SUMMARY This document describes the structure of the CMCC Seasonal Prediction System (hereafter CMCC-SPS) developed at CMCC, the various initialization methods and the experiments performed with the most recent version. The CMCC-SPS already contributed to the European multi-model seasonal prediction system in the framework of the EU project ENSEMBLE-based predictions of climate changes and their impacts (ENSEMBLES) and has been used to perform the hindcast experiments for the international CliPAS (Climate Prediction and its Applications to Society) ISO (Intra-Seasonal Oscillation) and EU Climate change predictions in Sub-Saharan Africa: impacts and adaptations (CLIMAFRICA) projects. This report is focused on the current CMCC-SPS version that includes an atmospheric off-line initialization tool, other than the ocean initial state which had been initialized in the first version. A set of retrospective forecasts was performed for the period 1989-2010. Every year, the model is restarted four times with re-analysis initial conditions, and then integrated for 6 months. Here we provide some preliminary results from this forecasting experiment.
Contents

Introduction ................................................................. 4
The Prediction System ....................................................... 4
    Model components .................................................... 4
    Methodologies of initialization .................................. 5
Experimental setup .......................................................... 7
    Verification Data ..................................................... 7
Results ........................................................................ 7
Conclusions and future works .......................................... 9
Bibliography ................................................................ 16
INTRODUCTION

The basis for dynamical prediction is the assumption that large scale, long-lasting anomalies in surface fields will impart predictive skill to seasonal forecast. The atmosphere, interacting with the slowly varying component of the climate system (land surface and ocean) might enhance predictive skill at a regional scale. Also, atmospheric teleconnections are able to propagate climate variability on monthly and longer timescales between distant regions of the globe, and models may use this information for the purpose of predictability. This work summarizes the technical development of the CMCC Seasonal Prediction System, which is the evolution of the system described in Alessandri et al. (2010) [3]. The prediction system is described in section n. 2. The experimental set-up and validation data are illustrated in section n. 3. Results are provided in section n. 4 and, finally, section n. 5 summarize the conclusions and future plan.

THE PREDICTION SYSTEM

The seasonal prediction system documented in the present study represents the evolution of the system described in Alessandri et al. (2010) [3] and further developed in the framework of the EU project ENSEMBLES and for the international CliPAS ISO hindcast experiment. The initial conditions for the ocean-atmosphere system are prepared separately for the atmosphere and for the ocean in off-line mode. Figure 1) summarizes the structure of the latest version of the CMCC-SPS. The details are provided in the subsections 2.1 and 2.2, while section 3 reports the setup, integrations and validation data used for this study. In the following, subsection 2.1 describes the coupled model components, while subsection 2.2 reports the available initialization strategies.

MODEL COMPONENTS

The coupled model included in the CMCC-SPS is represented by the physical core (i.e. with the carbon cycle dynamics turned off) of the CMCC Earth System model documented in Fogli et al (2009), Vichi et al. (2011) and Alessandri et al. (2012) [9] [13] [4]. The coupled model is composed by three parts:

- Land-Atmosphere $\rightarrow$ ECHAM5-SILVA;
- Ocean/Sea-Ice $\rightarrow$ OPA-LIM;
- Coupler $\rightarrow$ OASIS3.

The atmospheric component is constituted by the atmospheric model ECHAM5.3 (Roekner et al., 1996, 2003, 2006) [15] [16] [17], a global spectral model truncated at zonal wave number 63. The associated Gaussian grid in which the physical fields are calculated has 192 points in the longitudinal direction and 96 points in the latitudinal direction. Therefore, the horizontal resolution of the atmospheric model is about 200 Km.

The vertical structure of the model is represented by 19 levels, which use sigma as the vertical coordinate.

ECHAM5.3 is coupled to the land-surface-vegetation model SILVA (Alessandri 2006, Alessandri et al., 2007, Alessandri et al., 2012) [1] [5] [4].

The ocean component is OPA8.2 (Madec et al., 1998) [14], an Ocean Global Circulation Model (OGCM) run on an ORCA2 grid, quasi-isotrope tri-polar grid (2 poles in the northern hemisphere, one over Canada and the other over Siberia). The horizontal resolution is variable, with a nominal resolution of 1.5° in latitude and 2° in longitude, with an increase to 0.5° latitude near the equator. The vertical structure of the model is represented by 31 levels with 10-m resolution in the top 100 m. OPA8.2 is coupled to the SEA-ICE model LIM (Timmermann
et al., 2005) [20] run at the same horizontal resolution. Fluxes between the atmosphere and the ocean (Scoccimarro et al., 2007; Fogli et al., 2011) [18] [9] are exchanged through the OASIS3 coupler (Valke et al., 2000) [21]. Coupling frequency - No adjustment.

**METHODOLOGIES OF INITIALIZATION**

Initial conditions for the various components are imposed to the fully coupled CMCC general circulation model. The various CMCC-SPS versions differ for the initialized components. (see also table 1) specifically:

- **CMCC-SPS1**: only the ocean state is initialized;
- **CMCC-SPS1.5**: both ocean and atmosphere are initialized with observed (analyses) 3-D fields;
- **CMCC-SPS2**: ocean, atmosphere and land surface are initialized;

Here we illustrate and discuss the results obtained from CMCC-SPS2 version of the seasonal prediction system, in which the atmospheric and land components are initialized with data from the ERA Interim reanalysis (Berrisford et al., 2009) [7]. In the CMCC-SPS1.5 and CMCC-SPS2, the multi-level fields (temperature, specific humidity, divergence and vorticity) are interpolated from the 60 hybrid levels of ERA-Interim to the 19 hybrid levels of ECHAM5 through INTERA (Kirchner, 2001) [10]. Surface pressure, surface geopotential,
surface temperature, land-sea mask and snow depth form the set of surface initial condition. In order to fit into the model, the atmospheric initial state needed to be horizontally interpolated from the ERA Interim resolution (T255) to ECHAM5 resolution (T63). In the following, we describe the way the interpolation was carried out:

1. SPECTRAL (HIGH RES) → GRID(HIGH RES)
   (a) wind field spectral fields
   (b) transform spectral fields

2. GRID(HIGH RES) → GRID(LOW RES)
   (a) Full field interpolation to grid space
   (b) humidity correction

3. GRID(HIGH RES) → TRUNCATION → GRID(LOW RES)
   (a) conversion of humidity fields to spectral space
      i. spectral transform
      ii. truncation
      iii. back transform
   (b) humidity correction
   (c) truncate spectral fields
   (d) calculate wind field at low resolution
   (e) spectral to grid transformation

4. VERTICAL INTERPOLATION
   (a) interpolation at ECHAM resolution grid
   (b) vertical interpolation ERAI (LOW RES) → ECHAM (LOW RES)
   (c) reset log surface pressure
   (d) form new spectral fields
   (e) calculate vorticity and divergence
   (f) humidity field unfiltered

5. FINAL CALCULATION
   (a) correction humidity fields
   (b) compose SST and land sea mask
   (c) correct surface temperature
   (d) second part of statistic energy correction of surface temperature t(lowestlevel) truncated back transform
(e) store the new lowest temperature
(f) interpolate snow
(g) interpolate sst field

6. WRITE INTERPOLATED FIELDS

(a) surface spectral fields
(b) hybrid level spectral fields
(c) surface Gaussian grid fields
(d) hybrid level Gaussian grid fields

The land surface-vegetation model SILVA is initialized with soil moisture and soil temperature data also taken from ERA Interim reanalysis. The ocean initialization of the CMCC-SPSs, uses data assimilation products made available by CMCC-INGV Global Ocean Data Assimilation (CIGODAS, Di Pietro and Masina, 2009, Bellucci et al. 2007) [8] [6]. CIGODAS is based on the assimilation of temperature and salinity profile and forced by ECMWF operational analysis using an optimal interpolation scheme. In the ocean analyses no observational information is included about the sea-ice. Therefore, the initial sea-ice cover is empirically diagnosed at the onset of the forecasts from the analyzed SST of the ocean model, while sea ice depth is set to a monthly climatology obtained through a long-lasting simulation of the CMCC-SPS. To this aim, the coupled model was run from 1960 to 2008 at a "present climate" state (i.e. the radiative forcing changes every year according to greenhouse gases concentrations), and the resulting sea ice depth climatology (calculated from a 7-day running mean) of the last 20 years was used. Hence, each start-date contains the sea-ice depth information as a modeled monthly climatology.

In the CMCC-SPS, the uncertainty that characterizes the initial state of the system is accounted for by using ensembles of perturbed atmospheric ICs. For CMCC-SPS1.5 and CMCC-SPS2, the perturbed atmospheric ICs are obtained by taking slightly lagged initial states from the reanalysis.

EXPERIMENTAL SETUP

A set of retrospective forecasts (hindcasts) was performed for 22 years (1989 through 2010). Every year, the model is initialized at 4 different start-dates (Fig. 2) (February 1st, May 1st, August 1st, November 1st) and then integrated for 6 months. In order to account for the uncertainty that characterizes the initial state of the system, an ensemble of nine perturbed atmospheric initial conditions (ICs) was prepared for each start date. Specifically, the atmospheric conditions are perturbed, by imposing restart files saved every 12 hours during the 4 days preceding the start date (see also Fig. 3). In this way, we obtain 9 different initial states from which the SPS evolves, producing a probability distribution of the forecast.

VERIFICATION DATA

The predictive skill of the model is assessed comparing the forecasts with analyses products. The ERA-Interim reanalysis (Berrisford et al. 2009 [7]) is used for verification of the forecasts. The model and the observed (reanalysis) anomalies are defined as deviations from the respective climatology for the period 1989-2009.

RESULTS

A detailed discussion of the results obtained from the CMCC-SPS1 version can be found in Alessandri et al. (2010) [3]. The sensitivity of the SPS1 forecast skill to the improvement of the oceanic ICs by including subsurface initialization in the ocean can be found in Alessandri
et al (2010, 2012) [3] [2]. The reader is further referred to two upcoming papers by Alessandri et al. (2013) and Materia et al. (2013), which discuss the sensitivity of the CMCC-SPS2 to the initialization of the atmosphere and land components.

In this section we briefly present and discuss the performance of the CMCC-SPS2. When not differently specified, the maps we show represent the ensemble mean of the nine members of the forecast. Figure 4 shows the systematic error for sea surface temperature (SST) and precipitation. Each plot refers to an average of the 3 months following the start date (i.e. for the start date 1st of February, the March-April-May mean is shown). Clearly, large biases show up in the polar regions, due to the lack of information about sea ice. Cold temperatures systematically affect the central equatorial Pacific, as well as the north eastern Pacific and the north Atlantic. Prediction for equatorial Atlantic are most of the times warmer than verification.

With regard to precipitation, the largest dry biases are shown in equatorial and southern Africa, and in southeastern Asia. Forecast for northern Sahel, Sahara region and Amazon basin, on the other hand, appear to be systematically wetter than verifications with the exception of the austral summer. To evaluate the skill of the model, we made use of the anomaly correlation (ACC) for each grid point between the predicted and observed anomalies, calculated as:

$$ACC = \frac{\sum_{m=1}^{M} x'_m o'_m}{\sqrt{\sum_{m=1}^{M} x'^2_m \sum_{m=1}^{M} o'^2_m}}$$  

(1)

where $x'$ and $o'$ are the anomalies of the model and the observations respectively. The anomalies are computed with respect to the reference period 1989-2009.

Figure (5) shows surface temperature anomaly correlation at lead time 1 (from 1 to 3 months). The highest skill is found in the Tropics, and in particular over central and eastern tropical Pacific (ENSO region). The ocean-atmosphere coupling in this area makes the climate system predictable up to lead 3 months, while the model performances degrade almost everywhere else as leadtime increases. Figure (6) shows the ACC between predicted and observed SST in the Niño3.4 region (area between 190°W and 240°W in longitude and between 5°S and 5°N in latitude). The ensemble mean shows the higher skill compared to any ensemble member.

Only a few ocean regions in the northern Atlantic and Pacific shows some predictability at lead time 2 and 3 months. In particular, CMCC-SPS shows considerable skill in the central sector of northern Atlantic during winter, and along the coasts of Alaska and western Canada throughout the year.

On land, ACC is rather weak, and predictability skill degrades consistently as the run proceeds, moving away from the initial condition. Reasonable predictability is detected at high latitudes during northern winter, thanks to the snow depth initialization that regulates temperature anomalies through albedo effect. Also, tropical regions characterized by strong coupling with the ocean, like the Amazon basin and part of Australia, associated with the ENSO signal, show a moderate predictability.

Figure (7) shows the surface temperature Root Mean Square Error (RMSE) of the system (2) at lead time 1, that is the season initiating one month after the start date (e.g. June-July-August for the start date of May). Discrepancy from the observations (ERA interim) is very much pronounced at high latitudes, where the presence of a modeled sea ice climatology affects sea surface temperatures and land temperature accordingly.
\[ \text{RMSE} = \sqrt{\frac{1}{M} \sum_{m=1}^{M} (x'_m - o'_m)^2} \quad (2) \]

The tropics generally show higher predictability with respect to mid and high latitudes, although North Atlantic and part of the southern Pacific have small RMSE in their respective cold semesters. Differences to observations are larger over the continents, since RMSE increases where the variance is higher. This is noticeable in the tropical Pacific, where the error in the area mainly interested by ENSO is always more pronounced than immediately north and south. The evolution of sea surface temperature RMSE in the El Niño 3.4 region (120°W-170°W and 5°S-5°N) is shown in Figure (8) for the four start dates. The thick blue line represents the ensemble mean of the nine members, shown as thinner blue lines. Clearly, the error quickly propagates as the system moves away from the initial condition, showing the decay of predictability with time. As expected, the ensemble mean forecasts have better skill than any ensemble member, since the average reduces the internal dynamic noise present in the individual forecasts (Kirtman and Shukla, 2002 [11]). The ACCs (Figure 6) and the RMSEs (Figure 8) have been computed relative to the ERA Interim SST and persistence forecasts are made by continuing the ERA Interim monthly anomaly from the month prior to the start date of the model forecasts. For example, SST persistence forecasts for the period May to September 1991 have been made by continuing the SSTA found for the observed April 1991.

Another analysis to test the skill of the CMCC-SPS was performed on the indexes Niño3.4 and WAMI. As mentioned above, Niño3.4 is the average sea surface temperature anomaly in the region bounded by 5°N to 5°S, from 170°W to 120°W. WAMI (West African Monsoon Index) is computed as the standardized difference between standardized wind modulus at 925 hPa and standardized zonal wind at 200 hPa in the Sahel (3°N-13°N, 20°W-20°E). Figure (9) shows the spaghetti plot of Niño34 index for all the start-dates. On figure 9(a-b) the onset of El Nino 97-98 was predicted quite well. The CMCC-SPS shows systematic high uncertainty for neutral episodes, while is confident for the strong-one. The correlation of each single event with observation is high confirming the results in figure 6. Figure (10) represents the WAMI for the start date of May, covering the period June-July-August-September. While the monsoon simulated by the CMCC-SPS is generally weaker comparing to observations, the correlation is rather high (about 66%) and the interannual variability is well predicted. The large distance between the ensemble members (shown as blue crosses in figure 10) points out that predictability of winds in Africa is lower than predictability of SSTs in the ENSO region, and that skill in Sahel is not as high as it is in the tropical Pacific.

CONCLUSIONS AND FUTURE WORKS

The CMCC seasonal prediction system shows good skill over Tropical ocean, in particular in the ENSO region (Niño3.4) and the northeastern Pacific in general. Other areas characterized by rather strong predictability are the northeastern Pacific and, in the winter, the north Atlantic. Over continents, the skill is generally low except in regions with a strong coupling land surface-oceans, and at high latitudes during winter, most likely due to the snow cover initialization from ERA interim reanalysis. The next step will be the full system initialization. In particular, with respect to the version in current use, land surface and vegetation will be initialized with an observational data set: in particular, initial condition of soil moisture and
temperature and leaf area index will be imposed to the CMCC-SPS. The "soil moisture memory" (Koster et al, 2006 [12]) can be as long as 6 months/1 year, affecting climate at a seasonal time scale. We expect major enhancement of the system’s skill, since many authors have pointed out that small perturbation in the land surface condition can propagate vertically and impact regional atmospheric circulation (e.g. Steiner et al., 2009 [19]).
Seasonal Prediction System at CMCC

Figure 2:
Experimental set-up of Seasonal Prediction System

Figure 3:
The perturbed initial conditions of Seasonal Prediction System for each start-date
Figure 4: Systematic error (BIAS) of seasonal mean predicted SST [K] (panels a, c, e, g) and precipitation [mm/day] (panels b, d, f, h) for lead time month 1 (target months from 2 to 4). SST forecasts with starting dates in (a) 1st February, (c) 1st May, (e) 1st August and (g) 1st November. Precipitation forecasts with starting dates in (b) 1st February, (d) 1st May, (f) 1st August, (h) 1st November. The error is defined as the difference between the 1989-2009 climatologies obtained from the forecast ensemble means and from the ERA-Interim. For SST blue (red) shading indicates values below -0.5 (above 0.5) K. For precipitation blue (red) shading indicates value above 0.5 (below -0.5) mm/day.
Figure 5:
Ensemble mean forecasts vs ERA-Interim surface air temperature anomalies: point-by-point correlation of lead time month 1 (target months 2-to-4). Forecasts with starting dates (a) 1st February, (b) 1st May, (c) 1st August and (d) 1st November. The grid points in which correlations are positive (yellow, orange and red shading) have a good skill. While if correlations are negative (cyan and blue shading) have a bad skill.

Figure 6:
ACCs between predicted and ERA-Interim SST anomalies over the Niño3.4 region (5°S-5°N, 190°E-240°E). The 1989-2009 averaged ACCs for the forecasts with starting dates (a) 1st February, (b) 1st May, (c) 1st August, (d) 1st November are plotted as a function of the leadtime month. Solid thick blue line is for ensemble means, while thin lines show the results for each ensemble member forecast. The dashed lines stand for the persistence forecasts, obtained by continuing the monthly anomaly observed during the month prior to the start date of the model forecasts.
Figure 7: Ensemble mean forecasts vs ERA-Interim surface air temperature anomalies: point-by-point root mean square error [K] of lead time month 1 (target months 2-to-4). Forecasts with starting dates (a) 1st February, (b) 1st May, (c) 1st August and (d) 1st November. The grid points in which root mean square errors are near 0-0.5 K (green and light yellow shading) have a good skill. While if root mean square errors are greater than 1 K (yellow to red shading) have a bad skill.

Figure 8: RMSEs between predicted and ERA-Interim SST anomalies over the Niño3.4 region (5°S-5°N, 190°E-240°E). The 1989-2009 averaged RMSEs for the forecasts with starting dates (a) 1st February, (b) 1st May, (c) 1st August, (d) 1st November are plotted as a function of the leadtime month. Solid thick blue line is for ensemble means, while thin lines show the results for each ensemble member forecast. The dashed black line stand for the persistence forecasts, obtained by continuing the monthly anomaly observed during the month prior to the start date of the model forecasts, while dashed blue line indicated the ensemble spread. Both persistence and spread are the upper and lower limits of the forecast error.
Figure 9: Spaghetti plot of Niño3.4 index for 1st February start-date (a), 1st May start-date (b), 1st August start-date (c) and for 1st November start-date (d). Black line stand for HadISST data, while red thin lines represent forecasts. Each forecast is for 6 months integration.

Figure 10: West African Monsson index (WAMI): the red line is the ERA-Interim index, while blue line is the ensemble mean forecast. Blue crosses represent forecast ensemble members. WAMI was evaluated for JJAS season on May start date.
Bibliography


