The NCEP Climate Forecast System Version 2

(http://cfs.ncep.noaa.gov)

Suranjana Saha[1], Shrinivas Moorthi[1], Xingren Wu[2], Jiande Wang[4], Sudhir Nadiga[2], Patrick Tripp[2], David Behringer[1], Yu-Tai Hou[1], Hui-ya Chuang[1], Mark Iredell[1], Michael Ek[1], Jesse Meng[2], Rongqian Yang[2], Huug van den Dool[3], Qin Zhang[3], Wanqiu Wang[3], Mingyue Chen[3].

Submitted to the Journal of Climate

November 22, 2012

[1] Environmental Modeling Center, NCEP/NWS/NOAA
[2] I. M. Systems Group, Inc
[3] Climate Prediction Center, NCEP/NWS/NOAA
[4] Science Systems and Applications, Inc.

Corresponding Author: Dr. Suranjana Saha,
NOAA Center for Weather and Climate Prediction (NCWCP)
5830 University Research Court, College Park, MD 20740, USA
Suranjana.Saha@noaa.gov

Abstract

The second version of the NCEP Climate Forecast System (CFSv2) was made operational at NCEP in March 2011. This version has upgrades to nearly all aspects of the data assimilation and forecast model components of the system. A coupled Reanalysis was made over a 32 year period (1979-2011), which provided the initial conditions to carry out a comprehensive Reforecast over 29 years (1982-2011). This was done to obtain consistent and stable calibrations, as well as, skill estimates for the operational sub seasonal and seasonal predictions at NCEP with CFSv2. The operational implementation of the full system ensures a continuity of the climate record and provides a valuable up-to-date dataset to study many aspects of predictability on the seasonal and sub seasonal scales. Evaluation of the reforecasts show that the CFSv2 increases the length of skillful MJO forecasts from 6 to 17 days (dramatically improving sub-seasonal forecasts), nearly doubles the skill of seasonal forecasts of 2 meter temperatures over the U.S. and significantly improves global SST forecasts over its predecessor. The CFSv2 not only provides greatly improved guidance at these time scales, it also creates many more products for sub-seasonal and seasonal forecasting with an extensive set of retrospective forecasts for users to calibrate their forecast products. These retrospective and real time operational forecasts will be used by a wide community of users in their decision making processes in areas such as water management for rivers and agriculture, transportation, energy use by utilities, wind and other sustainable energy, and seasonal prediction of the hurricane season.

1. Introduction

In this paper, we describe the development and performance of NCEP's Climate Forecast System version 2 (CFSv2). The first CFS, retroactively called CFSv1, was implemented into operations at NCEP in August 2004 and was the first quasi-global, fully coupled atmosphere-ocean-land model used at NCEP for seasonal prediction (Saha *et al*.,2006). Earlier coupled models at NCEP had full ocean coupling restricted to only the tropical Pacific Ocean. CFSv1 was developed from four independently designed pieces of technology, namely the R2 NCEP/DOE Global Reanalysis (Kanamitsu *et al*., 2002) which provided the atmospheric and land surface initial conditions, a global ocean data assimilation system (GODAS) operational at NCEP in 2003 (Behringer, 2007) which provided the ocean initial states, NCEP's Global Forecast System (GFS) operational in 2003 which was the atmospheric model run at a lower resolution of T62L64, and the MOM3 ocean forecast model from GFDL. The CFSv1 system worked well enough that it became difficult to terminate it, as it was used by many in the community, even after the CFSv2 was implemented into operations in March 2011. It was finally decommissioned in late September 2012.

Obviously CFSv2 has improvements in all four components mentioned above, namely the two forecast models and the two data assimilation systems. CFSv2 also has a few novelties: an upgraded four level soil model, an interactive three layer sea-ice model, and historical prescribed (i.e. rising) CO_2 concentrations. But above all, CFSv2 was designed to improve consistency between the model states and the initial states produced by the data assimilation system. It took nearly seven years to complete the following aspects:

(1) Carry out extensive testing of a new atmosphere-ocean-sea-ice-land model configuration including decisions on resolution, etc;

(2) Make a coupled atmosphere-ocean-seaice-land Reanalysis from 1979-2011 with the new system (resulting in the Climate Forecast System Reanalysis, CFSR) for the purpose of creating initial conditions for CFSv2 retrospective forecasts;

(3) Make retrospective forecasts with the new system using initial states from CFSR from 1982-2011 and onward to calibrate operational subsequent real time subseasonal and seasonal predictions;

(4) Operational implementation of CFSv2.

Items (1) and (2) have already been described in Saha *et al*., 2010, and aspect (4) does not need to be treated in any great detail in a scientific paper, other than to mention that CFSv2 is run in near real time with a very short data cut-off time, thereby increasing its applicability to the shorter time scales relative to CFSv1, which was late by about 36 hours after real time. So, in this paper, we mainly describe the CFSv2 model, the design of the retrospective forecasts, and some results from these forecasts.

The performance of the CFSv2 retrospective forecasts can be split into four time scales.

- The shortest time scale of interest is the subseasonal, mainly geared towards the prediction of the Madden Julian Oscillation (MJO) and more generally forecasts for the week 2 to week 6 period over the United States (or any other part of the globe).
- The next time scale is the ‘long-lead’ seasonal prediction, out to 9 months, for which these systems are ostensibly designed. For both the subseasonal and seasonal, we have a very precise comparison between skill of prediction by the CFSv1 and CFSv2 systems evaluated over exactly the same hindcast years.

- The final two time scales are decadal and centennial. Here the emphasis is less on forecast skill, and more on the general behavior of the model in extended integrations for climate studies.

Structurally, this paper makes a number of simple comparisons between aspects of CFSv1 and CFSv2 performance, and discusses changes relative to CFSv1. For the background details of most of these changes, we refer to the CFSR paper (Saha *et al*., 2010) where all model development over the period 2003-2009 has been laid out. In addition, some new changes were made relative to the models used in CFSR. These changes to the atmospheric and land model in the CFSR were deemed necessary when they were used for making the CFSv2 hindcasts. For instance, changes had to be made to combat a growing warm bias in the surface air temperature over land, or a decrease in the tropical Pacific sea surface temperature in long integrations. A model that displays good performance in a 9 hour guess field, may not be satisfactory when making a 9 month or 9 year integration.

The lay out of the paper is as follows. Section 2 deals with changes in model components relative to CFSR. In Section 3 the design of the hindcasts are described. Model performance in terms of forecast skill for intraseasonal to long lead seasonal prediction is given in section 4. Section 5 describes other aspects of performance, including diagnostics of the land surface and sea-ice. Model behavior in very long integrations, both decadal and centennial, is described in Section 6. Conclusions and some discussion are presented in Section 7. We also include four appendices that include the restrospective forecast calendar, reforecast and operational configuration of the CFSv2, and most importantly a summary of the availability of the CFSv2 data (absolutely free of charge).

2. Overview of the Coupled Climate Forecast System Model

The coupled forecast model used for the seasonal retrospective and operational forecasts is different from the model used for obtaining the first guess forecast for CFSR and operational CDAS analyses (CDAS is the real time continuation of CFSR). The ocean and sea-ice models are identical to those used in CFSR (Saha *et al*., 2010). The atmospheric and the land surface components, however, are somewhat different and these differences are briefly described below.

The atmospheric model has a spectral triangular truncation of 126 waves (T126) in the horizontal (equivalent to nearly a 100 Km grid resolution) and a finite differencing in the vertical with 64 sigma-pressure hybrid layers. The vertical coordinate is the same as that in the operational CDAS. Differences between the model used here and in CFSR are mainly in the physical parameterizations of the atmospheric model and some tuning parameters in the land surface model and are as follows:

- We use virtual temperature as the prognostic variable, in place of enthalpy that was used in major portions of CFSR. This decision was made with an eye on unifying the GFS (which uses virtual temperature) and CFS, as well as the fact that the operational CDAS with CFSv2 currently uses virtual temperature.
- We also disabled two simple modifications made in CFSR to improve the prediction of marine stratus (Moorthi *et al.,* 2010, Saha *et al*., 2010, Sun *et al*., 2010). This was done because including these changes resulted in excessive low marine clouds, which led to increased cold sea surface temperatures over the equatorial oceans in long integrations of the coupled model.
- We added a new parameterization of gravity wave drag induced by cumulus convection based on the approach of Chun and Baik (1998) (Johansson, 2009, personal

communication). The occurrence of deep cumulus convection is associated with the generation of vertically propagating gravity waves. While the generated gravity waves usually have eastward or westward propagating components, in our implementation only the component with zero horizontal phase speed is considered. This scheme approximates the impact of stationary gravity waves generated by deep convection. The base stress generated by convection is parameterized as a function of total column convective heating and applied at the cloud top. Above the cloud top the vertically propagating gravity waves are dissipated following the same dissipation algorithm used in the orographic gravity wave formulation.

- As in CFSR, we use the Rapid Radiative Transfer Model (RRTM) adapted from AER Inc. (e.g. Mlawer *et al*., 1997; Iacono *et al*., 2000; Clough *et al*., 2005). The radiation package used in the retrospective forecasts is similar to the one used in the CFSR but with important differences in the cloud-radiation calculation. In CFSR, a standard cloud treatment is employed in both the RRTM longwave and shortwave parameterizations, that layers of homogeneous clouds are assumed in fractionally covered model grids. In the new CFS model, an advanced cloud-radiation interaction scheme is applied to the RRTM to address the unresolved variability of layered cloud. One accurate method would be to divide the clouds in a model grid into independent sub-columns. The domain averaged result from those individually computed sub-column radiative profiles can then represent the domain approximation. Due to the exorbitant computational cost of a fully independent column approximation (ICA) method, an alternate approach, which is a Monte-Carlo independent column approximation (McICA) (Barker *et al*., 2002, Pincus *et al.,* 2003), is used in the new CFS model. In McICA, a random column cloud generator samples the model layered cloud into sub-columns and pairs each column with a pseudo-monochromatic calculation in the radiative transfer model. Thus the radiative computational expense does not increase, except for a small amount of overhead cost attributed to the random number generator.

- In calculating cloud optical thickness, all the cloud condensate in a grid box is assumed to be in the cloudy region. So the in-cloud condensate mixing ratio is computed by the ratio of grid mean condensate mixing ratio and cloud fraction when the latter is greater than zero.
- The CO_2 mixing ratio used in these retrospective forecasts includes a climatological seasonal cycle superimposed on the observed estimate at the initial time.
- The Noah land surface model (Ek *et al*., 2003) used in CFSv2 was first implemented in the GFS for operational medium-range weather forecast (Mitchell *et al*., 2005) and then in the CFSR (Saha *et al*., 2010). Within CFSv2, Noah is employed in both the coupled land-atmosphere-ocean model to provide land-surface prediction of surface fluxes (surface boundary conditions), and in the Global Land Data Assimilation System (GLDAS) to provide the land surface analysis and evolving land states. While assessing the predicted low-level temperature, and land surface energy and water budgets in the CFSRR reforecast experiments, two changes to CFSv2/Noah were made. First, to address a low-level warm bias (notable in mid-latitudes), the CFSv2/Noah vegetation parameters and rooting depths were refined to increase evapotranspiration, which, along with a change to the radiation scheme (RRTM in GFS and CFSR, and now McICA in CFSv2), helped to improve the predicted 2-meter air temperature over land. Second, to accommodate a change in soil moisture climatology from GFS to CFSv2, Noah land

surface runoff parameters were nominally adjusted to favorably increase the predicted runoff (see section 5 for more comments).

3. The Design of the Retrospective and Real Time Forecasts: Considerations for operational implementation

3a. 9-month retrospective predictions:

- The earliest release of CPC operational seasonal prediction is on Thursday the 15th of a month. In this case, products must be ready by Friday the 9th of the month. For these products to be ready, the latest CFSv2 run that can be admitted is from the 7th of each month. These considerations are adhered to in the hindcasts.
- The retrospective 9-month forecasts have initial conditions of the 0, 6, 12 and 18Z cycles for every 5th day, starting from 1 Jan 0Z of every year, over the 29-year period 1982-2010. There are 292 forecasts for every year for a total of 8468 forecasts (see Appendix A). Selected data from these forecasts may be downloaded from the NCDC web servers (see Appendix D)
- The retrospective forecast calendar (Appendix B) outlines the forecasts that are used each calendar month, to estimate proper calibration and skill estimates, in such a way to mimic CPC operations.
- This results in an ensemble size of 24 forecasts for each month, except November which has 28 forecasts.
- Smoothed calibration climatologies have been prepared from the forecast monthly means and time series of selected variables and is available for download (see Appendix D)
- Having a robust interpolated calibration for each cycle, each day and each calendar month, allows CPC to use real time ensemble members (described in section 3c) as close as possible to release time.

3b. First season and 45-day Retrospective forecasts.

- These retrospective forecasts have initial conditions from every cycle (0, 6, 12 and 18Z) of every day over the 12-year period from Jan 1999-Dec 2010. Thus, there are approximately 365*4 forecasts per year, for a total of 17520 forecasts. The forecast from the 0Z cycle was run out to a full season, while the forecasts from the other 3 cycles (6, 12 and 18Z) were run out to exactly 45 days (see Appendix A for the reforecast configuration). Selected data from these forecasts may be downloaded from the NCDC (see Appendix D)
- Smoothed calibration climatologies have been prepared from the forecast time series of selected variables (http://cfs.ncep.noaa.gov/cfsv2.info/CFSv2.Calibration.Data.doc) and is available for download (see Appendix D). It is essential that some smoothing is done when preparing the climatologies of the daily timeseries, which are quite noisy.
- Having a robust calibration for each cycle, each day and each calendar month, allows CPC to use ensemble members very close to the release time of their 6-10day and week 2 forecasts. They are also exploring the possibility of using the CFSv2 predictions in the week3-week6 range.

3c. Operational configuration:

The initial conditions for the CFSv2 retrospective forecasts are obtained from the CFSR, while the real time operational forecasts obtain their initial conditions from the real time operational CDASv2. Great care was made to unify the CFSR and CDASv2 in terms of cutoff times for data input to the atmosphere, ocean and land surface components in the data assimilation system. Therefore, there is greater utility of the new system, as compared to CFSv1 (which had a lag of a few days), since the CFSv2 initial conditions are made completely in real time. This makes it possible to use them for the subseasonal (week1-week6) forecasts. There are

16 CFSv2 runs per day in operations; four out to 9 months, three out to 1 season and nine out to 45 days (see Appendix C). Operational real time data may be downloaded from the official site (see Appendix D).

4. Results in terms of skill

4a. Sub seasonal prediction

Figure 1 shows the skill, as per the bivariate anomaly correlation BAC (Lin *et al*., 2008, equation 1), of CFSv2 forecasts in predicting the MJO, as expressed by the Wheeler and Hendon (2004) WH index, using two EOFs of combined zonal wind and outgoing longwave radiation (OLR) at the top of the atmosphere. The period is 1999-2009. On the left is CFSv2, on the right is CFSv1. Both are subjected to systematic error correction (SEC). The BAC stays above the 0.5 level (the black line) for two to three weeks in the new system, while it was at only one week in the old system. Both models show a similar seasonal cycle in forecast skill with maxima in May-June and Nov-Dec respectively, and minima in between. Correlations were calculated as a function of lead for each starting day, i.e. for any given lead, there were only 11 cases, one case for each year. Figure 1 (both panels) was then plotted with day of the year along the vertical axis (months are labeled for reference) and forecast lead along the horizontal axis, with the correlation*100 being contoured. To suppress noise, a light smoothing was applied in the vertical (i.e. over adjacent starting days). The right panel in Figure 1 for CFSv1 would have holes, because no CFSv1 forecasts originated from 4th through the 8th, 14th through 18th and 24th through 28th of each month. In the CFSv1 graph, the smoothing also serves to mask these holes. Note that, consistent with CPC operations, we verify both CFSv1 and CFSv2 against R2 based observations of RMM1 and RMM2, using an observed climatology (1981-2004) based on R2 winds and satellite OLR.

It is quite clear that CFSv2 has much higher skill than CFSv1 throughout the year which reaches out to 30 days. In fact, this is the improvement made by half a generation (~15 years) of work by many in both data assimilation and modeling fields (taking into account that CFSv1 has rather old R2 atmospheric initial conditions as its weakest component). One rarely sees such a demonstration of improvement. This is because operational atmospheric NWP models are normally abandoned when a new model comes in. But in the application to seasonal climate forecasting, systems tend to have a longer lifetime. This gave us a rare opportunity to compare two frozen models that are about 15 years apart in vintage.

The causes for the enormous improvement seen in Figure 1 are probably very many, but especially the improved initial states in the tropical atmosphere and the consistency of the initial state and the model used to make the forecasts play a role. Further research should bring out the importance of coupling to the ocean and its quantitative contribution to skill. Further results and discussion on MJO in CFSv1/v2 can be found in Zhang and Van den Dool(2012), hereafter ZV.

We studied the results with and without the benefit of systematic error correction (SEC) for both CFSv1 and CFSv2. We found that SEC results in improvements for either CFS over raw forecasts, more often than not, and overall the improvement in CFSv2 is between 5 and 10 points (see Fig.2 in ZV), which could be the equivalent of several new model implementations. This is a strong justification for making hindcasts.

As is the case with CFSv2, version 1 did benefit noticeably from the availability of its hindcasts. While the distribution of the improvement with lead and season is different for CFSv1, the overall annual mean improvement is quite comparable, see Fig.3 in ZV. Both CFSv1 and CFSv2 appear to gain about 2-3 days of prediction skill by applying an SEC. Obviously, the model and data assimilation improvements between 1995 and 2010 count for much more than the availability of the hindcasts, but the latter do correspond to a few years of model improvement.

4b. Seasonal prediction out to 9 months

The anomaly correlation of three-month mean sea surface temperature (SST) forecasts is shown in Figure 3 for 3-month and 6-month lead times. The forecasts are verified against a weekly OIv2 SST (Reynolds *et al,* 2002). A lagged ensemble mean of 20 members from each starting month is used to compute the correlation. Similar spatial distributions of the correlation are seen in both CFS versions, with relatively higher skill in the tropical Pacific than the rest of the globe. Overall, the skill for CFSv2 is improved in the extratropics with an average anomaly correlation poleward of 20S and 20N of 0.34(0.27) for 3-month lead (6-month lead) compared to the corresponding CFSv1 anomaly correlation of 0.31 (0.24). In the tropical Pacific, the CFSv2 skill is slightly lower than that of CFSv1. This lower CFSv2 skill is related to the climatology shift with significantly warmer mean predicted SST in the tropical Pacific after 1999, compared to that before 1999, which is likely due to the start of assimilating the ATOVS satellite observations in the CFSR initial conditions in 1999.

Figure 3 compares the amplitude of interannual variability between the SST observation and forecasts at 3-month and 6-month lead times. The largest variability over the globe is related to the ENSO variability in the tropical Pacific. The variability of the forecast is computed as the standard deviation based on anomalies of individual members (rather than the ensemble mean). Both CFSv1 and CFSv2 are found to generate stronger variability than observed over most of the globe. In particular, the forecast amplitude is larger than the observed in the tropical Indian Ocean, eastern Pacific and northern Atlantic. Compared to CFSv1, CFSv2 produced more reasonable amplitude. For examples, the strong variability in CFSv1 in the tropical Pacific is substantially reduced, and the variability in CFSv2 in the northern Pacific is comparable to the observation (Figures 3b and 3c), while the CFSv1 variability in this region is too strong (Figures 3d and 3e).

Figure 4 provides a grand summary of the skill of monthly prediction as a function of target month (horizontal axis) and lead (vertical axis). For precipitation and 2 meter temperature the area is all of NH extra-tropical land, and the measure is the anomaly correlation evaluated over all years (1982-2010). We compare CFSv2 directly to CFSv1, over the same years. One may also compare this to Figures 1 and 7 in Saha *et al*., 2006 for CFSv1 alone (and 6 fewer years). The top panels of Figure 4 show that prediction of temperature (top row) has substantially improved from CFSv1 to CFSv2. We believe this is caused primarily by increasing CO_2 in the initial conditions and hindcasts[1], and possibly eliminating some soil moisture errors (and too cold temperatures) that have plagued CFSv1 in real time in recent years. The positive impact of increasing CO_2 was to be expected (Cai *et al*., 2009) especially at long leads. Still, skill is only modest, a 0.20 correlation.

While skill for 2m temperature is modest, skill for precipitation forecasts (middle panels of Figure 4) for monthly mean conditions over NH land remains less than modest. Except for the first month (lead 0), which is essentially weather prediction in the first 2 weeks, there is no skill at all (over 0.1 correlation) which is a sobering conclusion. CFSv2 is not better than CFSv1. Although, these systems have skill in precipitation prediction over the ocean (in conjunction with ENSO), the benefit of ENSO skill in precipitation over land appears small or washed away by other factors.

The bottom panels of Figure 4 shows that both systems have decent skill in predicting the SST at grid points inside the Nino3.4 box (170W-120W, 5S-5N). Skill for the Nino34 area,

[1] CO_2 is not increased during a particular hindcast, but through the initial conditions, hindcasts for say 2010 are run at much higher CO_2 (which is maintained throughout the forecast) than for hindcasts in 1982. In CFSv1, a single CO_2 value valid in 1988 was used for all years.

overall, has not improved for CFSv2 versus CFSv1, but the seasonality has changed. Skill has become lower at long lead for winter target months and higher for summer target months, thereby decreasing the spring barrier. In general, CFSv2 is better in the tropics than CFSv1 for SST prediction (see Figure 2), but Nino3.4 is the only area where this is not so.

5. Diagnostics

While section 4 contains results of CFSv2 (vs. CFSv1) in terms of forecast skill, we also need to report on some diagnostics that describe model behavior. Even without strict verification, one may judge models as being 'reasonable' or not. In particular the surface water budget, which was mentioned in section 2 as being the subject of tuning, is discussed in section 5a. We also present some results on sea-ice prediction (without a strict verification) since this is an important emerging aspect of global coupled models. CFSv1 had an interactive ocean only up to 65^0 North and 75^0 South latitudes, with climatological sea-ice in the polar areas. The aspect of a global ocean and interactive sea-ice model in the CFSv2 is new in the seasonal modeling context at NCEP.

5a. Land Surface

Table 1 shows a comparison of surface water budget terms averaged over the Northern Hemisphere land between CFSv1 and CFSv2 and with CFSR. The quantities in CFSv1 and CFSv2 are computed from seasonal ensemble means covering a 29-yr period (1982-2010), where the CFSv1 is based on seasonal predictions from 15 ensemble members whose initial conditions are from Mid-April to early May (April 9-13, 19-23, and April 29-May 3 at 00Z) for the summer season (JJA), and from Mid-October to early November (October 9-13, 19-23, and October 29 – November 3) for the winter season (DJF), while the CFSv2 is based on 24 ensemble members (initial conditions from 4 cycles of the 6 days between April 11 and May 6 with 5 days apart) for summer and 28 ensemble members (initial conditions from 4 cycles of 7 days between October 8 and November 7 with 5 days apart) for winter season, respectively.

Compared to the CFSR, precipitation (snow in winter) in the CFSv1 is higher in both seasons, which yields higher values for both evaporation and runoff. The higher evaporation in the summer season in the CFSv1 yields a much larger seasonal variation in soil moisture (though lower absolute values) than in both CFSR and CFSv2. In contrast, precipitation in the CFSv2 is considerably lower than in both CFSv1 and CFSR, consistent with lower evaporation in the CFSv2. While less than the CFSv1, runoff in the CFSv2 is more than in CFSR, indicating that soil moisture is a more important source for surface evaporation in the CFSv2; this higher runoff in winter season leads to a damped seasonal variation in soil moisture since soil moisture is re-charged in winter when evaporation is at its minimum. The increases in both surface evaporation from root-zone soil water and runoff production are consistent with the changes made to vegetation parameters and rooting depths in CFSv2 (see comments in section 2) to address high biases in predicted T2m, and the accommodated changes in soil moisture climatology and surface runoff parameters. The good agreement in soil moisture between CFSR and CFSv2 is expected because they use the same Noah land model.

5b. Sea Ice

Sea ice prediction is challenging and relatively new in the context of seasonal climate prediction models. Sea ice can form or melt and can move with wind and/or ocean current. Sea ice interacts with both the air above and the ocean beneath and it is influenced by, and has impact on, the air and ocean conditions. The CFSv2 sea ice component includes a dynamic/thermodynamic sea ice model and a simple "assimilation" scheme, which are described in details in Saha *et al.* (2010). One of the most important developments in CFSv2, compared to CFSv1, is the extension of the

CFS ocean domain to the global high latitudes and the incorporation of a sea ice component. The ice initial condition (IC) for the CFSv2 hindcasts is from CFSR as described in Saha *et al.* (2010). For sea ice thickness, there is no data available for assimilation, and we suspect there is a significant bias of sea ice thickness in the CFSv2 model, which causes the sea ice to be too thick in the IC. For the sea ice prediction, sea ice appears too thick and certainly too extensive in the spring and summer. Figure 5 shows the mean September sea ice concentration from 1982 to 2010, and the bias in the predicted mean condition at lead times of 1-month (August 15 IC), 3-month (June 15 IC), and 6-month (March 15 IC). The model shows a consistent high bias in its forecasts of September ice extent. The corresponding predicted model variability at the 3 different lead times is shown in Figure 6. The variability from the model prediction is underestimated near the mean September ice pack and overestimated outside the observed mean September ice pack. Although the CFSv2 captured the observed seasonal cycle, long-term trend and interannual variability to some extent, large errors exist in its representation of the observed mean state and anomalies, as shown in Figures 5 and 6. Therefore in the CFSv2, when the sea ice predictions are used for practical applications, bias correction is necessary. The bias can be obtained from the hindcast data for the period 1982-2010, which are available from NCDC. In spite of the above reported shortcomings, when the model was used for the prediction of the September minimum sea ice extent organized by SEARCH (Study of Environmental Arctic Change) during 2009 and 2011, CFSv2 (with bias correction applied) was among the best prediction models. In the future we plan to assimilate the sea ice thickness data into the CFS assuming that would reduce the bias and improve the sea ice prediction.

6. Model behavior in very long integrations.

6a. Decadal prediction

The protocol for the next IPCC (Inter Governmental Panel for Climate Change) model runs, called AR5, recommended the making of decadal predictions to assist in the study of climate change, see: http://www.ipcc.ch/activities/activities.shtml#.UGyOHpH4Jw0 These decadal runs may bring in elements of the initial states in terms of land, ocean, sea ice and atmosphere and thus perhaps add information in the first 10 years, in addition to the general warming that most models may predict when greenhouse gases (GHG) increase. Following this recommendation, sixty 10-year runs were made from initial conditions on Nov 1, 0Z, 6Z, 12Z and 18Z cycles (i.e. 4 'members'), for the following years: 1980, 1981, 1983, 1985, 1990, 1993, 1995, 1996, 1998, 2000, 2003, 2005, 2006, 2009 and 2010 (every 5^{th} year from 1980 to 2010, as well as some interesting intermediate years). Each run was 122 months long (the first 2 months were not used to avoid spin-up). The forcing for these decadal runs included both shortwave and longwave tropospheric aerosol effects and is from a monthly climatology that repeats its values year after year (described in Hou *et al,* 2002). Also, included in the runs are historical stratospheric volcanic aerosol effects on both shortwave and longwave radiation, which end in 1999, after which a minimum value of optical depth=1e-4 was used (Sato *et al*, 1993). The runs also used the latest observed CO_2 data when available (WMO Global Atmospheric Watch (http://gaw.kishou.go.jp) and an extrapolation was done into the future with a fixed growth rate of 2ppmv.

Results using only monthly mean data from the 60 decadal runs are presented in this paper. Variable X in an individual run can be denoted as $X_{j,m}$, where j and m is the target year and month. How 'anomalies' are obtained is not obvious in these type of decadal runs. We proceeded as follows: first a 60 run mean was formed, i.e. $<X_{j,m}>$, where j=1, 10 and m=1, 120. Averaging across all years, we get $<<X_m>>$. The anomaly is then computed as $X_{j,m} - <<X_m>>$.

Figure 7a (top panel) shows the global mean SST anomalies (here X is SST). There are 60 yellow traces, each of 10 year length. The observations (Reynolds *et al*, 2007) are shown as the full black line, and the monthly anomaly is formed as the departure from 1982-2010 climatology. One can conclude that the observations are in the cloud of model traces produced by CFSv2, especially after 1995 and before 1987 when the observations are near the middle of the cloud. The model appears somewhat cold in the late eighties and early nineties. Figure 7b (bottom panel) shows the same thing, but for global mean land temperature. The black line, from GHCN-CAMS (Fan and Van den Dool, 2008), is comfortably inside the cloud of model traces, except around 1993 when perhaps the model overdid the aerosol impact of the Pinatubo volcanic eruption. The spread produced by the model is much higher in Figure 7b than in Figure 7a, not only because the land area is smaller than the oceanic area, but also because the air temperature is much more variable to start with. This model, never before exposed before to such long integrations, passed the zeroth order test, in that it produced some warming over the period from 1980 to the present and has enough spread to cover what was observed (essentially a single model trace). In this paper there is no attempt to address any model prediction skill over and beyond a capability to show general warming and uncertainty.

Some monthly mean and 3-hourly time series data from the NCEP decadal runs is available for download (see Appendix D)

6b. Long 'free' runs

On the very long time-scales, a few single runs were made lasting from 43 to 100 years, which were designated as 'CMIP' runs. There is nothing that reminds these runs of the calendar years they are in, except for GHG levels which are prescribed when available (see section 4c), and in case of CO_2 is projected to increase by 2ppm in future years. Here, we are interested in behavioral aspects, including a test as to whether the system is even stable or drifting due to assorted technical issues. The initial conditions were chosen for Jan of three years, namely 1987, 1995, and 2001 (similar runs were made with the first version of the CFS). Allowing for a spin up of 1 year, data was saved for 1988-2030 (43 years), 1996-2047 (52 years) and 2002-2101 (100 years) from these three runs, one of which is truly centennial. None of these runs became unstable or produced completely unreasonable results. A common undesirable feature was a slow cooling of the upper ocean for the first 15-20 years. Only after this temperature decline stabilized, a global warming of the sea surface temperature was seen starting 25-35 years after initial time. In contrast, the water at the bottom of the ocean showed a small warming from the beginning to end, which is unlikely to be correct.

An important issue was to examine the onset and decay of warm and cold events (El Ninos and La Ninas) and ascertain how regular they were. The CFSv1 was found to be too regular and very close to being periodic in its CMIP runs (Penland and Saha, 2006) when diagnosed via a spectral analysis of Nino3.4 monthly values. Figure 8 shows the spectra of Nino3.4 for the observations from 1950-2011 (upper left) and the three CFSv2 CMIP runs. A harmonic analysis was conducted on monthly mean data with a monthly climatology removed. Raw power was estimated as ½ of the amplitude (of the harmonic) squared. The curves shown were smoothed by a 1-2-1 filter. The variance of all the CMIP runs is higher than observed by at least 25%, therefore the integral under the blue (model) and black (observed) curves differs. The model variance being too large was already noted in Figure 3 for leads of 3 and 6 months. The observations have a broad spectral maximum from 0.15 to 0.45 cycles per year (cpy). The shortest of the CMIP runs (upper right) resembles the broad spectral maximum quite well, the longer runs are somewhat more sharply peaked but are not nearly as periodic as in CMIP runs made by CFSv1, especially when T62 resolution was used (Penland and Saha 2006). On the whole, the behavioral aspects of ENSO (well beyond prediction) appear acceptable. One may also consider the possibility that certain segments of 43 years from the 100 year run may look like the upper right entry. Or by the same token, that the behavior of observations for 1951-2011 are not necessarily reproduced exactly when a longer period could be considered, or a period without mega-events like the 1982/83 and 1997/98 ENSO events. Some data from these CMIP runs are available for download from the CFS website (see Appendix D).

7. Concluding Remarks

This paper describes the transition from the CFSv1 to the CFSv2 operational systems. The Climate Forecast System (CFS), retroactively named version 1, was operationally implemented at NCEP in August 2004. The CFSv1 was described in Saha et al 2006. Its successor, named CFSv2, was implemented in March 2011 even though version 1 was only decommissioned in October 2012. The overlap (1.5 years) was needed, among other things, to give users time to make their transition between the two systems. In contrast to most implementations at NCEP, the CFS is accompanied by a set of retrospective forecasts that can be applied by the user community to calibrate subsequent real time operational forecasts made by the same system. Therefore, a new CFS takes time to develop and implement both on the part of NCEP and on the side of the user. One element that took a lot of time at NCEP to complete, was a new Reanalysis (the CFSR), that was needed to create the initial conditions for the coupled land-atmosphere-ocean-seaice CFSv2 retrospective forecasts. Every effort was made to create these initial conditions (for the period 1979-present) with a forecast system that was as consistent as possible with the model used to make the long range forecasts, whether it be for the retrospective forecasts or the operational forecasts going forward in real time.

For convenience, the evolution of the model components between CFSv1 and CFSv2 has been split into two portions, namely the very large model developments between CFSv1 and CFSR, and the far smaller model developments between CFSR and CFSv2. The development of model components between the time of CFSv1 (of 1996-2003 vintage) and CFSR (of 2008-2010 vintage) to generate the background guess in the data assimilation has already been documented in Saha et al (2010). Therefore, in the present paper, we only describe some further adjustments/tunings of the land surface parameters and clouds in the equatorial SST (in section 2).

The paper describes the design of both the long lead seasonal (out to 9 months) and shorter lead intraseasonal predictions (out to 45 days) for the retrospective forecasts and the real-time operational predictions going forward. This information is essential for any user who may want to use these forecasts. The retrospective forecasts are important for both calibration and skill estimates of subsequent real time prediction. The size of the hindcast data set is very large, since it spans forecasts from 1982-present for long lead seasonal range (4 runs out to 9 month, every 5th day), and forecasts from 1999-present for intraseasonal range (3 runs ***each day*** out to 45 days, plus one run each day out to 90 days), with all model forecast output data archived at 6 hour intervals for each run.

The paper also describes some of the results, in terms of the forecast skill, determined from the retrospective forecasts, for the prediction of the intraseasonal component (MJO in particular), and the seasonal prediction component (in section 4). This is done by comparing, very precisely, the CFSv2 predictions to exactly-matching CFSv1 predictions. There is no doubt that CFSv2 is superior to CFSv1 on the intraseasonal time scale; in fact the improvement is impressive from 1 week to more than 2 weeks (at the 0.5 level of anomaly correlation) for MJO prediction. For seasonal prediction, we note a substantial improvement in 2 meter temperature prediction over global land. This is mainly a result of successfully simulating temperature trends (which are large over the 1980-2010 period and thus an integral part of any verification) by increasing the amount of prescribed greenhouse gases in the model (a feature that was missing in CFSv1). For precipitation over land, the CFSv2, unfortunately, is hardly an improvement over CFSv1. This is perhaps due to the predictability ceiling being too low to expect big leaps forward in prediction. The SST prediction has been improved modestly over most of the global oceans and extended in CFSv2 to areas where CFSv1 had prescribed SST and/or sea-ice, as well as over the extra-tropical oceans. In the tropics, SST prediction has also improved, but least so in the much-focused-on Nino34 area, where the subsurface initial states of CFSR show unrealistic warming after 1998, due to the introduction of the ATOVS satellite data.

Being a community model to some extent, the CFSv2 has been (and will be) applied to decadal and centennial runs. These have not been typical NCEP endeavors in the past, so we have tested the behavior of this new model in integrations beyond the operational 9-month runs. Some results are described in section 6. The decadal runs appear reasonable in that, in the global mean, reality is within the cloud of the 65 decadal runs, both for 2 meter temperature over land and for SST in the ocean. The three centennial runs did not de-rail (a minimal test passed), and show both reasonable and unreasonable behavior. Unreasonable, we believe, since there is a small but steady cooling of the global ocean surface that lasts about 15 years before GHG forced warming sets in. Equally unreasonable may be a small warming of the bottom layers of global oceans from start to finish. The better news is that the ENSO spectrum in these free runs is far more acceptable in CFSv2, in contrast to CFSv1. When run in its standard resolution of T62L64, the CFSv1 produced too regular and almost periodic ENSO in its free runs, lasting up to a century.

A few diagnostics (presented in section 5) were made in support of the need for tuning some of the land surface parameters when going from CFSR to CFSv2. The main concern was the fact that the NH mean precipitation in summer over land reduced from 3.2 mm/day in CFSR to 2.7 mm/day in CFSv2 which posed a real problem for improved prediction of evaporation, runoff and surface air temperature. Some diagnostics are also presented for the emerging area of coupled sea-ice modeling, imbedded in a global ocean. Although this topic is important for monthly seasonal prediction, it has taken on new urgency due to concerns over shrinking sea-ice coverage (and thickness) in the Arctic. It is easy to identify some large errors in sea-ice coverage and variability and it is obvious that a lot more work needs to be done in this area of seaice modeling.

This paper is mainly to describe CFSv2 as a whole, from inception to implementation. There are many subsequent papers in preparation (or submitted/published) about detailed studies of CFSv2 prediction skill and/or diagnostics of some of the parts of CFSv2, whether it be the stratosphere, troposphere, deep oceans, land surface, etc.

While there are many users for the CFS output (sometimes one finds out how many only by trying to discontinue a model), the first line user is the Climate Prediction Center at NCEP. The CFSv2 plays a substantial role in the seasonal prediction efforts at CPC, both directly and through joint efforts such as National and International Multi-Model Ensembles.[2] CFSv2 is also used in the sub seasonal MJO prediction, and in a product called international hazards assessment. Because CFSv2 runs practically in real time (compared to CFSv1 which was about 36 hours later than real time), it plays a role in the operational 6-10day and week 2 forecasts and conceivably in the future prediction of the week 3 – week 6 forecasts for the US, which is on the

[2] We should point out that what we call the International Multi-Model Ensembles (IMME) has its counterpart called Eurosip in Europe. CFSv2 has been admitted as a member in the Eurosip ensemble which consists of the ECMWF, UK Met Office and Meteo France.

drawing board at CPC. The appropriate forcing fields extracted from CFSv2 predictions, such as daily radiation, precipitation, wind, relative humidity, etc. are used to carry the Global Land Data Assimilation Systems (GLDAS) forward, yielding an ensemble of drought related indices over the US and soon globally.

Acknowledgements

The authors would like to recognize all the scientists and technical staff of the Global Climate and Weather Modeling Branch of EMC for their hard work and dedication to the development of the GFS. We would also like to extend our thanks to the scientists at GFDL for their work in developing the MOM4 ocean model. George Vandenberghe, Carolyn Pasti and Julia Zhu are recognized for their critical support in the smooth running of the CFSv2 retrospective forecasts and the operational implementation of the CFSv2. We also thank Ben Kyger, Allan Darling, Dan Starosta, Christine Magee and Becky Cosgrove from the NCEP Central Operations (NCO) for the timely operational implementation of the CFSv2 in March 2011.

Appendix A: Reforecast Configuration of the CFSv2 (Figure A1)

- 9-month hindcasts were initiated from every 5th day and run from all 4 cycles of that day, beginning from Jan 1 of each year, over the full 29 year period from 1982-2010. This is required to calibrate the operational CPC longer-term seasonal predictions (ENSO, etc) (full lines in Figure A1).
- There was also a single 1 season (123-day) hindcast run, initiated from every 0 UTC cycle between these five days, but only over the 12 year period from 1999-2010. This is required to calibrate the operational CPC first season predictions for hydrological forecasts (precip, evaporation, runoff, streamflow, etc) (dashed lines in Figure A1)
- In addition, there were three 45-day hindcast runs from every 6, 12 and 18 UTC cycles, over the 12-year period from 1999-2010. This is required for the operational CPC week3-week6 predictions of tropical circulations (MJO, PNA, etc) (dotted lines in Figure A1)
- **Total number of years of integration = 9447 years!!!!!**

APPENDIX B: Retrospective Forecast Calendar (292 runs per year) organized by date of release of the official CPC seasonal prediction every month

As outlined in Appendix A, four 9-month retrospective forecasts are made every 5th day over the period 1982-2010. The calendar always starts on January 1 and proceeds forward in the same manner each year. Forecasts are always made from the same initial dates every year. This means that in leap years, Feb 25 and March 2 are separated by 6 days (instead of 5). Table A1 describes the grouping of the retrospective forecasts in relation to CPC's operational schedule (all forecast products must be available a week before the official release on the third Thursday of each month). For instance, for the release of the official forecast in the month of February, all retrospective forecasts made from initial conditions over the period from 11th January through Feb 5th for all previous years can be used for calibration and skill estimates, which constitute a lagged ensemble of 24 members. Obviously one can use more (going back farther), or less (since older forecasts may have much less skill).

All real time forecasts that are available closest to the date of release are used (see Appendix C).

Placeholder: Table A1 about here.

Appendix C: Operational Configuration of the CFSv2 for a 24-hour period (Figure A2)

- There are 4 control runs per day from the 0, 6, 12 and 18 UTC cycles of the CFSv2 real-time data assimilation system, out to 9 months (full lines in Fig A2)
- In addition to the control run of 9 months, there are 3 additional runs at 0 UTC out to one season. These 3 perturbed runs are initialized as in current operations (dashed lines in Figure A2)
- In addition to the control run of 9 months at the 6, 12 and 18 UTC cycles, there are 3 additional perturbed runs, out to 45 days. These 3 runs per cycle are initialized as in current operations (dotted lines in Figure A2)
- There are a total of 16 CFS runs every day, of which four runs go out to 9 months, three runs go out to 1 season and nine runs go out to 45 days.

APPENDIX D: Availability of CFSv2 data

- Real time operational data: Users must maintain their own continuing archive by downloading the real time operational data from the 7-day rotating archive located at: http://nomads.ncep.noaa.gov/pub/data/nccf/com/cfs/prod/
  This site includes both the initial conditions and forecasts made at each cycle of each day. Monthly means of the initial conditions are posted once a month and can be downloaded from a 6-month rotating archive at the same location given above.
- Selected data from the CFSv2 retrospective forecasts (both seasonal and sub seasonal) for the forecast period 1982-2010, may be downloaded from the NCDC web servers at: (http://nomads.ncdc.noaa.gov/data.php?name=access#cfs)
- Smoothed calibration climatologies have been prepared from the forecast monthly means and time series of selected variables and is available for download from the CFS website (http://cfs.ncep.noaa.gov). Please note that two sets of climatologies have been prepared for calibration, for the full period (1982-2010) and the later period (1999-2010). We highly recommend that the climatology prepared from the later period be used when calibrating real time operational predictions for variables in the tropics, ***such as SST and precipitation over oceans***. For skill estimates, we recommend that split climatologies be used for the two periods when removing the forecast bias.
- A small amount of CFSv2 forecast data for 2011-present may be found at the CFS website at http://cfs.ncep.noaa.gov/cfsv2/downloads.html
- Decadal runs : Some monthly mean and 3-hourly time series data from the NCEP decadal runs may be obtained from the ESGF/PMDI website at http://esgf.nccs.nasa.gov/esgf-web-fe/
- CMIP runs : Monthly mean data from the 3 CMIP runs is available for download from the CFS website at: http://cfs.ncep.noaa.gov/pub/raid0/cfsv2/cmipruns

References

Barker, H. W., R. Pincus, and J-J. Morcrette, 2002: The Monte Carlo Independent Column Approximation: Application within large-scale models. *Extended Abstracts, GCSS-ARM Workshop on the Representation of Cloud Systems in Large-Scale Models,* Kananaskis, AB, Canada, GEWEX, 1–10. [Available online at http://www.met.utah.edu/skrueger/gcss-2002/Extended-Abstracts.pdf.].

Cai, Ming, Chul-Su Shin, H. M. van den Dool, Wanqiu Wang, S. Saha, A. Kumar, 2009: The Role of Long-Term Trends in Seasonal Predictions: Implication of Global Warming in the NCEP CFS. *Wea. Forecasting*, **24**, 965–973. doi: 10.1175/2009WAF2222231.1

Chun, H.-Y, and J.-J. Baik, 1998: Momentum Flux by Thermally Induced Internal Gravity Wave and its Approximation for Large-Scale Models. *Journal of the Atmospheric Sciences,* **55,** 3299-3310.

Clough, S.A., M.W. Shephard, E.J. Mlawer, J.S. Delamere, M.J. Iacono, K. Cady-Pereira, S. Boukabara, and P.D. Brown, 2005: Atmospheric radiative transfer modeling: a summary of the AER codes, *J. Quant., Spectrosc. Radiat. Transfer*, 91, 233-244.

Behringer, D. W. 2007. The Global Ocean Data Assimilation System at NCEP. *11th Symposium on Integrated Observing and Assimilation Systems for Atmosphere, Oceans and Land Surface, AMS 87th Annual Meeting, San Antonio, Texas, 12pp*

Ek, M., K. E. Mitchell, Y. Lin, E. Rogers, P. Grunmann, V. Koren, G. Gayno, and J. D. Tarpley, 2003: Implementation of Noah land-surface model advances in the NCEP operational mesoscale Eta model. J. Geophys. Res., 108(D22), 8851, doi:10.1029/ 2002JD003296.

Fan, Y., and H. van den Dool (2008), A global monthly land surface air temperature analysis for 1948—present, *J. Geophys. Res.*, 113, D01103, doi:10.1029/2007JD008470.

Hou, Y., S. Moorthi and K. Campana, 2002: Parameterization of Solar Radiation Transfer in the NCEP Models. NCEP Office Note 441. http://www.emc.ncep.noaa.gov/officenotes/newernotes/on441.pdf

Iacono, M.J., E.J. Mlawer, S.A. Clough, and J.-J. Morcrette, 2000: Impact of an improved longwave radiation model, RRTM, on the energy budget and thermodynamic properties of the NCAR Community Climate Model, CCM3, *J. Geophys. Res*., 105, 14873-14890, 2000.

Kanamitsu, M., W. Ebisuzaki, J. Woollen, S.K. Yang, J.J. Hnilo, M. Fiorino, and G.L.Potter, 2002: NCEP–DOE AMIP-II Reanalysis (R-2). *Bull. Amer. Meteor. Soc*., **83,** 1631–1643.

Lin, H., G. Brunet, and J. Derome (2008), Forecast skill of the Madden–Julian oscillation in two Canadian atmospheric models. *Mon. Wea. Rev.,* **136,** 4130–4149.

Mitchell, K. E., H. Wei, S. Lu, G. Gayno and J. Meng, 2005: NCEP implements major upgrade to its medium-range global forecast system, including land-surface component. GEWEX newsletter, May 2005.

Mlawer E. J., S. J. Taubman, P. D. Brown, M.J. Iacono and S.A. Clough, 1997: radiative transfer for inhomogeneous atmosphere: RRTM, a validated correlated-K model for the longwave. *J. Geophys. Res*., 102(D14), 16,663-16,6832.

Moorthi, S., R. Sun, H. Xia, and C. R. Mechoso, 2010: Low-cloud simulation in the Southeast Pacific in the NCEP GFS: role of vertical mixing and shallow convection. NCEP Office Note 463, 28 pp [Available online at http://www.emc.ncep.noaa.gov/officenotes/FullTOC.html#2000]

Pincus, R., H.W. Barker, and J.-J. Morcrette, 2003: A fast, flexible, approximate technique for computing radiative transfer in inhomogeneous cloud fields. J. Geophys. Res., 108(D13), 4376, doi:10.1029/2002JD003322.

Penland and Saha, 2006: El Nino in the Climate Forecast System: T62 vs T126. Poster 1.3 at Climate Diagnostic and Prediction Workshop #30. Available online at http://www.cpc.ncep.noaa.gov/products/outreach/proceedings/cdw30_proceedings/P1.3.pdf

Reynolds, R. W., N. A. Raynor, T. M. Smith, D. C. Stokes, and W. Wang, 2002: An improved in situ and satellite SST analysis for climate. *J. Climate,* **15***,* 1609-1625.

Reynolds, R. W., T. M. Smith, C. Liu, D. B. Chelton, K. S. Casey, and M. G. Schlax, 2007: Daily high-resolution blended analyses for sea surface temperature. *J. Climate,* **20***,* 5473–5496.

Saha, S. and Coauthors, 2006: The NCEP climate forecast system, *J. Climate*, **19**, 3483-3517.

Saha, S. and Coauthors, 2010: The NCEP climate forecast system reanalysis. *Bull. Amer. Meteor. Soc.,* ***91***, 1015-1057.

Sato, M., Hansen, J.E., McCormick, M.P., and Pollack, J.B., 1993, Stratospheric aerosol optical depths, 1850-1990: Journal of Geophysical Research, v. 98, p. 22987-22994.

Sun, R., S. Moorthi and C. R. Mechoso, 2010: Simulaton of low clouds in the Southeast Pacific by the NCEP GFS: sensitivity to vertical mixing. Atmos. Chem. Phys., **10**, 12261-12272.

Wheeler, M, and H. H. Hendon (2004), An all-season real-time multivariate MJO index: Development of an index for monitoring and prediction. *Mon. Wea. Rev.,* **132,** 1917–1932.

Zhang, Qin, Huug van den Dool, 2012: Relative Merit of Model Improvement versus Availability of Retrospective Forecasts: The Case of Climate Forecast System MJO Prediction. *Wea. Forecasting*, **27**, 1045–1051. doi: http://dx.doi.org/10.1175/WAF-D-11-00133.1

Table 1: **Surface Water Budget Comparison of CFSv1, CFSR and CFSv2 for summer (JJA) and winter (DJF). Values are averages for NH land. Units are mm/day.**

	CFSv1 (JJA/DJF)	**CFSR (JJA/DJF)**	**CFSv2 (JJA/DJF)**
Precipitation (mm/day)	3.3/1.6	3.2/1.4	2.7/1.3
Evaporation (mm/day)	2.5/1.1	2.2/0.89	2.1/0.71
Run off (mm/day)	0.56/0.16	0.16/0.04	0.22/0.06
Soil moisture (mm)	441/476	510/514	502.43/501.37
Snow water (mm)	0.09/4.1	0.02/4.2	0.01/6.5

Table A1 CFSv2 Retrospective Calendar
(organized by date of release of the official CPC seasonal prediction every month)

MID JANUARY RELEASE (24 members)
12 December at 0, 6 12 and 18 Z
17 December at 0,6,12 and 18 Z
22 December at 0,6,12 and 18 Z
27 December at 0,6,12 and 18 Z
1 January at 0,6,12 and 18 Z
6 January at 0,6,12 and 18 Z

MID FEBRUARY RELEASE (24 members)
11 January at 0, 6 12 and 18 Z
16 January at 0,6,12 and 18 Z
21 January at 0,6,12 and 18 Z
26 January at 0,6,12 and 18 Z
31 January at 0,6,12 and 18 Z
5 February at 0,6,12 and 18 Z

MID MARCH RELEASE (24 members)
10 February at 0, 6 12 and 18 Z
15 February at 0,6,12 and 18 Z
20 February r at 0,6,12 and 18 Z
25 February at 0,6,12 and 18 Z
2 March at 0,6,12 and 18 Z
7 March at 0,6,12 and 18 Z

MID APRIL RELEASE (24 members)
12 March at 0, 6 12 and 18Z
17 March at 0,6,12 and 18 Z
22 March at 0,6,12 and 18 Z
27 March at 0,6,12 and 18 Z
1 April at 0,6,12 and 18 Z
6 April at 0,6,12 and 18 Z

MID MAY RELEASE (24 members)
11 April at 0, 6 12 and 18 Z
16 April at 0,6,12 and 18 Z
21 April at 0,6,12 and 18 Z
26 April at 0,6,12 and 18 Z
1 May at 0,6,12 and 18 Z
6 May at 0,6,12 and 18 Z

MID JUNE RELEASE (24 members)
11 May at 0, 6 12 and 18 Z
16 May at 0,6,12 and 18 Z
21 May at 0,6,12 and 18 Z
26 May at 0,6,12 and 18 Z
31 May at 0,6,12 and 18 Z
5 June at 0,6,12 and 18 Z

MID JULY RELEASE (24 members)
10 June at 0, 6 12 and 18 Z
15 June at 0,6,12 and 18 Z
20 June at 0,6,12 and 18 Z
25 June at 0,6,12 and 18 Z
30 June at 0,6,12 and 18 Z
5 July at 0,6,12 and 18 Z

MID AUGUST RELEASE (24 members)
10 July at 0,6,12 and 18 Z
15 July at 0,6,12 and 18 Z
20 July at 0,6,12 and 18 Z
25 July at 0,6,12 and 18 Z
30 July at 0,6,12 and 18 Z
4 August at 0,6,12 and 18 Z

MID SEPTEMBER RELEASE (24 members)
9 August at 0,6,12 and 18 Z
14 August at 0,6,12 and 18 Z
19 August at 0,6,12 and 18 Z
24 August at 0,6,12 and 18 Z
29 August at 0,6,12 and 18 Z
3 September at 0,6,12 and 18 Z

MID OCTOBER RELEASE (24 members)
8 September at 0,6,12 and 18 Z
13 September at 0,6,12 and 18 Z
18 September at 0,6,12 and 18 Z
23 September at 0,6,12 and 18 Z
28 September at 0,6,12 and 18 Z
3 October at 0,6,12 and 18 Z

MID NOVEMBER RELEASE (28 members)
8 October at 0,6,12 and 18 Z
13 October at 0,6,12 and 18 Z
18 October at 0,6,12 and 18 Z
23 October at 0,6,12 and 18 Z
28 October at 0,6,12 and 18 Z
2 November at 0,6,12 and 18 Z
7 November at 0,6,12 and 18 Z

MID DECEMBER RELEASE (24 members)
12 November at 0,6,12 and 18 Z
17 November at 0,6,12 and 18 Z
22 November at 0,6,12 and 18 Z
27 November at 0,6,12 and 18 Z
2 December at 0,6,12 and 18 Z
7 December at 0,6,12 and 18 Z

Figure legends

Figure 1. The bivariate anomaly correlation (BAC)x100 of CFS in predicting the MJO for period 1999-2009, as expressed by the Wheeler and Hendon (WH) index (two EOFs of combined zonal wind and OLR). On the left is CFSv2 and on the right is CFSv1. Both are subjected to Systematic Error Correction. The black lines indicate the 0.5 level of BAC.

Figure 2. Anomaly correlation of three-month-mean SST between model forecasts and observation. (a) 3-month lead CFSv2, (b) 6-month lead CFSv2, (c) 3-month lead CFSv1 and (d) 6-month lead CFSv1. Contours are plotted at an interval of 0.1.

Figure 3. Standard deviation of three-month-mean SST forecasts (K). (a) Observation (b) 3-month lead CFSv2 minus observation, (c) 6-month lead CFSv2 minus observation, (d) 3-month lead CFSv1 minus observation, and (e) 6-month lead CFSv1 minus observation. Contours are plotted at an interval of 0.2 from 0.2 to 1.6 in (a) and from -0.5 to 0.5 in (b), (c), (d) and (e).

Figure 4. Evaluation of anomaly correlation as a function of target month (horizontal axis) and forecast lead (vertical axis). On the left is CFSv1, on the right CFSv2. Top row shows monthly 2-meter temperature over NH land, middle row shows monthly precipitation over NH land and the bottom row shows the SST in the Nino3.4 area. The scale is the same for all 6 panels. Except for the years added, the CFSv1 entries in this figure (left column) should correspond to the figures in Saha et al (2006).

Figure 5. The mean September sea ice concentration from 1982 to 2010 from CFSR (top left), and the bias from the predicted mean condition for the September sea ice concentration with a lead time of 1-month (top right, August 15 IC), 3-month (bottom left, June 15 IC), and 6-month (bottom right, March 15 IC).

Figure 6. The standard deviation of the September sea ice concentration from 1982 to 2010 from CFSR (top left), and the difference of the standard deviation between the model

prediction and that from the CFSR for the September sea ice concentration with a lead time of 1-month (top right, August 15 IC), 3-month (bottom left, June 15 IC), and 6-month (bottom right, March 15 IC).

Figure 7. Top panel (a) shows the globally averaged SST anomaly in NCEP decadal integrations. Sixty two 10 year integration were made and they are plotted as yellow traces. The observed single trace of 30+ years is given in black. Units along the Y-axis are in Kelvin. The definition of anomaly is given in the text. Bottom panel (b) shows the same, except for the globally averaged 2 meter temperature anomaly over land.

Figure 8: Power spectra of time series of monthly anomalies of the Nino34 index (average SST from 170W to 120W, and 5S to 5N). Upper left is for the observation while the other three panels are for CMIP runs of 43, 52 and 100 years respectively.

Figure A1: Reforecast configuration of the CFSv2.

Figure A2: Operational configuration of the CFSv2.

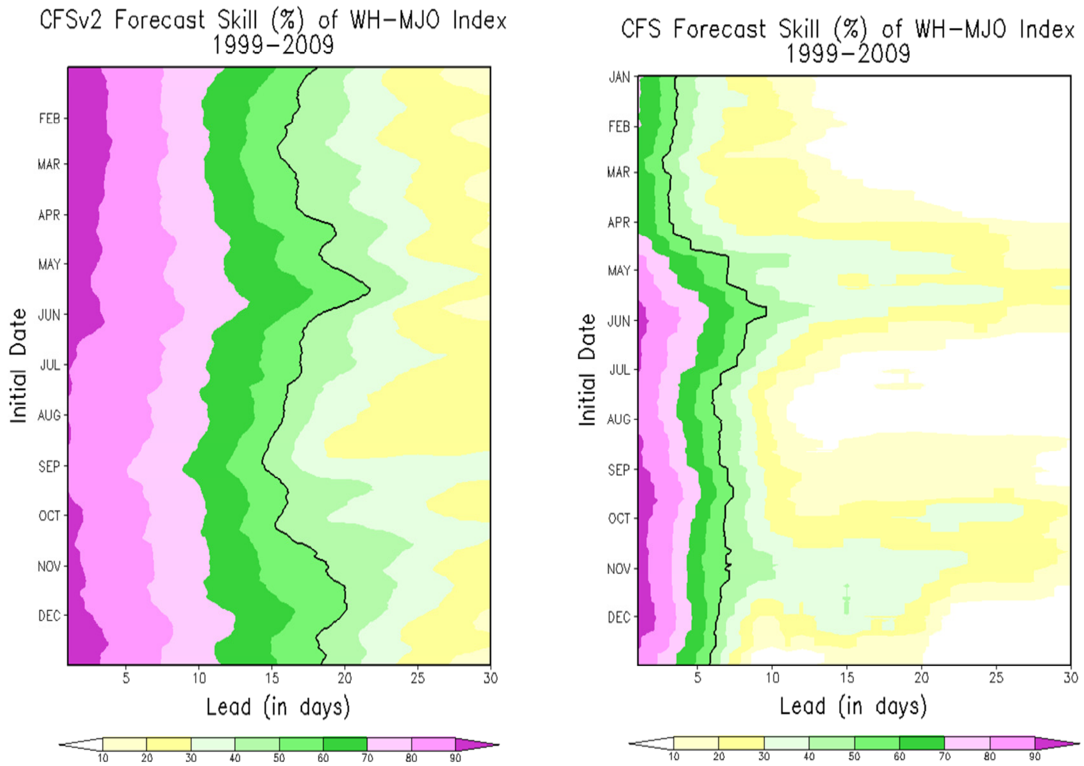

Figure 1. The bivariate anomaly correlation (BAC)x100 of CFS in predicting the MJO for period 1999-2009, as expressed by the Wheeler and Hendon (WH) index (two EOFs of combined zonal wind and OLR). On the left is CFSv2 and on the right is CFSv1. Both are subjected to Systematic Error Correction. The black lines indicate the 0.5 level of BAC.

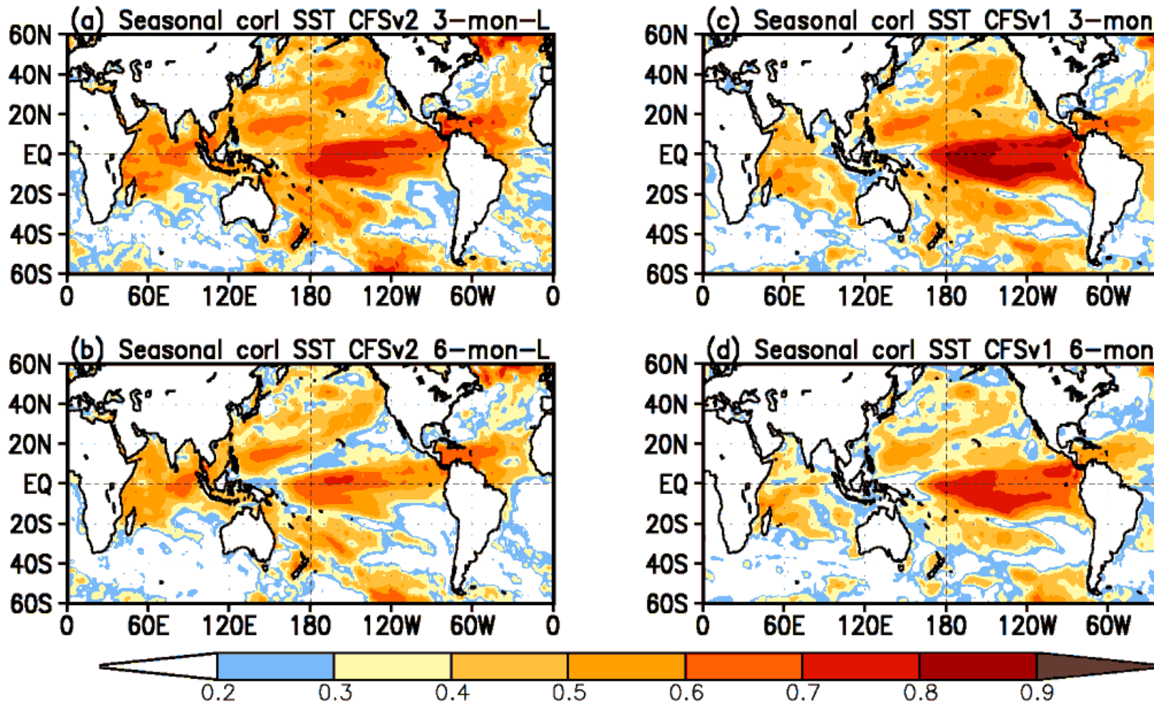

Figure 2. Anomaly correlation of three-month-mean SST between model forecasts and observation. (a) 3-month lead CFSv2, (b) 6-month lead CFSv2, (c) 3-month lead CFSv1 and (d) 6-month lead CFSv1. Contours are plotted at an interval of 0.1.

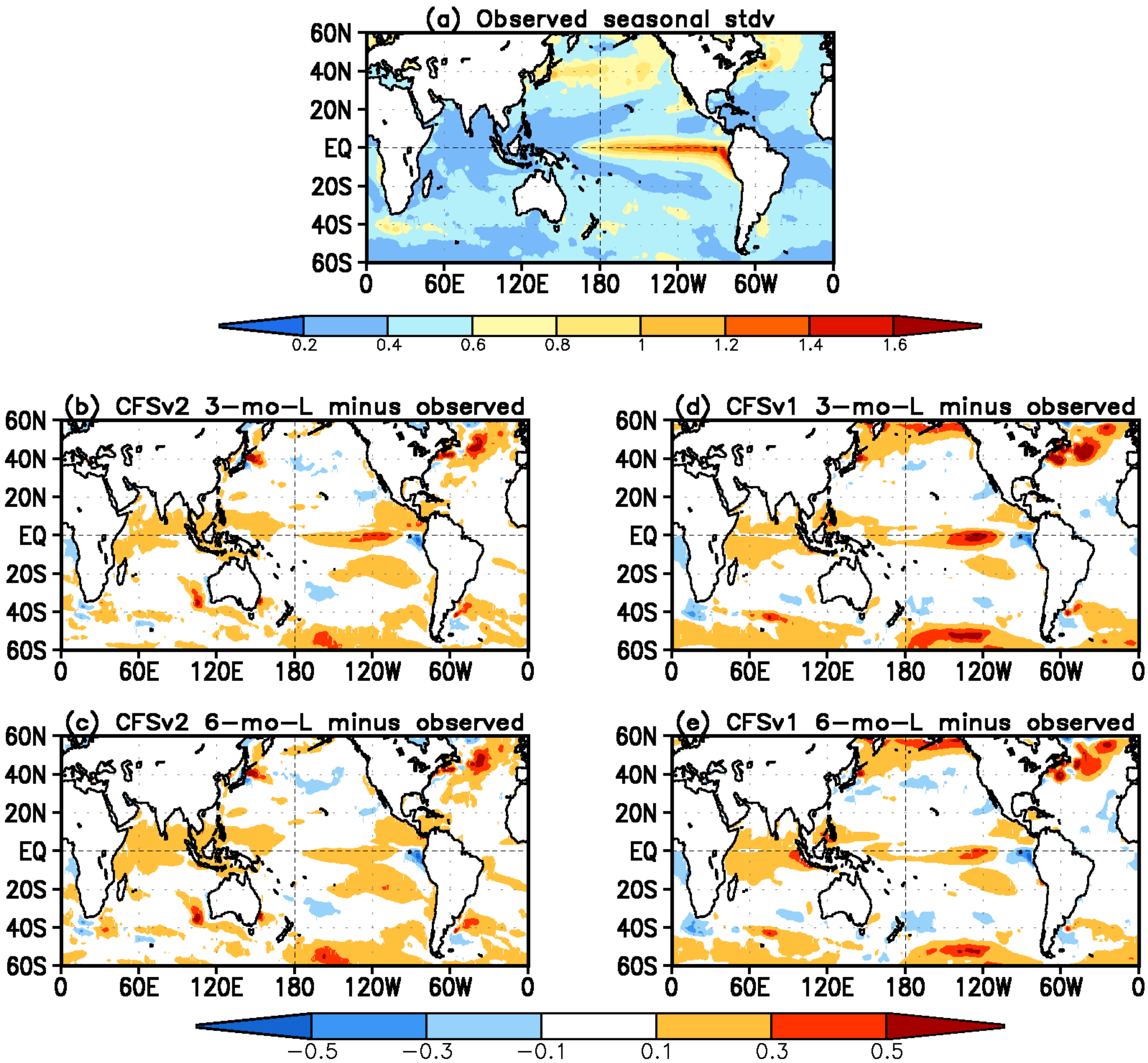

Figure 3. Standard deviation of three-month-mean SST forecasts (K). (a) Observation (b) 3-month lead CFSv2 minus observation, (c) 6-month lead CFSv2 minus observation, (d) 3-month lead CFSv1 minus observation, and (e) 6-month lead CFSv1 minus observation Contours are plotted at an interval of 0.2 from 0.2 to 1.6 in (a) and from -0.5 to 0.5 in (b), (c), (d) and (e).

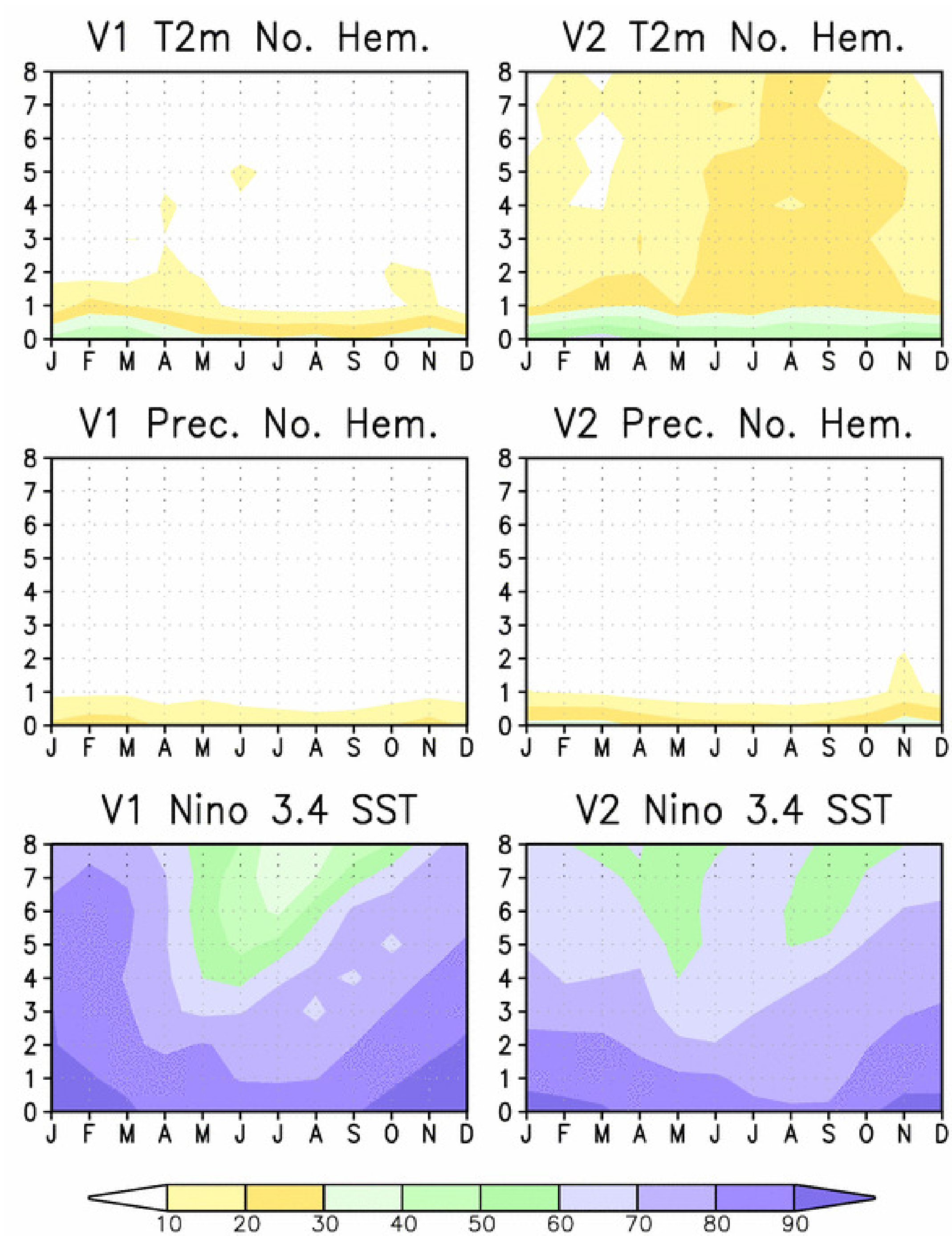

Figure 4. Evaluation of anomaly correlation as a function of target month (horizontal axis) and forecast lead (vertical axis). On the left is CFSv1, on the right CFSv2. Top row shows monthly 2-meter temperature over NH land, middle row shows monthly precipitation over NH land and the bottom row shows the SST in the Nino3.4 area. The scale is the same for all 6 panels. Except for the years added, the CFSv1 entries in this figure (left column) should correspond to the figures in Saha et al (2006).

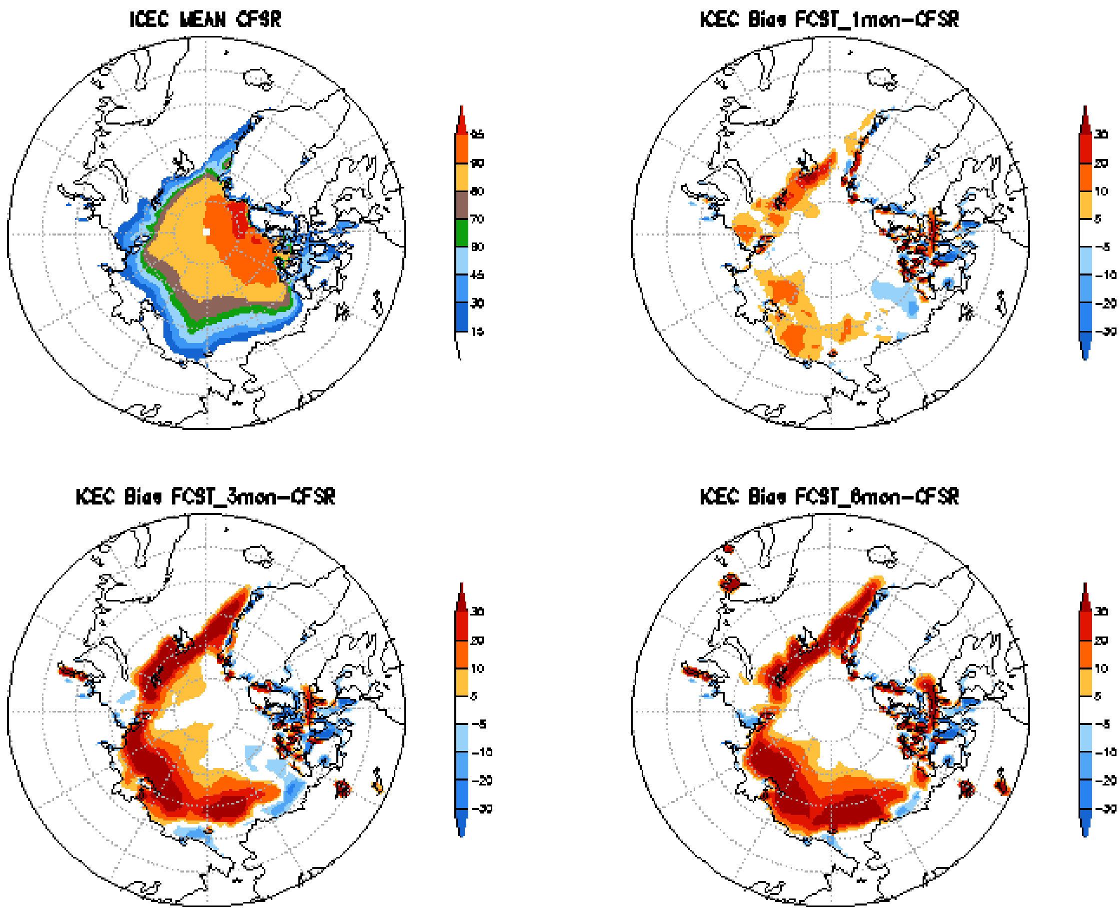

Figure 5; The mean September sea ice concentration from 1982 to 2010 from CFSR (top left), and the bias from the predicted mean condition for the September sea ice concentration with a lead time of 1-month (top right, August 15 IC), 3-month (bottom left, June 15 IC), and 6-month (bottom right, March 15 IC).

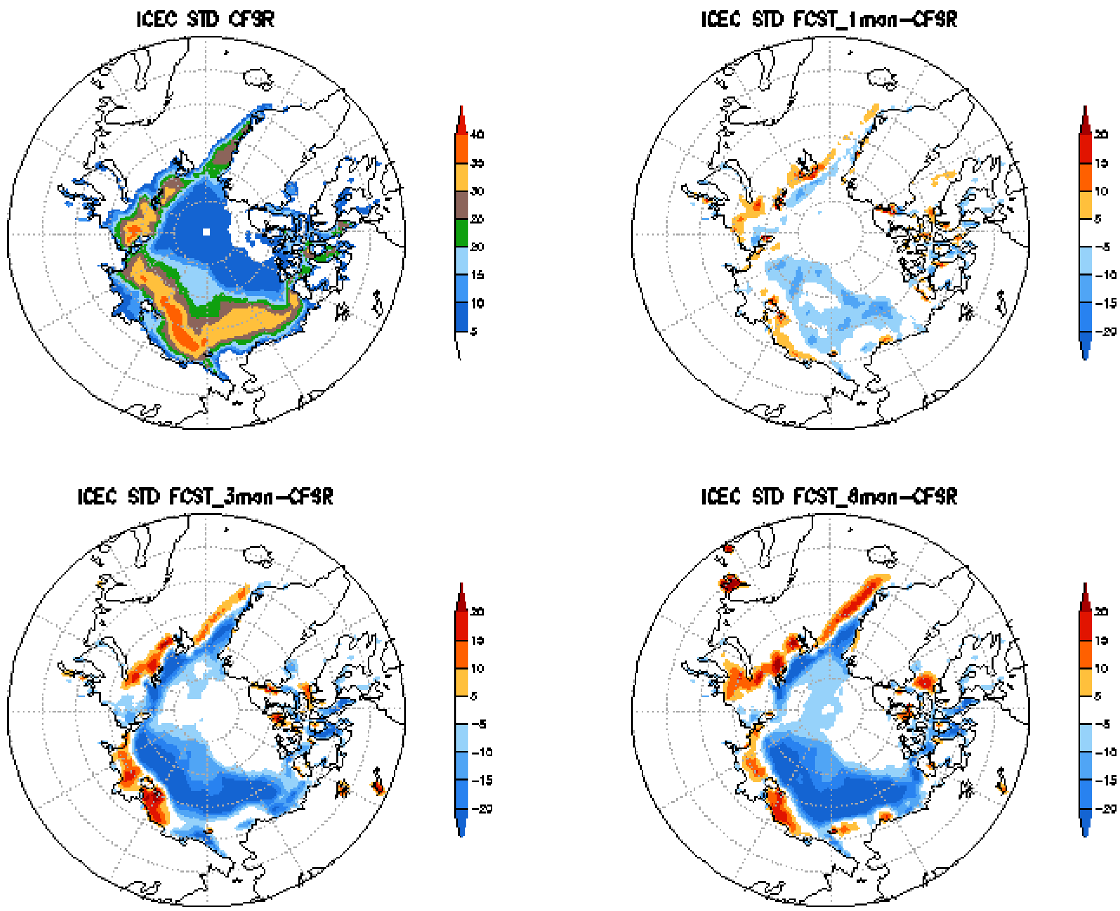

Figure 6; The standard deviation of the September sea ice concentration from 1982 to 2010 from CFSR (top left), and the difference of the standard deviation between the model prediction and that from the CFSR for the September sea ice concentration with a lead time of 1-month (top right, August 15 IC), 3-month (bottom left, June 15 IC), and 6-month (bottom right, March 15 IC).

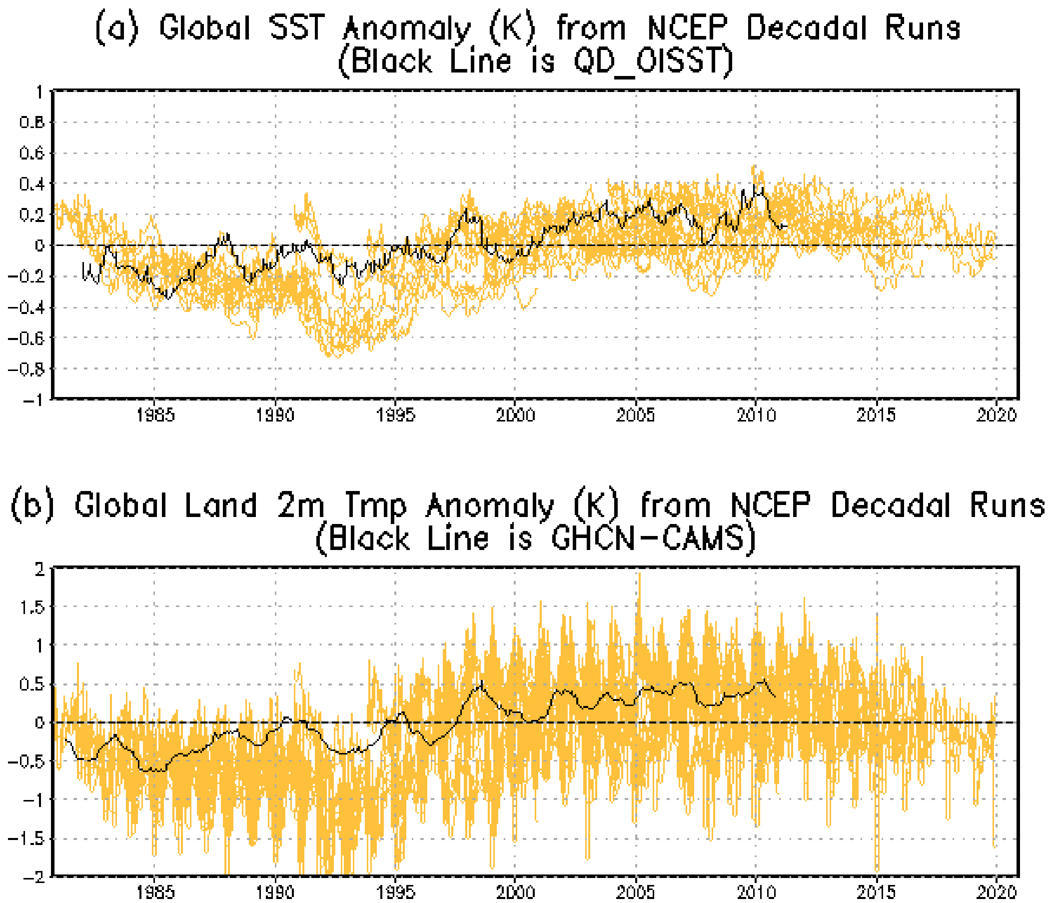

Figure 7. Top panel (a) shows the globally averaged SST anomaly in NCEP decadal integrations. Sixty two 10 year integration were made and they are plotted as yellow traces. The observed single trace of 30+ years is given in black. Units along the Y-axis are in Kelvin. The definition of anomaly is given in the text. Bottom panel (b) shows the same, except for the globally averaged 2 meter temperature anomaly over land.

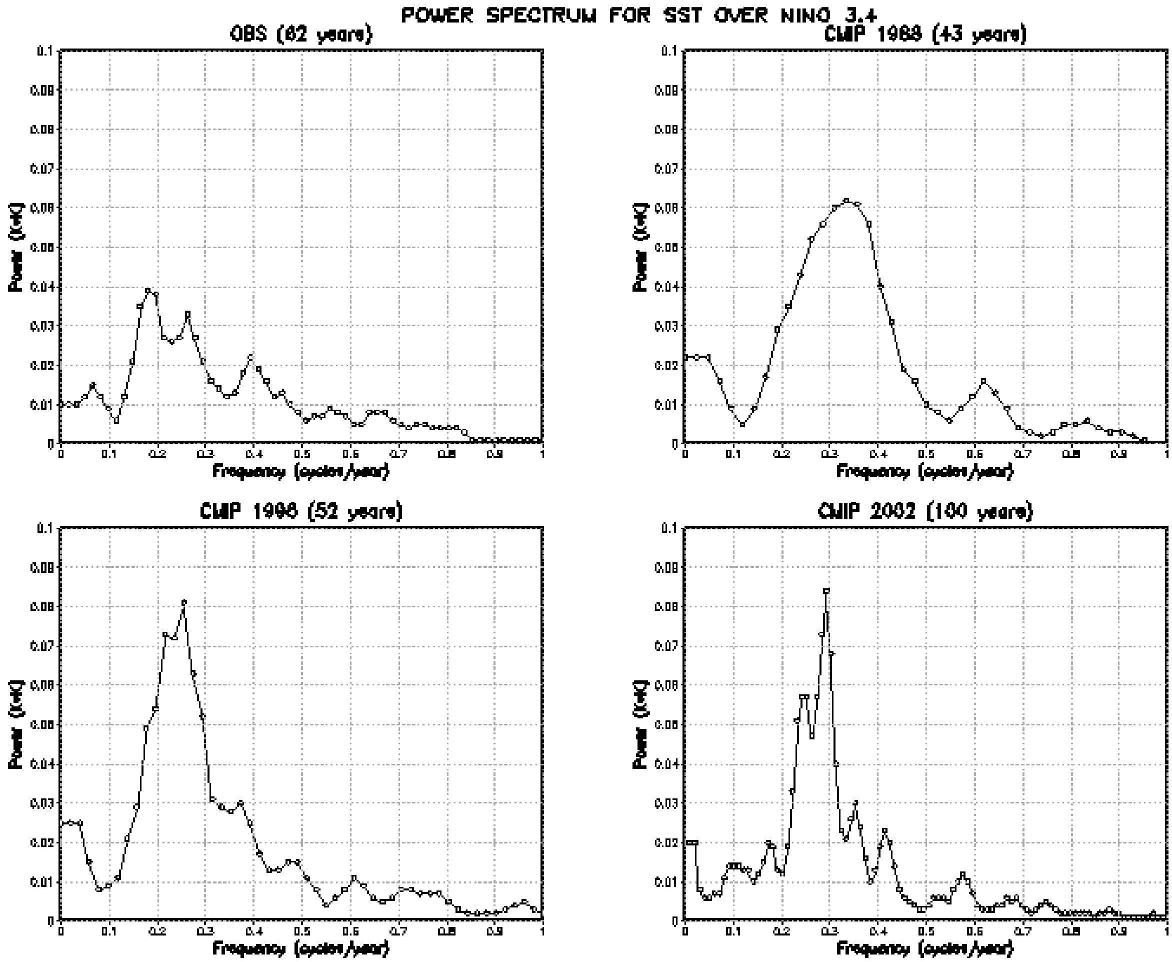

Figure 8: Power spectra of time series of monthly anomalies of the Nino34 index (average SST from 170W to 120W, and 5S to 5N). Upper left is for the observation while the other three panels are for CMIP runs of 43, 52 and 100 years respectively.

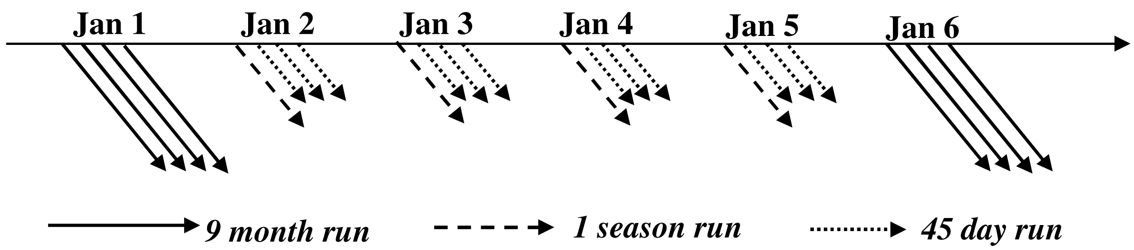

Figure A1: Reforecast Configuration of the CFSv2

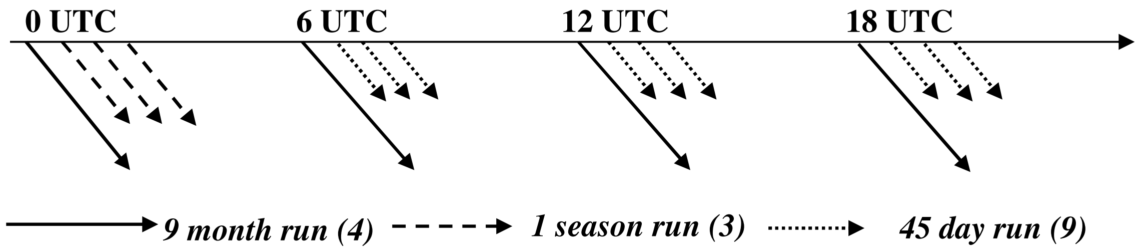

Figure A2: Operational Configuration of the CFSv2