evaluate all DEMETER single models as well as the multi-model DEMETER ensemble system has been set up at ECMWF. The system runs periodically to monitor hindcast production, to quality control the data (and correct archival) and to calculate a common set of diagnostics. The basic set of diagnostics (performed in cross-validation mode) comprises: global maps and zonal averages of the single-model bias, time series of specific climate indices, standard deterministic and probabilistic measures of forecast quality and a comparison of single-model ensembles skill with that of multi-model ensembles.

3. Results

Sea surface temperature skill assessment over the tropical Pacific suggests that both, the multi-model ensemble and the single models perform at levels comparable to dedicated ENSO prediction models and much better than persistence. In general, the identity of the most skilful single model varies with

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

[2] Monthly data can be retrieved in GRIB and NetCDF from http://www.ecmwf.int/research/demeter/data

region and year. Deterministic skill measures indicate that, in most years, the multi-model ensemble skill is close to the best single-model skill and is the most skilful when performance is averaged over all years. However, a key result was that multi-model ensemble probability scores were generally better than those from any of the single-model ensembles. The greater probabilistic skill of the multi-model ensemble compared to the single-model skill leads to an increased potential economic value (Richardson, 2000). For instance, it has been found that, for predictions of positive tropical winter (December to February, November start date) precipitation anomalies, the multi-model ensemble improves potential economic value from 15% to 80%, depending on the single model taken as reference (not shown).

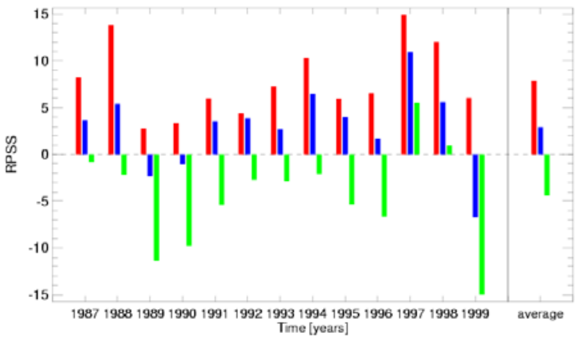

Figure 1: Ranked probability skill score for the 1-month lead tropical summer (JJA) precipitation 1987-1999 for the 54-member DEMETER multi-model ensemble (red), and the single ECMWF ensemble with 54 (blue) and 9 members (green). Average results are shown on the right end of the plot.

In spite of the clear improvement of the multi-model ensemble performance an important question arises. This improvement could be due either to the multi-model approach itself or to the increased ensemble size resulting from collecting all members of the single-model ensembles, or both. In order to separate the multi-model benefits that derive from combining models of different formulation to those derived simply from the accompanying increase in ensemble size, a 54-member ensemble hindcast has been generated with the ECMWF model alone for the period 1987-1999 using a single start date (May, boreal summer). As forecast quality measure, the ranked probability skill score for tropical summer (June to August) precipitation positive anomalies for a 54-member multi-model ensemble (red bars) and the ECMWF model (blue bars) is shown in Figure 1. Tercile categories have been used. The multi-model ensemble for this example was constructed by randomly selecting 54 members out of the 63 available in the multi-model hindcasts. Values for the ECMWF model 9-member ensemble are also shown (green bars). The single-model skill increases with ensemble size, though it turns out that the multi-model outperforms the single-model skill regardless of the ensemble size. Similar results are found for other variables and regions. It has been found that the largest contribution to the multi-model ensemble skill improvement is due to an increase in reliability.

4. Conclusions

As part of a European Union-funded DEMETER project, a multi-model ensemble system based on 7 European global coupled ocean-atmosphere models has been described and validated in hindcast mode using the ECMWF Re-Analysis ERA-40 data. Results indicate that the multi-model ensemble is a viable pragmatic approach to the problem of representing model uncertainty in seasonal-to-interannual prediction, and will lead to a more reliable forecasting system than that based on any one single model. As a result of the success of DEMETER, real-time multi-model ensemble forecasting is now being established as part of the operational seasonal forecast suite at ECMWF.

The DEMETER project has applications partners in agronomy and in tropical disease prediction. These models have been directly linked to the output of individual members of the multi-model ensemble, after correction of the bias and downscaling onto a finer grid than the one used in the coupled models. As such, the design of DEMETER was based on the concept of an “end-to-end” system, in which users can feed information back to the forecast producers. Results from the application models show that multi-model seasonal forecasts have useful economic value.

Acknowledgements

The DEMETER project has been funded by the European Union under the contract EVK2-1999-00197.

References

Richardson, D. S., 2000. Skill and relative economic value of the ECMWF ensemble prediction system. *Quart. J. Roy. Meteor. Soc.*, **126**, 649-668.

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, 2003. Development of a European multi-model ensemble system for seasonal to inter-annual prediction (DEMETER), submitted to *Bull. Amer. Meteor. Soc.*