Minutes of the preliminary ENSEMBLES RT1 meeting (ECMWF, Reading, UK, 3rd September 2004)

Attendants: Myles Allen, Magdalena Balmaseda, Matt Collins, Francisco Doblas-Reyes, Matt Huddleston, Malcolm McVean, James Murphy, Tim Palmer, Mark Rodwell, Lenny Smith.

The meeting was organized to prepare the ENSEMBLES Kick-Off meeting that will take place in Hamburg on 14-16 September 2004. The main goal was to decide the details concerning the integrations to be carried out in the initial 18 months of the project. Only UK-based partners were invited to attend to reduce costs, although the rest of the RT1 participants were informed of the meeting. It started at 11:00 and closed at 15:00.

The discussions were organized in two parts, following the two time scales under consideration: centennial and seasonal to decadal. The time scale separation is due more to practical reasons rather than to an essential difference in the way the forecast uncertainty problem should be tackled. The main difference resides in the verification of the predictions. Seasonal-to-decadal predictions are susceptible of verification and the reliability of the different forecast systems can be assessed. This is not the case with centennial predictions, for which the reliability assessment is based upon the performance of the different systems in reproducing present-day climate and climates changes in the past decades.

The centennial integrations will be carried out by both the Met Office and the University of Oxford using slightly different versions of HadCM3. The perturbed parameters method will be employed to estimate forecast uncertainty. Myles Allen suggested that large ensembles should be constructed to investigate several issues, such us the predictability of extreme events. He proposed a first experiment in which a large ensemble would be created by perturbing initial conditions for one particular version of HadCM3, based on an observed initial state obtained from the Met Office, in order to sample to sample uncertainty in, for instance, the THC. However, it remains to be decided specifically how big a large initial-condition ensemble should be. Although this is not one of the RT1 deliverables, it fits well with the activities planned within WP1.2 and WP1.6. This experiment will be considered in Hamburg as an extra activity. It has been suggested that the centennial simulations should use initial conditions similar to those of the multi-annual predictions. This would allow us to tighten links between both sets of predictions and move towards a generalised seasonal to centennial ensemble prediction system.

A close co-ordination with the activities in RT2A will be required. In particular, RT2A is generating a set of multi-model centennial integrations to be used for the 4th Assessment Report of the IPCC. A comparison of these simulations with the ones to be performed in RT1 during the first 18 months would contribute to the preliminary assessment of the relative merits of the perturbed parameter approach to representing forecast uncertainty (M1.3). This topic will be discussed in Hamburg during the joint RT discussions. Given that the multi-model approach is an ad-hoc method and that the perturbed parameter approach is still under development, the recommendations issued from this preliminary assessment of the best method to represent forecast uncertainty

is expected to be backed up by the RT1 and RT2A integrations carried out beyond month 18. It was also suggested that an ensemble of centennial simulations using the stochastic physics approach should be considered as another possible method, so ECMWF will study the possibility of performing such an experiment. It was speculated that a multi-model perturbed parameter approach would be an obvious recommendation for improving the centennial prediction system during the second phase of ENSEMBLES, although the possibility of using persisted systematic perturbations in the physics should also be considered.

An open question is the methodology used to assess the forecast quality of the ensemble prediction systems. In the seasonal-to-decadal time scale several scoring metrics of probabilistic predictions (Brier score, ROC, potential value, LEPS, …) are commonly used. These metrics are applied to different variables over different regions, and can give in specific cases very different answers to the question of which is the best method to estimate forecast uncertainty. However, the forecast quality of centennial integrations cannot be directly verified (or falsified). Therefore, different methods of assessing the quality of a centennial prediction system are required, such as ensuring that the system samples the widest possible range of modelling uncertainties and that it samples the range of values of observables that can be shown to be a relevant constraint on its predictions of long-term climate change. The relationship between verifiable probabilities produced by seasonal to decadal predictions and unfalsifiable probabilities produced by centennial predictions needs to be investigated to encourage a consistent approach across all time scales. More discussion seems to be required on this topic and a link to RT5 needs to be established.

The seasonal-to-decadal integrations for the first 18 months will be carried out following the experimental setup described in the appendix. To perform these experiments, the multi-model system (ECMWF, Met Office-GloSea, CNRM, IfM, CERFACS) will have to be installed at ECMWF (D1.4). The atmospheric initial conditions will be taken from ERA-40, while the ocean initial conditions are expected to be made available by the ENACT project. Therefore, an update of the present status of ENACT and its schedule until the end of the project is required by the RT1 partners. The issues of installing the models (when necessary), designing the common archiving strategy (that should take into account the extension of the integrations when compared to DEMETER) and preliminary forecast quality assessment are a priority. A preliminary working document on the common list of variables to be archived has already been prepared by RT2A and will be circulated within RT1 too. After 18 months, it is expected to provide a preliminary answer to the question of what is the best method to estimate forecast uncertainty (M1.2). The question of how relevant the conclusions of the seasonal-to-decadal experiment will be for the centennial problem was raised and will need further discussion.

Only two or three multi-annual ensembles are expected within the first 18 months. The proposed starting dates are 196x and 199x (to be discussed on 16/9/04 in Hamburg). The initialization of these simulations is still an open issue, as is the question of how to represent variations in external radiative forcing (arising anthropogenic changes in greenhouse gas and aerosol concentrations, solar variations and volcanic eruptions). A document describing the decadal prediction activities at the

Met Office has been prepared by James Murphy and Doug Smith and will be distributed to foster the discussions.

Although ENACT is an excellent starting point to address the problem of generating ocean initial conditions, some additional research will be carried out on this topic. In particular, it has been suggested to sample ocean uncertainty in extra-tropical areas. In addition, some work needs to be done to assess the relevance of introducing soil moisture perturbations, for which a strategy similar to the one used to sample wind stress uncertainty has been suggested.

An RT1 website (D1.0) will be created and maintained at ECMWF.

Appendix: List of seasonal-to-decadal integrations to be carried out during the first 18 months

1. Multi-model ensemble

Two 6-month ensemble hindcasts per model per year, initialised from 1 May and 1 November from (at least) 1991-2001. Each ensemble will have 9 members, started from slightly different initial conditions.

One 12-month ensemble per model per year, again 9 members. This may be done by extending one of the 6-month ensembles, or by doing a separate experiment, at the discretion of individual partners.

Two (or maybe three) 9-member decadal ensembles per model (duration somewhere between 5 and 10 yrs). Initial dates to be confirmed following discussion with the modelling groups, although a first proposal is 1964 and 1994. These simulations could also be done by either extending existing runs or by doing separate experiments.

The initial conditions will be taken from ENACT, whenever possible. Both wind and SST perturbations will be used to generate the ensemble.

2. Stochastic physics ensemble

A set of additional 9-member ensembles will be done with the ECMWF model using an updated and more physically sound stochastic physics scheme, following the above design.

3. Perturbed physics HadCM3 ensemble

Sixteen versions of HadCM3 with parameter perturbations will be available from the QUMP project. Ideally, 9-member initial condition ensembles for all of these will be run, for the 22 start dates for 10 years. However, for the first 18 months the intention is to run a 9-member perturbed physics ensemble consisting of the "standard" version plus eight members containing parameter perturbations. In addition, a 9-member ensemble of the standard version distinguished by different initial conditions will be performed. This will be done for all 22 initialisation dates (May and November 1991-2001), giving 17x22 integrations, each of 10 years duration.

4. Perturbed physics GloSea contribution

A set of nine 6-month perturbed parameter GloSea hindcasts twice a year using the same parameter perturbations as in the 9-member perturbed physics ensemble of HadCM3. No initial condition perturbations will be used.

The differences in the two MetOffice systems (HadCM3 and GloSea) are the ocean analysis, anomaly coupling and model resolution. A mean of the available hindcasts for each month for each GloSea version will be used to estimate the model drift.