Project no. GOCE-CT-2003-505539

Project acronym: ENSEMBLES

Project title: ENSEMBLE-based Predictions of Climate Changes and their Impacts

Instrument: Integrated Project

Thematic Priority: Global Change and Ecosystems

**Deliverable D1.4**

A new multi-model coupled model ensemble system for seasonal to decadal forecasts will be created and installed at ECMWF, with capabilities to run, in addition, perturbed parametrizations, and stochastic physics.

Due date of deliverable: month 18
Actual submission date: 2 March 2006

Start date of project: 1 September 2004  Duration: 60 Months

Organisation name of lead contractor for this deliverable: ECMWF

<table>
<thead>
<tr>
<th>Dissemination Level</th>
<th>PU</th>
<th>PP</th>
<th>RE</th>
<th>CO</th>
</tr>
</thead>
<tbody>
<tr>
<td>Public</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Restricted to other programme participants (including the Commission Services)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Restricted to a group specified by the consortium (including the Commission Services)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Confidential, only for members of the Consortium (including the Commission Services)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
ENSEMBLES

Deliverable D1.4

WP number: 1.4
Participants: ECMWF, METO-HC, IfM-Kiel, CNRM, CERFACS
Due: month 18

A new multi-model coupled model ensemble system for seasonal to decadal forecasts will be created and installed at ECMWF, with capabilities to run, in addition, perturbed parametrizations, and stochastic physics.

One of the objectives in RT1 is the development and assessment of methods to represent forecast uncertainty arising from initial conditions and model error. Using ensembles of different initial conditions is a methodology to account for the first problem. Model error in climate forecasts occurs because climate models cannot in principle simulate every single aspect of the climate system with arbitrary detail. For example, the way in which processes on spatial scales smaller than the resolved grid scales are represented in climate models is uncertain. Likewise, many of the physical parameters used in climate models either don’t have a direct equivalent in the real climate system, or their numerical values are not precisely known.

One way to represent model uncertainty is to incorporate, within the ensemble of initial conditions, different models themselves and thus to construct a multi-model ensemble of quasi-independent models. Other possible ways are to explicitly account for the effect of the unresolved scale on the larger scale motion by introducing a stochastic component to the model parameterisations, or to perturb physical model parameters within the range of their uncertainties. All three approaches are applied in RT1.
**Task 1.4.1**
The latest versions of the ECMWF, METO-HC, CNRM, CERFACS and MPIfM global coupled climate models will be installed on the IBM supercomputer at ECMWF, with initialisation procedures taken from the FP5 ENACT programme.

A multi-model ensemble system for climate prediction on seasonal, interannual and decadal time scales has been developed and installed on a single supercomputer. This allows for efficient implementation of computer-intensive ensemble experiments, while ensuring model diversity. The system design includes a set of common experiments in order to assess the benefits of the different approaches to model uncertainty.

The ENSEMBLES seasonal-to-decadal multi-model system is based on the experiences of the FP5 DEMETER system and the operational multi-model seasonal forecast system (EUROSIP) installed at ECMWF. Since the end of the DEMETER project, a significant amount of work has been devoted to upgrade many of the climate model components and to port the models to the newer supercomputer architecture (from Fujitsu VPP platforms to IBM mainframes). EUROSIP comprises the operational seasonal forecast models from the UK Met Office, Météo-France and ECMWF and has benefited from the effort mentioned above.

The ENSEMBLES seasonal-to-decadal multi-model ensemble is built from six coupled climate model systems, namely IFS/HOPE (ECMWF), ARPEGE4.5/OPA (CNRM), ARPEGE3/OPA (CERFACS), GloSea (METO-HC), DePreSys (METO-HC), and ECHAM5/MPI-OM1 (IfM-Kiel). See Table 1 at the end of this report for a brief summary of the individual model component details.

In four of these systems, model uncertainties in the initial state are represented through an ensemble of nine different ocean initial conditions. This is achieved by creating three different ocean analyses. A control ocean analysis is forced with momentum, heat and mass flux data from ERA-40 (40 years of ECMWF Re-Analysis, Uppala et al., 2005; see also http://www.ecmwf.int/research/era/), and two perturbed ocean analyses are created by adding daily wind stress perturbations to the ERA-40 momentum fluxes. The wind stress perturbations are randomly taken from a set of monthly differences between the analyses from CORE (Coordinated Ocean Reference Experiments; see also http://data1.gfdl.noaa.gov/nomads/forms/mom4/CORE.html) and ERA-40, which are considered as quasi-independent datasets. In addition, in order to represent the uncertainty in SSTs, four SST perturbations are added to or subtracted from the initial field at the start of the hindcasts. As in the case of the wind perturbations, the SST perturbations are based on differences between two quasi-independent SST analyses, here Reynolds2DVAR and ReynoldsOlv2. More details describing how the wind stress and SST perturbations are estimated can be found on the ENSEMBLES RT1 web site at http://www.ecmwf.int/research/EU_projects/ENSEMBLES/documents/docu_perturbations.pdf. The atmospheric and soil initial conditions come from ERA-40. Two models use a different initialisation strategy, as described below.

Figure 1 shows a schematic representation of the ensemble generation and hindcast production strategy for the majority of the models. Dashed lines represent the three continuous runs of ocean analyses forced by ERA-40 data, the control analysis without any wind stress perturbations (grey) and two additional analyses with positive/negative (green/red) daily wind stress perturbations applied. In order to generate nine different initial conditions for the coupled hindcasts, four SST perturbations (represented by the ellipses) are added to (blue ellipse) or subtracted from (yellow ellipse) to the ocean analyses. Thus, there is one member with no wind stress perturbations.
stress or SST perturbations applied and eight perturbed ensemble members. This procedure is performed at every start date of the hindcasts.

**Figure 1**: Schematic representation of the ensemble generation and hindcast production strategy for two different start dates: May and November 1991.

The common hindcast production period for the stream 1 seasonal simulations is 1991 to 2001; stream 2 will cover the period 1960 to 2001. For each of these years, in stream 1 two sets of simulations are launched, one starting on May 1 at 00 GMT and the other one on November 1. The May start date runs cover at least 7 months, whereas the November start dates are hindcasts over at least 14 months. Furthermore, multi-annual/decadal simulations have been carried out with each model, with a minimum of one start date in November 1994.

At ECMWF the latest version of the IFS atmospheric model component (CY29R1) coupled to the HOPE ocean model has been successfully installed on ECMWF’s HPC IBM supercomputer. The model includes land surface modules and climatological sea-ice cover. For the initialisation of the atmosphere, perturbations based on singular vectors have been applied in a similar way as in the operational medium range ensemble forecasts. As boundary forcings the evolution of greenhouse gases (CO$_2$, CH$_4$, N$_2$O, CFC12, CFC11), a climatological annual cycle of 5 types of aerosol (sea-salt, desert dust, organic matters, black carbon) and interannual solar activity have been used.

ECMWF performed 2 seasonal-to annual simulations per year over the 11-year period from 1991-2001, as described above. Additionally, two decadal runs over 10 years starting in November 1965 and in November 1994 have been completed.

Two forecasting systems have been installed as the Met Office (METO-HC) contribution to D1.4.

Firstly, the METO-HC GloSea forecasting system as initially developed in DEMETER has been further enhanced to include the evolution of: greenhouse/trace gases (CO$_2$, CH$_4$, N$_2$O, CFC11, CFC12, CFC113, HCFC22), ozone concentrations, volcanic aerosols and the interannual solar cycle. For initialisation, an ensemble of ocean analyses was used which is generated by using prescribed wind stress and SST perturbations, see Figure 1.
The standard experiment design has been followed as described above for both seasonal and decadal runs with additional seasonal hindcast experiments being done using lagged ocean analyses rather than perturbed analyses. A further set of hindcasts has been initialised 2 weeks prior to the standard start dates to enable studies on information retention and decay in the forecasting system.

Secondly, the METO-HC DePreSys system is based on the HadCM3 climate model. The version installed on the ECMWF supercomputer is similar to that run previously at the Hadley Centre (Smith et al., 2006), except that the ENSEMBLES version includes an enhanced representation of the atmospheric sulphur cycle, and flux adjustments are used to restrict the development of regional biases in sea surface temperature and salinity (Collins et al., 2006). The historical and future forcing datasets (described below) also differ in some respects from Smith et al. (2006). Two variants of DePreSys are described in this report. The first, described here, is set up to generate perturbed initial condition ensembles using the "standard" model version with parameter settings following those of Gordon et al. (2000): it is contributed as a member of the multi-model ensemble. The second variant, described in Task 1.4.2 below, is set up to generate ensembles sampling modelling uncertainties by perturbing poorly constrained model parameters.

In order to create initial conditions for the hindcasts, HadCM3 was run in assimilation mode from December 1989 to November 2001. During this integration, the atmosphere and ocean were relaxed towards analyses of observations, which were assimilated as anomalies with respect to the model climate in order to minimise climate drift when the assimilation is switched off. The assimilation integration was itself started from an initial state taken from a simulation of 20th century climate including variations in radiative forcing derived from observed changes in well-mixed trace gases (CO₂, CH₄, N₂O, CFC11, CFC12, CFC113 and HCFC22), ozone, sulphate aerosol, total solar irradiance and volcanic aerosol. The forcings in the assimilation and hindcast integrations are the same, switching from values based on observations to values based on the SRES A1B emissions scenario after the year 2000. One exception is that during hindcasts solar irradiance was estimated by repeating the previous 11 year solar cycle and volcanic aerosol was specified to decay exponentially from the initial value with a time scale of one year. This was done to avoid assuming accurate knowledge of future variations in these forcings.

The stream 1 hindcasts consist of nine member ensembles obtained by starting hindcasts from initial states taken from the assimilation integration 0,1,2,...,8 days prior to each hindcast start date (see above). All hindcasts will be run for 10 years (unlike the other multi-model partners), started from May and November 1991 to 2001. These simulations will be run during April 2006, following completion of the perturbed parameter simulations of Task 1.4.2.

The ARPEGE/OPA model at CNRM uses the most recent cycle 24T1 of ARPEGE-IFS for the atmosphere and OPA8.2 as the ocean model. The coupling between ocean and atmosphere is done using the coupler OASIS (version 3). The atmospheric part is very similar to the one used in DEMETER. Detailed information about the model dynamics and physics can be found in http://www.cnrm.meteo.fr/hiretys/div/arp4cae.pdf. OPA8.2 uses a different grid than the ORCA model in DEMETER and a free surface dynamics. Both atmosphere and ocean are coupled with the GELATO sea-ice model (Salas Melia, 2002).
CNRM performed 2 seasonal-to-annual simulations per year over the 11-year period from 1991-2001, as described above. Additionally, two decadal runs over 10 years starting in November 1965 and in November 1994 have been completed.

For stream 1, CERFACS has run the DEMETER system initialised with new initial conditions issued from an ocean reanalysis using an early version of the 3D-variational system developed in ENACT. The ensembles were initialised using the same approach as DEMETER, with the addition of perturbed temperature observations (this is described in Rogel et al., 2005). Additional runs including decadal ensembles are under way.

For stream 2, CERFACS is committed to produce ocean initial conditions using an updated 3D-Var version and decadal integrations using a coupled model with the same ARPEGE version as CNRM. This model has been extensively tested at CERFACS and will be installed at ECMWF for production. The generation of ocean initial conditions ensembles will take advantage of an ensemble method currently developed at CERFACS to design the variational system.

IfM Kiel is using the latest version of the Max-Planck-Institute for Meteorology (MPIfM) climate model to perform the seasonal-to-decadal hindcasts. The model has been successfully installed on the ECMWF supercomputer. The initialisation method is different to that described above: ocean and atmospheric initial conditions are generated from three coupled runs for the period 1950 till present, in which model SSTs are strongly damped to observed (NCEP) SST. The 9-member ensemble is generated by different combinations of ocean and atmosphere initial conditions from these three runs. As boundary forcings the evolution of greenhouse gases (CO₂, CH₄, N₂O, O₃, CFC12, CFC11, and other CFC species), natural and anthropogenic sulphate aerosol, and interannual solar activity have been used. The effect of volcanic aerosols is also included by varying optical depth.

IFM Kiel has performed 2 seasonal and annual simulations per year over the 11-year period from 1991-2001, as described above. The two decadal runs over 10 years starting in November 1965 and in November 1994 are underway and will be completed soon.

**Tasks 1.4.2**

In addition, the Met Office model will be installed with perturbed parametrisation schemes, and the ECMWF model will be installed with stochastic physics.

**METO-HC**

The DePreSys system has been extended to sample modelling uncertainties, by perturbing uncertain surface and atmospheric parameters in HadCM3. Nine-member ensembles consist of the version with standard parameter settings plus eight versions with simultaneous perturbations to 29 parameters. The perturbations were selected in two stages from a larger ensemble of 128 atmospheric model versions designed to sample a wide range of parameter and climate sensitivity values while providing credible simulations of present day multiannual mean fields when coupled to a mixed layer ocean (Webb et al., 2006).

**Stage 1**

From these 128 model versions, a subset of 16 was selected to simulate transient climate change (in ENSEMBLES WP 1.6) when coupled to the dynamical ocean module of HadCM3. These were selected by choosing the member with the best simulation of present day climate, followed by 15 further members designed to
maximise a non-dimensional measure of the average distance between members in terms of both climate sensitivity and model parameter values. This procedure was different from that followed by Collins et al. (2006), and gave a different selection of parameter combinations leading to a wider range of climate sensitivities.

Stage 2
For seasonal-to-decadal prediction, the 16 HadCM3 versions were reduced to eight by picking members sampling a wide range of climate sensitivities (from 2.6-7.1°C) and a wide range of ENSO amplitudes (diagnosed from the simulated standard deviation of monthly NINO3 SST anomalies, which range from 0.5-1.2°C compared to the observed value of 0.8°C).

Each of the perturbed versions of HadCM3 used identical ocean physics, but required a separate set of flux adjustment fields due to the effects of the atmospheric perturbations on simulated surface heat and water fluxes. These were calibrated as described by Collins et al. (2006). Initial conditions were created by following the procedure described in Task 1.4.1. Specifically, a separate assimilation integration was run for each perturbed version in which the model was relaxed towards time series of ocean and atmospheric analyses expressed as anomalies relative to the model climatology (estimated for each version from a simulation of 20th century climate carried out for WP 1.6).

The stream 1 ensemble hindcasts are in progress, and will be completed by April 2006. As in Task 1.4.1, nine-member ensembles of decadal hindcasts will be run for each May and November start date from 1991-2001.

This method of sampling modelling uncertainties is distinct from both the multi-model ensemble of Task 1.4.1 (in that it samples perturbations to uncertain parameters in a single-model framework rather than sampling perturbations to the framework itself), and also from the stochastic parameterisation approach described below (in that it is designed to produce sustained rather than stochastic perturbations to the outputs of parameterisation schemes).

ECMWF
Stochastic parametrisation can potentially lead to both a reduction of systematic error through a noise-induced drift effect, and to a more complete representation of model uncertainty in ensemble prediction systems. It is used here to study its impact on seasonal-to-decadal forecasts quality and to assess the benefits compared with other approaches to represent model uncertainty, like multi-model or perturbed physical parametrisation ensembles.

The stochastic parametrisation scheme (CASBS, Cellular Automaton Stochastic BackScatter) used has been recently developed for the medium-range ensemble forecasting system at ECMWF. It is based on the idea of backscatter of kinetic energy from unresolved spatial scales (Shutts, 2005). Here, the level of dissipation associated with parametrisation of convection, orographic wave drag and numerical dissipation (horizontal diffusion and semi-Lagrangian interpolation error) is calculated. A fraction of it is re-injected into the model near the truncation scale to account for energy transfer out of the sub-grid scale and back to the resolved scale. The scales onto which this energy is backscattered are determined by a simple cellular automaton model, which essentially plays the role of a stochastic number generator.
CASBS has been successfully applied for the first time to seasonal-to-decadal simulations. For the stream 1 hindcast production period 1991-2001, a total of 22 simulations with start dates in May and November have been completed using CASBS version 1.1. Additionally, two 9-member ensembles of 10-year long simulations starting in November 1965 and in November 1994 are available. For those runs CASBS version 1.2 has been applied, which uses a new algorithm to calculate the dissipation from deep convection based on mass flux.

Task 1.4.3
Unified output and archival routines will be developed, so that atmosphere and ocean data can be output into ECMWF MARS archival.

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 output and archiving strategy for all models. Special attention was given to the time-consuming task of ensuring that all model output complies with agreed data formats and units. Figure 2 shows the strategy for the seasonal-to-decadal output, storage and dissemination of model data.

![Figure 2: Overview of the seasonal-to-decadal output, storage and dissemination strategy](image)

A large subset of atmosphere and ocean variables, both daily data and monthly means, has been stored (common output in Fig. 2). A complete list of the atmosphere and ocean archived variables can be found on the ENSEMBLES RT1 website [http://www.ecmwf.int/research/EU_projects/ENSEMBLES/news/common_variables.html](http://www.ecmwf.int/research/EU_projects/ENSEMBLES/news/common_variables.html). These variables are a minimum common set to ensure that some analyses, downscaling exercises and end-user applications can be carried out for all the coupled models. A longer list is also available upon request from the individual modelling centres.

Atmospheric fields from all seasonal-to-decadal models are stored in the Meteorological Archival and Retrieval System (MARS), the main repository of meteorological data at ECMWF. MARS data is freely available to registered users in the Member States and Cooperating States. The fields are archived following a set of atmospheric conventions (for details, see RT1 website), based on the FP5
programme DEMETER and the operational European multi-model seasonal forecasts (EUROSIP). The format of the archived atmospheric output is GRIB.

Ocean data are archived in the ECMWF File Storage system (ECFS), the Centre's archive system for user files. Monthly means for both analyses and forecasts are available in the common ocean archiving. Both 3-D fields and a small set of 2-D fields are archived in NetCDF format. Vertical sections can be derived from the 3-D fields. The ocean encoding is carried out using the common grid from the FP5 programme ENACT conventions, with storage of CF-compliant NetCDF files into ECFS. The ENSEMBLES s2d ocean conventions allow the preparation of the NetCDF files and their archiving.

Within the set of stream-1 seasonal-to-interannual simulations, three partners (METO-HC, CNRM and ECMWF) have agreed to additionally store seasonal model level data as input for downscaling activities (additional output in Fig. 2). These data will be available for the hindcast period 1991-2001, using May and November as start dates for runs over 6 months. All model level data will be archived in MARS. The ENSEMBLES RT1 web site provides information on how to interpolate model level to pressure level data. This might not only be useful as input for regional models, but also for the partners to create additional sets of pressure-level data for more specific diagnostics.

All data are initially stored in MARS and ECFS at ECMWF and, later, a major subset will be publicly disseminated through public data and OPeNDAP servers in both, GRIB and NetCDF formats. In collaboration with WP2A.4 “Storage, extraction and creation of distributed databases for provision of the results” a public dissemination strategy for the output of the seasonal-to-decadal simulations has been developed (see Fig. 3). It involves a 5 TByte public data server provided by ECMWF, which will have copies of all common atmospheric and ocean output. A MARS client and an OPeNDAP server will be installed and provide the gateway to the public. Furthermore, the KNMI Climate Explorer software (http://climexp.knmi.nl) will be linked to the public data server in collaboration with WP5.3. It not only allows access to a large amount of datasets including station data, climate indices, reanalyses, but also provides tools for easy extraction and plotting of fields and time series and standard diagnostics like correlations, EOFs and probabilistic forecast verification estimates.
Figure 3: Overview of the dissemination strategy used for seasonal-to-decadal simulations.
### Table 1: Combinations of atmosphere and ocean models used by the modelling groups. The resolution of the models and the initialization strategy is outlined as well.

<table>
<thead>
<tr>
<th></th>
<th>CERFACS</th>
<th>ECMWF</th>
<th>CNRM</th>
<th>METO-HC DePreSys</th>
<th>METO-HC GloSea</th>
<th>IFM-Kiel</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>atmosphere component</strong></td>
<td>ARPEGE</td>
<td>IFS</td>
<td>ARPEGE</td>
<td>HadAM3</td>
<td>HadAM3</td>
<td>ECHAM-5</td>
</tr>
<tr>
<td><strong>resolution</strong></td>
<td>T63</td>
<td>T95</td>
<td>T63</td>
<td>2.5° x 3.75°</td>
<td>2.5° x 3.75°</td>
<td>T63</td>
</tr>
<tr>
<td><strong>atmosphere initial conditions</strong></td>
<td>ERA-40</td>
<td>ERA-40</td>
<td>ERA-40</td>
<td>ERA-40 anomalies</td>
<td>ERA-40</td>
<td>coupled run relaxed to observed SSTs</td>
</tr>
<tr>
<td><strong>ocean component</strong></td>
<td>OPA 8.2</td>
<td>HOPE-E</td>
<td>OPA 8.2</td>
<td>HadCM3 OGCM</td>
<td>GloSea OGCM, based on HadCM3</td>
<td>MPI-OM1</td>
</tr>
<tr>
<td><strong>resolution</strong></td>
<td>2.0° x 2.0°</td>
<td>1.4° x 0.3°-1.4°</td>
<td>1.25° x 1.25°</td>
<td>1.25° x 0.3°-1.25°</td>
<td>1.5° x 1.5°</td>
<td>40 levels</td>
</tr>
<tr>
<td><strong>ocean initial conditions</strong></td>
<td>ocean analyses forced by ERA-40</td>
<td>ocean analyses forced by ERA-40</td>
<td>ocean analyses forced by ERA-40</td>
<td>coupled run relaxed to analyses of temperature and salinity anomalies</td>
<td>ocean analyses forced by ERA-40</td>
<td>coupled run relaxed to observed SSTs</td>
</tr>
<tr>
<td><strong>ensemble generation</strong></td>
<td>windstress and SST perturbations</td>
<td>windstress and SST perturbations</td>
<td>windstress and SST perturbations</td>
<td>lagged initial conditions and perturbed model parameters</td>
<td>windstress and SST perturbations</td>
<td>ensemble coupled initialisation runs</td>
</tr>
</tbody>
</table>
References


