NOAA Climate Test Bed
Center for Ocean-Land-Atmosphere Studies
Earth System Science Interdisciplinary Center
UM Department of Atmospheric and Oceanic Science

Research to Operation and Operation to Research

Office of Science and Technology

NOAA's National Weather Service

2009

"This Climate Test Bed (CTB) Joint Seminar Series is going to focus on the CTB strategic priorities; specifically composed of the Research to Operation (R2O) component, i.e. NCEP Climate Forecast System (CFS) improvement, multi-model ensemble, climate forecast products, and the Operation to Research (O2R) component of using CFS for scientific research."

— Opening address by Fiona Horsfall,
Director of NOAA Climate Test Bed

Table of Content

Regional and Seasonal Improvements in the Skill and Value of CPC 3-Month Outlooks

Edward A. O'Lenic, David A. Unger, Kenneth S. Pelman, and Mike Halpert

Climate Prediction Center, NOAA/NWS/NCEP

O'Lenic *et al.* (2008) reported on the use of an objective technique which, on dependent data from 1995 through 2004, improved the average Heidke skill score of CPC's operational ½-month-lead 3-Month temperature and precipitation outlooks from 22 to 24, for temperature, and from 8.8 to 12.1, for precipitation.

This paper extends that work by calculating the increase in the percentage of the time non-EC (equal chances) probabilities are predicted, over the 1995-2004 dependent data period. The fraction of the map covered by non-EC probabilities is one metric of usefulness which users are quite sensitive to. Reducing the area covered by EC (33.33…% for each of the three tercile categories) is highly desirable.

Each map in the top row of Fig. 1 shows the average Heidke skill score (HSS, lines, see Appendix) over 1995-2004 for 30 official (OFF) forecasts of ½-month lead precipitation, made in real-time.
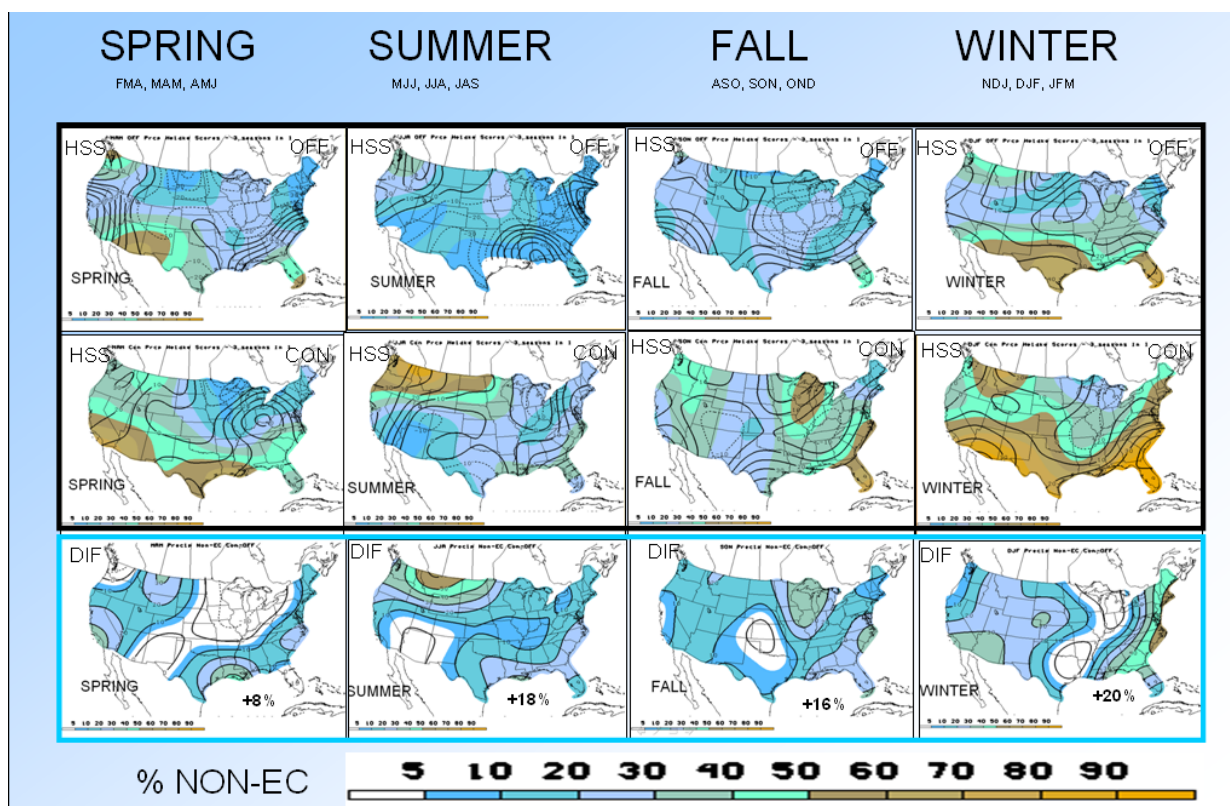
Fig. 1. Average Heidke skill score (HSS, lines) over 1995-2004 for 30 official (OFF) forecasts (top row) made in real-time and retrospective objective consolidation (CON) forecasts (middle row) of ½-month lead 3-month precipitation. The colored shading indicates the percent of the time that forecasts with other than equal chances forecasts were made. The scale is indicated by the bar at the bottom of the diagram. Lines and colors in the bottom row of maps indicate the difference between CON and OFF % non-EC forecasts.

Correspondence to: Edward A. O'Lenic, Climate Prediction Center, NOAA/NWS/NCEP;
 E-mail: Ed.Olenic@noaa.gov

Fig. 2. Same as Figure 1 except for the temperature forecast.

The second row from the top shows the same information as the top row, except for forecasts made using an objective technique (Unger *et al.*, 2009) to numerically combine the identical set of forecast tools which were available to the forecasters in real-time during 1995-2004. This is referred to as the consolidation (CON). These tools include the 2-tier NCEP coupled model described in Ji and Leetmaa (1995), the optimum climate normals (OCN, or trend) (Court, 1967-68), the canonical correlation analysis (CCA) (Barnston, 1997), and the screening multiple linear regression (SMLR). The latter two tools are multi-variate statistical techniques.

The lines and colors in the set of maps on the third row show the arithmetic difference between the CON and OFF maps in rows 1 and 2. A number near the bottom right side of each map shows the map average of this difference. The CON technique averages a higher percentage of non-EC forecasts than did the OFF by 8%, in spring, 18%, in summer, 16%, in fall, and 20%, in winter. Since the trend is relatively small for 3-month precipitation, these improvements are smaller than those for temperature, but still large enough to be noticeable to users.

The results for temperature are shown in Fig. 2.

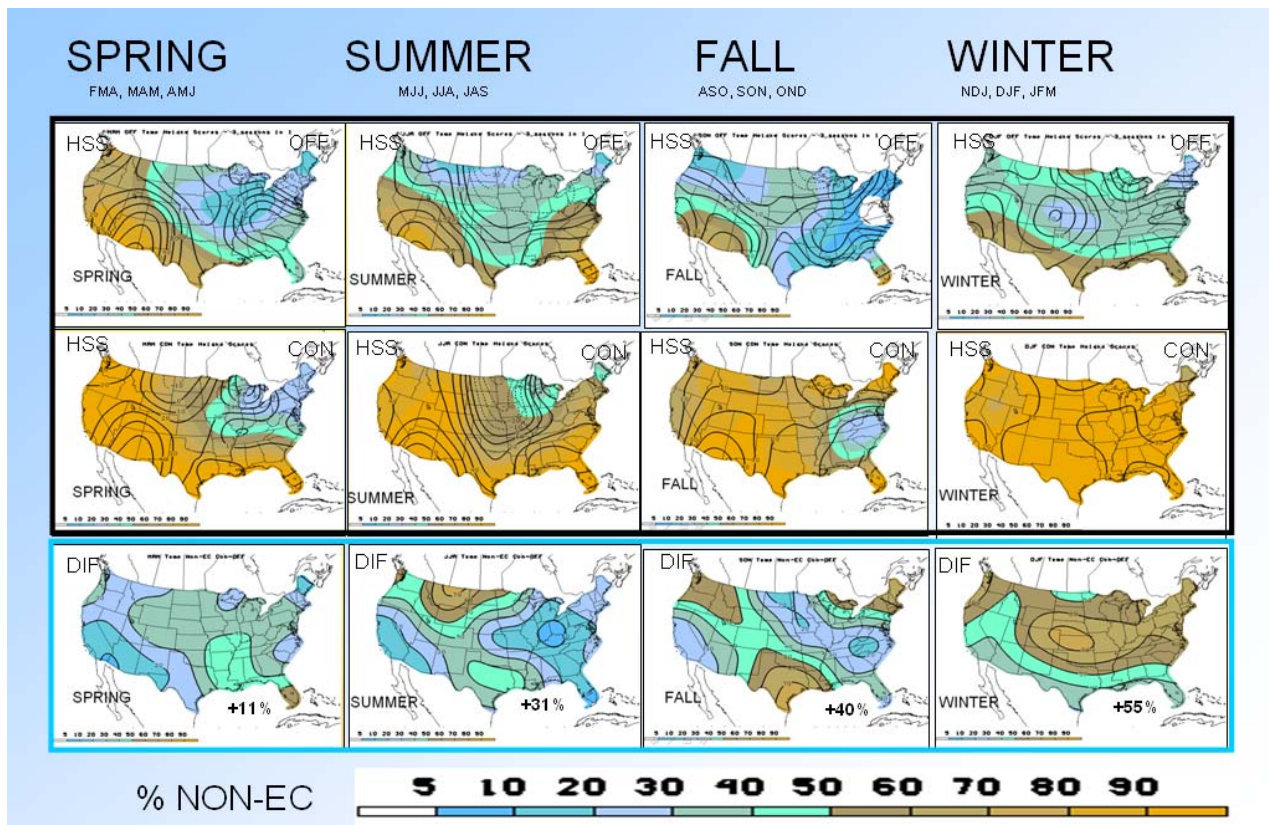The consolidation technique produces large increases in non-EC forecast percentages for temperature (Fig. 2): 11%, for spring, 31%, for summer, 40%, for fall, and 55%, for winter. This large increase is due, in part, to the fact that the trend is more strongly reflected in temperature than it is in precipitation (Fig. 1).

Since these results are on dependent data, we show, in Fig. 3, the result of using the CON technique in operational ½-month lead 3-Month temperature outlooks since 2006. There is a clear break in the time series of 48-month running mean HSS which commences when the CON was implemented into CPC operations in early 2006. This performance is evidence that the CON technique is a reasonable way to present forecasters with an accurate first-guess from which to begin forecasting.

Appendix

The Heidke skill score (HSS) is a categorical score which compares the accuracy of a forecast of interest (e.g., 3-month outlooks) with that of a reference forecast, such as climatology (random) forecasts:

$$HSS = (c-e)/(t-e)*100\%$$

where c = # gridpoints forecast correctly
       e = # gridpoints expected correct randomly
       t = # gridpoints in total

In a 3-class, tercile system, $-50 \leq HSS \leq 100$.

Fig. 3. 48-month running mean of the Heidke skill score of real-time, operational ½-month lead 3-month temperature outlooks from September 2002 through September 2008 (blue line) (GPRA score). The red line is the GPRA goal.

References

Barnston, A. G., 1994: Linear statistical short-term climate predictive skill in the Northern Hemisphere, *J. Climate*, **7**, 1513-1564.

Court, A., (1967-68): Climate normals as predictors: Parts I-IV. Science Reports, Air Force Cambridge Research Laboratory, Bedford MA, Contract AF19(628)-5176.

Ji, M. A., A. Kumar, and A. Leetmaa, 1994: A multi-season climate forecast system at the National Meteorological Center. *Bull. Amer. Meteor. Soc.*, **75**, 569-577.

O'Lenic, E. A., D. A. Unger, M. S. Halpert, and K. S. Pelman, 2008: Developments in operational long-range prediction at CPC. *J. Weather and Forecasting*, **23**, 496-515.

O'Lenic, E.A., D.A. Unger, M.S. Halpert, and K. S. Pelman, 2008: Corrigendum. *J. Weather and Forecasting*, **23**, 1044.

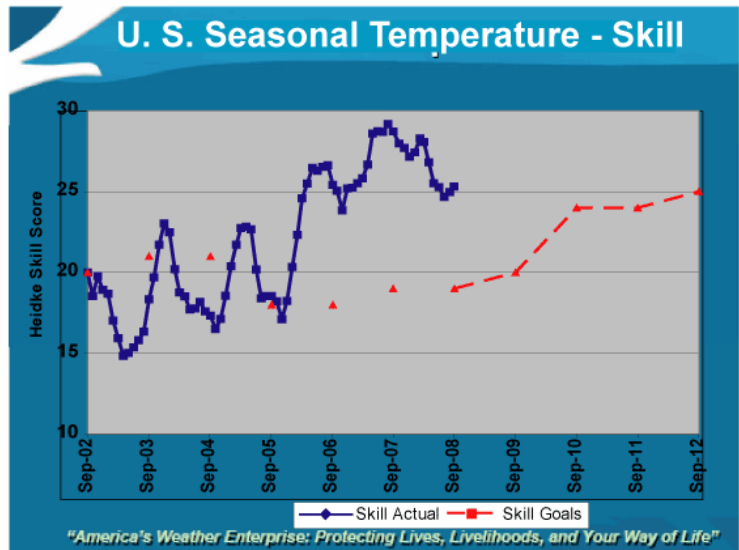Unger, D., H. van den Dool, E. OL'enic, and D. Collins, 2009: Ensemble regression. *Mon. Wea. Rev.*, in press.

Annual Cycle and Prediction of Interannual Variability

Zhaohua Wu

Department of Meteorology & Center for Ocean-Atmospheric Prediction Studies

Florida State University, Tallahassee, FL

1. Some Issues of Climate Prediction

There are four major issues that we concerned in this talk. The first one is the reference frame for climate anomaly. Current climate prediction of interannual variability starts mostly with what is called climate anomaly, which is the deviation of a climate variable from its reference frame, often its annual cycle. Traditionally, this annual cycle is taken to be an exact repeat of itself year after year. Such defined reference frame for climate anomaly has several drawbacks: (1) The repetitive annual may not reflect well the intrinsic nonlinearity of the climate system, especially under external forcing, for we are concerning the annual cycle of a climate variable which is the response of nonlinear climate system to external forcing rather than the periodicity of the external forcing itself; (2) The traditional annual cycle defined uses an a priori determined episode of a climate data, e.g., from 1960 to 1990. This a priori determined episode is not backed by any physical reason; rather, it is just for convenience. This 'convenience' unfortunately brings inconvenience, for we can define annual cycles of other episodes with equal validation. Following the way of defining traditional annual cycle, we can define numerous versions annual cycles as well as numerous versions of anomaly; and (3) When the anomaly with a traditional annual cycle is examined more carefully, we would find that the anomaly may still contain "annual cycle", which leads to the conceptual inconsistency for the anomaly.

Recently, Wu et al. (2008) has proposed an alternative reference frame for climate anomalies, the modulated annual cycle (MAC) that allows the annual cycle to change from year to year, for defining anomalies. The MAC is temporally locally determined so that it is unique and has some "absoluteness" for any given climate time series. Moreover, the temporally locality of MAC also bypasses stationarity and linear assumptions of a climate time series often used in the majority of climate studies, and hence, the anomaly being defined is also unique. The MAC reference frame has been demonstrated to have many advantages in our understanding of climate system by Wu et al. (2008). For examples, the re-emergence mechanism may be alternatively interpreted as an explanation of the change of the annual cycle instead of an explanation of the interannual to interdecadal persistence of SST anomalies; the ENSO phase locking can largely be attributed to the residual annual cycle (the difference of the MAC and the corresponding traditional annual cycle) contained in the traditional anomaly, and, therefore, can be alternatively interpreted as a part of the annual cycle phase locked to the annual cycle itself. It was also shown in Wu et al. (2008) that using MAC as a reference framework for anomaly can bypass the difficulty brought by concepts such as "decadal variability of summer (or winter) climate" for understanding the low-frequency variability of the climate system.

The second major issue is what the predictable part of a climate variable, e.g., the sea surface temperature at Niño3.4 region, is for short term (for example, one year) climate prediction. A time series of climate variable contains noise, sub-annual high frequency variability, changing annual cycle, interannual and longer timescale variability, and secular trend. For variability of different timescales, the physical reasons that lead to the variability of particular timescales may be different. A handy example is the SST at Niño3.4 region, of which the sub-annual variability and the changing annual cycle may have more to do with (in large-scale view) the direct response of the tropical upper ocean to solar radiation while the interannual variability has more to do with the atmosphere-ocean coupled instability and equatorial wave dynamics, a relatively slowly evolving physical process.

To further explain the concept of the predictability of climate system, we use string-mass-pendulum system under random forcing, which is sketched in Figure 1. Suppose that the cube of large mass move leftward or rightward slowly and the small ball swings fast, the actual position of the ball of the pendulum can

---

Correspondence to: Zhaohua Wu, Department of Meteorology & Center for Ocean-Atmospheric Prediction Studies, Florida State University, Tallahassee, FL; E-mail: zwu@fsu.edu

be determined by the position of the cube and the position of ball with respect to the balance point of the pendulum. In the case of no random forcing, the positions are quite predictable for both the cube and the small ball with respect to the balance point of the pendulum, and thereby the actual position of the small ball. When there is moderate random forcing, due to the large inertia of the cube, its left-rightward oscillation would not be affected much, therefore, the position of the cube is still predictable. However, the position of small ball with respect to the balance point of the pendulum is affected significantly by the random forcing, and, therefore, is not predictable.

A relevant question is whether an accurate prediction model of the small ball based on the historical record of the position of the small ball can be constructed. Due to different physical processes (string-mass system and pendulum system) hidden in the record of the small ball and to the lack of a priori knowledge of the hidden string-mass system and pendulum system and the random forcing, a simple oscillatory model may not explain the position data and can not predict the position of small ball well. The inference of this argument is that we should be less ambitious: instead of trying to predict the exact location of small ball, we rather focus on the prediction of the location of the predictable part, the location of the cube. The implication of this analogue for short term climate prediction is: We should try to isolate the predictable part, the relatively low-frequency part of the climate data and make prediction of that, and leave out the hardly predictable high frequency part.

Fig. 1. A string-mass system analogue of climate data.
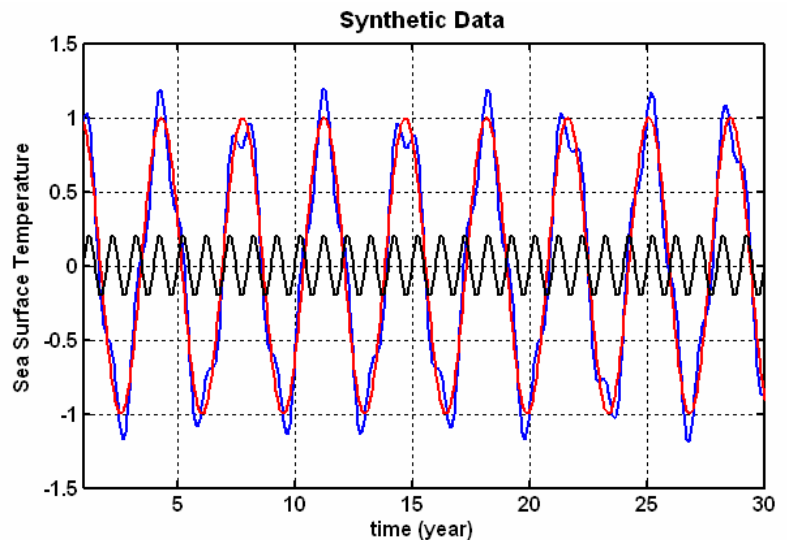
Fig. 2. The schematics of a synthetic climate data (blue line) and its components (red and black lines, two sinusoidal waves of different frequency).

The third major issue is from data analysis perspective, as displayed in Figure 2. Suppose that the blue line is a climate index. If we construct a statistical model, e.g., a Markov model, directly based on this index and use that to predict the exact value of the index for future, we can not predict well. However, if we separate the two components, which have exact functional forms (with the simple physical mechanisms), we can predict both the functional forms separately with accuracy. As a consequence, we can predict the blue line accurately. This synthetic example implies that for a climate time series, we may obtain a better prediction based on the predictions of individual components of the time series.

The last major issue is related to the non-stationarity of the climate data. In many statistical prediction models, which are constructed based on the data over the whole temporal domain, stationarity is a pre-requirement. However, if the climate data was generated by non-stationary processes, the stationarity assumption used in the construction of a model would lead to significant error of prediction. In such a case, a non-stationary approach may lead to a significantly improved prediction.
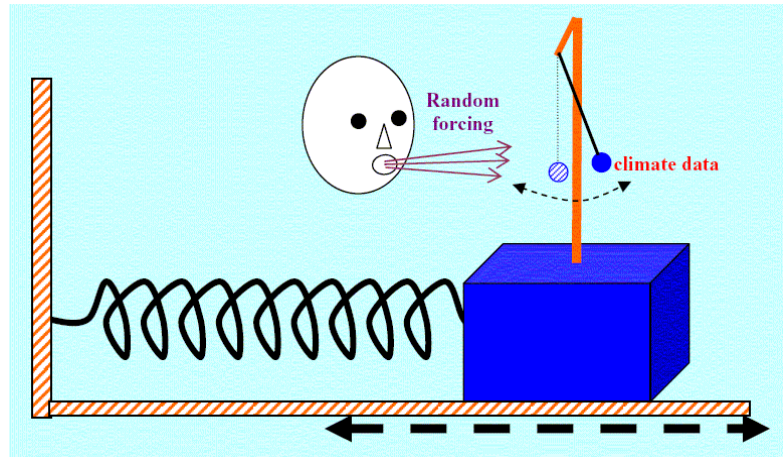
2. An Alternative Prediction Scheme

Based on the discussions in the previous section, we constructed an alternative type of prediction scheme to the traditional statistical models. The new scheme is illustrated in Figure 3. In Figure 3, the black line is the interannual and lower frequency components of cold tongue index (CTI), an index defined based on equatorial SST of central and eastern Pacific. We use the Ensemble Empirical Mode Decomposition (Huang et al. 1998, Huang and Wu 2008, Wu and Huang 2009) to decompose the black line into oscillatory components of different frequency and obtain both the instantaneous amplitude and the instantaneous frequency for each component. Suppose that we want to predict the index for 1989, we first predict the instantaneous amplitude of each component for the year by using cubic spline extrapolation based on the instantaneous amplitude of previous years. By using the instantaneous frequency of the Dec. 1988 for each component, we predict the oscillatory component one by one.

Based on this scheme, we make prediction of interannual timescale and longer timescale part of CTI for year 1967 to 1999. Figure 4 presents the result of the retrospective prediction from each month

Fig. 3. The part of interannual and longer timescales of the cold tongue index (black line) and its oscillatory components (blue lines and magenta line). The red lines are upper and lower envelopes. The green dashed line corresponding to Jan. 1989.

Fig. 4. The retrospective prediction of the part of cold tongue index of interannual and longer timescales (red line). The predictions from each month are displayed by green lines.

during the period. In each prediction, the prediction duration is one year. In general, the prediction results are quite promising, which can be validated by the closeness of the green lines to the red line (which is the part of interannual and longer timescales of CTI). The predictions seem to have systematic errors: for certain years, the predictions are consistently higher than the index values while in the other periods, the predictions are consistently lower than index values. Such systematic errors may be reduced if other potential correction schemes are designed and used.

References

Huang, N. E. , Z. Shen, and S. R. Long, M. C. Wu, E. H. Shih, Q. Zheng, C. C. Tung, and H. H. Liu, 1998: The empirical mode decomposition method and the Hilbert spectrum for non-stationary time series analysis, *Proc. Roy. Soc. London,* ***454A***, 903-995.

Wu, Z., E. K. Schneider, B. P. Kirtman, E. S. Sarachik, N. E. Huang, and C. J. Tucker, 2008: Amplitude-frequency modulated annual cycle: an alternative reference frame for climate anomaly. *Climate Dynamics*. **DOI** 10.1007/s00382-008-0437-z.

Huang, N. E, and Z. Wu., 2008: A review on Hilbert-Huang Transform: the method and its applications on geophysical studies. *Rev. Geophys.*, **46**, RG2006, doi:10.1029/2007RG000228.

Wu, Z., and N. E Huang, 2009: Ensemble Empirical Mode Decomposition: a noise-assisted data analysis method. *Advances in Adaptive Data Analysis*. **1**, 1-41.
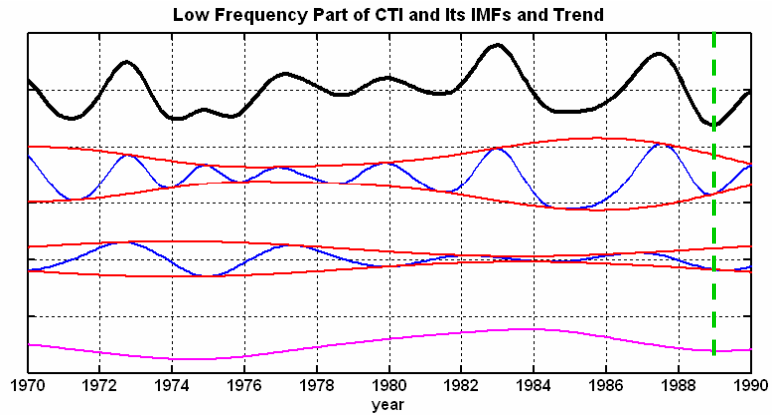
Some Ideas for Ensemble Kalman Filtering

Eugenia Kalnay

Department of Atmospheric and Oceanic Science
University of Maryland, College Park, MD

ABSTRACT

In this seminar we show clean comparisons between EnKF and 4D-Var made in Environment Canada, briefly describe the Local Ensemble Transform Kalman Filter (LETKF) as a representative prototype of Ensemble Kalman Filter, and give several examples of how advanced properties and applications that have been developed and explored for 4D-Var can be adapted to the LETKF without requiring an adjoint model. Although the Ensemble Kalman Filter is less mature than 4D-Var, its simplicity and its competitive performance with respect to 4D-Var suggest that it could become the method of choice.

---

1. Prelude

The WMO/THORPEX Workshop on Intercomparisons of 4D-Var and EnKF that took place in Buenos Aires, Argentina, 10-13 November 2008 was widely attended, with most major operational and research centers throughout the world sending several participants. Presentations are available at http://4dvarenkf.cima.fcen.uba.ar/. Mark Buehner et al., 2008, from Environment Canada, presented a very clean comparison of their operational 4D-Var and EnKF using the same model resolution for the inner loop as in the ensemble, and the same observations (Fig. 1). The results show that the two methods are giving comparable results, with a slight edge favorable to EnKF. In the SH, including a background error covariance based on the EnKF into the 4D-Var improved the 5-day forecast by about 10 hours (not shown).

2. Brief review of the Local Ensemble Transform Kalman Filter algorithm (Hunt et al., 2007)

This description is written as if all the observations are at the analysis time (*i.e.*, for the 3D-LETKF), but the algorithm is the same for the 4D-LETKF (Hunt *et al.*, 2007). In this case the observations are in a time interval that includes the analysis time and *H* is evaluated at the observation time.

a) LETKF **forecast step** (done globally) for each ensemble member *k*:

$$\mathbf{x}_{n,k}^b = M_{t_{n-1},t_n}\left(\mathbf{x}_{n-1,k}^a\right) \quad k = 1,...K$$

b) LETKF **analysis step** (at time *t_n*, so the subscript *n* is dropped):

$$\mathbf{X}^b = \left[\mathbf{x}_1^b - \overline{\mathbf{x}}^b, ..., \mathbf{x}_K^b - \overline{\mathbf{x}}^b\right];$$

$$\mathbf{y}_k^b = H(\mathbf{x}_k^b); \mathbf{Y}^b = \left[\mathbf{y}_1^b - \overline{\mathbf{y}}^b, .., \mathbf{y}_K^b - \overline{\mathbf{y}}^b\right]$$

These computations can also be done locally or globally, which is more efficient. Here the overbar represents the ensemble average, and *M* and *H* are the nonlinear model and observation operators respectively.

Localization: choose for each grid point the observations to be used, and compute the local analysis error covariance and analysis perturbations in ensemble space:

$$\hat{\mathbf{P}}^a = \left[(K-1)\mathbf{I} + \mathbf{Y}^{bT}\mathbf{R}^{-1}\mathbf{Y}^b\right]^{-1}$$

$$\mathbf{W}^a = [(K-1)\hat{\mathbf{P}}^a]^{1/2}$$

The square root required for the matrix of analysis perturbations in ensemble space $\mathbf{W}^a$ is computed using the symmetric square root (Wang *et al.* 2004). This square-root has the advantage of having a zero mean and

---

Correspondence to: Eugenia Kalnay, Department of Atmospheric and Oceanic Science, University of Maryland, College Park, MD; E-mail: ekalnay@atmos.umd.edu

EnKF mean analysis vs. 4D-Var Bnmc (vs raobs)

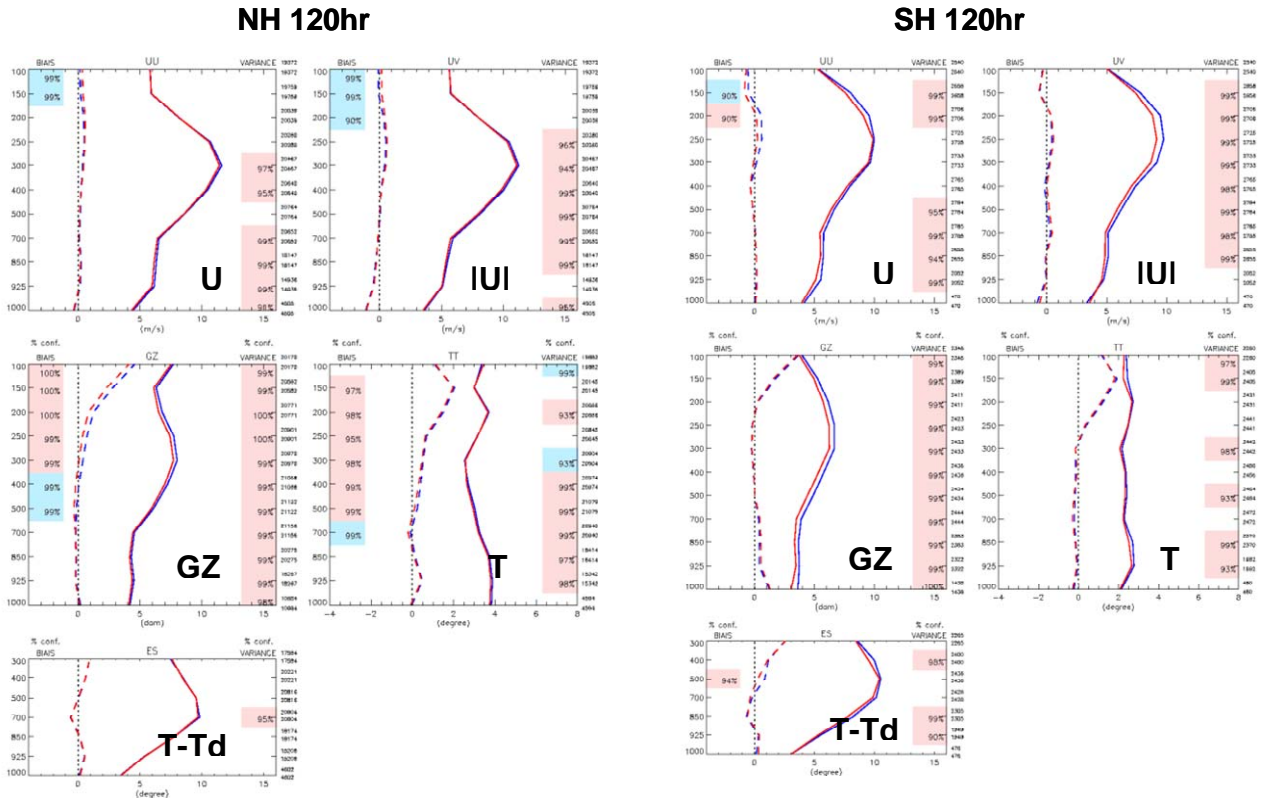### NH 120hr   SH 120hr

Fig. 1. Comparison of bias and standard deviation of 5-day forecasts for February 2007 in the NH and SH verified against rawinsondes for zonal and total wind speed, geopotential height, temperature and dew point depression. Blue and pink colors on the left and right side of each panel indicate the results are better for 4D-Var and for EnKF, respectively, with a level of significance of at least 95%. From Buehner et al., 2008, http://4dvarenkf.cima.fcen.uba.ar/Download/Session_7/Intercomparison_4D-Var_EnKF_Buehner.pdf

being closer to the identity than the square-root matrix obtained by Cholesky decomposition. As a result the analysis perturbations (chosen in different ways in different EnKF schemes) are also closest to the background perturbations (Ott *et al*. 2002). Note that W^α can also be considered a *matrix of weights* since multiplying the forecast ensemble perturbations at each grid point by W^α gives the grid point analysis ensemble perturbations.

Local analysis in ensemble space:

$$\overline{w}^a = \hat{P}^a Y^{bT} R^{-1}(y^o - \overline{y}^b)$$

Note that $\overline{\mathbf{w}}^a$, in the analysis ensemble space, is a vector of weights, which when multiplied by the matrix X^b of forecast perturbations gives the grid point analysis increment.

$$\mathbf{W}^a \leftarrow \mathbf{W}^a \oplus \overline{\mathbf{w}}^a$$

Here the analysis $\overline{\mathbf{w}}^a$ is added to each column of W^α to get the analysis ensemble in ensemble space. The new ensemble analyses are the *K* columns of

$$\mathbf{X}^a = \mathbf{X}^b \mathbf{W}^a + \overline{\mathbf{x}}^b$$

Global analysis ensemble:

The analysis ensemble columns for each grid point are gathered together to form the new global analysis ensemble $\mathbf{x}^a_{n,k}$, and the analysis cycle can proceed.

3. Adaptation of 4D-Var techniques into EnKF

4DVar and EnKF are essentially solving the same problem since they minimize the same cost function using different computational methods. These differences lead to several advantages and disadvantages for each of the two methods (see, for example, Lorenc 2003; Table 7 of Kalnay et al. 2007a; discussion of Gustafsson 2007; response of Kalnay et al. 2007b).

A major difference between 4D-Var and the EnKF is the dimension of the subspace of the analysis increments (analysis minus background). 4D-Var corrects the background forecast in a subspace that has the dimension of the linear tangent and the adjoint models used in the minimization algorithm, and this subspace is generally much larger than the local subspace of corrections in the EnKF determined by the ensemble size K-1. It would be impractical to try to overcome this apparent EnKF disadvantage by using a very large ensemble size. Fortunately, the localization of the error covariances carried out in the EnKF in order to reduce long distance covariance sampling errors, substantially addresses this problem by greatly increasing the number of degrees of freedom available to fit the data. As a result, experience so far has been that the quality of the EnKF analyses with localization increases with the number of ensemble members, but that there is little further improvement when the size of the ensemble is increased beyond about 100. The observation that 50-100 ensemble members are sufficient for the EnKF seems to hold for atmospheric problems ranging from the storm- and meso-scales to the global-scales.

There are a number of additional attractive advantages of 4D-Var. They include the ability to assimilate observations at their right time (Talagrand and Courtier 1987), the fact that within the data assimilation window 4D-Var acts as a smoother (Thépaut and Courtier 1991), ability to use the adjoint model to estimate the impact of observations on the analysis (Cardinali et al. 2004) and on the forecasts (Langland and Baker 2004), the ability to use long assimilation windows (Pires et al. 1996), the computation of outer loops correcting the background state when computing nonlinear observation operators, the ability to use a lower resolution simplified model in the inner loop (see discussion of Fig. 3 later), and the possibility of accounting for model errors by using the model as a weak constraint (Trémolet 2007). In the rest of this section we discuss how these advantages that have been developed for 4D-Var systems can also be adapted and used in the LETKF, a prototype of EnKF.

Fig. 2. Schematic showing that the 4D-LETKF finds the linear combination of the ensemble forecasts at t_n that best fits the observations *throughout* the assimilation window $t_{n-1} - t_n$. The white circles represent the ensemble of analyses (whose mean is the analysis $\overline{\mathbf{x}}^a$), the full lines represent the ensemble forecasts, the dashed line represents the linear combination of the forecasts whose final state is the analysis, and the grey stars represent the asynchronous observations. The cross at the initial time of the assimilation window t_{n-1} is a *no-cost Kalman smoother*, i.e., an analysis at t_{n-1} improved using the information of "future" observations within the assimilation window by weighting the ensembles at t_{n-1} with the weights obtained at t_n. The *smoothed analysis ensemble* at t_{n-1} (not shown in the schematic) can also be obtained at no cost using the same linear combination of the ensemble forecasts valid at t_n given by $\mathbf{W}^a$. (Adapted from Kalnay *et al.* 2007b).

a) 4D-LETKF and no-cost smoother

As indicated by Figure 2, the same weighted combination of the forecasts with weights given by the vector $\overline{\mathbf{w}}^a$ is valid at any time of the assimilation interval. This provides a smoothed analysis mean that (like in 4D-Var) is more accurate than the original analysis because it uses all the future data available throughout the assimilation window (Kalnay *et al.* 2007b; Yang *et al.* 2008a). It should be noted that, as in 4D-Var, the smoothed analysis at the beginning of the assimilation window is an improvement over the filtered analysis computed using only past data. At the end of the assimilation interval only past data is used so that (like in 4D-Var) the smoother coincides with the
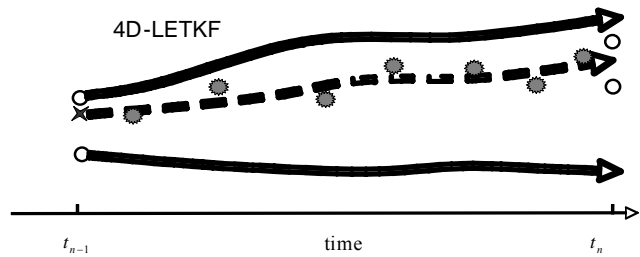
analysis obtained with the filter. Similarly we can use the matrix W^α and apply it to the forecast perturbations $X^b W^\alpha$ to provide an associated uncertainty evolving with time (Ross Hoffman, pers. comm., 2008). The updating of the uncertainty is critical for the "Running in Place" method described next, but the uncertainty is not updated in the "outer loop" approach.

b) Application of the no-cost smoother to the acceleration of the spin-up

4D-Var has been observed to spin up faster than EnKF (e.g. Caya *et al.* 2005), presumably because of its smoothing properties that allow finding the initial conditions at the beginning of the assimilation window that will best fit all the observations. The fact that we can compute a no-cost smoother allows the development of an efficient algorithm, called *Running in place* by Kalnay and Yang (2008), that should be useful in rapidly evolving situations.   For example, when radar measurements first detect the development of a severe storm, then the current EnKF estimate of the atmospheric state and its uncertainty are no longer useful. In other words, *while formally the EnKF members and their average are still the most likely state and best estimate of the uncertainty given all the past data, these EnKF estimates are no longer likely at all*. At the start of severe storm convection, the dynamics of the system change substantially, and the statistics of the processes become non-stationary. In this case, as in the spin-up case in which there are no previous observations available, the running in place algorithm ignores the rule "use the data and then discard it" and recycles a few times the new observations.

Running in place algorithm:

This algorithm is applied to each assimilation window during the spin-up phase. The LETKF is "cold-started" with any initial ensemble mean and perturbations at t_0. The "running in place" loop at time t_n (initially t_0) is as follows: 1. Integrate the ensemble from t_n to t_{n+1}, perform a standard LETKF analysis and obtain the analysis weights for the interval $[t_n, t_{n+1}]$, saving the mean square observations minus forecast (OMF) computed by the LETKF; 2. Apply the no-cost smoother to obtain the smoothed analysis ensemble at t_n by using these weights; 3. Perturb the smoothed analysis ensemble with small zero-mean random Gaussian perturbations, a method similar to additive inflation. Typically the perturbations have amplitudes equal to a small percentage of the climate variance; 4. Integrate the perturbed smoothed ensemble to t_{n+1}. While the forecast fit to the observations continues to improve according to a criterion such as

$$\frac{OMF^2(iter) - OMF^2(iter+1)}{OMF^2(iter)} > \varepsilon,$$

go to step 2 and perform another iteration. If not, replace t_n with t_{n+1} and go to step 1.

Fig. 3.   Comparison of the spin-up of a quasi-geostrophic model simulated data assimilation when starting from random initial conditions. Observations (simulated radiosondes) are available every 12 hours, and the analysis RMS errors are computed comparing with a nature run. Black line: original LETKF with 40 ensemble members, and no prior statistical information. Blue line: 4D-Var with optimal background error covariance. Red line: LETKF "running in place" with $\varepsilon = 5\%$ and 40 ensemble members. Green line: as the red line but with 20 ensemble members.

Running in place was tested with the LETKF in a quasi-geostrophic, QG, model (Fig. 3, adapted from Kalnay and Yang 2008). When starting from a 3D-Var analysis mean, the LETKF converges quickly (not shown), but from random initial states it takes 120 cycles (60 days) to reach a point in which the ensemble perturbations represent the "errors of the day" (black line in Fig. 3). From then on the ensemble converges quickly, in about 60 more cycles (180 cycles total).
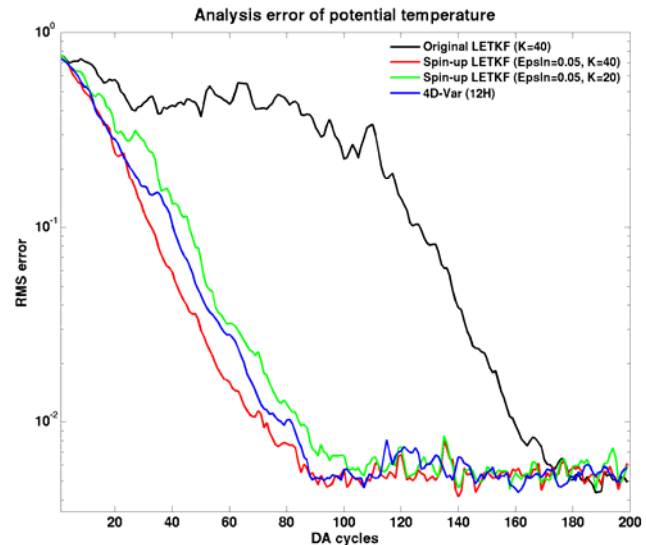
By contrast, the 4D-Var started from the same initial mean state, but using as background error covariance the 3D-Var B scaled down with an optimal factor, converges twice as fast, in about 90 cycles (blue line in Fig. 3). The running in place algorithm with $\varepsilon = 5\%$ (red line) converges about as fast as 4D-Var, and it only takes 2-4 iterations per window even though it does not have the benefit of any prior statistical information.

c) "Outer loop" and dealing with nonlinear ensemble perturbations

A disadvantage of the EnKF is that the Kalman Filter equations used in the analysis assume that the ensemble perturbations are Gaussian, so that when windows are relatively long and perturbations become nonlinear, this assumption breaks down and the EnKF is not optimal. By contrast, 4D-Var is recomputed within an assimilation window until the initial conditions that minimize the cost function for the nonlinear model integration in that window are found. In many operational centres (*e.g.* the National Centers for Environmental Prediction, NCEP, and the European Centre for Medium-range Weather Forecasts, ECMWF) the minimization of the 3D-Var or 4D-Var cost function is done with a linear "inner loop" that improves the initial conditions minimizing a cost function that is quadratic in the perturbations. In the 4D-Var "outer loop" the nonlinear model is integrated from the initial state improved by the inner loop, and the linearized observational increments are recomputed for the next inner loop (Fig. 4).
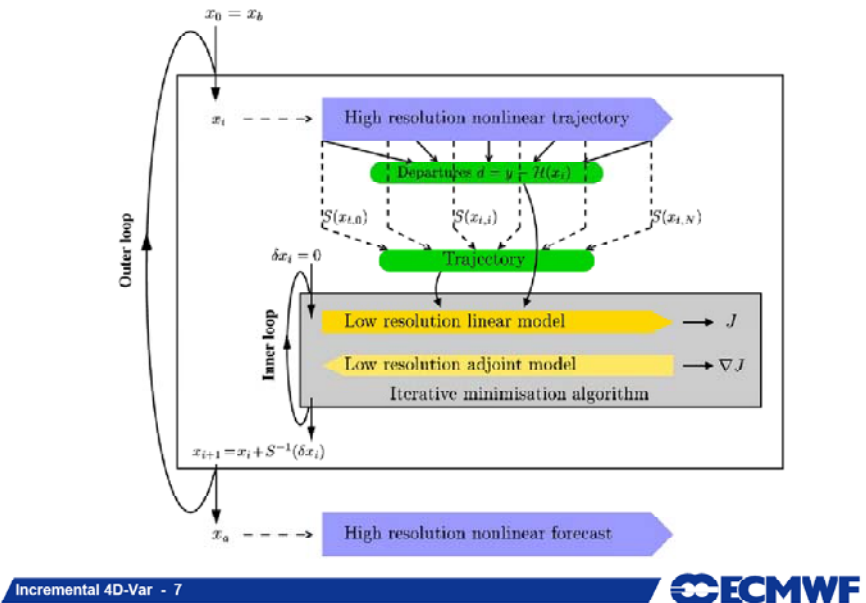

Fig. 4. Schematic of how the 4D-Var cost function is minimized in the ECMWF system. (From Yannick Trémolet, August 2007 class on Incremental 4D-Var at University of Maryland summer Workshop on Applications of Remotely sensed data to Data Assimilation).

The ability of including an outer loop increases significantly the accuracy of both 3D-Var and 4D-Var analyses (Arlindo da Silva, pers. comm., 2006), so that it would be important to develop the ability to carry out an equivalent "outer loop" in the LETKF. This can be done by considering the LETKF analysis for a window as an "inner loop", and, using the no-cost smoother, adapting the 4D-Var outer loop algorithm to the EnKF. The method was tested with the Lorenz (1963) model with short and long windows as in Kalnay et al. 2007a. The results (Table 1) suggest that it should be possible to deal with nonlinearities and obtain results comparable or better than 4D-Var by methods such as an outer loop and running in place.

d) Adjoint forecast sensitivity to observations without adjoint model

Langland and Baker (2004) proposed an adjoint-based procedure to assess the observation impact on short-range forecasts without carrying out data-denial experiments. This adjoint-based procedure can evaluate the impact of any or all observations assimilated in the data assimilation and forecast system on a selected measure of short-range forecast error. In addition, it can be used as a diagnostic tool to monitor the quality of observations, showing which observations make the forecast worse, and can also give an estimate of the relative importance of observations from different sources. Following a similar procedure, Zhu and Gelaro (2008) showed that this adjoint-based method provides accurate assessments of the forecast sensitivity with respect to most of the observations assimilated, and detected that the way certain AIRS

RMSE analysis error	4D-Var	LETKF (inflation factor)	LETKF with less than 3 "outer loop" iterations
Window=8 steps (perturbations are approximately linear)	0.31	0.30 (1.05)	0.27 (1.04)
Window=25 steps (perturbations are nonlinear)	0.53	0.66 (1.28)	0.48 (1.08)

Table 1. Comparison of 4D-Var and LETKF for the Lorenz (1963) 3-variable model. 4D-Var has been simultaneously optimized for the window length (Kalnay et al., 2007a, Pires et al. 1996) and the background error covariance, and the full nonlinear model is used in the minimization. LETKF is performed with 3 ensemble members (no localization is needed for this problem), and inflation is optimized. For the 25 steps case, running in place further reduces the error to a remarkably low value of about 0.39.

humidity channels were used actually made forecasts worse. Unfortunately, this powerful method to estimate observation impact requires the adjoint of the forecast model which is complicated to develop and not always available.

Liu and Kalnay (2008) proposed an ensemble-base sensitivity method able to assess the same forecast sensitivity to observations as in Langland and Baker (2004), but without using the adjoint model.
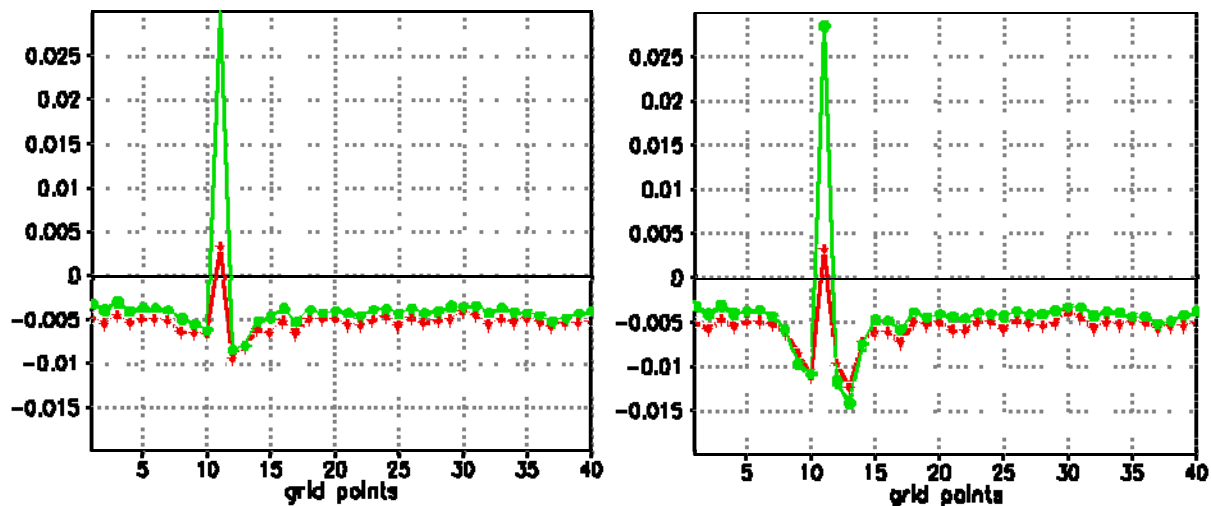
Fig. 5. Time average (over the last 7000 analysis cycles) of the contribution to the reduction of forecast errors from each observation location. Left: the observation at the 11th grid point has 0.8 random errors rather than the specified value of 0.2. Right: the observation at the 11th grid point has random errors as specified but it has a bias of 0.5 rather than 0.0, as specified. Ensemble sensitivity method: green line with closed circles; adjoint method: red line with plus signs. Adapted from Liu and Kalnay (2008).

Figure 5 shows the result of applying this method to the Lorenz (1996) 40-variables model. In this case there were observations at every point every 6 hours created from a "nature" run by adding Gaussian observational errors of mean zero and standard deviation 0.2. At the location 11, however, the standard deviation of the errors was increased to 0.8, (Fig. 5, left-hand panel) without "telling" the data assimilation system about the observation problem in this location. In Fig. 5 (right-hand panel), the standard deviation was kept at its correct value, but a bias of 0.5 was added to the observation at the 11th grid point, still assuming that the bias was zero in the data assimilation. As shown in the figure, both the adjoint and the ensemble-based sensitivity were able to identify that the observations at grid point 11 had a deleterious impact on the forecast. They both show that the neighboring points improved the forecasts more than average by partially correcting the effects of the 11th point observations.

Although the cost function in this example was based on the Eulerian norm, appropriate for a univariate problem, the method can be easily extended to an energy norm, allowing the comparison of the impact of winds and temperature observations (or any other type of observation such as radiances) on the forecasts.

e) Use of a lower resolution analysis

The inner/outer loop used in 4D-Var was introduced in subsection c), where we showed that a similar outer loop can be carried out in EnKF. We now point out that it is common practice to compute the inner loop minimization, shown schematically in Figure 4, using a simplified model (Lorenc, 2003), which usually has lower resolution and simpler physics than the full resolution model used for the nonlinear outer loop integration. The low-resolution analysis correction computed in the inner loop is interpolated back to the full resolution model (Figure 4). The use of lower resolution in the minimization algorithm of the inner loop results in substantial savings in computational cost compared with a full resolution minimization, but it also degrades the analysis.

Yang *et al.* (2008b) took advantage that in the LETKF the analysis ensemble members are a weighted combination of the forecasts, and that the analysis weights W^α are much smoother (they vary on a much larger scale) than the analysis increments or the analysis fields themselves. They tested the idea of interpolating the weights but using the full-resolution forecast model on the same quasi-geostrophic model discussed before. They performed full resolution analyses and compared the results with a computation of the LETKF analysis (*i.e.*, the weight matrix W^α) on coarser grids, every 3 by 3, 5 by 5 and 7 by 7 grid points, corresponding to an analysis grid coverage of 11%, 4% and 2% respectively, and with the interpolation of the analysis increments. They found that interpolating the weights did not degrade the analysis compared with the full resolution, whereas

Fig. 6. The time series of the RMS analysis error in terms of the potential vorticity from different DA experiments. The LETKF analysis from the full-resolution is denoted as the black line and the 3D-Var derived at the same resolution is denoted as the grey line. The LETKF analyses derived from weight-interpolation with different analysis coverage are indicated with blue lines. The LETKF analyses derived after the first 20 days from increment-interpolation with different analysis coverage are indicated with the red lines. Adapted from Yang *et al.* (2008b).

interpolating the analysis increments resulted in a serious degradation (Fig. 6). The use of a symmetric square root in the LETKF ensures that the interpolated analysis has the same linear conservation properties as the full resolution analysis. The results suggest that interpolating the analysis weights computed on a coarse grid without degrading the analysis can substantially reduce the computational cost of the LETKF. Although the full resolution ensemble forecasts are still required, they are also needed for ensemble forecasting in operational centers. We note that the fact that the weights vary on large scales, and that the use of a coarser analyses with weight interpolation actually improves slightly the analysis in data sparse regions, suggesting that smoothing the weights is a good approach to filling data gaps such as those that appear in between satellite orbits. (Yang *et al.* 2008b, Lars Nerger, *pers. comm.* 2008). Smoothing the weights, both in the horizontal and in the vertical may also reduce sampling errors and increase the accuracy of the analyses.
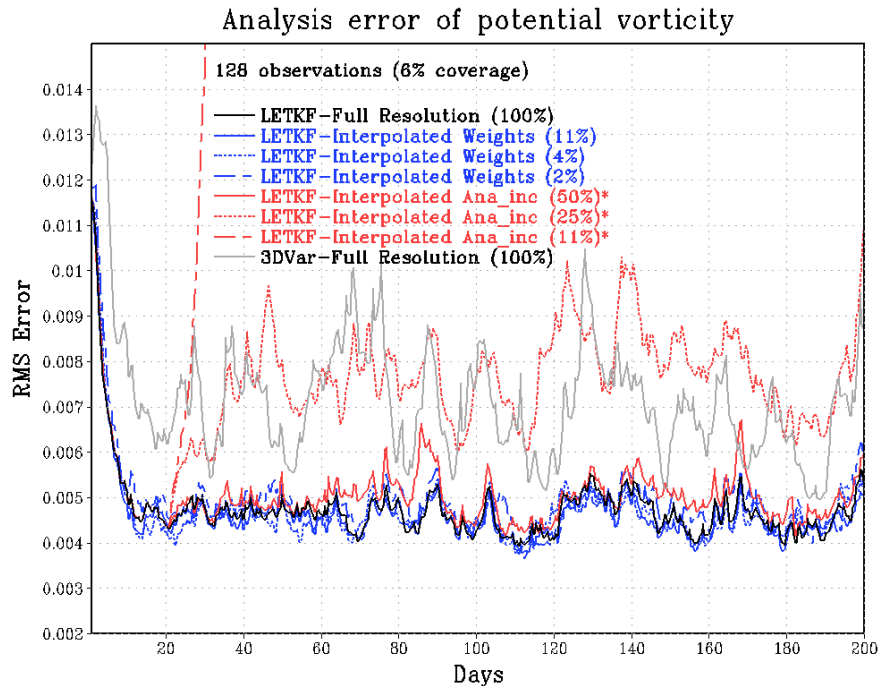
f) Model error

Model error can seriously affect the EnKF because, among other reasons, the presence of model biases cannot be detected by the EnKF original formulation, and the ensemble spread is the same with or without model bias (Li 2007). For this reason, the most widely used method for imperfect models is to increase the multiplicative or additive inflation (*e.g.* Whitaker *et al.* 2007). Model biases can also be taken into account by estimating the bias as in Dee and da Silva (1998) or its simplified approximation (Radakovich *et al.* 2001). More recently, Baek et al. (2007) pointed out that model bias could be estimated accurately augmenting the model state with the bias, and allowing the error covariance to eventually correct the bias. Because the bias was assumed to be a full resolution field, this required doubling the number of ensemble members in order to reach convergence.

In the standard 4D-Var, the impact of model bias cannot be neglected within longer windows because the model (assumed to be perfect) is used as a strong constraint in the minimization (*e.g.* Andersson *et al.* 2005). Trémolet (2007) has developed several techniques allowing for the model to be a weak constraint in order to estimate and correct model errors. Although the results are promising, the methodology for the weak constraint is complex, and still under development.

Li (2007) compared several methods to deal with model bias (Fig. 7), including a "Low-dimensional" method based on an independent estimation of the bias from averages of 6 hour estimated forecast errors started from a reanalysis (or any other available good quality analysis). This method was applied to the SPEEDY (Simplified Parameterizations primitivE-Equation Dynamics) model assimilating simulated observations from the NCEP-NCAR (National Centers for Environmental Prediction-National Center for Atmospheric Research) Reanalysis, and it was found to be able not only to estimate the bias,

Fig. 7. Comparison of the analysis error averaged over two months for the zonal velocity in the SPEEDY model for several simulations with the radiosonde observations available every other point. The yellow line corresponds to a perfect model simulation with the observations extracted from a SPEEDY model "nature run". The red is the control run, in which the observations were extracted from the NCEP-NCAR Reanalysis, but the same multiplicative inflation was used as in the perfect model case. The blue line and the black solid lines correspond to the application of optimized multiplicative and additive inflation respectively. The long-dashed line was obtained correcting the bias with the Dee and DaSilva (1998) method, and combining it additive inflation. The short-dashed is as the long-dashed but using the Danforth *et al.* (2007) low dimensional method to correct the bias, and the green line is as the long-dashed line but using the simplified Dee and DaSilva method. Adapted from Li (2007).

but also the errors in the diurnal cycle and the model forecast errors linearly dependent on the state of the model (Danforth *et al.* 2007; Danforth and Kalnay 2008).

The results obtained by Li (2007) accounting for model errors within the LETKF, presented in Figure 7 indicate that: a) additive inflation is slightly better than multiplicative inflation, b) methods to estimate and correct model bias (e.g. Dee and da Silva 1998; Danforth *et al.* 2007) should be combined with inflation, which is more appropriate in correcting random model errors. The combination of the Low-Dimensional method with additive inflation gave the best results, and was substantially better than the results obtained assuming a perfect model (Fig. 7). We note that the approach of Baek *et al.* (2007) of correcting model bias by augmenting the state vector with the bias can be used to determine other parameters, such as surface fluxes, observational bias, nudging coefficients, etc. It is similar to increasing the control vector in the
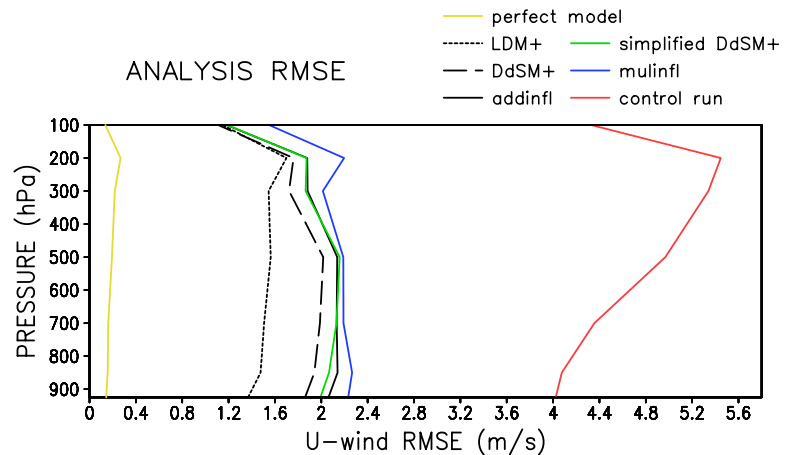
variational approach, and is only limited by the number of degrees of freedom that are added to the control vector and sampling errors in the augmented background error covariance.

4. Summary and discussion

4D-Var and the EnKF are the most advanced methods for data assimilation. 4D-Var has been widely adopted in operational centres, with excellent results and much accumulated experience. EnKF is less mature, and has the disadvantage that the corrections introduced by observations are done in a space of lower resolution, since they depend on the ensemble size, although this problem is ameliorated by the use of localization. The main advantages of the EnKF are that it provides an estimate of the forecast and analysis error covariances, and that it is much simpler to implement than 4D-Var. Recent "clean" comparisons between the operational 4D-Var and EnKF systems in Environment Canada, using the same model resolution and observations, indicated that the forecasts had essentially identical scores, whereas the 4D-Var using a background error covariance based on the EnKF gave a 10-hour improvement in the 5-day forecasts in the Southern Hemisphere (Buehner *et al.* 2008).

It is frequently stated that the best approach should be a hybrid that combines "the best characteristics" of both EnKF and 4D-Var (e.g. Lorenc 2003). Unfortunately this would also bring the main disadvantage of 4D-Var to the hybrid system, *i.e.*, the need to develop and maintain an adjoint model. This makes the hybrid approach attractive to operational centres that already have appropriate linear tangent and adjoint models, but not otherwise.

In this review we have instead focused on the idea that the advantages and new techniques developed over the years for 4D-Var, can be adapted and implemented within the EnKF framework, without requiring an adjoint model. The LETKF (Hunt *et al.* 2007) was used as a prototype of the EnKF. It belongs to the square root or deterministic class of the EnKF (*e.g.* Whitaker and Hamill 2002) but simultaneously assimilates observations locally in space, and uses the ensemble transform approach of Bishop et al. (2001) to obtain the analysis ensemble as a linear combination of the background forecasts. We showed how the LETKF could be modified to include some of the most important 4D-Var advantages: a no-cost smoothing algorithm, useful not only to use "future" (as in reanalysis) but also to accelerate spin-up and handle nonlinear, non-Gaussian ensemble perturbations, and how to implement an "outer loop" within the LETKF. Taking advantage that the LETKF calculates the analysis weights valid throughout the data assimilation window that linearly combine the forecast perturbations to compute the analysis ensemble, we computed the LETKF on coarse grids and interpolated the weights to the full resolution grid. Yang *et al.* (2008) found that the weight interpolation from a coarse resolution grid did not degrade the analysis, suggesting that the weights vary on large scales and smoothing them can increase the accuracy of the analysis, and that weight interpolation is ideal to fill up analysis data voids.

One of the most powerful applications of the adjoint model is the ability to estimate the impact of classes of observations on the short range forecast (Langland and Baker 2004), and we showed how this "adjoint sensitivity" can be computed within the LETKF without an adjoint model (Liu and Kalnay, 2008). Finally, Li (2007) compared several methods used to correct model errors and showed that it is advantageous to combine methods that correct the bias, such as that of Dee and da Silva (1998) and the low-dimensional method of Danforth *et al.* (2007), with methods like inflation that are more appropriate to account for random model errors. This is an alternative to the weak constraint method (Trémolet 2007) to deal with model errors in 4D-Var, and involves the addition of a relatively small number of degrees of freedom to the control vector. Li *et al.* (2008) also developed a method to optimally estimate both the inflation coefficient for the background error covariance and the actual observation error variances (not shown here).

In summary, we have emphasized that the EnKF can profit from the methods and improvements that have been developed in the wide research and operational experience acquired with 4D-Var. Given that operational tests comparing 4D-Var and the LETKF indicate that the performance of these two methods is already very close (*e.g.* Miyoshi and Yamane 2007, Buehner *et al.* 2008), and that the LETKF and other EnKF methods are much simpler to implement, their future looks bright.

References

Bishop, C.H., B.J. Etherton and S.J. Majumdar, 2001. Adaptive Sampling with Ensemble Transform Kalman Filter. Part I: Theoretical Aspects. *Mon. Weather Rev.*, **129**, 420-436.

Buehner, M., C. Charette, Bin He, Peter Houtekamer, Herschel Mitchell, 2008. Intercomparison of 4-D Var and EnKF systems for operational deterministic NWP. Available from
http://4dvarenkf.cima.fcen.uba.ar/Download/Session_7/Intercomparison_4D-Var_EnKF_Buehner.pdf

Burgers, G., P.J. van Leeuwen and G. Evensen, 1998. On the analysis scheme in the Ensemble Kalman Filter. *Mon. Weather Rev.*, **126**, 1719-1724.

Caya, A., J. Sun and C. Snyder, 2005. A comparison between the 4D-VAR and the ensemble Kalman filter techniques for radar data assimilation. *Mon. Weather Rev.*, **133**, 3081-3094

Danforth, C.M., E. Kalnay and T. Miyoshi, 2006 Estimating and Correcting Global Weather Model Error. *Mon. Weather Rev.*, **134**, 281-299.

Danforth, C.M. and E. Kalnay, 2008. Using Singular Value Decomposition to parameterize state dependent model errors. *J. Atmos. Sci.*, **65**, 1467-1478

Dee, D.P. and A.M. da Silva, 1998. Data assimilation in the presence of forecast bias. *Q. J. R. Meteorol. Soc.*, **124**, 269-295.

Evensen, G., 2003. The ensemble Kalman Filter: theoretical formulation and practical implementation. *Ocean Dyn.*, **53**, 343-367.

Fertig, E., J. Harlim and B. Hunt, 2007a. A comparative study of 4D-Var and 4D Ensemble Kalman Filter: Perfect Model Simulations with Lorenz-96. *Tellus*, **59**, 96-101.

Fisher, M., M. Leutbecher and G. Kelly, 2005. On the equivalence between Kalman smoothing and weak-constraint four-dimensional variational data assimilation. *Q. J. R. Meteorol. Soc.*, **131**, 3235-3246.

Gustafsson, N., 2007. Response to the discussion on "4-D-Var or EnKF?". Tellus A, 59, 778-780.

Hamill, T.M., J.S. Whitaker and C. Snyder, 2001. Distance-dependent filtering of background error covariance estimates in an ensemble Kalman filter. *Mon. Weather Rev.*, **129**, 2776–2790.

Houtekamer, P.L. and H.L. Mitchell, 1998. Data Assimilation Using an Ensemble Kalman Filter Technique. *Mon. Weather Rev.*, **126**, 796-811.

Houtekamer, P.L. and H.L. Mitchell. 2001. A Sequential Ensemble Kalman Filter for Atmospheric Data Assimilation. *Mon. Weather Rev.*, **129**, 123–137.

Houtekamer, P.L., H.L. Mitchell, G. Pellerin, M. Buehner, M. Charron, L. Spacek and B. Hansen, 2005. Atmospheric Data Assimilation with an Ensemble Kalman Filter: Results with Real Observations. *Mon. Weather Rev.*, **133**, 604-620.

Houtekamer, P.L. and H.L. Mitchell, 2005. Ensemble Kalman filtering. *Q. J. R. Meteorol. Soc.*, **131**, 3269-3290.

Hunt, B.R., E. Kalnay, E.J. Kostelich, E. Ott, D.J. Patil, T. Sauer, I. Szunyogh, J.A. Yorke and A.V. Zimin, 2004. Four-dimensional ensemble Kalman filtering. *Tellus*, **56A**, 273-277.

Hunt, B.R., E.J. Kostelich and I. Szunyogh, 2007. Efficient data assimilation for spatiotemporal chaos: a Local Ensemble Transform Kalman Filter. *Physica D*, **230**, 112-126.

Ide, K., P. Courtier, M. Ghil and A. Lorenc, 1997. Unified notation for data assimilation: operational, sequential and variational. *J. Meteor. Soc. Japan*, **75**, 181-189.

Kalnay, E., H. Li, T. Miyoshi, S.-C. Yang and J. Ballabrera-Poy, 2007a. 4D-Var or Ensemble Kalman Filter? *Tellus A*, **59**, 758-773.

Kalnay, E., H. Li, T. Miyoshi, S.-C. Yang and J. Ballabrera-Poy, 2007b. Response to the Discussion on "4D-Var or EnKF?" by Nils Gustaffson. *Tellus A*, **59**, 778-780.

Kalnay, E. and S-C. Yang, 2008. Accelerating the spin-up in EnKF. *Arxiv: physics:Nonlinear/0.806.0180v1*.

Langland, R.H. and N.L. Baker, 2004. Estimation of observation impact using the NRL atmospheric variational data assimilation adjoint system. *Tellus*, **56A**, 189-201.

Li, H., 2007. Local ensemble transform Kalman filter with realistic observations. Ph. D. thesis. Available at http://hdl.handle.net/1903/7317

Li, Hong, E Kalnay and T Miyoshi, 2008. Simultaneous estimation of covariance inflation and observation errors within an ensemble Kalman filter. . *Q. J. R. Meteorol. Soc.,* **134**, in press.

Liu, J., and E Kalnay, 2008. Estimating observation impact without adjoint model in an ensemble Kalman Filter. *Q. J. R. Meteorol. Soc.,* **134**, 1327-1335.

Lorenc, A.C., 2003. The potential of the ensemble Kalman filter for NWP – a comparison with 4D-Var. *Q. J. R. Meteorol. Soc.*, **129**, 3183-3203.

Lorenz, E., 1963. Deterministic Non-periodic Flow, *J. Atmos. Sci.*, **20**, 130-141.

Mitchell, H.L., P.L. Houtekamer and G. Pellerin, 2002. Ensemble Size, Balance, and Model-Error Representation in an Ensemble Kalman Filter *Mon. Weather Rev.*, **130**, 2791–2808.

Miyoshi, T., 2005. Ensemble Kalman filter experiments with a Primitive-Equation global model. Doctoral dissertation, University of Maryland, College Park, 197pp. Available at https://drum.umd.edu/dspace/handle/1903/3046.

Miyoshi, T. and S. Yamane, 2007. Local ensemble transform Kalman filtering with an AGCM at a T159/L48 resolution. *Mon. Weather Rev.*, **135**, 3841-3861.

Molteni, F., 2003: Atmospheric simulations using a GCM with simplified physical parameterizations. I: model climatology and variability in multi-decadal experiments. *Clim. Dyn.*, **20**, 175-191.

Pires, C., R. Vautard and O. Talagrand, 1996. On extending the limits of variational assimilation in chaotic systems. *Tellus*, **48A**, 96-121.

Rabier, F., H. Järvinen, E. Klinker, J.-F. Mahfouf and A. Simmons, 2000. The ECMWF operational implementation of four-dimensional variational physics. *Q. J. R. Meteorol. Soc.*, **126**, 1143-1170.

Radakovich, J.D., P.R. Houser, A.M. da Silva and M.G. Bosilovich, 2001. Results from global land-surface data assimilation methods. *Proceeding of the fifth symposium on integrated observing systems,* 14-19 January 2001, Albuquerque, NM. 132-134.

Rotunno, R. and J.W. Bao, 1996. A case study of cyclogenesis using a model hierarchy. *Mon. Weather Rev.*, **124**, 1051-1066.

Szunyogh, I., E.J. Kostelich, G. Gyarmati, D.J. Patil, B.R. Hunt, E. Kalnay, E. Ott, and J.A. Yorke, 2005. Assessing a local ensemble Kalman filter: Perfect model experiments with the NCEP global model. *Tellus*, **57A**, 528-545.

Szunyogh, I., E. Kostelich, G. Gyarmati, E. Kalnay, B.R. Hunt, E. Ott, E. Satterfield and J.A. Yorke, 2008. Assessing a local Ensemble Kalman Filter: Assimilating Real Observations with the NCEP Global Model. *Tellus*, in press.

Talagrand O. and P. Courtier, 1987. Variational assimilation of meteorological observations with the adjoint vorticity equation I: theory. *Q. J. R. Meteorol. Soc.*, **113**, 1311–1328.

Tippett, M.K., J.L. Anderson, C.H. Bishop, T.M. Hamill and J.S. Whitaker, 2003. Ensemble Square Root Filters. *Mon. Weather Rev.*, **131**, 1485-1490.

Trémolet, Y., 2007. Model-error estimation in 4D-Var. *Q. J. R. Meteorol. Soc.*, **133**, 1267–1280.

Whitaker, J.S. and T.M. Hamill, 2002. Ensemble data assimilation without perturbed observations. *Mon. Weather Rev.*, **130**, 1913–1924.

Yang, S-C, M. Corazza, A. Carrassi, E. Kalnay and T. Miyoshi, 2008. Comparison of ensemble-based and variational-based data assimilation schemes in a quasi-geostrophic model. *Mon. Weather Rev*., under revision.

Yang, S-C, E. Kalnay, B. Hunt and N. Bowler, 2008. Weight interpolation for efficient data assimilation with the Local Ensemble Transform Kalman Filter. *Q. J. R. Meteorol. Soc.*, in press.

Zupanski, M., 2005. Maximum likelihood ensemble filter: theoretical aspects. *Mon. Weather Rev.*, **133**, 1710-1726.

Subseasonal Variability of Hurricane Activity

Kathy Pegion[1*], Philip Pegion[2*], Tim DelSole[1], and Mihai Sirbu[1]

[1]*Center for Ocean-Land-Atmosphere Studies, Calverton, Maryland*
[2]*Climate Prediction Center, Camp Springs, Maryland*
[*]*Current Affiliation: NOAA/ESRL/PSD, CIRES Climate Diagnostics Center, Boulder, Colorado*

1. Introduction

Several studies have suggested a relationship between tropical cyclone activity in the eastern Pacific and Atlantic basins and the Madden-Julian Oscillation (MJO; Maloney and Hartmann 2000a, b; Higgins and Shi 2001). This relationship is attributed to the ability of the MJO to impact the atmospheric circulation, making the large-scale environment more or less favorable for tropical cyclone development. In two separate studies, Maloney and Hartmann (2000a, b) find a strong relationship between the MJO and tropical cyclone activity in the eastern Pacific, Gulf of Mexico, and Caribbean. They found that tropical cyclone activity is increased significantly in these regions when the MJO-related convection is in the Indian Ocean and associated westerly low-level wind anomalies are present across the eastern equatorial Pacific. These anomalous winds set up a region of large-scale cyclonic low-level vorticity and reduced wind shear over these regions, two conditions that are favorable for the development of tropical cyclones. The opposite conditions exist when MJO-related convection is located over the western equatorial Pacific and associated easterly anomalies in low-level winds are present over the eastern equatorial Pacific resulting in less favorable conditions for tropical cyclone development.

A study by Higgins and Shi (2001) investigates the variability of Northern Hemisphere warm season precipitation and corresponding large-scale circulation in the CPC merged analysis of precipitation and NCEP/NCAR Reanalysis data. In this study, they find that there is evidence that the MJO modulates tropical cyclone activity by collocating tropical storm and hurricane origination points during the peak months of the Atlantic hurricane season (JAS) with 200 hPa velocity potential composited by MJO phase. This analysis indicates that the strongest tropical cyclones tend to form when the MJO favors enhanced precipitation in that region. This is characterized by large-scale divergence associated with large-scale ascent.

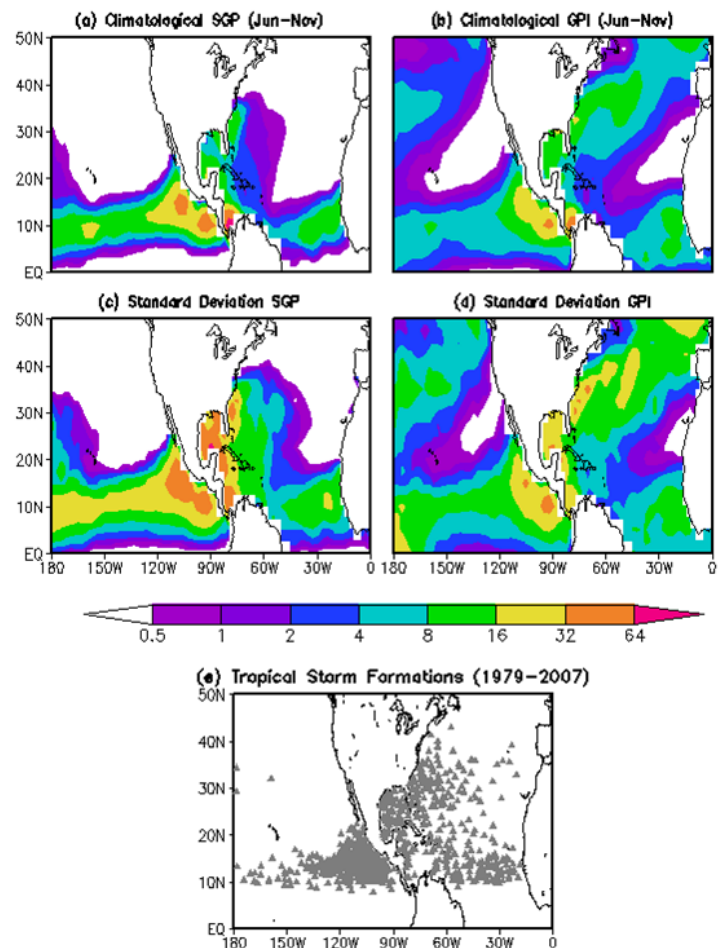
Fig. 1 (a)-(d) NCEP Reanalysis-2 Jun-Nov climatology and standard deviation of SGP & GPI. (e) All tropical storm formations from the Hurricane Best Track Database 1979-2007. by their respective standard deviations, based on 18 years of data, at each grid point.

Correspondence to: Kathy Pegion, NOAA/ESRL/PSD, CIRES Climate Diagnostics Center, Boulder, Colorado
E-mail: Kathy.Pegion@noaa.gov

These studies indicate that the MJO impacts the large-scale circulation, making the environment more or less favorable for the development of tropical cyclones on subseasonal timescales, specifically in terms of the dynamics (*i.e.* low-level vorticity and vertical wind shear). In the current study, we further investigate this relationship by using a commonly used index of MJO activity and indices that quantify whether the environment is favorable or unfavorable for tropical cyclone activity. Furthermore, we quantify this relationship using the technique of mutual information. Additionally, we investigate whether the NCEP operational climate model can capture the observed relationship and offer implications for prediction of tropical cyclone activity on subseasonal timescales.

2. Data and indices

a) MJO index

The all season real-time multivariate MJO (RMM) index of Wheeler and Hendon (2004) is used as a measure of MJO activity. These data were obtained from the Bureau of Meteorological Research Centre website and are available for the dates June 6, 1974 - December 31, 2007. The index is based on combined EOFs of equatorial zonal wind at 200 hPa, zonal wind at 850 hPa, and outgoing longwave radiation. These

Fig. 2   Composite OLR anomalies by MJO phase for observations (left) and CFS (right).

data are limited by the extent of the OLR data from the NOAA AVHRR satellite. We use the version of the index in which the interannual variability related to ENSO has been removed. The first two principal component timeseries (hereafter referred to as RMM1 and RMM2) are in quadrature and together represent a measure of MJO both in terms of phase and amplitude.

b) Hurricane indices

Two indices are used to represent hurricane activity: the seasonal genesis parameter (SGP; Gray 1979) and the genesis potential index (GPI; Emanual and Nolan 2004). These indices were originally designed to represent the large-scale, climatological locations favorable for tropical cyclone development. However, it is reasonable to assume that they are capable of representing large-scale regions where hurricane activity is favorable on other timescales provided the data used contains the necessary variability. Here, we apply the SGP and GPI to subseasonal variability of hurricane activity by calculating both indices using daily data from NCEP/NCAR Reanalysis (R2) for the years 1979-2007.

The SGP is the product of low-level (925hPa) absolute vorticity, inverse of the shear between 925 hPa and 200 hPa, relative humidity between 700 hPa and 500 hPa, moist stability in the layer between 1000 hPa and 500 hPa, and thermal energy in the layer between the surface and 60 meters based on a threshold temperature of 26 degrees C. The climatological values of SGP are calculated for the Atlantic and eastern
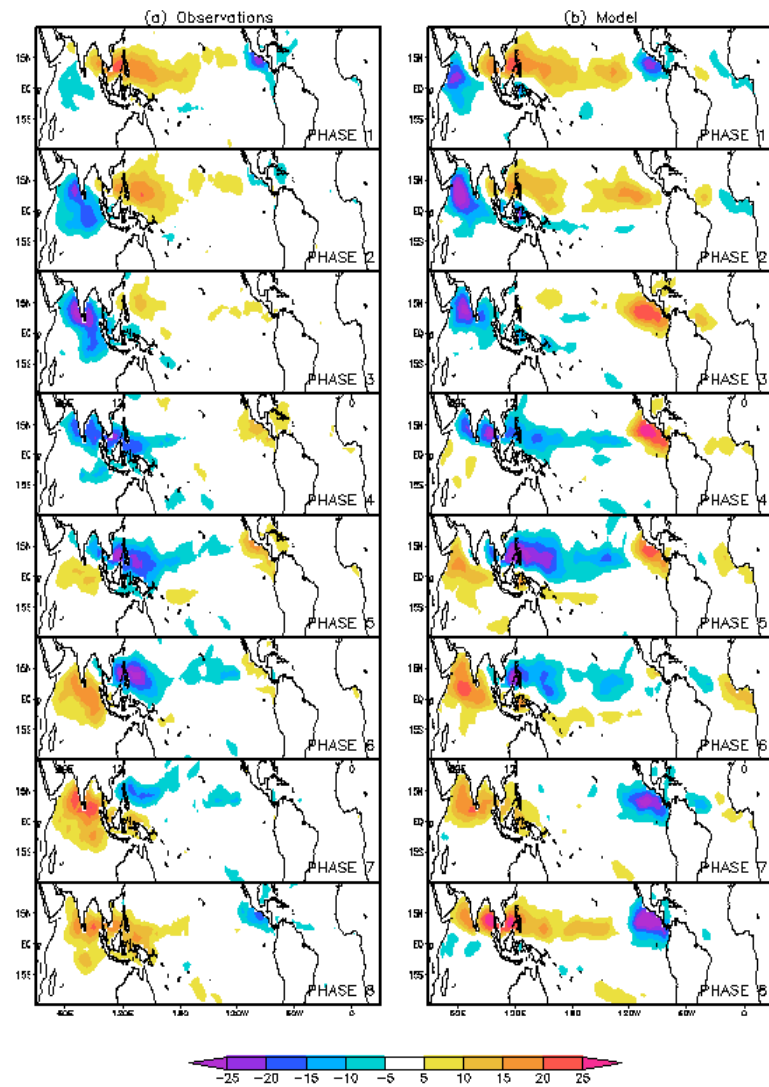
Pacific and averaged over the hurricane season of June through November (Figure 1a). The actual initiation points of tropical storms formed based on the best track hurricane database for the years 1974-2007 are also shown (Figure 1e). The climatological SGP shows the well-known climatologically favored regions for tropical cyclone development. For example, the larger values of SGP indicate that the eastern Pacific along the coast of Central America is the most favored region for tropical cyclones to develop. This is consistent with the actual formation of tropical cyclones seen in Figure 1e. Additionally, the region referred to as the main development region (MDR) in the Atlantic is also apparent and consistent with the tropical storm formation in the eastern Atlantic. It is also apparent that the Caribbean, Gulf of Mexico, southeastern coast of the United States, and Pacific out to the dateline are climatologically favored regions for tropical cyclone development based on both the SGP and the historical tropical cyclone formations. The standard deviation of SGP (Figure 1c) shows that the regions of highest SGP are also the regions with the largest variability. The GPI is the product of low-level (850 hPa) absolute vorticity, relative humidity at 700 hPa, shear between 200 hPa and 850 hPa, and the potential wind intensity that could be obtained if a tropical system

Fig. 3 Composite SGP (left) and GPI (right) by MJO phase calculated from NCEP Reanalysis. Locations of tropical storm formations are also shown (green triangles).

developed. The details of the calculation of potential intensity can found in Emanual and Nolan (2004). The potential intensity combines much of the same information contained in the ocean thermal energy and moisture stability terms in the SGP, but avoids the need to use a fixed threshold value for temperature. The climatological values of GPI and its standard deviation are calculated and averaged over the hurricane season (Figure 1b,d). The climatologically favored regions for tropical cyclone development are consistent with those calculated by SGP and by historical development including the dominance of the eastern Pacific, the MDR in the Atlantic, the Gulf of Mexico and east coast of the United States. The standard deviation again shows that the regions with the largest mean also have the largest variability. The most obvious difference between SGP and GPI are the non-zero values of GPI that extend from 30°N to 50°N in both basins. These differences are due primarily to the 26°C threshold in temperature used in the SGP calculation since SSTs are typically colder than 26°C this far north, the SGP is zero in these regions.

3. Observed relationship between the MJO and hurricane activity
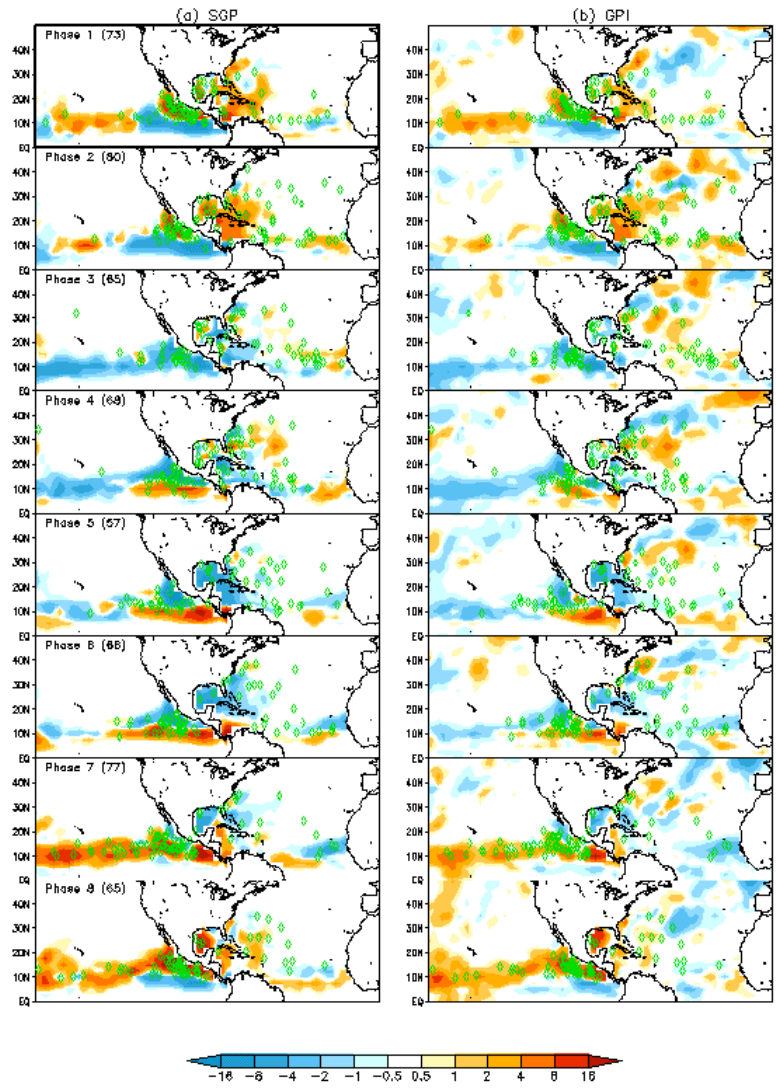
In this section, we investigate this relationship using the SGP and GPI indices calculated from Reanalysis and compare with historical hurricane activity. Specifically, we attempt to quantify the relationship between

the MJO and tropical cyclone activity in terms of the hurricane indices in order to assess some implications for using the MJO to predict periods of increased or decreased probability of tropical cyclone activity.

In order to provide a reference for identifying the phase of the MJO and the relationship with hurricane activity, we first calculate composite MJO events by phase for the observed NOAA/AVHRR OLR during the Atlantic hurricane season (Jun-Nov). First, a five-day running mean is applied to the OLR data to remove short timescale weather variations. Then, the anomalies are determined by removing an average annual and semi-annual cycle. A MJO event is determined to occur when the amplitude of RMM1 and RMM2 is greater than one. The anomalous OLR is averaged for each phase of the MJO to calculate a composite MJO event (Figure 2a). In phase one, enhanced convection indicated by negative OLR anomalies are present over the Western Hemisphere and off the coast of Africa in the westernmost part of the Indian Ocean. The anomalies propagate eastward and increase in amplitude in the Indian Ocean during phases two and three, then across the Maritime Continent during phases four and five. In phases six, the convection moves into the western Pacific Ocean, then north and eastward in phase seven. By phase eight, the convection is located over central America. The corresponding region of suppressed convection is also present and propagates eastward ahead of the region of enhanced convection. It is also important to point out the convective anomalies near the Americas associated with the MJO. For example,

Fig. 4 Mutual information for SGP (top), random SGP mutual information (middle), and 95% significance for SGP mutual information (bottom).

phases 1 and 8 have enhanced convective anomalies in this region while phases 4 and 5 have suppressed convective anomalies in this region.

We now assess the SGP and GPI indices by MJO phase in both the Atlantic and eastern Pacific in order to understand the relationship between the MJO and tropical cyclone activity. SGP and GPI are both calculated using full fields that have had a 5-day running mean applied. Additionally, the linear regression with Nino 3.4 has also been removed in order to remove the impacts of ENSO on the calculations. The anomalous SGP and GPI are determined by removing the average combined annual and semi-annual cycle for the years 1979-2007. The composite anomalies of SGP and GPI are calculated over the eight MJO phases for the months of Jun-Nov (Figure 3). Positive (negative) values indicate that the region is more (less) favorable for development of tropical cyclones. Both the SGP and GPI appear to have a strong signal based on MJO phase in the eastern Pacific and Caribbean and a weaker signal in the eastern Atlantic.

Apparently, conditions are more (less) favorable for development during phases eight, one, and two (three, four, five, and six) in the eastern Pacific and Caribbean. It is also apparent that the relationship between the MJO and the conditions for hurricane activity in the eastern Pacific is consistent with the relationship in the Caribbean and Gulf of Mexico. These results are consistent with those of Maloney and Hartmann (2000a,b). The historical development of tropical storms is also shown (green diamonds). These data are generally consistent with the SGP and GPI regions of more or less favorable conditions based on
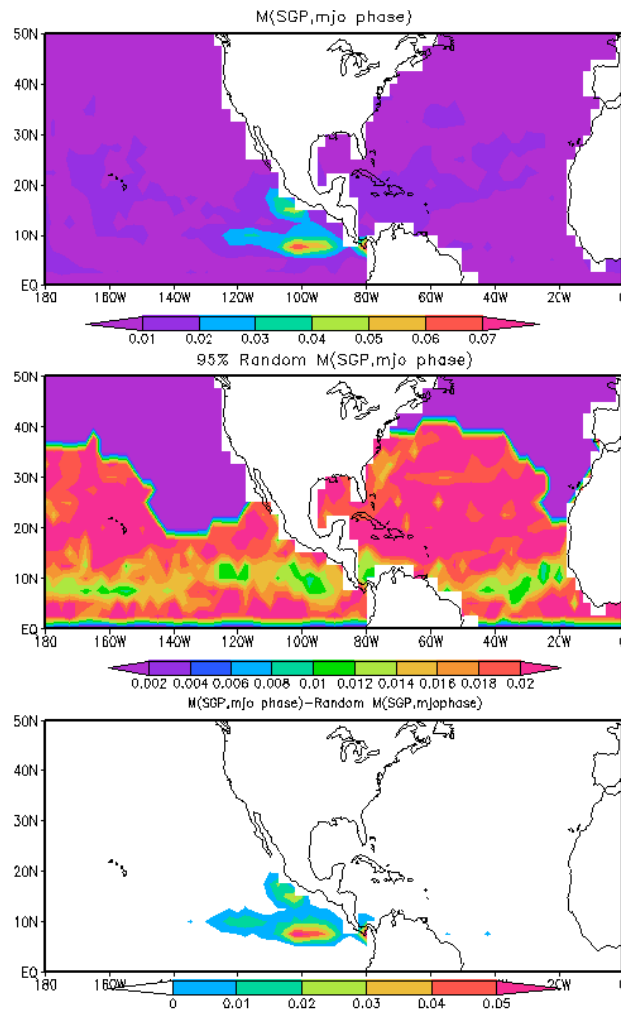
MJO phase with the exception of some regions of negative SGP and GPI anomalies in the eastern Pacific in phases three, four, and five. Perhaps these regions are strongly climatologically favored and the anomalies associated with the MJO have little impact.

Next, we quantify this relationship using the method of mutual information. The mutual information between a hurricane index and the MJO index is calculated as a measure of the null hypothesis that the hurricane index and the MJO index are independent. The formula for the calculation is given as:

$$M = \sum_{HurIndex\,=\,3,\,MJO\,=\,9} P_{HurIndex,\,MJO} \log \frac{P_{HurIndex,\,MJO}}{P_{HurIndex}\,P_{MJO}}.$$

If the hurricane index and the MJO index are independent, then the mutual information between the two will be zero. The RMM MJO index consists of eight MJO phases and a null phase for a total of nine possible phases. The hurricane indices are divided into terciles. Therefore, the summation in the calculation of mutual information occurs over twenty-seven possible cases. The mutual information between the SGP and the MJO index are shown in Figure 4 (top) and for GPI and the MJO index in Figure 5 (top). For the SGP, the largest values for the mutual information are located in the eastern Pacific. This is consistent with the changing values of SGP by MJO phase in the composites in this region. For the GPI, the values are

Fig. 5 Same as Fig 4, but for GPI.

almost constant throughout the region. It is surprising that the mutual information values are so different for the GPI compare to the SGP considering that the composites are very similar for the two. To determine if the value calculated for the mutual information is statistically significant from zero, the mutual information is recalculated with the joint probability between the hurricane index and the MJO index calculated by taking a different, random year for the hurricane index. This calculation is performed one-thousand times and the values are ordered. The 95th percentile for the mutual information is shown in Figure 4 (middle panels). Where the value for the mutual information exceeds the 95th percentile, it is considered to be statistically significant from zero and the null hypothesis can be rejected (Figure 4, bottom panels). Clearly for SGP, the mutual information values are significant in the eastern Pacific. For GPI, significant values are found throughout the region, but there is little spatial structure present.

4. Simulation of the relationship between MJO and hurricane activity in the CFS

The ability of the NCEP operational climate model to simulate the relationship between MJO and tropical cyclone activity is also investigated. Here, we compare composites of SGP and GPI from the CFS with those from the reanalysis products discussed in section 3 and calculate the mutual information between the hurricane indices and the MJO to see if the model captures the observed relationship.

a) Model description

The CFS is the fully coupled atmosphere-ocean general circulation model used operationally by NCEP for climate forecasts. It is composed of the NCEP Global Forecast System 2003 (GFS) as the atmospheric component and the Geophysical Fluid Dynamics Laboratory (GFDL) Modular Ocean Model version 3
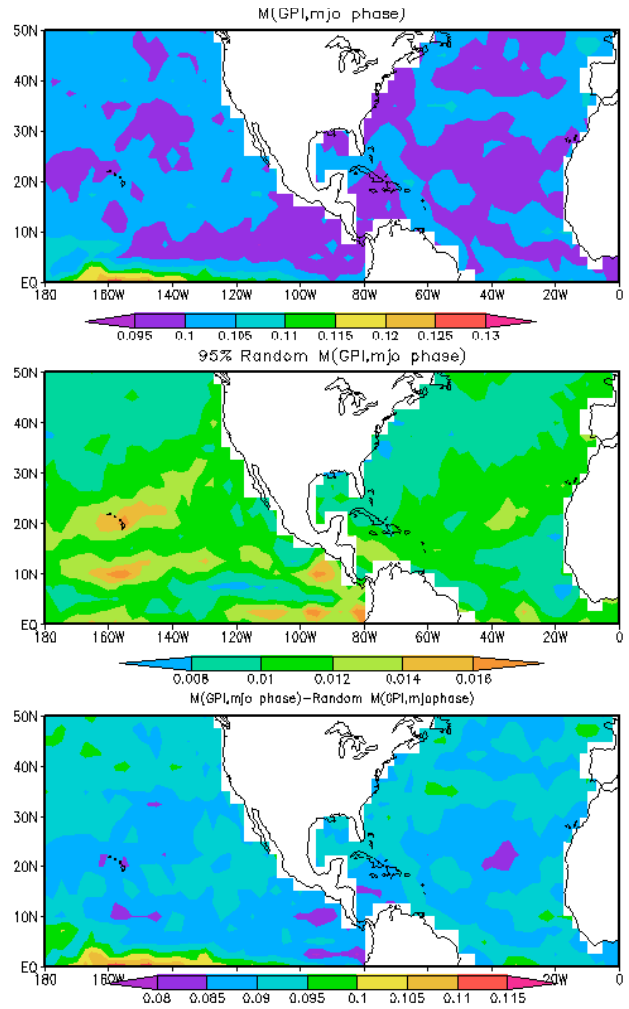
(MOM3; Pacanowski and Griffies 1998) as the oceanic component. The version of the GFS used in this study has a resolution of T62 in the horizontal and 64 layers in the vertical. The ocean model has a quasiglobal domain ranging from 74°S to 64°N latitude. It has 40 layers in the vertical and a resolution of $1/3°x1°$ in the tropics and $1°x1°$ in the extratropics. The atmosphere and ocean exchange fluxes and sea surface temperatures once per day without flux correction. The sea ice extent is taken as climatology. For a more detailed description of this version of the CFS, see Saha et al. (2006) and Wang et al. (2005).

Since the model is relatively coarse resolution, there is no expectation that the model could accurately develop hurricanes. Therefore, we use the SGP and GPI indices as indicators of when the model simulates conditions that are favorable for the development of tropical cyclones. The RMM index, SGP and GPI are calculated using daily data from a 52-year, freely coupled simulation of the CFS. The simulation is initialized on Jan 1, 1985, and run for 52 years. The initial conditions for the atmosphere are from the NCEP Reanalysis-2. The ocean is initialized from the Global Ocean Data Assimilation System (GODAS).

Fig. 6 Same as figure 3, but for CFS.

b) Simulation of the MJO by the CFS

The simulation of the MJO by the CFS has been assessed extensively in Pegion and Kirtman 2008. They found that the model is generally able to simulate regions of enhanced and suppressed convective anomalies and the eastward propagation. However, the model has difficulty with the exact locations of the maximum and minimum anomalies, tends to be too strong, and propagates the anomalies slower than observed. Here, we summarize the simulation and emphasize the associated enhanced and suppressed convective anomalies in the eastern Pacific and Atlantic basins.

The RMM MJO index has been calculated for the 52-year CFS simulation following the same method as Wheeler and Hendon (2004). Composite OLR anomalies by the model MJO phase are shown in Figure 2b. Consistent with the results of Pegion and Kirtman 2008, large-scale OLR anomalies propagate eastward from the Indian Ocean and out to the dateline. Strong anomalies are also apparent in the eastern Pacific, off the coast of Mexico and Central America. These anomalies are generally much stronger than those from the observations and the phasing appears to be off somewhat in the CFS. Despite the deficiencies in the simulation, the CFS is able to capture the overall features of the OLR associated with the MJO. Therefore, it is possible that the model can simulate the relationship between the MJO and the hurricane indices.

c) SGP and GPI

The hurricanes indices, SGP and GPI, have been calculated for the CFS simulation using the same method described in section 2. The climatological values of SGP and GPI for the CFS (not shown) have
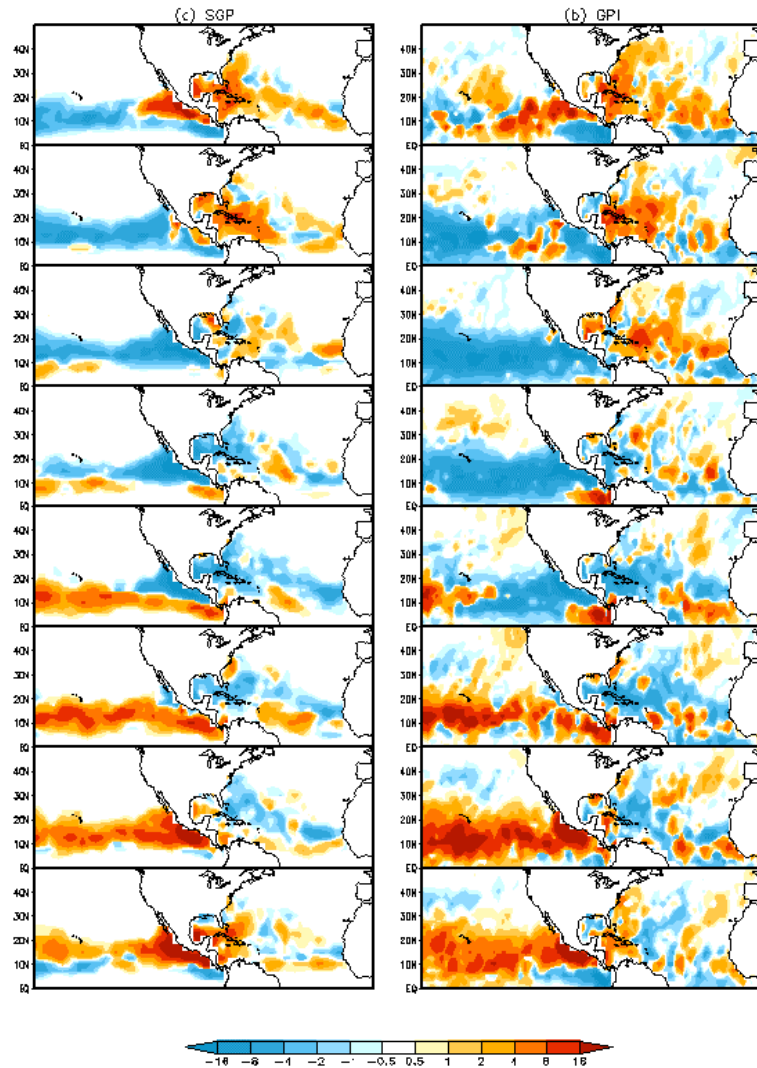
patterns generally consistent with those calculated from R2 and with the historical locations of tropical cyclone formations; however the CFS is generally too strong. The standard deviations (not shown) for both indices are also consistent in pattern, but too strong compared to R2.

Composites of SGP and GPI are also calculated by model MJO phase (Figure 6). The model composites are consistent with the composites calculated using R2 (Figure 3) and ERA40 (not shown). For example, the SGP and GPI indices are positive in the Caribbean and Gulf of Mexico regions during phases eight, one, and two and generally negative in these regions during phases four and five. The SGP and GPI values are generally larger in the CFS than in R2.

Furthermore, we also calculate the mutual information between the hurricane indices and the MJO in the model following the same method described previously. The mutual information between the MJO index and the hurricane indices along with the 95th percentile calculations of the mutual information are shown in Figures 7 and 8. The mutual information between SGP and MJO index and GPI and MJO index are significantly different from zero in the eastern Pacific region. This indicates that the model is able to capture the observed relationship, but that a significant relationship is only present in the eastern Pacific. This is somewhat disappointing considering that there also appeared to be a relationship in the Gulf of Mexico region as well. Future work will assess whether these indices should be modified to be more appropriate for subseasonal timescales.

Fig. 7 Same as Fig 4, but for CFS.

5. Summary and implications for subseasonal forecasts of tropical cyclone activity

The relationship between the MJO and tropical cyclone activity is assessed in terms of the ability of the MJO to change the large-scale environment, making it more or less favorable for tropical cyclone development. This relationship is assessed using two hurricane indices that have traditionally been used to for seasonal and climatological tropical cyclone development. Additionally, we have used these indices to assess whether the NCEP operational climate model can capture this relationship. We use the method of mutual information to quantify whether there is a relationship. Our results indicate that there is a statistically significant relationship in the eastern Pacific. Future work will attempt to determine if an index more appropriate for subseasonal variability can be determined. Furthermore, this work only quantifies the relationship between the hurricane indices and MJO. A relationship between the MJO index or the hurricane indices with tropical cyclones has not yet been established and is the subject of future work.

Because of the slowly evolving nature of the MJO, the composites of SGP and GPI may be useful for prediction purposes, specifically in the eastern Pacific where a statistically significant relationship between the hurricane indices and MJO index has been found. If it is known that the MJO is active and what phase it is in, the fact that it propagates eastward slowly allows these composites to be used to make forecasts for the next couple of phases. This kind of forecasting is already being used at the Climate Prediction Center to make forecasts on one to two week timescales as part of the Global Tropical Hazards product. However, since these
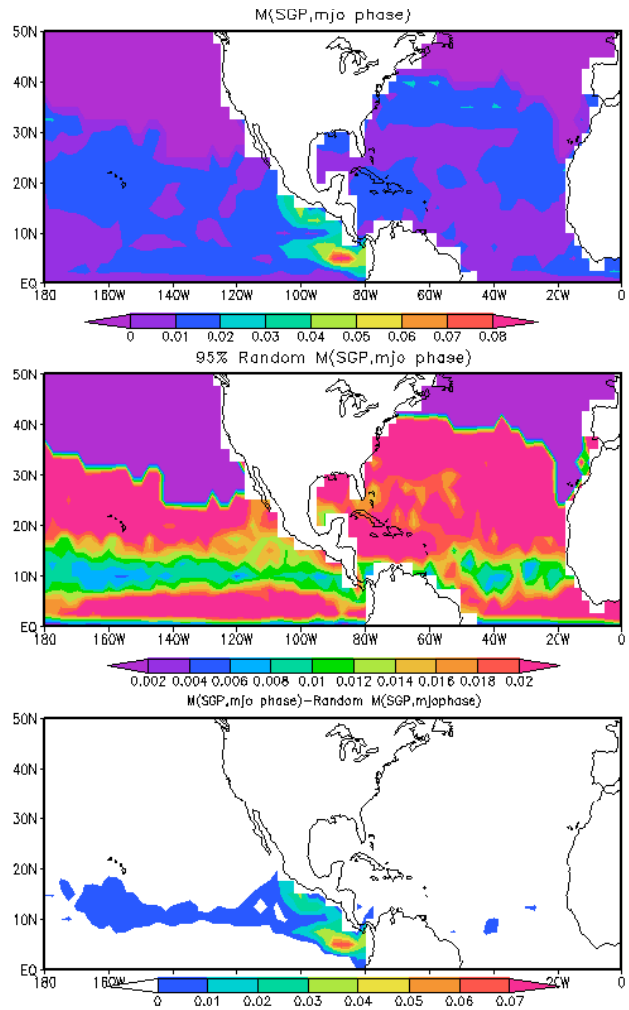
composites only represent some averaged common behavior of the SGP and GPI related to the MJO and do not contain the interaction with other types of variability, it would be ideal to be able to use dynamical models to make actual forecasts of SGP and GPI. This work has established that the CFS is able to simulate the observed relationship between the hurricane indices and the MJO index relationship. Future work will assess its skill in forecasting these indices.

References

Emanuel, Kerry A., and David S. Nolan, 2004: Tropical cyclone activity and the global climate system. Preprints, 26th AMS Conference on Hurricanes and Tropical Meteorology, Miami Beach, May, 2004.

Gray, W.M., 1979: Hurricanes: Their formation, structure and likely role in the general circulation. *Meteorology over the Tropical Oceans.* Royal Meteorological Society, 155−218.

Higgins, R. W., and W. Shi, 2001: Intercomparison of the principal modes of interannual and intraseasonal variability of the North American monsoon system. *J. Climate*, **14**, 403-417.

Maloney, E. D., and D. L. Hartmann, 2000a: Modulation of hurricane activity in the Gulf of Mexico by the Madden-Julian Oscillation. *Science*, **287**, 2002-2004.

Maloney, E. D. and D. L. Hartmann, 2000b: Modulation of eastern North Pacific hurricanes by the Madden-Julian Oscillation. *J. Climate*, **13**, 1451-1460.

Pacanowski, R. C, and S. M. Griffies, 1998: MOM3 manual, NOAA/Geophysical Fluid Dynamics Laboratory, Princeton, USA, 08542.

Pegion, K., and B. P. Kirtman, 2008: The impact of air-sea interactions on the simulation of tropical intraseasonal variability. *J. Climate*, **21**, 616-6635.

Saha, S., S. Nadiga, C. Thiaw, J. Wang, W. Wang, Q. Zhang, D. Behringer, W. Ebisuzaki, S. Lord, S. Moorthi, H. L. Pan, P. Peng, D. Stokes, H. M. van den Dool, G. White, P. Xie, 2006: The NCEP Climate Forecast System. *J. Climate*, **19**, 2483- 3516.

Wang W, Saha S, Pan HL, Nadiga S, White G, 2005: Simulation of ENSO in the new NCEP coupled forecast system model (CFS03). *Mon. Wea. Rev.*, **133**, 1574–1593.

Wheeler, M. C., 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.
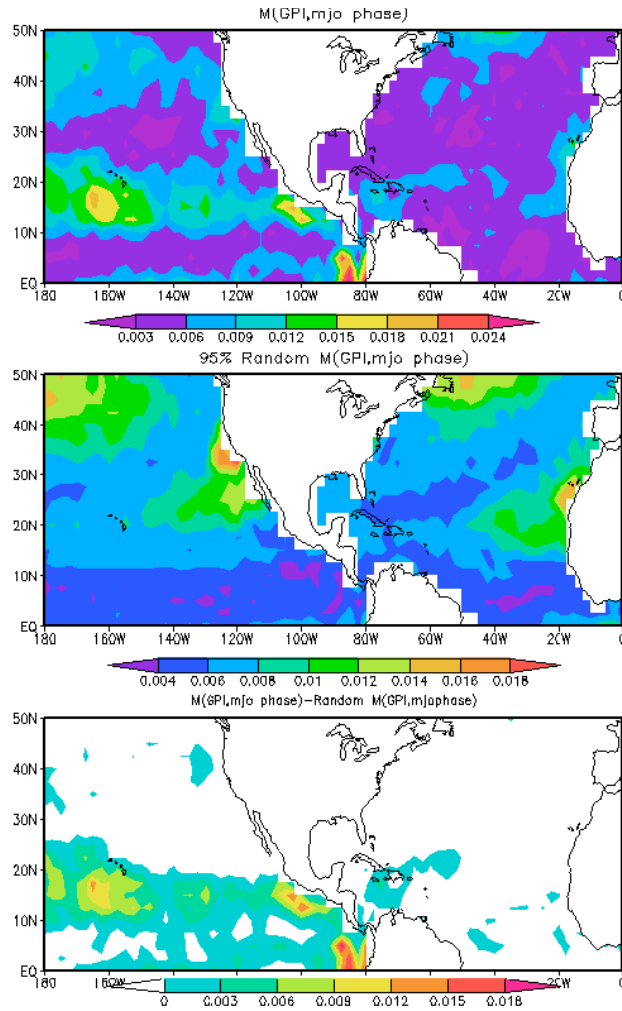
Dynamic Hurricane Season Prediction Experiment with the NCEP CFS CGCM

Jae-Kyung E. Schemm and Lindsey Long

Climate Prediction Center, NOAA/NWS/NCEP, Camp Springs, MD, 20746

A series of experimental forecast runs with the NCEP Climate Forecast System (CFS) coupled GCM was made to examine the feasibility of dynamical hurricane season prediction as one of the Climate Test Bed internal projects in 2007/2008. A series of 7-month forecast experiments with the initial conditions in mid-April during 1981-2008 were made in T382 spectral resolution to evaluate tropical storm statistics in the CFS at the highest possible spatial resolution. This is a preliminary report based on the experimental runs and a set of CFS runs in the T62, T126 and T382 resolutions initialized at 0Z, May 15 for 1981-2008.

Tropical storms in the CFS runs were identified using the tropical storm detection method devised by Carmago and Zebiak (2002). Storms depicted in the CFS have very realistic tracks in all four basins in the Northern Hemisphere (Fig. 1) and a robust seasonal cycle (Fig. 2a). Comparisons of interannual variability in storm activities (Fig. 2b) indicate that the CFS has a reasonable skill and captures the shift to more active storm era in the Atlantic basin during the post-1995 period. Two environmental variables that control interannual variability in storm activity over the Atlantic basin are sea surface temperature (SST) and vertical wind shear over the main development region. The T382
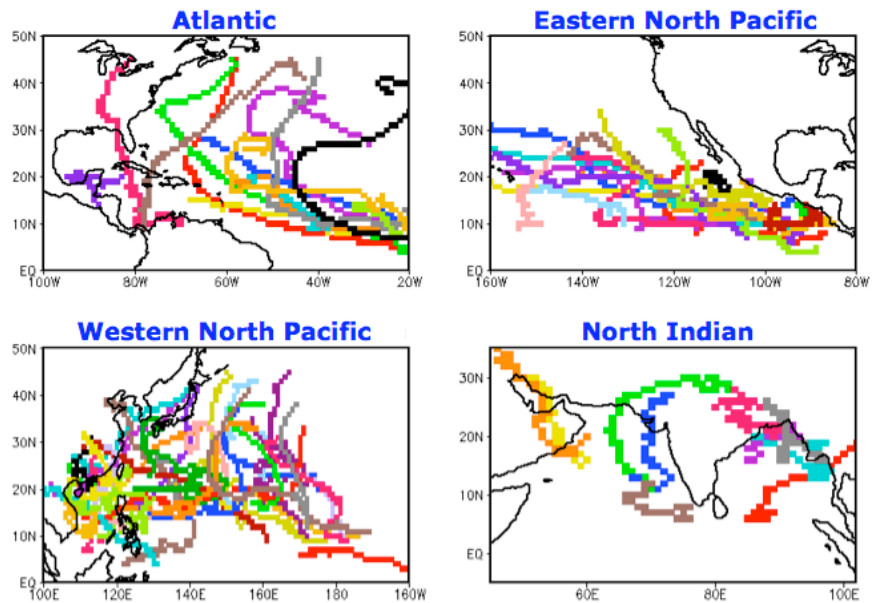
Fig. 1 Examples of tropical storm tracks over the four basins in the Northern Hemisphere depicted in one of CFS hindcast runs. Each color represents an individual storm.
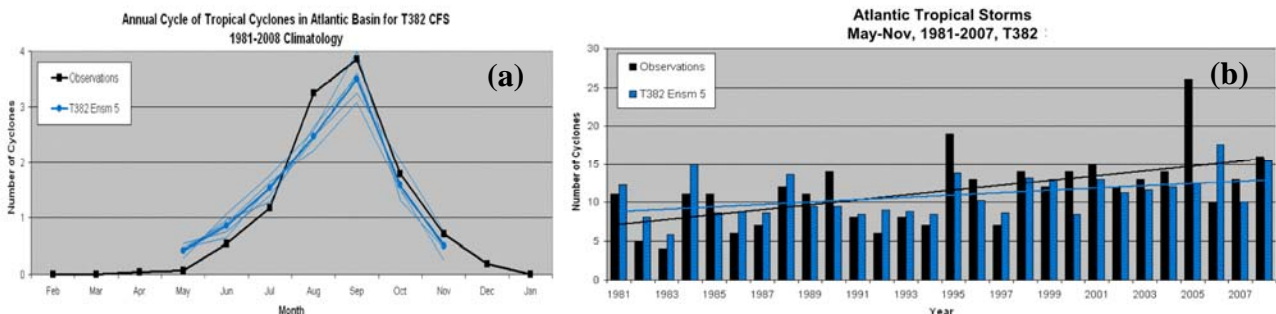
Fig. 2 Climatological seasonal cycle (left panel, a) and interannual variability of tropical storms (right panel, b) for the Atlantic basin. Black lines and black bars correspond to seasonal cycle and number of storms based on observations over the 1981-2008 period, and blue lines and blue bars are for those from CFS hindcast ensemble runs. Straight lines in the right panel represent linear trends over the period.
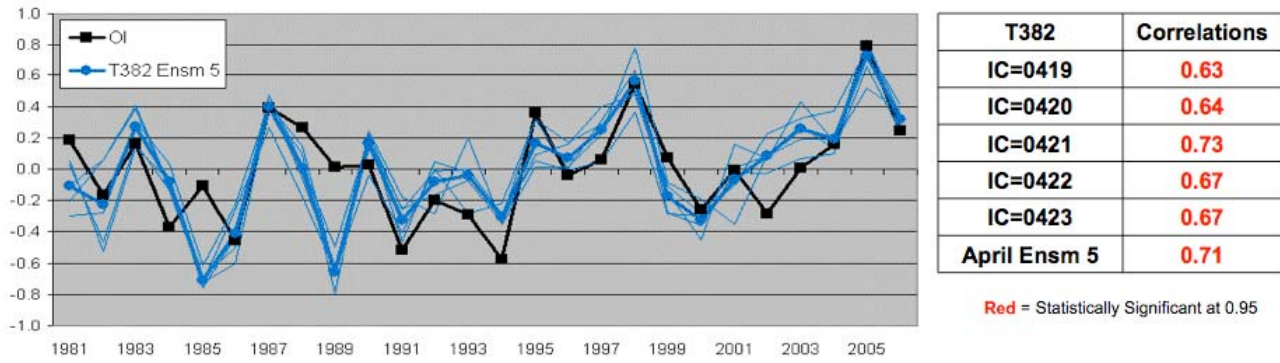
Fig. 3 Interannual variability of JJA SST index over the Atlantic Main Development Region (20ºW-80ºW, 10ºN-20ºN) and anomaly correlation scores of each ensemble member and the ensemble mean. Black line corresponds to the SST index based on the NCEP OI SST analysis and blue lines to the index from CFS hindcast runs.
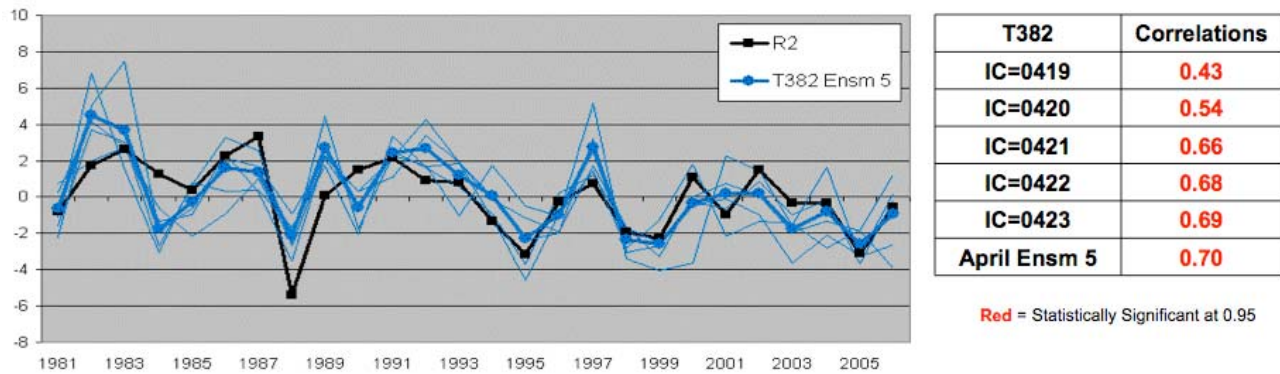
Fig. 4 Interannual variability of JJA wind shear index over the Atlantic main development Region (20ºW-80ºW, 10ºN-20ºN) and anomaly correlation scores of each ensemble member and the ensemble mean. Black line corresponds to the shear index based on the NCEP/DOE Reanalyis-2 and blue lines to the index from CFS hindcast runs.
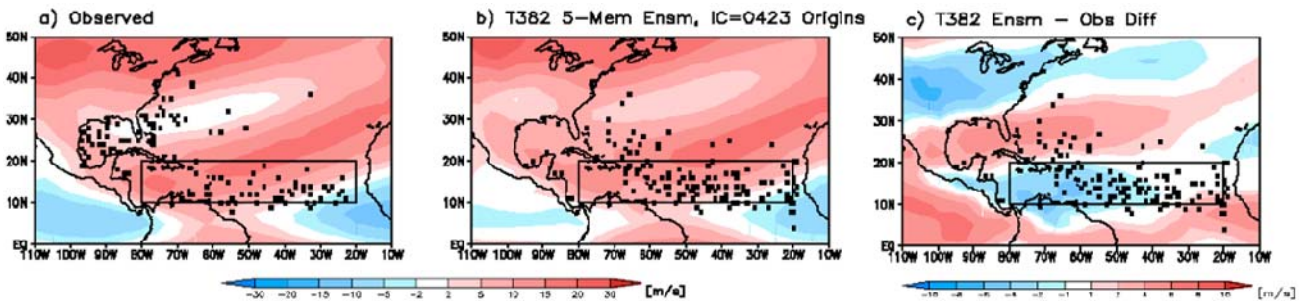
Fig. 5 JJA tropical storm origins (black square) and climatological vertical wind shear (shaded) from observations (a) and CFS hindcast runs (b). Right panel shows wind shear bias and the storm origins in the CFS (c).

CFS has shown a fair level of skill in predicting these environmental conditions over the region (Figs. 3 and 4). A preliminary analysis of bias of storm origins in the CFS hindcasts indicates that over the Atlantic basin, more storms tend to form over the Main Development Region (MDR, 20W-80W, 10N-20N) and that fewer storms form over the Gulf of Mexico and along the east coast of US (Fig. 5). The primary factor controlling storm formation is found to be the vertical wind shear over the basin. Compared to observations, the CFS hindcasts have weaker shear bias over the MDR and stronger bias over the Gulf of Mexico and the Atlantic north of 20N (Fig. 5c).

Companion experiments with the CFS in T62 and T126 resolutions were also conducted to investigate the impact of spatial resolution on storm statistics. The structure and intensity of storms in the lower resolution runs did not compare so well to the observed as in the T382 cases. However, the computing resources needed for routine operation of the T382 CFS are very large in the current computing environment. The T126 resolution might be a good compromise in view of the relatively comparable number of storms generated and the advantage of adopting a multi-member ensemble approach.

Reference

Carmago, S. J. and S. E. Zebiak, 2002: Improving the detection and tracking of tropical cyclones in atmospheric general circulation models. *Weather and Forecasting*, **17**, 1152-1162.

Two Flavors of El Niño and Its Predictability

Emilia K. Jin

Center for Ocean-Land-Atmosphere Studies, Calverton, Maryland
Department of Atmospheric, Ocean and Earth Sciences, George Mason University, Fairfax, Virginia

1. Introduction

The systematic errors of CGCMs have a profound influence on the capability of theses climate models to simulate the fluctuations of the tropical climate. Therefore, the characteristics of systematic errors are a fundamental issue in studies of the limit of predictability of the coupled ocean–atmosphere system. Forecast errors depend on a given model's characteristics, in particular, after the influence of the initial conditions fades out with respect to lead time in a forecast (Jin *et al.* 2008; Jin and Kinter 2009). Focusing on the tropical SST predictability, model errors associated with the El Niño and the Southern Oscillation (ENSO) mechanism may have a strong impact.

Different from the definition of conventional El Niño which is as a phenomenon in the equatorial Pacific Ocean characterized by a positive sea surface temperature departure from normal in the Niño -3.4 region (*i.e.*, 5°S–5°N, 170°–120°W) greater than or equal in magnitude to 0.5°C averaged over three consecutive months (National Oceanic and Atmospheric Administration), there have been several studies to define the different flavors of El Niño (or ENSO) (Trenberth and Stepaniak, 2001; Larkin and Harrison 2005; Ashok *et al.* 2007; Guan and Nigam, 2008; Kao and Yu 2009). Even though there are differences among studies, the distinctive interannual SST variation over the central Pacific which becomes more active in recent years and significantly different global impact from conventional El Niño are common features. Recently, Kug *et al.* (2009) shows that the transition mechanisms and dynamical structure of two-types of El Niño are significantly different.

The main objective of this study is to investigate the predictability of different flavors of ENSO in the state-of-the-art CGCMs. Based on previous definitions, CGCM's ability to predict the distinguishable characteristics of two types of El Niño is investigated using two state-of-the-art CGCMs retrospective forecasts dataset. The ensemble forecasts of the tropical Pacific in 2 CGCMs have been compared with each other and with observations.

2. Data and model

Two retrospective forecast data set of NCEP CFS (Saha *et al.* 2006) and FRCGC/SINTEX-F (Luo *et al.* 2005) are used. A set of retrospective ensemble forecast data set of NCEP CFS was created by running a 9-month integration of 15 members for each of the 12 calendar months in the 27 years from 1981 to 2007. A set of ensemble forecast with 9 members of FRCGC/SINTEX-F was created by running a12-month integration for each of the 12 calendar months in the 26 years from 1982 to 2007. Note that forecast data used here is reconstructed with respect to lead time using all data starting from 12 calendar months to focus on the change of predictability with respect to lead month. The initialization processes of two models are independent. With these retrospective forecasts, a 52-year of NCEP CFS and 200-year of SINTEX-F long run were analyzed to investigate the characteristics of model error.

In this study, SST is mainly used as the variable which represents the coupled system. For comparison with observation, the Optimum Interpolation Sea Surface Temperature (OISST) analyses dataset (Reynolds and Smith 1994) created by the Climate Prediction Center (CPC) of the National Centers for Environmental Prediction (NCEP) is used.

3. Two flavors of El Niño and its predictability

The definition of two types of El Niño is as follows. El Niño events show stronger SST anomalies over the eastern Pacific, and it is elongated to the central Pacific, we will refer these El Niño events to Cold tongue

_____

Correspondence to: Emilia K. Jin, Center for Ocean-Land-Atmosphere Studies, 4041 Powder Mill Road, Suite 302, Calverton, MD 20705; E-mail: kchin@gmu.edu

(CT) El Niño. The SST pattern of CT El Niño is quite similar to that of conventional El Niño (Rasmusson and Carpenter 1982; Harrison and Larkin 1998; Kug et al. 2009). Unlike CT El Niño events, some El Niño events have larger SST anomalies over the central Pacific, while the eastern Pacific SST is small but still positive. Hereafter, we will call these El Niño events as Warm Pool (WP) El Niño events. Both El Niño shows different rainfall patterns which induce different global impact. The CT El Niño is characterized by relatively large SST anomalies in the NIÑO3 region (5ºS–5ºN, 150º–90ºW), while the WP El Niño is associated with SST anomalies mostly confined to the NIÑO4 region (5ºS–5ºN, 160ºE–150ºW). During 1981 to 2006, this is the individual case of three categories. The 82-83 and 1997-98 events are CT El Niño and the 1990-91, 1994-95, 2002-03, and 2004-05 events are WP El Niño. There are three more events, which have features between the CT and WP El Niño events: the 1986–88 and 1991–92 events.

The predictability of individual case of CT and WP El Niño is considered. In the case of forecast lead month 1 (not shown), both models show quite good accordance with observed pattern. Comparing two

Figure 1. Observed and simulated DJF SST anomalies of WP and CT El Niño cases at the forecast lead month 6. (a) CFS WP El Niño, (b) CFS CT El Niño, (c) SINTEX-F WP El Niño and (d) SINTEX-F CT El Niño. Solid line denotes observation and shading denotes model.
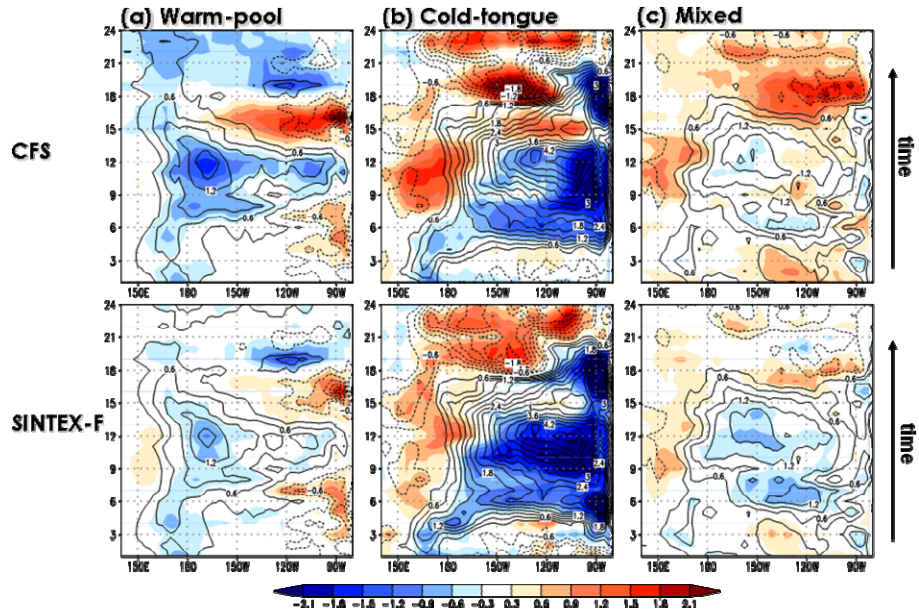
Figure 2. WP, CT and mixed El Niño composite of SST anomalies along the equator at the forecast lead month 7. Solid line denotes observation and shading denotes composite bias of model forecast by subtracting observation from model forecast. (a) CFS and SINTEX-F WP El Niño, (b) CFS and SINTEX-F CT El Niño and (c) CFS and SINTEX-F mixed El Niño.

models, CFS tend to underestimate the magnitude of anomalies more than SINTEX. Figure 1 shows the individual case of CT and WP El Niño at the forecast lead month 6. In this plot, the contour is for observation and the shading is for model forecast. It looks that models reproduce the CT El Niño better than WP El Niño.
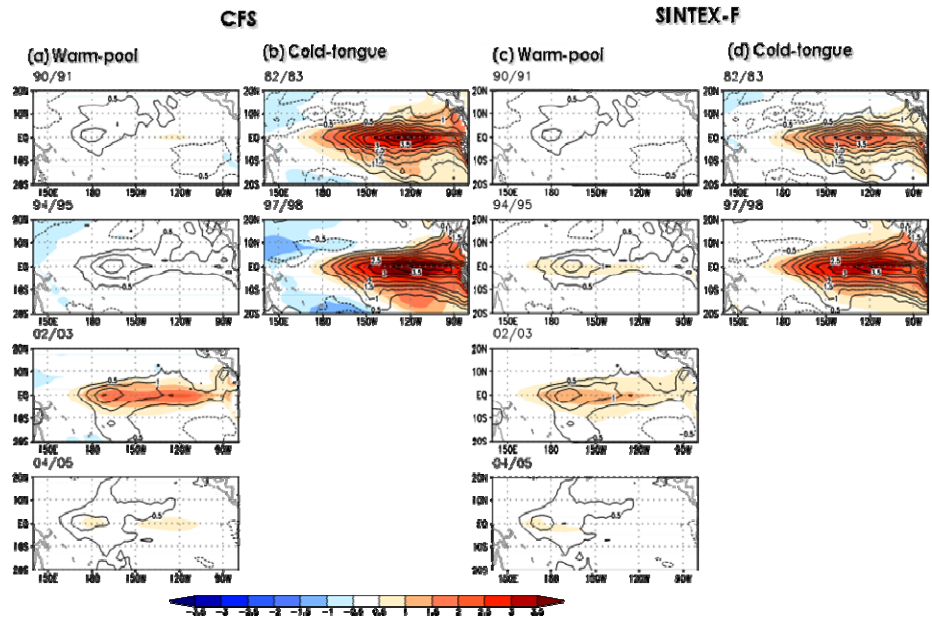
The differences of magnitude of SST anomalies can be one factor because CT El Niño is stronger than WP El Niño in general.

To describe the distinctions in characteristics of CT and WP El Niño events and its predictability, a composite analysis is performed. The ensemble mean of forecast data is reconstructed with respect to lead month and then WP, CT and mixed El Niño events composites are calculated. Figure 2 shows the bias composite along the equator at the forecast lead month 7. The contour denotes observed composite and the shading denotes error composite of model forecast from observation. Observed composites show distinctive centers of action for CT, WP and mixed El Niño, respectively. Interestingly, both models commonly underestimate the SST anomalies over the center of action, where the maximum positive SST anomaly is shown in observation. Negative bias is shown in the warm pool region for WP El Niño and negative bias is shown in the cold tongue region for CT El Niño. As a result, the sign of model error is opposite of that of observed SST anomaly. The errors of mixed case of El Niño are relatively small. Regardless of the independency of dynamics, physics and initialization process of two models, similarity of forecast errors at long forecast lead month in CT and WP El Niño is very intriguing.

Figure 3. The scatter diagram of normalized NIÑO3 and NIÑO4 index. Upper panels show CFS and SINTEX-F at the forecast lead month 1 and lower panels show CFS and SINTEX-F at the forecast lead month 7. Black circle is for observation, red circle is for model and WP and CT cases denotes as cross and x, respectively.

Focusing on the NIÑO indices, the normalized interannual variability of NIÑO3 and NIÑO4 index with respect to lead month is calculated (not shown). Even though the detailed forecast skill has differences in two models, their tendency looks similar. With respect to increase of lead month, models tend to simulate the regular amplitude of two indices and the difference of two indices gets smaller.

Figure 3 is the scatter diagram of normalized NIÑO3 and NIÑO4 index. Black circle is for observation, red is for model forecast, and WP and CT cases are shown as cross and x, respectively. The X axis is NIÑO3 SST anomalies and the y-axis is NIÑO4 SST anomalies. The dashed green line is the indication of linearity between two indices. In observation, it is shown that CT El Niño events are clear outliers from linear relationship and WP El Niño events also show somewhat nonlinear relationship between NIÑO3 and NIÑO4 SST anomalies. This is reasonable considering the definition of CT and WP El Niño. In the case of forecast lead month 1 (upper panels), both models show good accordance with observed relationship as expected. At forecast lead month 7 (lower panels), red circles shows that the nonlinear relationship between two indices almost gets disappeared. Different from observation, their relationship looks very linear in model forecast.

The correlation coefficient between NIÑO3 and NIÑO4 index is 0.69 in observation. At the forecast lead month 1, the correlation coefficient in SINTEX-F indicates 0.73 and CFS indicates 0.77 and it is well
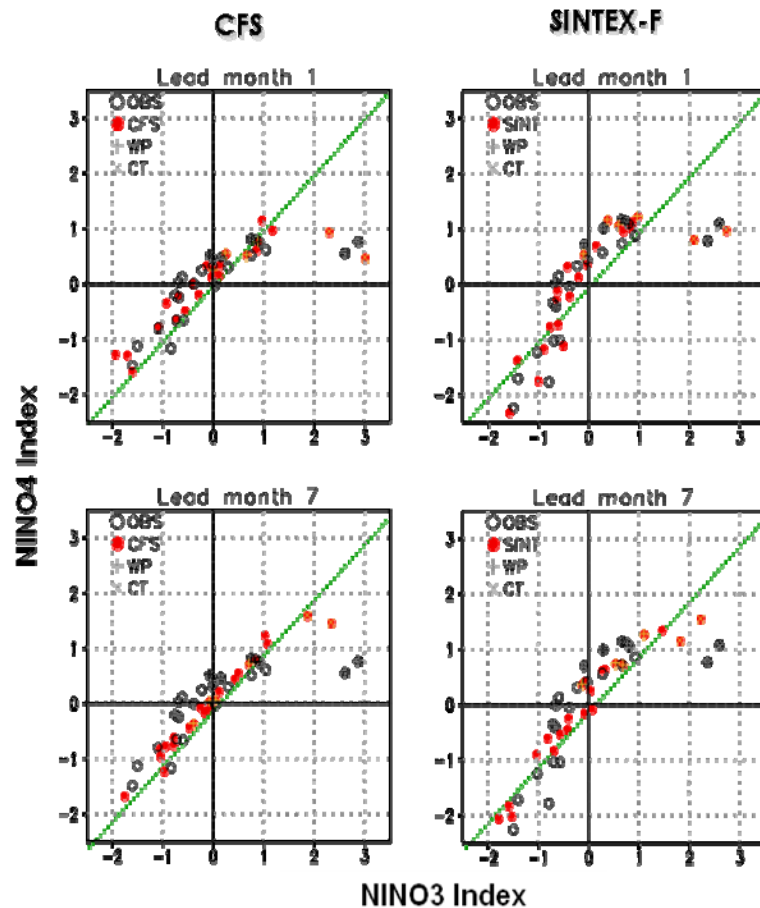
matched with the fact that SINTEX-F shows small forecast error. However, models show increase of relationship with respect to increase of lead month. Models show the increase of relationship to 0.9 and it is associated with drop of forecast skill.

These results indicate that the centers of action of two models in each case of CT and WP El Niño tend to move close to the mixed mode with respect to the increase of lead month. In the previous study, we showed that model's error associated with ENSO dynamics which is different from observation degrades the ENSO forecast skill despite of the advantageous impact of initial condition (Jin and Kinter 2009). Hence, common ENSO forecast errors of two CGCMs at the long lead month can be associated with common errors of ENSO dynamics in CGCMs. To distinguish a given model's problematic features away from the influence of initial conditions, the analysis of ENSO characteristics in long simulations made with the coupled GCMs that are used for operational SST forecasting can be useful. To investigate the relationship between long run behavior and model forecast error, 52-year long run of CFS and 200-year long run of SINTEX-F are used.
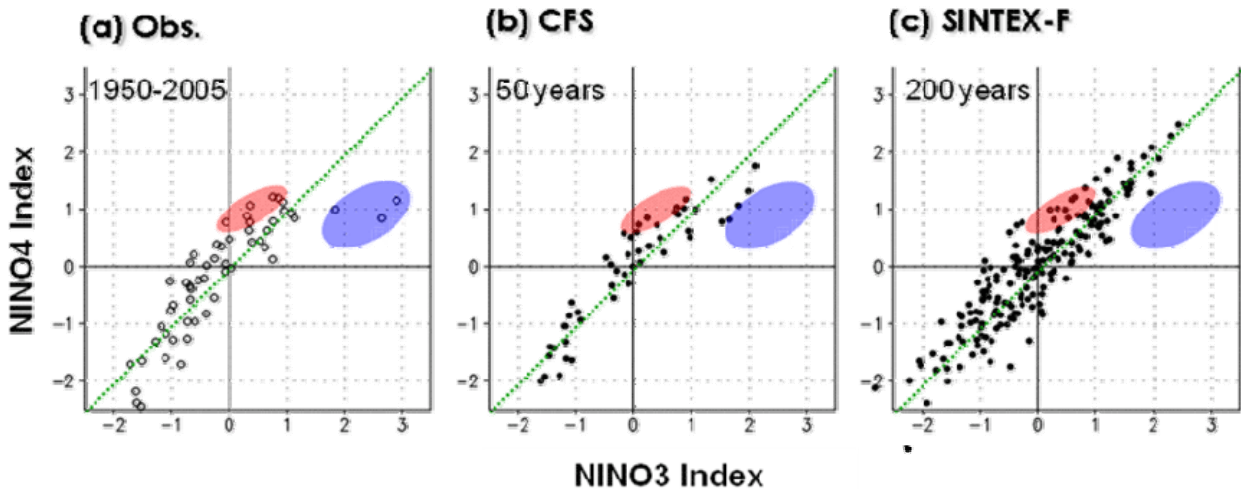
Figure 4. The scatter diagram of normalized NIÑO3 and NIÑO4 index. Right panel shows observation (ERSST and OISST), middle panel shows CFS long run, and left panel shows SINTEX-F long run.

The scatter diagram of normalized NIÑO3 and NIÑO4 index of long run is shown in Figure 4. Blue shading area denotes CT El Niño and red shading area denotes WP El Niño, respectively. In both models, most of El Niño events occurred show linear relationship between NIÑO3 and NIÑO4 index suggesting that most of El Niño events are close to the mixed mode. Overall, the correlation coefficient between NIÑO3 and NIÑO4 index is higher than observation as 0.82 in CFS and 0.86 in SINTEX-F. It suggests that CGCMs have common flaw having monotonic flavor of El Niño and fail to reproduce the complexity in nature. This defect is also associated with the failure of distinguished forecast of different flavors of El Niño, in particular at the long lead month.

4. Concluding remarks

In two state-of-the-art CGCMs, the forecast skill of El Niño is investigated focusing on two flavors of El Niño, which are cold-tongue (CT) and warm-pool (WP). As the lead month of forecast increases, the models fail to distinguish between two flavors of El Niño. Both models have difficulties to reproduce the nonlinear relationship between NIÑO3 and NIÑO4 SST anomalies at the long lead forecast month. This problematic feature is related with the forecast skill of ENSO.

Among several factors to limit the predictability of ENSO in coupled models, model flaw is one of most dominant problems to degrade the forecast skill. From the long run, it is found that models commonly tend to simulate monotonic flavor of El Niño which is close to the mixed mode rather than CT and WP El Niño. This is one of the common errors of two CGCMs associated with drop of ENSO forecast skill at the long lead month of retrospective forecasts. This common flaw in models suggests that the distinctive transition

mechanism associated with the spatial structure of SST and the relative importance of advective and thermocline feedbacks between the two types of El Niño in nature (Kug et al., 2009) is not reproduced in models. Further analysis will be needed to verify this point of view.

The predictability of El Niño on seasonal time scales is important because of the associated global-scale climate anomalies of precipitation and near-surface air temperature (Ropelewski and Halpert 1987; Trenberth et al., 1998; Mason and Goddard 2001). The different pattern of the anomalous convection can lead to difference of the atmospheric circulation, and one may expect distinctive teleconnection of two El Niño events because the tropical precipitation is a key source of the extra-tropical teleconnections. Further research will focus on the tropical precipitation and teleconnection anomalies associated with two flavors of El Niño and its predictability.

References

Ashok, K., S. K. Behera, S. A. Rao, H. Weng, and T. Yamagata, 2007: El Niño Modoki and its possible teleconnection. J.Geophys. Res., 112, C11007, doi:10.1029/2006JC003798.

Harrison, D. E., and N. K. Larkin, 1998: El Niño and the Southern Oscillation sea surface temperature and wind anomalies, 1946-1993. *Rev. Geophys.*, 36(3), 353–399.

Kug, J.-S., F.-F. Jin and S.-I. An, 2009: Two types of El Niño events: cold tongue El Niño and warm pool El Niño. *J. Clim.,* **22**, 1499-1515.

Jin, E. K., and J. L. Kinter III, 2009: Characteristics of tropical Pacific SST predictability in coupled GCM forecasts using the NCEP CFS. *Climate Dynamics,* **32**, 675-691, doi:10.1007/s00382-008-0418-2.

Kao, H.-Y., and J.-Y. Yu, 2009: Contrasting eastern-Pacific and central-Pacific types of ENSO. *J. Clim.,* **22**, 615–632.

Larkin, N. K., and D. E. Harrison, 2005a: On the definition of El Niño and associated seasonal average U.S. weather anomalies. *Geophys. Res. Lett.*, **32**, L13705, doi:10.1029/2005GL022738.

Luo, J. J., S. Masson, S. Behera, S. Shingu and T. Yamagata T, 2005: Seasonal climate predictability in a coupled OAGCM using a different approach for ensemble Forecasts. *J. Clim.,* **18**, 4474–4497.

Mason, S. J., and L. Goddard, 2001: Probabilistic precipitation anomalies associated with ENSO. *Bull. Amer. Meteor. Soc.,* **82,** 619–638.

Rasmusson, E. M., and T. H. Carpenter, 1982: Variations in tropical sea surface temperature and surface wind fields associated with the Southern Oscillation/El Niño. *Mon. Wea. Rev.*, 110, 354–384.

Ropelewski, C. F., and M. S. Halpert, 1987: Global and regional scale precipitation associated with El Niño/Southern Oscillation. *Mon. Wea. Rev.,* **115,** 1606–1626.

Saha, S., S. Nadiga, C. Thiaw, J. Wang, W. Wang, Q. Zhang, H. M. van den Dool, H.-L. Pan, S. Moorthi, D. Behringer, D. Stokes, G. White, S. Lord, W. Ebisuzaki, P. Peng, and P. Xie, 2006 : The NCEP Cliamte Forecast System. *J. Clim.,* **15**, 3483-3517.

Trenberth, K. E., G. W. Branstator, D. Karoly, A. Kumar, N.-C. Lau, and C. Ropelewski, 1998: Progress during TOGA in understanding and modeling global teleconnections associated with tropical sea surface temperature anomalies. *J. Geophys. Res.,* **103,** 14 291–14 324.

Trenberth, K. E. and D. P. Stepaniak, 2001: Indices of El Niño Evolution. *J. Clim.,* **14**, 1697-1701.

Seamless Prediction of Weather and Climate:
A New Paradigm for Modeling and Prediction Research

J. Shukla

George Mason University, Fairfax, Virginia
Institute for Global Environment and Society, Calverton, Maryland

Abstract

This paper describes the concept of seamless prediction and its evolution during the establishment of the COPES (Coordinated Observation and Prediction of the Earth System) framework of the World Climate Research Program (WCRP), where the concept was first presented in 2005.

1. Introduction

There is considerable historical evidence that major scientific and technical discoveries are often followed by the creation of institutions that can take advantage of those discoveries for the betterment of society. The breakthrough in our understanding of atmospheric dynamics that was developed during and after the Second World War, accompanied by the technological breakthrough of fast automatic computing devices, led to the rapid development of numerical weather prediction, a capability that has been institutionalized by many governments around the world and commercialized into a multi-billion dollar enterprise worldwide.

A second example is the development of our scientific understanding of chaos in nonlinear dynamical systems and the potential for predictability at seasonal time scales in the midst of that chaos. The application of that capability for seasonal climate prediction led to the creation of a number of institutions for dynamical seasonal prediction and research.

Now we have before us, thanks to the Intergovernmental Panel on Climate Change (IPCC), a third discovery: humans are affecting the Earth's climate. Beyond IPCC, this discovery will inevitably lead to the establishment of new institutions, or a transformation of current institutions, whose goal will be to help the peoples, governments and corporations of the world manage the consequences of climate change wisely, economically and effectively.

In this paper we present an outline of the emerging paradigm of seamless prediction of weather and climate and present a strategy to revolutionize weather and climate prediction. There is no scientific basis to draw artificial boundaries between meso-scale prediction, synoptic scale prediction, seasonal prediction, ENSO prediction, decadal prediction and climate change. However, practical considerations of computing and of model complexity may require different prediction systems for different time scales. The simulation and prediction of meso-scale systems, synoptic scale disturbances, intra-seasonal, seasonal and inter-annual variations are intimately linked, and therefore, it is suggested that future research on prediction of weather and climate be carried out in a unified framework.

For reliable prediction of regional climate change it is essential that climate models accurately simulate the modes of natural variability from diurnal to seasonal and decadal. Utilization of the insights gained from operational weather and seasonal prediction, and of the synergy between the weather and climate prediction communities is essential for the development of next-generation seamless prediction systems.

One of the scientific implications of the seamless framework is that decadal and multi-decadal prediction using IPCC class models should correctly initialize the state of the ocean-land- atmosphere-cryosphere system. Just as the 1 day NWP forecast is critical in determining the 10 day forecast, it is likely that 10-30 year forecasts will be determined by one season to one year forecast. Institutionally the seamless framework requires the weather and climate communities to work as an integral part of a single scientific enterprise.

2. Weather prediction

It was nearly 100 years a go that V. Bjerknes made a prophetic statement that by using the "knowledge of the state of the atmosphere at the initial time," and "knowledge of the laws according to which one state of the

---

Correspondence to: J. Shukla, Institute for Global Environment and Society, 4041 Powder Mill Road, Suite 302, Calverton, MD 20705; E-mail: shukla@iges.org

atmosphere develops another," if we were able to make accurate forecasts for the future, "meteorology would then have become an exact science, a true physics of the atmosphere." The seminal work of V. Bjerknes (1904) laid the foundation for mathematical modeling of atmospheric dynamics based on the laws of physics. Subsequent works by Richardson, Rossby, Charney and Phillips lead to a successful demonstration of prediction of short-term changes in the large scale atmospheric flow pattern by using dynamical models and heralded the modern era of numerical prediction of weather. Needless to say, that the advances in observational technologies which can accurately describe the state of the atmosphere, data analysis and data assimilation methodologies, and, advances in modeling of the atmosphere have made the dream of V. Bjerknes a reality. The steady advance in the accuracy of numerical weather prediction during the past 30 years is one of the most outstanding examples of scientific progress in the twentieth century. Global society is reaping huge benefits from improved weather forecasts. The number of deaths caused by weather and climate disasters has been reduced. In the past 20 years, the forecast error at day one has been reduced by more than 50%. This has occurred due to steady progress in improving the initial conditions of the atmosphere (which is mainly due to better data quality procedures, better assimilation techniques, and higher resolution models) and improving the accuracy and parameterizations in weather prediction models.

Having already made steady progress during the past 30 years, there is no indication yet that we have reached the limit of weather prediction. After examining the predictability of one of the operational weather prediction models of the European Centre for Medium Range Weather Forecasts (ECMWF), Lorenz (1982) made the following remarks:

> "Better than guesswork forecasts of instantaneous weather patterns nearly two weeks in advance appear to be possible, and efforts to establish numerical-prediction models which are potentially capable of making such forecasts, and observing systems which enable to models to realize their potentialities, should continue."

This statement is as valid today as it was 27 years ago. After publication of Lorenz's paper, it could have been argued that as we continue to increase the resolution of models and resolve fast growing instabilities, the growth rate of initial errors becomes so large that no additional advantage will be gained by increasing the resolution of models. This did not happen at all. The skill of weather forecasts has continued to increase. While it is true that as the one-day forecast error has been reduced by 50%, the growth rate of initial error has increased from the doubling time of about 2 days to 1.5 days, however there still remains unrealized weather predictability for days 1-10.

In a classic paper Lorenz (1969) demonstrated how in a two-dimensional barotropic system an initial error in small scales induces error growth in larger scales. He showed that the predictability limit depended sensitively on the equilibrium energy spectrum. For a -5/3 slope, predictability is intrinsically limited no matter how small the initial error. However, for spectral slope of -3 or greater, increases in predictability were possible with the reduction in the scale and size of the initial error. Experience with operational numerical weather prediction (NWP) during the past 25 years has shown that as we continue to reduce the resolution of NWP models and improve the initial conditions of the atmosphere with better data analysis and assimilation schemes, we continue to improve medium range weather forecasts. Does this mean that we are still in the -3 regime? Or, is it possible that although our current high resolution models are beginning to resolve the upper portions of the -5/3 spectral range, the reason we continue to improve forecasts is because the scaling arguments on which intrinsic limits of predictability were based are not valid for the small scale of the real atmosphere? Recent observations of the atmospheric energy spectrum and numerical experiments with extremely high resolution models suggest that further improvements in medium range weather forecasts with improved high resolution models and improved observations is indeed possible.

A few examples of very high-resolution regional weather prediction models clearly show that local weather phenomena forced by local orography can be resolved and simulated better by higher resolution models. There are also a few examples of global very high-resolution models that show a far more realistic simulation of global weather by eliminating the parameterization of moist convection. There is also well-documented evidence to show that ensemble forecasting has improved the reliability of weather forecasts. In summary, there is overwhelming evidence to suggest that substantial improvements in the accuracy and the reliability of medium range weather forecasts beneficial to society is still possible if we can improve the

initial state by using accurate observations and extremely high resolution weather prediction models that can explicitly resolve deep convective cloud systems. The improvements in the medium-range weather forecasts will directly translate into improvements in seasonal forecasts and vice versa.

3. Dynamical seasonal prediction

Twenty-five years ago, dynamical seasonal prediction was inconceivable. Today, it is a reality. This has become possible because of another major breakthrough in our field with the realization that, although the atmosphere is a chaotic dynamical system, there is predictability in the midst of chaos (Shukla 1998). Certain regions of the global atmosphere are so strongly determined by the underlying boundary conditions (viz sea surface temperature) that they do not show sensitive dependence on initial conditions of the atmosphere, and it should be possible to predict the atmospheric circulation (and rainfall) for as long as the lower boundary forcing can be predicted. Researchers have demonstrated that interactions among the global atmosphere, global oceans and global land surface, can produce fluctuations that are potentially predictable beyond the deterministic limit of predictability for weather. Researchers have further demonstrated that the short-term fluctuations of climate at seasonal and interannual timescales are indeed predictable. Just like the advances in weather prediction, this has also been made possible because of steady improvements in the prediction models which could be integrated for hundreds of days, and could simulate the observed covariability between the slowly varying boundary conditions at the earth's surface and the atmosphere, and, improvements in coupled models which could predict these boundary conditions.

However, operational dynamical seasonal prediction is still in its infancy. Differences in the estimates of seasonal predictability by various "state-of-the-art" models in the world today are quite large, just as large, or even larger than the difference in the estimates of weather predictability 30 years ago. There is no evidence to suggest that the current limitations in the skill of dynamical seasonal forecasts is due to some fundamental predictability limit, on the other hand, there is overwhelming evidence that improvements in models and the initial conditions of ocean, land and atmosphere will improve accuracy of dynamical seasonal predictions.

When several state-of-the-art models were forced with the same SST, they produced very large differences in the tropical convection and the associated seasonal mean precipitation. It is our considered opinion that difference in the parameterizations of moist convection is the main reason for differences in the tropical precipitation and the associated global circulation. Therefore, in order to make better estimate of predictability of seasonal mean circulation and rainfall, it is essential to utilize high resolution models which can resolve and describe organized cloud systems with deep moist convection in the atmosphere, small scale orographic and landscape features overland, and energetic eddies in the oceans.

4. Sub-seasonal prediction

In spite of significant progress in weather forecasting for days 1-10, and some progress in dynamical seasonal prediction for days 1-100 averages, there is no comparable progress in understanding and predicting evolution of weather events during a season. It is well known that there are large week-to-week variations within a season, and even if it is not possible to predict the actual sequence of day-to-day weather, if it were possible to predict statistics of intra-seasonal variability, it would be highly beneficial to society. Examples of Intraseasonal variability which affect seasonal mean as well as produce large societal impacts during a season are given below:

1. Active and break cycles of Asian summer monsoon: the large weekly and monthly variations of regional monsoon rainfall are important for agriculture, energy and water supply planning.
2. Madden-Julian Oscillations: they produce a large tropical as well as extra-tropical variations during a season.
3. Tracks of tropical disturbances: even if individual hurricanes, typhoons and other tropical disturbances cannot be predicted beyond a few days, if it were possible to predict the statistics of these disturbances (changes in tracks, number and intensity), it would be extremely beneficial to society.
4. Extra-tropical storm tracks modulated by slowly varying large-scale flow.
5. Statistics of extreme events within slowly varying large scales flow.

In order to address the question of predictability of the statistics of high impact local weather events it is required that the models used for seasonal prediction resolve and simulate the local weather events. It is therefore essential that the models for predicting daily, weekly, monthly and seasonal variation be unified. In other words, very high-resolution models without parameterizations of deep moist convection must be extended at least up to a season to predict the statistics of high impact weather events within a season. It can be conjectured that since parameterizations of moist convection in coarse-resolution models is the main source of uncertainty in simulation and predicting seasonal means, models that can explicitly simulate organized could systems with moist convection will enhance the predictability of weekly and monthly averages during a season. A recent paper by Miura et al. (2007) supports this conjecture.

5. Decadal prediction and projection of climate change

Thanks to the IPCC, there is vast literature on projections of climate change. It is well known that the most recent assessment of the IPCC (2007) shows a large model dependent range in the projected global warming. We do not know, and there is no simple way to find out, that even for identical external forcing what parts of these differences is due to deficiencies in model physics and model resolution and what part is due to the intrinsic chaotic nature of the coupled climate system. A comparison of the model's ability to simulate the climate of the twentieth century and global warming in the twenty-first century showed (Shukla et al., 2006) that global warming in the twenty-first century was large (4°-5°C) for those climate models that had the least error in the simulation of the twentieth century climate, and models with the largest error in simulating the current climate had small (2°-3°C) warming. Such results, although by no means definitive, strongly suggest the need to improve the fidelity of climate models.

There is also a great deal of scientific and societal interest in quantifying the predictability of decadal variations. Climate models should be able to simulate the regional weather fluctuations, intra-seasonal, seasonal and interannual variations as well as the decadal and multi-decadal variations to be able to be a reliable tool to predict changes in regional climate for 10-30 year time scales. The challenges of making reliable predictions of decadal variations require that the models are able to distinguish between the "unforced" decadal variations due to the internal dynamics of the coupled climate system and the "human forced" external component due to changes in greenhouse gases and land surface conditions. Decadal variations can be caused by changes in the amplitude and frequency of El Niño events, which can be caused by changes in the intra-seasonal (MJO, westerly wind bursts etc) variations which can, in turn, be caused by changes in the organized tropical convection. It is therefore important that climate models can resolve and simulate the meso-scale cloud systems with organized deep convection, intra-seasonal variations, ENSO, as well as decadal meridional oceanic overturning. This does not mean that climate models must correctly predict the phase evolution of each phenomenon for a long time (that will be impossible because of chaos; each phenomena will have its own intrinsic limit of deterministic prediction), but it does mean that the statistics of each phenomenon is correctly simulated by the climate models. If a model cannot simulate a phenomena, it cannot predict that phenomena.

6. Evolution of the concept of seamless prediction in WCRP

It is in the context described above that the author proposed to the WCRP in 2002 (WCRP, 2007) that a World Climate Experiment be launched to assess the predictability of climate system at all time scales. In response, the Joint Scientific Committee of WCRP established a task force (members: B. Hoskins, J. Church, J. Shukla), which, for the first time, introduced the concept of seamless prediction to WCRP. The task force on predictability assessment was later expanded by the WCRP to a much larger task force to develop a strategic framework for the future of WCRP. The newly developed WCRP strategic framework for 2005-2015 was named: Coordinated Observation and Prediction of the Earth System (COPES). The main aim of COPES is to facilitate analysis and prediction of Earth system variability and change for use in an increasing range of practical application of direct relevance, benefit and value to society.

The concept of seamless prediction as articulated by the WCRP-COPES strategic framework is quoted below (WCRP, 2005).

1. *"There is now a new perspective of a continuum of prediction problems, with a blurring of the distinction between shorter-term predictions and longer-term climate projections. Increasingly,*

decadal and century-long climate projection will become an initial-value problem requiring knowledge of the current observed state of the atmosphere, the oceans, cryosphere, and land surface (including soil moisture, vegetation, etc.) in order to produce the best climate projections as well as state-of-the-art decadal and interannual predictions."

2. *"The shorter time-scales and weather are known to be important in influencing the longer-time-scale behavior. In addition, the regional impacts of longer-time scale changes will be felt by society mainly through the resulting changes in the character of the shorter time-scales, including extreme events. In recognition of this, climate models are being run with the highest possible resolution; resolutions there were employed in the best weather forecast models only a few years ago."*

3. *"Even though the prediction problem itself is seamless, the best practical approach to it may be described as unified: models aimed at different time-scales and phenomena may have large communality but place emphasis on different aspects of the system."*

The WCRP strategic framework and the WCRP Modeling Panel adopted the concept of seamless prediction as the organizing principle for the future of WCRP modeling.

Since climate in a region is an ensemble of weather events, understanding and prediction of regional climate variability and climate change, including changes in extreme events, will require a unified initial value approach that encompasses weather, blocking, intra-seasonal oscillations, MJO, PNA, NAO, ENSO PDO, THC, etc. and climate change, in a seamless framework.

The seamless prediction concept implies seamlessness across space and time scales (multi-scale interactions); across scientific disciplines (physical climate system, biogeochemical cycles, socioeconomic systems); across institutions (academic, government, corporations); and across geographical boundaries (local, state, national, international).

Considering that the existing institutions have been established separately for weather prediction and climate research, it will be indeed a challenge to integrate and synthesize the activities of different weather and climate institutions in a seamless framework.

7. A proposal to revolutionize weather and climate prediction

About 150 scientists from the major modeling centers of the world gathered at the World Modelling Summit for Climate Prediction (held in Reading, England [6-9 May, 2008]) organized by the World Climate Research Program to discuss ways to revolutionize weather and climate prediction.

The deliberations at the Summit lead to a statement published in BAMS (Shukla et al., 2009). An important statement of the Summit was the need for a world climate research facility for climate prediction:

"The central component of this world facility will be one or more dedicated high-end computing facilities that will enable climate prediction at the model resolutions and levels of complexity considered essential for the most advanced and reliable representations of the climate system that technology and our scientific understanding of the problem can deliver. This computing capability acceleration, leading to systems at least a thousand times more powerful than the currently available computers, will permit scientists to strive towards kilometer-scale modeling of the global climate system which is crucial to more reliable prediction of the change of convective precipitation especially in the tropics."

In this paper, I take the Summit statement one step further and propose the establishment of an international center for climate prediction consisting of three interconnected advanced computing facilities for climate research and prediction. This international center will help global society for adaptation, mitigation and sustainable development.

8. International center for climate prediction

The world recognizes that humans are contributing to climate change. The impending threat of global change is one of the most important and urgent problems facing humanity. The nations of the world are engaged in serious discussions about the most desirable and practical national and international strategies to

achieve the dual objectives of reducing the emissions of greenhouse gases and ensuring sustainable development.

One important aspect of the problem that is not getting sufficient attention, especially in international negotiations, is the inevitability of climate change due to emissions that have already taken place, and the emissions that will certainly take place in the coming decades. Global society, especially the developing world, will require accurate, reliable and quantitative prediction of inevitable regional climate change for sustainable and resilient development. To adapt to, and to cope with, the dire consequences of inevitable climate change will require investments of trillions of dollars.

Although the current generation of climate models has shown convincingly that human activities can produce large changes in the climate, these models are not adequate to make accurate and reliable predictions of regional climate change and extreme events that are required for science-based adaptation strategies. The major effort of the international scientific community during the past 20 years has been focused on convincing a handful of powerful skeptics that global warming is real, rather than conducting the appropriate research necessary to build sophisticated climate models suitable for providing useful guidance to the policy makers for adaptation and mitigation. Therefore a large number of members of the scientific community, including climate scientists in India, are highly skeptical, and correctly so, about the current projections of regional climate change in India. We do not have confidence in the estimates of projected climate change for different regions of India, for example, identifying which regions will be affected by severe droughts or severe floods, the number and intensity of tropical cyclones, etc.

It is therefore necessary to develop and build accurate and advanced global models for climate prediction so that reliable and quantitative predictions of regional climate change can be provided to global society for science-based adaptation strategies. It is both possible and necessary to have a revolution in climate prediction. It is possible because of the major scientific and technological advances, and it is necessary because adaptation strategies in response to climate change require the most accurate and reliable regional predictions of climate. Sustainable development and the security of the future of life on Earth over the next century demand the best possible information both on the planet's life-support system, in particular the availability of water and food, and possible socio-economic catastrophes due to climate change. The necessary scientific expertise and the computational-technological capabilities to produce reliable regional climate predictions are not available in any single nation.

Regional climate prediction does not mean we should use regional models to predict climate change because regional models require lateral boundary conditions which are currently obtained from low-resolution climate models. It has been pointed out by the WCRP Modeling Panel that, "Use of high resolution regional models to downscale regional climate change is questionable if the global models from which lateral boundary conditions for regional models are prescribed do not have reliable simulation of planetary waves and statistics of storms and blocking." Global climate models at ultra fine resolution will require a huge increase in computational capability. Table 1 gives some estimates of computing capability required for different model grid resolutions and throughput rate (simulated days per wall clock hour). We argue that comprehensive multi-national efforts are essential to provide reliable climate predictions, and assessing their impact, with the level of confidence required by society.

Climate change is a global problem and the solution to the problem of providing governments and society with reliable climate predictions should also be addressed at the global level. Many countries of the world will be making adaptation decisions costing sizable fractions of GDP in the billions (perhaps trillions) of dollars. It is highly desirable that these decisions are based on the most accurate climate predictions made by the most advanced scientific tools and infrastructure possible. It is proposed that an International Center for Climate Prediction be established to provide all the nations of the world the most accurate and reliable description of the current global environment, and prediction of climate change and its impacts based on the best science and the most advanced technology. The international center will foster research and training for building global capacity, developing a trained scientific workforce and engaging the global user community.

There are several examples of successful international institutions. The Consultative Group on International Agricultural Research (CGIAR) is credited with launching the global Green Revolution towards sustainable food security and poverty reduction. The Hubble Telescope, the Human Genome Project,

ECMWF and the high-energy particle accelerator at CERN are outstanding successful examples of internationally funded infrastructures, and the ITER nuclear fusion facility may prove to be another. We assert that the problem of climate prediction is so important and urgent that it needs international collaboration and international infrastructure to provide society accurate and reliable climate information.

Peak Rate:	10 TFLOPS	100 TFLOPS	1 PFLOPS	10 PFLOPS	100 PFLOPS
Cores	1,400 (2006)	12,000 (2008)	80-100,000 (2009)	300-800,000 (2011)	6,000,000? (20xx?)
Global NWP[0]: 5-10 days/hr	18 - 29	8.5 - 14	**4.0** - 6.3	**1.8 - 2.9**	**0.85 - 1.4**
Seasonal[1]: 50-100 days/day	17 - 28	8.0 - 13	**3.7** - 5.9	**1.7 - 2.8**	**0.80 - 1.3**
Decadal[1]: 5-10 yrs/day	57 - 91	27 - 42	12 - 20	5.7 - 9.1	**2.7** - 4.2
Climate Change[2]: 20-50 yrs/day	120 - 200	57 - 91	27 - 42	12 - 20	5.7 - 9.1

Table 1 Computing Capability & Model Grid Size (km). These estimates are based on the performance of the WRF atmospheric model of the grid size (km) that could be used at various computational capability levels for problems ranging from numerical weather prediction (NWP) to climate change projection. A minimum threshold of time-to-solution is assumed in each category. The first column gives the throughput rate (simulated days/years per wall clock hour/day).

Range: Assumed efficiency of 10-40%

[0] - Atmospheric General Circulation Model (AGCM; 100 levels)

[1] - Coupled Ocean-Atmosphere-Land Model (CGCM; ~ 2X AGCM computation with 100-level OGCM)

[2] - Earth System Model (with biogeochemical cycles) (ESM; ~ 2X CGCM computation)

* Core counts above O(104) are unprecedented for weather or climate codes, so the last 3 columns require getting 3 orders of magnitude in scalable parallelization (scalar processors assumed; vector processors would have lower processor counts)

We propose that the international center will consist of at least three international nodes of advanced computing facilities for climate prediction. Considering the probabilistic nature of climate prediction and the need for creative competition among the scientific groups, and the large investments needed to create such facilities, the establishment of three independent but inter-connected facilities was considered to be optimal. Once the three international facilities are in place they should be connected to multiple regional climate prediction and adaptation research centers worldwide which will take the prediction from the international center and add value for local adaptation. The three interconnected computing and prediction facilities will engage scientists from all over the world. In particular, it must guarantee that climate-modeling groups from the developing countries contributed with predictions for future climate change. We propose that the new climate prediction facilities should be truly international and include developing countries. Each of the three international facilities will have a core scientific and technical staff of about 300 persons, and computer capability of about 20 Petaflops in the near future and about 200 Petaflops by the end of the next decade. Each facility will work closely with more than 500 scientists who will have high-level connectivity and access to the facility. The cost of such a single facility will be about $200 million per year; about $100 million per year for computational and data facilities, and about $100 million per year for scientific and technical staff, capacity building and research. The funding for the center can come from both adaptation funds under the

UNFCCC and national contributions. The benefits to society will be enormous; some of the benefits are listed below:

1. To provide reliable quantitative predictions of regional climate change required for developing cost-effective adaptation and mitigation strategies, and to cope with the dire consequences of climate change.

2. All the nations of the world will have access to the most accurate and reliable climate information to plan their respective national adaptation strategies.

3. Build global capacity, develop a trained scientific workforce and engage the global user community.

4. Provide computational capacity to scientists worldwide that is unavailable at the national level. This has the potential to revolutionize climate prediction worldwide.

5. Make fundamental advances in climate modeling and prediction which is currently not possible because of insufficient computational capability and lack of a critical mass of scientific workforce.

References

Bjerknes, V., 1904: The problem of weather forecasting as a problem in mechanics and physics. *Meteorologische Zeitschrift*, **21**, 1-7.

IPCC, 2007: Climate change 2007: The physical science basis. *Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change.*

Lorenz, E.N., 1969: The predictability of a flow which possesses many scales of motion. *Tellus,* **21**, 289-307.

Lorenz, E.N., 1982: Atmospheric predictability experiments with a large numerical model. *Tellus*, **34**, 505-513.

Miura, H., and Coauthors, 2007: A Madden-Julian Oscillation event realistically simulated by a global cloud-resolving model. *Science*, **318**, 1763.

Palmer. T.N. and P.J. Webster, 1995: Towards a unified approach to climate and weather prediction. *Global Change: Proc. of the First Demetra Meeting Held at Chianciano Terme Italy*, European Commission, EUR 15158 EN, 427pp. ISBN 92-826-7757-5.

Shukla, J., 1998: Predictability in the midst of chaos: A scientific basis for climate forecasting. *Science*, **282**, 728-731.

Shukla, J., R. Hagedorn, B. Hoskins, J. Kinter, J. Marotzke, M. Miller, T.N. Palmer, and J. Slingo, 2009: Strategies: Revolution in climate prediction is both necessary and possible: A declaration at the world modeling summit for climate prediction. *Bull. Amer. Meteor. Soc.*, **90**, 175-178.

Shukla, J., T. DelSole, M. Fennessy, J. Kinter and D. Paolino, 2006: Climate model fidelity and projections of climate change. *Geophys. Res. Lett.*, **33**, doi:10.1029-2005GL025579.

WCRP, 2005: The world climate research programme strategic framework 2005-2015. WMO/TD-No. 1291.

WCRP, 2007: Report of the twenty-eighth session of the joint scientific committee, Zanzibar, Tanzania, 26-30 March 2007. WMO/TD-No. 1395, 173-176.

Relationship of U.S. Summer Droughts with SST and Soil Moisture: Distinguishing the Time Scale of Droughts

Renguang Wu[1] and James L. Kinter III[1,2]

[1]*Center for Ocean-Land-Atmosphere Studies, Calverton, Maryland*
[2]*Department of Atmospheric, Ocean and Earth Sciences, George Mason University, Fairfax, Virginia*

1. Introduction

Droughts and floods are extreme climate events, among the most costly natural disasters. Many studies have demonstrated that droughts over U.S. are associated with cold eastern tropical Pacific (*e.g.*, Piechota and Dracup 1996; Trenberth and Guillemot 1996; Ting and Wang 1997; Rajagopalan *et al.* 2000; Hoerling and Kumar 2003; Schubert *et al.* 2004; Seager *et al.* 2005; Mo and Schemm 2008). Studies have also indicated the contribution of SST anomalies in other regions to U.S. droughts, such as the mid-latitude North Pacific and the tropical North Atlantic Ocean. Another important factor for droughts is a positive soil moisture-precipitation feedback. Initial soil moisture anomalies induced by precipitation variations associated with atmospheric circulation changes can positively feed back on the succeeding precipitation, thus prolonging droughts (Findell and Eltahir 1997; Eltahir 1998; Pal and Eltahir 2001).

The present study investigates the relationship of U.S. summer drought with SST and soil moisture. The analysis distinguishes droughts at different time scales based on the Standardized Precipitation Index (SPI). We are interested in determining (1) which region SST has the most significant relationship with the U.S. summer droughts, (2) which region summer droughts are mostly influenced by remote SST forcing, (3) the role of soil moisture in droughts, and (4) how is the long-term change in the relationship between droughts and SST forcing.

2. Datasets

We use both the Palmer Drought Severity Index (PDSI; Palmer 1965) and the Standardized Precipitation Index (SPI; McKee *et al.* 1993) for 344 US Climate Divisions over the period 1895-2007, which are obtained from the National Climatic Data Center. The PDSI is based on a water balance model. The PDSI has several limitations, one of which is that it is not spatially comparable across the contiguous U.S. (Alley 1984; Guttman *et al.* 1992). The SPI is based solely on the probability of precipitation for a given time period. The SPI allows for comparison of precipitation observations at different locations
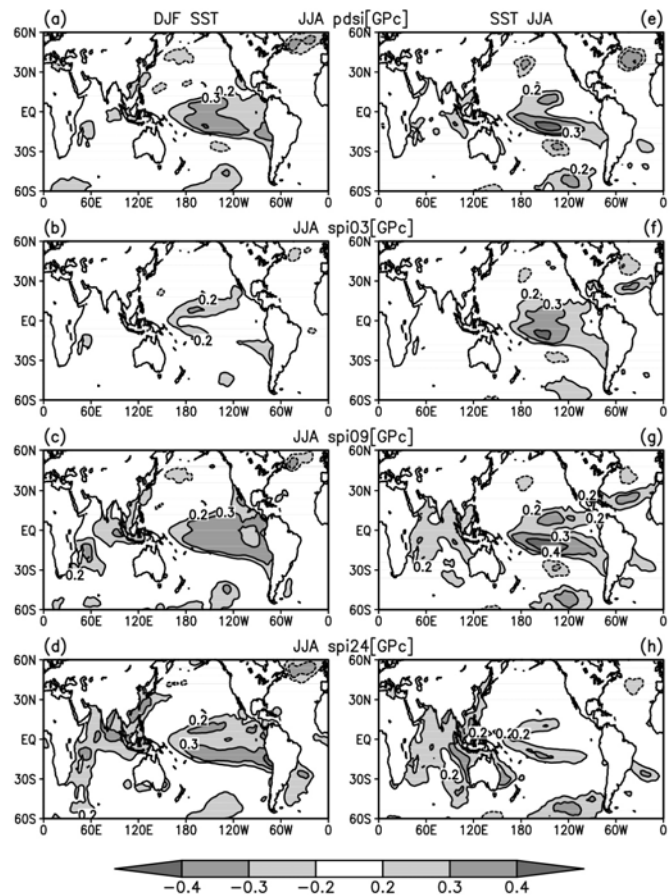
Fig. 1 Correlation of DJF and JJA SST with respect to JJA PDSI and SPI averaged over the Great Plains (30°-50°N, 95°-105°W) during 1897-2005. The contour interval is 0.1. Contours with correlation values smaller than 0.2 are suppressed. The correlation coefficient at the 1% significance level is about 0.25 (or 0.35 for SPI24).

Correspondence to: Renguang Wu, Center for Ocean-Land-Atmosphere Studies, 4041 Powder Mill Road, Suite 302, Calverton, MD 20705; E-mail: renguang@cola.iges.org

with markedly different climates. A key feature of the SPI is the flexibility to measure drought at different time scales, which allows separating and estimating contributions of different time scales to a specific drought event.

The present study uses monthly soil moisture and evaporation for 344 US Climate Divisions over the period 1932-2005 from the Climate Prediction Center, which is estimated with a one-layer hydrological model driven by observed precipitation and temperature (Fan and van den Dool 2004). Observations at 19 soil moisture stations in Illinois covering the period from 1981 to June 2004 are used for validating the soil moisture-drought relationship. The Illinois soil moisture data (Hollinger and Isard 1994) are obtained from the Global Soil Moisture Data Bank (Robock *et al.* 2000).

The present study uses the extended reconstruction of monthly mean SST version 2 (ERSST2) for the period 1854-2005 (Smith and Reynolds 2004). This SST dataset has a spatial resolution of 2°.

3. Global SST correlation with the Great Plains drought indices

Figure 1 shows the correlation of the June-August (JJA) drought indices in the Great Plains region (30°-50°N, 95°-105°W) with global SST in JJA and the preceding December-February (DJF). The SPI03, SPI09, and SPI24 indices are selected to represent short-term, medium-term, and long-term droughts, respectively. The correlation distribution in Fig. 1 broadly resembles the El Niño SST pattern, but with pronounced differences. There is a region of positive correlation over the central-eastern tropical Pacific and the tropical Indian Ocean, and weak negative correlation extends northeastward and southeastward from the tropical western Pacific. The largest positive correlation tends to be located on the north and south sides of the equatorial Pacific, which differs markedly from the ENSO SST pattern.

Overall, the correlation distribution for PDSI is most similar to that for SPI09. The correlation for SPI03 displays notable differences from that for SPI24, SPI09, and PDSI. The short-term droughts have a better correlation with simultaneous SST than the preceding DJF SST, with noticeable correlation limited to the tropical Pacific Ocean and the North Atlantic Ocean. The medium-term droughts have a good correlation with tropical Pacific SST in both JJA and DJF. The long-term droughts have a better correlation with the tropical Pacific SST in DJF than JJA. SPI09 and SPI24 display a positive correlation with SST in the tropical Indian Ocean in both DJF and JJA. In comparison, the PDSI has a weaker correlation in the tropical Indian Ocean. The positive correlation extends to the South China Sea and the East Asian coastal region. All the indices show a negative correlation with DJF SST in the middle-latitude North Atlantic Ocean. In the North Atlantic Ocean, the SPI03 and SPI09 have a similar correlation with JJA SST, with positive correlation in the subtropics and negative correlation in the middle latitudes; whereas the correlation for PDSI and SPI24 is only seen in the middle latitudes. There is a negative correlation in the middle latitudes of the North Pacific for PDSI, SPI03, and SPI09 in JJA and for PDSI and SPI09 in DJF, but the area covered is relatively small.

Figure 2 shows the lead-lag correlation of soil moisture and evaporation with reference to JJA
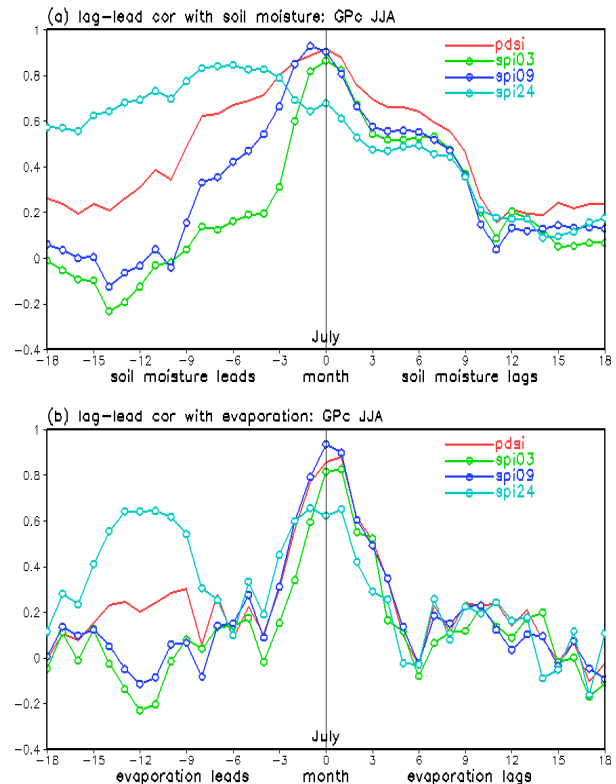

Fig. 2 Lag-lead correlation of soil moisture (a) and evaporation (b) with respect to PDSI and SPI averaged over the Great Plains (30°-50°N, 95°-105°W) during 1932-2005. The correlation coefficient at the 1% significance level is about 0.30 (or 0.40 for SPI24).

drought indices for the Great Plains region. The soil moisture-drought relationship has a strong dependence on the time scale of droughts. The difference is more prominent when the soil moisture leads. For SPI03, the soil moisture shows the highest correlation at lag 0, a quick drop of correlation in the first three lead months. The evaporation correlation is largest at lag 0-1 month. This indicates a sequence of precipitation, soil moisture, and evaporation changes for short-term droughts. For PDSI, the lag-lead correlation is fairly symmetric. For SPI09, the soil moisture has the largest correlation at 1-month lead. The evaporation correlation tends to be symmetric. Soil moisture and evaporation changes apparently lead SPI24. The large correlation in JJA common to different times scales suggests that the contribution of soil moisture to summer droughts may be through surface evaporation although the cause-effect relationship needs to be validated. In comparison, the lag-lead correlation suggests that for short-term droughts, the precipitation impacts are larger than the soil moisture impacts, whereas for medium-term and long-term droughts, the soil moisture impacts are pronounced.

4. U.S. summer drought correlation with the NINO3.4 SST

Figure 3 shows the correlation of JJA drought indices with respect to DJF and JJA NINO3.4 (5°S-5°N, 170°-120°W) SST. There are important differences between short-term and long-term droughts. The SPI03 has a higher correlation with JJA NINO3.4 SST than DJF NINO3.4 SST, whereas the PDSI, SPI09, and SPI24 have a larger correlation with DJF than JJA NINO3.4 SST. Notably, the JJA correlation is large for SPI03 in the Great Plains. The highest DJF correlation for PDSI, SPI09, and SPI24 is found in the Southwest with northward extension into the Great Plains. The magnitude of correlation is close for PDSI and SPI09, while the correlation for SPI24 is smaller.

The ENSO impacts on drought could be through soil moisture and evaporation changes induced by precipitation anomalies due to ENSO forced circulation changes. This is demonstrated by the lead-lag correlation of the drought indices averaged over the Southwest (30º-40ºN, 95º-115ºW) with respect to DJF NINO3.4 SST, which is shown in Fig. 4. The
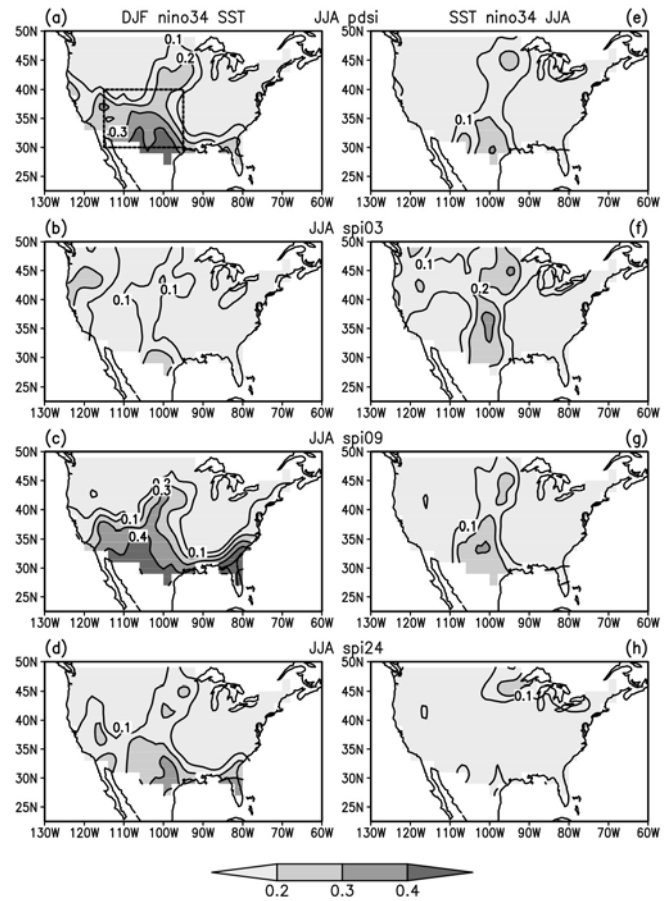
Fig. 3 Correlation of JJA PDSI and SPI with respect to DJF and JJA NINO3.4 SST during 1897-2005. The contour interval is 0.1. The correlation coefficient at the 1% significance level is about 0.25 (or 0.35 for SPI24). The box in (a) refers to the Southwest region (30°-40°N, 95°-115°W).
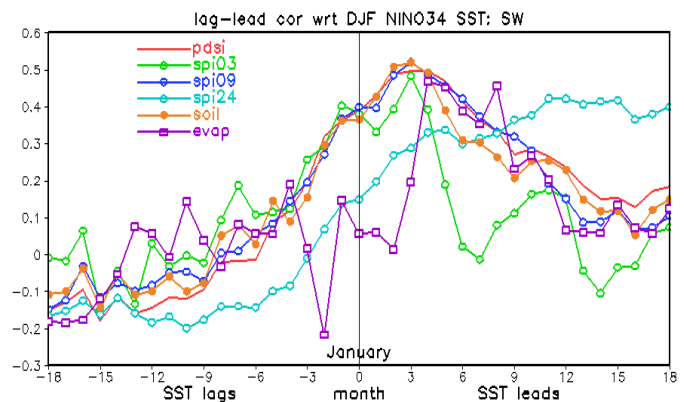
Fig. 4 Lag-lead correlation of U.S. Southwest (30°-40°N, 95°-115°W) PDSI, SPI, soil moisture, and evaporation with respect to DJF NINO3.4 SST during 1932-2005. The correlation coefficient at the 1% significance level is about 0.30 (or 0.40 for SPI24).

ENSO impacts on PDSI, SPI09, and soil moisture are apparent in DJF. The largest correlation is seen in MAM, lagging the ENSO mature phase by about one season. The induced droughts can persist into spring and summer. This persistence seems to be related to the enhanced evaporation in late spring and summer following the increase in soil moisture. This soil moisture-evaporation effect seems especially important for long-term drought that develops after DJF. Note that the evaporation correlation is small in DJF and becomes large in late-spring and summer, indicating a delayed feedback of soil moisture on droughts through evaporation, which may be related to the annual cycle of insolation and temperature. There is a similar feature in the Great Plains region except that the correlation is weaker (not shown).

5. Long-term changes in the SST-drought correlation

We have examined the long-term changes in the SST-drought relationship with the time. Shown in Fig. 5 is the sliding correlation between DJF NINO3.4 SST or JJA tropical Indian Ocean SST (15°S-15°N, 50°-100°E; TIO) and drought indices using a 31-year window.

The DJF NINO3.4 SST shows a

Fig. 5 Sliding correlation between DJF NINO3.4 SST or JJA TIO SST and JJA Southwest or Great Plains PDSI and SPI with a window of 31-years. The correlation shown in the figure is for the center year of the 31-year window. The correlation coefficient at the 5% significance level is about 0.36 (or 0.50 for SPI24).

robust positive correlation with PDSI, SPI09, and SPI24 in the Southwest during the whole period, but the magnitude of correlation varies largely (Fig. 5a). The correlation of SPI03 in the Southwest with DJF NINO3.4 SST displays an apparent contrast before and after the mid-1950s. For the Great Plains region, the correlation with the DJF NINO3.4 SST shows an overall weakening with time until the mid-1950s (Fig. 5b). After that, the correlation tends to increase until the early-1980s, which is then followed by a weakening in the correlation. Our results are consistent with Cole and Cook (1998). In comparison, the correlation change is relatively smaller for the Southwest than for the Great Plains.

The JJA TIO SST correlation shows a pronounced change with time. Before the late-1930s, the correlation is generally weak for both the Southwest and the Great Plains (Figs. 5c-d). The correlation underwent an obvious increase from the late-1930s to the mid-1940s. After a relatively steady period of about 10 years, the correlation displays an overall decrease with time for PDSI, SPI09, and SPI24 in the Southwest, and for PDSI and SPI09 in the Great Plains. The SPI24 in the Great Plains maintained a steady positive correlation from the 1960s to the mid-1980s (Fig. 5d).

6. Summary
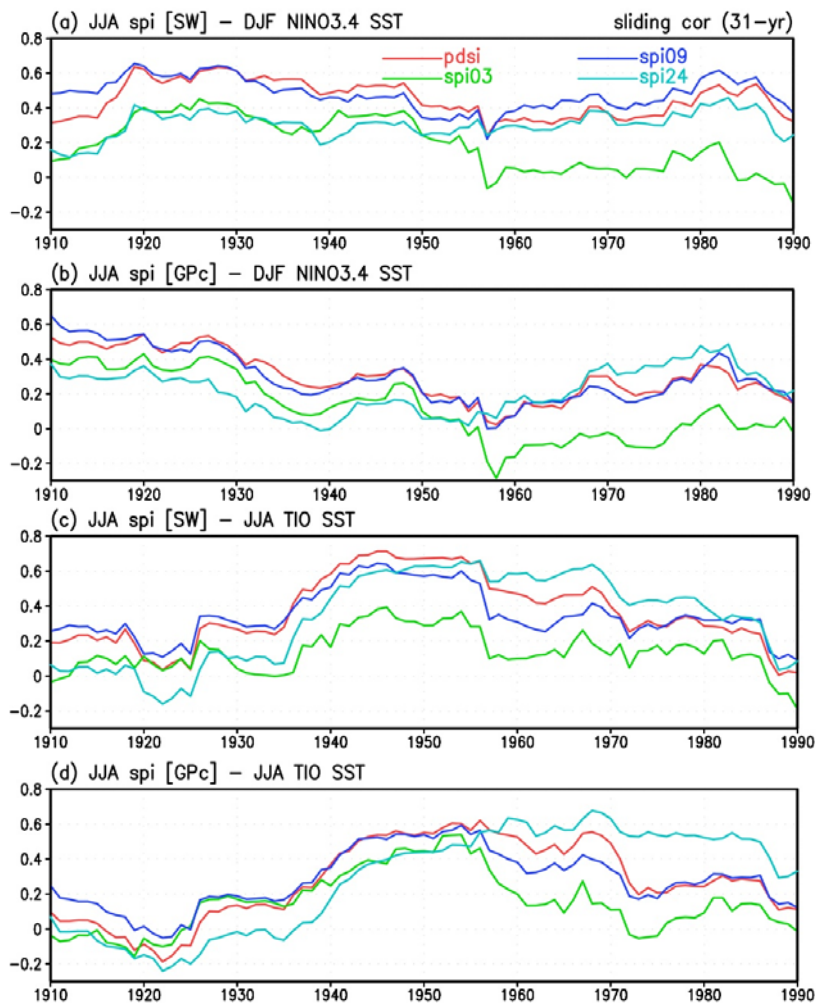
The main results are as follows:

- The relationship of boreal summer U.S. droughts with SST and soil moisture differs significantly between short-term and long-term droughts. The short-term droughts (<= 3 months) are mostly influenced by simultaneous SST forcing. The medium-term and long-term droughts (>= 6 months) are influenced by both preceding and simultaneous SST forcing. The soil moisture change shows obvious leading for medium-term and long-term droughts.

- A dominant remote forcing for U.S. droughts is tropical Pacific SST. Tropical Indian Ocean SST forcing has notable influence on medium-term and long-term droughts. Additional impacts for short-term and medium-term droughts are from the North Atlantic SST forcing.

- The most notable impacts of the tropical Pacific SST forcing on medium-term and long-term droughts are found in the Southwest with extension to the Great Plains. Anomalous soil moisture induced by remote ENSO forcing contributes to the persistence of droughts from winter to summer through anomalous evaporation during late spring to summer.

- The relationship between tropical Pacific SST and boreal summer U.S. droughts show obvious long-term changes. In comparison, the long-term change is more pronounced for the GP droughts than for the SW droughts. Obvious long-term changes are also found in the correlation of U.S. droughts with tropical Indian Ocean SST, especially for JJA SST.

References

Alley, W. M., 1984: The Palmer drought severity index: limitations and assumptions. *J. Clim. Appl. Meteorol.*, **23**, 1100-1109.

Cole, J., and E. Cook, 1998: The changing relationship between ENSO variability and moisture balance in the continental United States. *Geophys. Res. Lett.*, **25**, 4529-4532.

Eltahir, E. A. B., 1998: A soil moisture-rainfall feedback mechanism. 1: Theory and observations. *Water Resources Research*, **34**, 765-776.

Fan, Y., and H. M. van den Dool, 2004: Climate Prediction Center global monthly soil moisture data set at 0.5° resolution for 1948 to present. *J. Geophys. Res.*, **109**, D10102, doi:10.1029/2003JD004345.

Findell, K. L., and E. A. B. Eltahir, 1997: An analysis of the soil moisture-rainfall feedback, based on direct observations from Illinois. *Water Resources Research*, **33**, 725-735.

Guttman, N. B., J. R. Wallis, and J. R. M. Hosking, 1992; Spatial comparability of the Palmer Drought Severity Index. *Water Resour. Bull.*, **28**, 1111-1119.

Hoerling, M., and A. Kumar, 2003: The perfect ocean for drought. *Science*, **299**, 691-694.

Hollinger, S. E., and S. A. Isard, 1994: A soil moisture climatology of Illinois. *J. Climate*, **7**, 822-833.

McKee, T. B., N. J. Doesken, and J. Kleist, 1993: The relationship of drought frequency and duration to time scales. Preprints, *Eighth Conf, on Applied Climatology*, Anaheim, CA, Amer. Meteor. Soc., January 17-22, 179-184.

Mo., K. C., and J. E. Schemm, 2008: Drought and persistent wet spells over the United States and Mexico. *J. Climate*, **21**, 980-994.

Pal, J. S., and E. A. B. Eltahir, 2001: Pathways relating soil moisture conditions to future summer rainfall with a model of the land-atmosphere system. *J. Climate*, **14**, 1227-1242.

Palmer, W. C., 1965: Meteorological drought. *Research Paper No. 45*, U.S. Weather Bureau, 58pp.

Piechota, T. C., and J. A. Dracup, 1996: Drought and regional hydrologic variation in the United States: Associations with the El Niño-Southern Oscillation. *Water Resources Research*, **32**, 1359-1373.

Rajagopalan, B., E. Cook, U. Lall, and B. K. Ray, 2000: Spatialtemporal variability of ENSO and SST teleconnections to summer drought over the United States during the twentieth century. *J. Climate*, **13**, 4244-4255.

Robock, A., K. Y. Vinnikov, G. Srinivasan, J. K. Entin, S. E. Hollinger, N. A. Speranskaya, S. Liu, and A. Namkhai, 2000: The global soil moisture data bank. *Bull. Amer. Meteor. Soc.*, **81**, 1281-1299.

Schubert, S. D., M. J. Suarez, P. J. Pegion, R. D. Koster, and J. T. Bacmeister, 2004: Causes of long-term drought in the U.S. Great Plains. *J. Climate*, **17**, 485-503.

Seager, R., Y. Kushnir, C. Herweijer, M. Naik, and J. Velez, 2005: Modeling of tropical forcing of persistent droughts and pluvials over western North America: 1856-2000. *J. Climate*, **18**, 4065-4088.

Smith, T. M., and R. W. Reynolds, 2004: Improved extended reconstruction of SST (1854-1997). *J. Climate*, **17**, 2466-2477.

Ting, M., and H. Wang, 1997: Summertime U.S. precipitation variability and its relation to Pacific sea surface temperature. *J. Climate*, **10**, 1853-1873.

Trenberth, K. E., and C. J. Guillemot, 1996: Physical processes involved in the 1988 drought and 1993 floods in North America. *J. Climate*, **9**, 1288-1298.

Validation of Reanalysis Daily Precipitation over the Americas

V. B. S. Silva, V. E. Kousky, R. W. Higgins and Emily Becker

Climate Prediction Center, NOAA/NWS/NCEP

Camp Springs, MD, 20746

1. Introduction

An intercomparison of the statistics of daily precipitation over the Americas is carried out using gridded station data and the current generation of reanalysis products in use at the National Centers for Environmental Prediction (NCEP). Several simple measures are used to characterize relationships between the observations and reanalyses for the period of record, including difference in the means, ratio of variances, and correlation. Seasonality is accounted for by examining these measures on a monthly basis, using daily data in each case.

The intercomparison is motivated by Climate Prediction Center (CPC) plans to replace the current generation of reanalysis products in use for operational monitoring and prediction activities with a new generation of reanalysis products currently under development at the National Centers for Environmental Prediction (NCEP) Environmental Modeling Center (EMC) as part of the Coupled Forecast System Reanalysis and Reforecast (CFSRR) Project. A careful validation of the current generation of reanalysis products will provide a benchmark that can be used to confirm that the new generation of reanalysis products is an improvement.

2. Background

Since the mid-1990's the Climate Prediction Center (CPC) has used the National Centers for Environmental Prediction (NCEP) / National Center for Atmospheric Research (NCAR) reanalysis products (Kalnay et al. 1996; referred to as R1) and their real-time extension forward in time via the Climate Data Assimilation System (CDAS) for operational climate monitoring and prediction activities. The current generation of reanalysis products is among the most popular and widely used climate data sets currently in existence. The NCEP/DOE Reanalysis (R2), obtained using an updated forecast model and data assimilation system (Kanamitsu et al. 2002), is also used at CPC.

The NCEP is currently developing the next generation of reanalysis products as part of the Climate Forecast System (Saha et al. 2006) Reanalysis and Reforecast (CFSRR) project – a project that is driven by NCEP's intraseasonal-to-interannual prediction needs. The Environmental Modeling Center (EMC) plans call for the CFSRR to extend over the period 1981-present. CPC plans to extend the CFS reanalysis backward in time to 1948 and forward in real-time in order to satisfy operational climate monitoring and prediction needs. One of the advantages of the extension backward in time is that it will increase the number of cases of the low frequency modes of climate variability, such as ENSO, for a proper comparison of the CFSR to the current generation of reanalysis (R1 and R2). The CFS reanalysis products are expected to be a major improvement over the current global reanalysis, as they will be the result of a coupled Ocean-Atmosphere-Land system at higher spatial resolution.

3. Preliminary results for United States

Daily precipitation statistics were computed for the reanalyses (R1 and R2) and the observations (CPC gridded precipitation analysis). R1, R2 and OBS time series were constructed for the domain 20°N-60°N, 60°W-140°W, at a horizontal resolution of (lat/lon)=(2.5°x2.5°). Here we will show some results of three basic statistics: a) differences in probability of occurrence, b) ratio of variances, and c) correlation.

The probability of daily precipitation greater than 1 mm (Fig. 1) is less than observed in R1 and R2 over the Southeast during November to March (with slightly greater biases in R2) and less than observed over the Southwest during June to September (with greater biases in R1). Both R1 and R2 have greater than observed probabilities over the Southeast during June to September (with greater biases in R1).

Correspondence to: V. B. S. Silva, Climate Prediction Center, NOAA/NWS/NCEP, Camp Springs, MD, 20746;
E-mail: Viviane.Silva@noaa.gov

Fig. 1. Probability of daily precipitation greater than 1 mm in (a) OBS, and   expressed as the difference between (b) R1 and OBS, and (c) R2 and OBS. Results are shown for each month of the year and are based on daily data for the period 1979-2006.
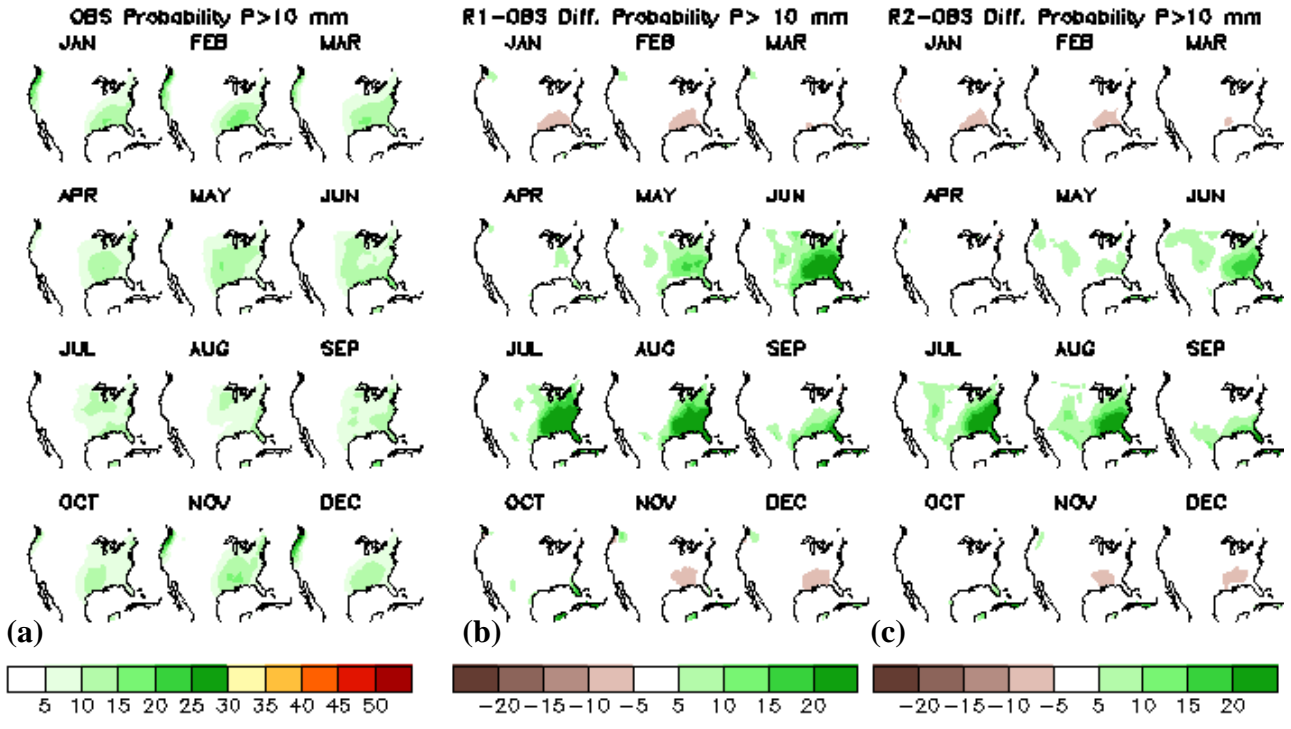
Fig. 2. Probability of daily precipitation greater than 10 mm in (a) OBS, and   expressed as the difference between (b) R1 and OBS, and (c) R2 and OBS. Results are shown for each month of the year and are based on daily data for the period 1979-2006.
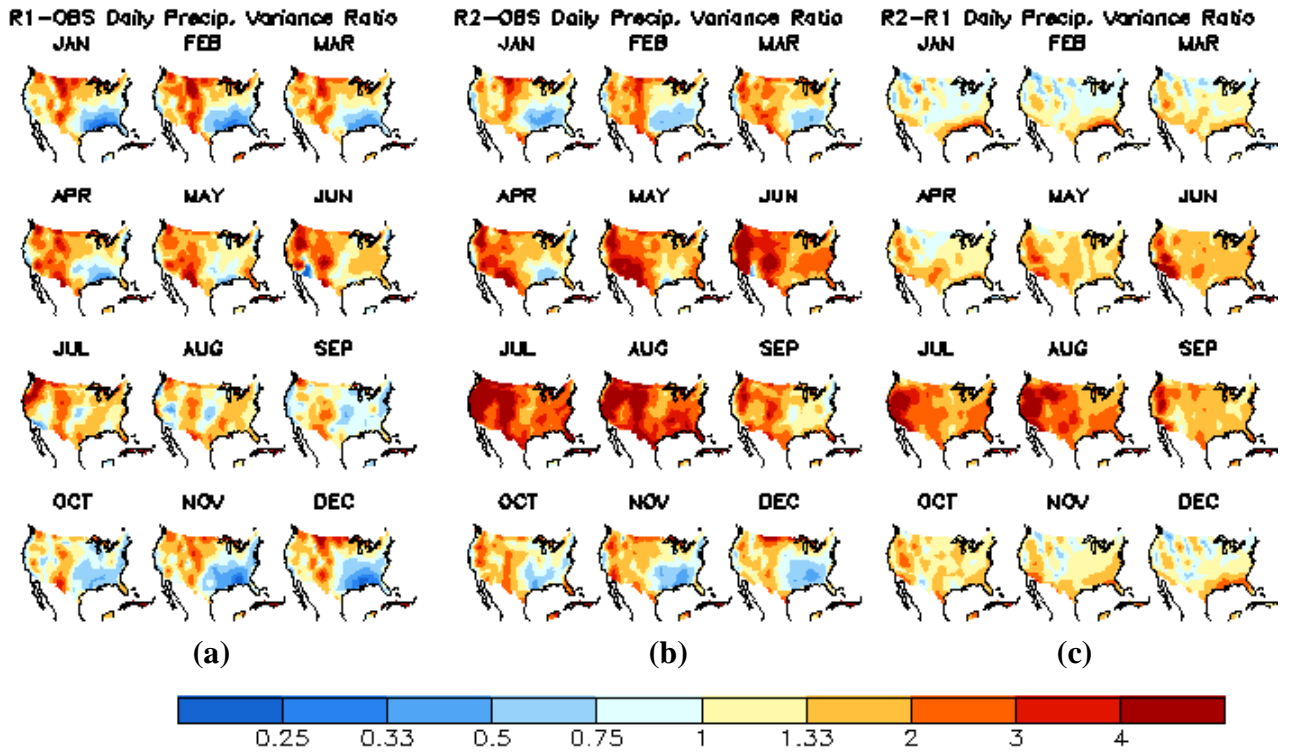
Fig. 3. Ratio of variance of daily precipitation: (a) R1 / OBS, (b) R2 / OBS and (c) R1 / R2. Results are shown for each month of the year and are based on daily data for the period 1979-2006.

Fig. 4. Spatial maps of the temporal correlation between (a) R1 and OBS, (b) R2 and OBS and (c) R1 and R2. Results are shown for each month of the year and are based on daily data for the period 1979-2006.

The probability of daily precipitation greater than 10 mm (Fig. 2) is less than observed in R1 and R2 over the Southeast during November to February. Both R1 and R2 have greater than observed probabilities over the Southeast during May to September (with greater biases in R1).   We note tha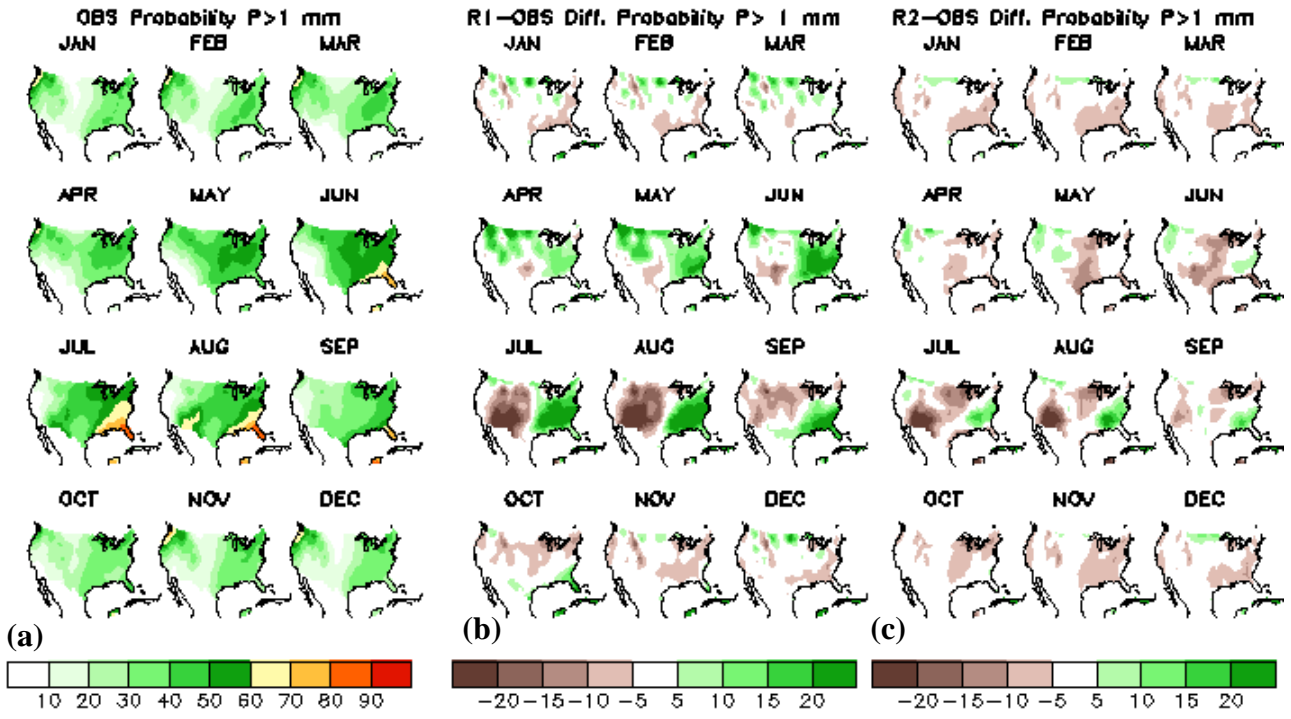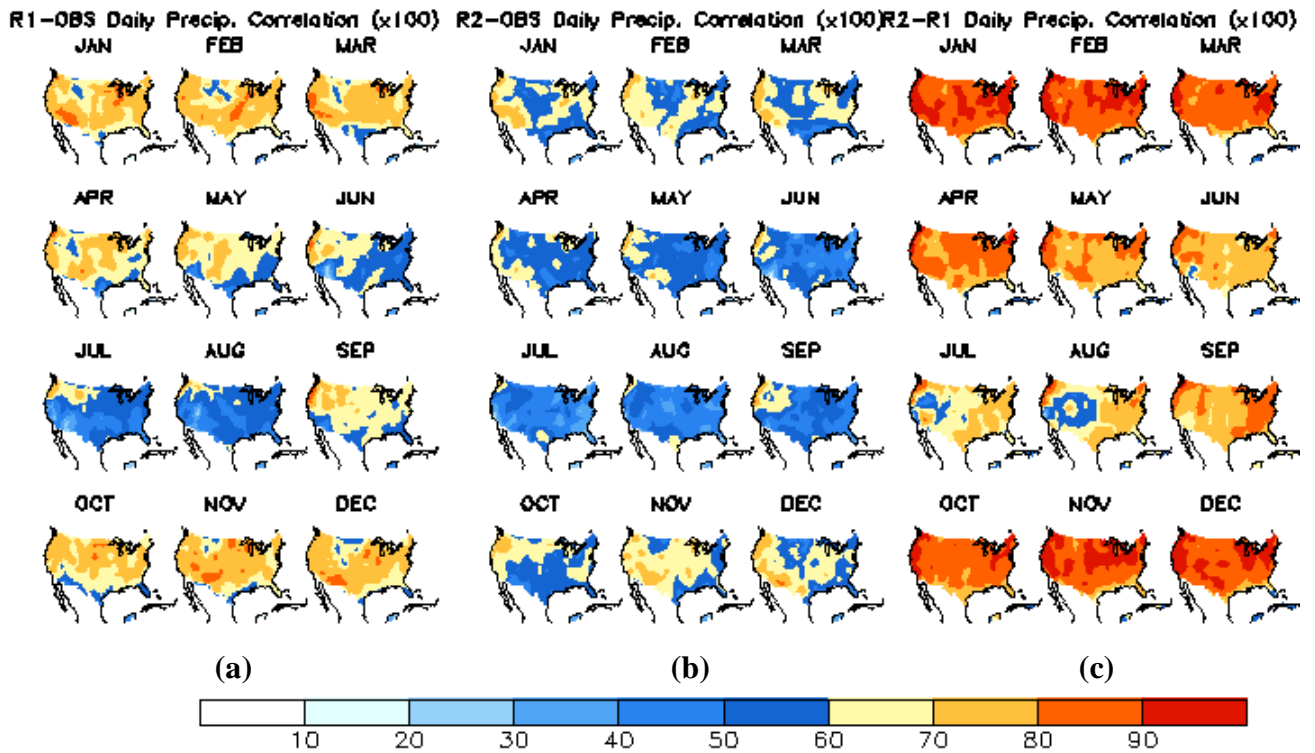t the differences in the Southeast during May to September are nearly as large as they are in Fig. 1, indicating that the overestimates in R1 and R2 are probably for the relatively heavy (convective) precipitation events.

An examination of the ratio of variance (Fig. 3) of daily precipitation between a) R1 and OBS and b) R2 and OBS by month shows that both R1 and R2 exhibit less variability than observations on daily timescales over eastern TX, the Gulf Coast states and Tennessee Valley during October-April (ratios are less than 1), while the reanalyses display more variability than observed over the West throughout the year (ratios are greater than 1). A notable exception is along the immediate West Coast where R1 and R2 are less variable than observations on daily time scales during much of the year (Figs 3a and 3b). Figure 3c shows that R2 variance is greater than R1 variance across the entire CONUS during May-October (greatest differences during July-September). In addition, R2 variance exceeds R1 variance along the Gulf Coast region during November-April. Since the R2 and R1 variances are less than OBS during November-April (see figures 3a and 3b) over the Gulf region, the R2/R1 ratio greater than 1 shown in figure 3c indicates that R2 is closer to OBS in this region.

Spatial maps of the temporal correlation between daily precipitation in R1 and OBS (R2 and OBS) (Fig. 4a and b) show generally high correlations in winter and low correlations in summer when convection is present. Throughout the year, the correlations between R2 and OBS are lower than those for R1 and OBS. So, in spite of improvements in the mean bias of daily precipitation in R2 (figure not shown), the correlation with observations is not as good.

4. Summary

Before CPC can confidently base its operational climate monitoring and prediction activities on a new generation of reanalysis and reforecast products, a comprehensive intercomparison of the old and new products is required. Although the next generation of reanalysis products is not available yet, a careful validation of the current generation of reanalysis products will serve as a benchmark from which an objective evaluation of improvements in the new generation of reanalysis products can be made. An important long term goal is to identify and correct biases in the statistics of daily precipitation within a season to improve CPC's current operational monitoring and outlook products and to develop new outlook products on intraseasonal and seasonal timescales.

References

Kalnay E., and co-authors, 1996: The NCEP/NCAR 40-year reanalysis project. *Bull. Amer. Met Soc.*, **77**, 437-471.

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

Saha, S., S. Nadiga, C. Thiaw, J. Wang, W. Wang, Q. Zhang, H. M. van den Dool, H.-L. Pan, S. Moorthi, D. Behringer, D. Stokes, M. Peña, S. Lord, G. White, W. Ebisuzaki, P. Peng, P. Xie, 2006: The NCEP Climate Forecast System. *J. Climate*, **19**, 3483-3517.

Drought Monitoring over the United States

Kingtse Mo

Climate Prediction Center, NOAA/NWS/NCEP
Camp Springs, MD, 20746

1. Introduction

The purpose of this presentation is to give an overview of the real time drought monitoring at CPC. We use the drought indices to monitor precipitation (P), soil moisture (SM) and runoff deficits. Soil moisture and runoff are based on the North American Land Data Assimilation system (NLDAS) products. The uncertainties of the NLDAS are assessed by intercomparing the NLDAS products from the NCEP and the University of Washington (UW).

2. Results

Drought indices derived from the NLDAS from NCEP and UW from 1979 to 2007 are intercompared and evaluated for their ability to assess drought severity over the United States. Each system has four models. The uniformly weighted ensemble means of four models are used for comparison. For meteorological drought, the Standardized Precipitation Index (SPI) is used to measure precipitation deficits. The Standardized Runoff Index (SRI) similar to the SPI is used to classify hydrological drought. Agricultural drought is measured by monthly mean soil moisture anomaly percentiles based on probability distributions (PDs). The PDs for total SM are regionally dependent and influenced by the seasonal cycle, but the PDs for SM monthly mean anomalies are unimodal and Gaussian.

Over the eastern United States (east of 95ºW), the indices derived from NCEP and UW are similar and they are able to detect the same drought events. Indices are also well correlated. For River Forecast Centers (RFCs) over the eastern United States, different drought indices are likely to select the same drought events.

The monthly mean SM percentiles and runoff indices between NCEP and UW have large differences over the western interior United States. For small areas with a horizontal resolution of 0.5 degrees on the time scales of one to three months, the differences of SM percentiles and SRI

Fig. 1 Monthly total P data counts averaged over grid points in 4 $^\circ$ x 5 $^\circ$ boxes.

between NCEP and UW are larger than the thresholds used to classify drought. For the western RFCs, drought events selected according to the SM percentiles or SRI derived from different NLDAS systems do not always overlap.

The largest differences came after 2004, when both systems went for near real time production. Figure 1 shows the monthly total data counts for boxes over the western region. After 2004, data counts dropped substantially. The differences in precipitation forcing cause large uncertainties in the NLDAS systems.

3. Concluding remarks

One of the major issues is the precipitation analysis. For real time operation, the P inputs over the western region are sparse. To improve P analysis is essential to improve NLDAS and drought monitoring.
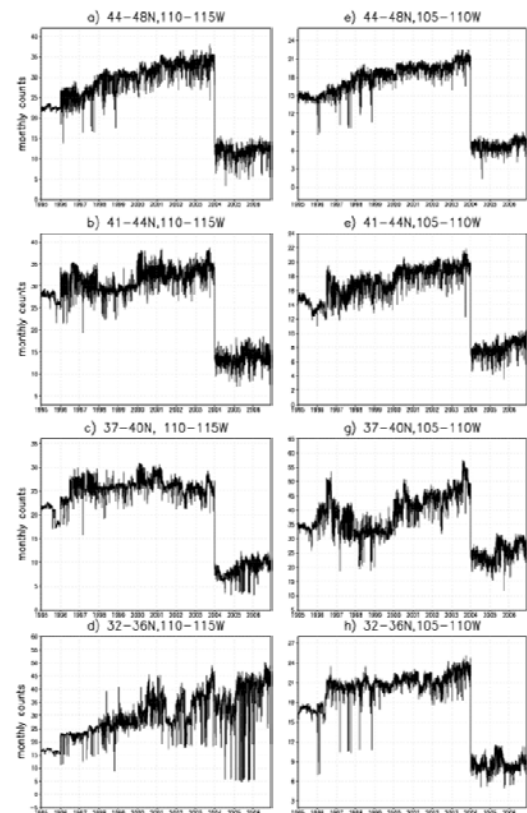
Correspondence to: Kingtse Mo, Climate Prediction Center, NOAA/NWS/NCEP, Camp Springs, MD, 20746;
E-mail: Kingtse.Mo@noaa.gov

Amazon Deforestation in CFS

Edwin K. Schneider

Department of Atmospheric, Ocean and Earth Sciences, George Mason University, Fairfax, Virginia
Center for Ocean-Land-Atmosphere Studies, Calverton, Maryland

1. Introduction

Many simulations of the potential effects of tropical deforestation on climate have been made using atmospheric general circulation models (AGCMs) coupled to land models and forced by *specified* SST (*e.g.* Dickinson and Henderson-Sellers, 1988; Nobre *et al.*, 1991). Recently, CGCM (coupled general circulation models, which include dynamical oceans) simulations of tropics-wide (Voldoire and Royer 2005, referred to as VR05) and Amazon (Schneider *et al.* 2006, referred to as S06) deforestation have been made. We extend these CGCM results by examining the effects of Amazon deforestation on the coupled ocean-atmosphere-land climate system using the NCEP CFS. A primary advantage of CFS is that it has a much more realistic simulation of current Amazon climate than the CGCMs used in the earlier studies.

Century length control and deforestation simulations are carried out with CFS. The results suggest that the impact of Amazon deforestation would be a warmer and drier Amazon, as well as a warmer tropical Pacific and tropical North Atlantic. However, these changes are small. ENSO is not noticeably affected. Sensitivities to changes in the land surface processes are diagnosed using uncoupled AGCM simulations, using GFS, the atmospheric model component of CFS. The GFS simulations suggest that albedo changes are the controlling influence for the Amazon deforestation effects found in CFS, due to the mechanism outlined by Charney (1975).

An unexpected warming occurs in the northern North Atlantic region in the deforestation simulation. We examine the Meridional Overturning Circulation and other quantities in the simulations in an attempt to understand the origins of this change, which appears to be related to the physically inconsistent treatment of sea ice.

2. Models and data

Two 100 year simulations were carried out with CFS, a current climate control simulation (CONTROL) and a deforestation simulation (DEFOREST) in which tropical rainforest in the Amazon region (vegetation type 1) was replaced with perennial ground cover (vegetation type 7). In the deforested region, the albedo and surface roughness were also changed to values appropriate for areas of vegetation type 7 found near the Amazon.

Fig. 1 Annual cycle of monthly means of precipitation (left) and surface temperature (right) averaged over land in the box from 80°W to 40°W and 15°S to 8°N for the COLA CGCM (top) and CFS (bottom). Observations/analysis are the black curves, the CONTROL simulations are green, and the DEFOREST simulations are red.

The vegetation change reduces the resistance to surface evaporation. The main effects of the deforestation on the specified surface properties were an increase in the shortwave beam albedo to about 30% from about 23% and a decrease in the roughness length to 0.1 m from values larger than 2 m. The simulations were started from analysis with identical Jan. 1, 1985 initial and boundary conditions.
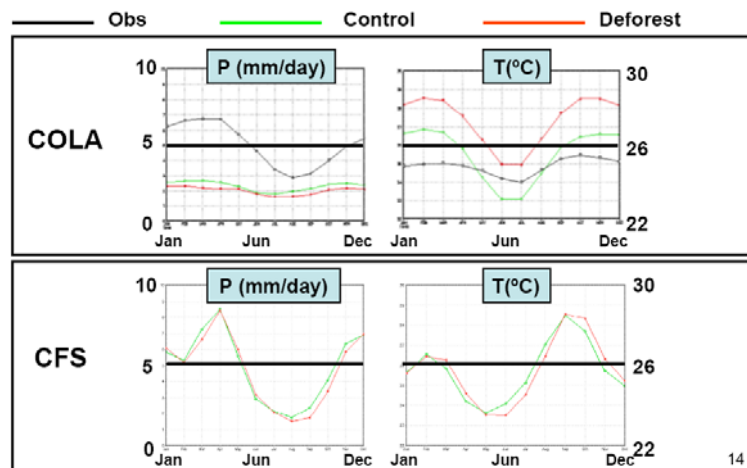
_____

Correspondence to: Edwin K. Schneider, Center for Ocean-Land-Atmosphere Studies, 4041 Powder Mill Road, Suite 302, Calverton, MD 20705; E-mail: schneide@cola.iges.org

DJFMAM

P CFS Wet Season 1986–2006

JJASON

P CFS Dry Season 1986–2006

ERBE OLR Amazon Wet Season (DJFMAM)
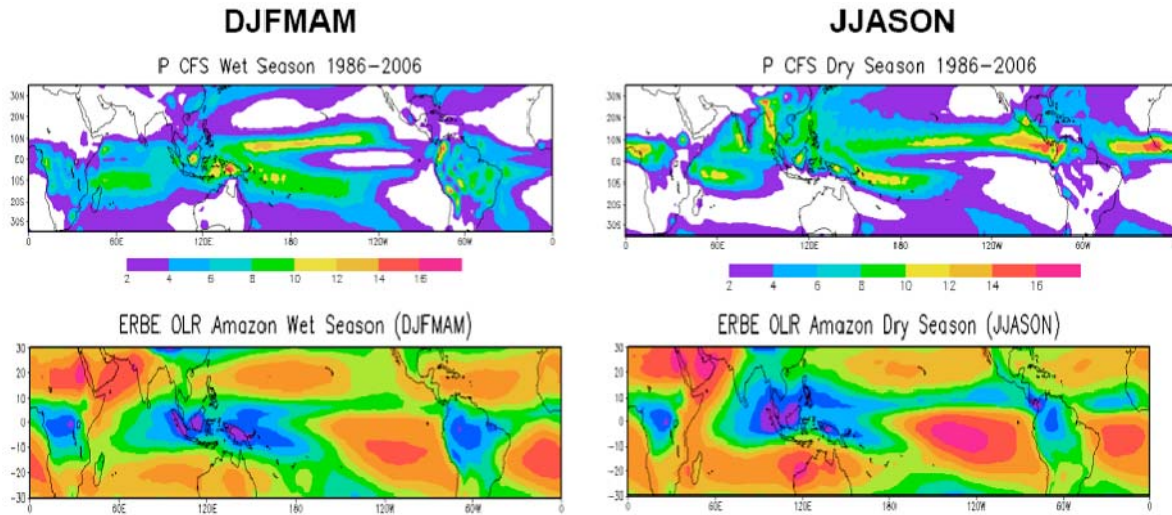
ERBE OLR Amazon Dry Season (JJASON)

Fig. 2 Precipitation for CONTROL (top, mm/day) and ERBE OLR (bottom, contour interval 10 W/m^2; cool/warm colors in OLR correspond to high/low precipitation) for the Amazonian wet (December through March, left) and dry (June through November) seasons.

A caveat concerning these simulations is that CFS is not designed for long climate change simulations. For one thing, the CO_2 concentration is constant in time. Additionally, the domain for ocean-atmosphere interaction is non-polar, and the sea ice distribution is specified. Therefore, the freezing and melting of sea ice are not directly tied to the ocean temperature, which can lead to energetic and physical inconsistencies.

3. Results

a. Simulation of current climate in Amazonian region

Figure 1 shows the climatology of the area averaged Amazon precipitation and surface temperature for observations/analysis, CFS, and the S06 COLA CGCM (consisting of the COLA V2 AGCM, SSiB land, and MOM3 OGCM with anomaly coupling). The model climatologies are taken over the 100 simulated years. The observed annual mean precipitation is about 5 mm day^{-1}, which is also the value simulated by CFS CONTROL. However, the COLA CGCM control simulation produces an annual mean precipitation of about 2 mm day^{-1}, only about 40% of the observed value, and much too small to support a rain forest. Similarly, the control simulation from the CGCM of VR05 (consisting of the ARPEGE-climate AGCM, ISBA land model, and LODYC OGCM) produces a very dry Amazon, with an annual mean rainfall of 3.5 mm day^{-1}. The more realistic performance in the simulating the climatological precipitation was the main motivation for us to adopt CFS for the experiments.
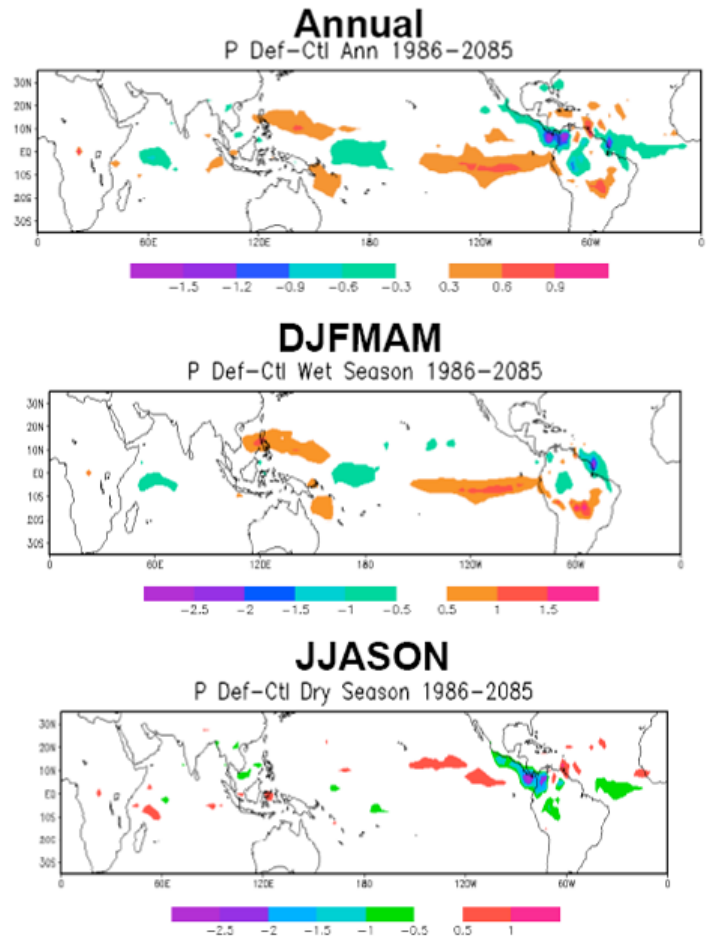
Annual

P Def–Ctl Ann 1986–2085

DJFMAM

P Def–Ctl Wet Season 1986–2085

JJASON

P Def–Ctl Dry Season 1986–2085

Fig. 3 Precipitation difference (mm/day) DEFOREST minus CONTROL, averaged from 1986 to 2085.

There are, however, biases in CFS that are obvious from Fig. 1. The amplitude of the annual cycle of precipitation is too large, as is the amplitude of the annual cycle of surface temperature. Additionally, the CFS surface temperature is about 1°C too high, and the COLA control simulation is closer to observations in this quantity. However, CFS has a drift in the surface temperature (see Fig. 6). When averaged over the first 10 years of the simulations, CFS annual mean land surface temperature in the region of Fig. 1 is 25.2°C, which is cooler than the 100 year mean by about 1°C and close to the analysis, while precipitation is not changed much.

Figure 2 compares the wet and dry season 1986-2005 climatological precipitation for CONTROL with the climatology of ERBE OLR (OLR is a commonly used a proxy for precipitation from deep convection). The precipitation distribution over the Amazon region in CONTROL is reasonably smoothly spatially distributed, although not as smooth as the OLR. There are some orographically-tied features that appear to be associated with the Andes and which are not seen in the OLR. The CONTROL Amazon precipitation distribution appears to have a much more realistic spatial distribution than that found in S06 for the COLA AGCM. The rainfall deficit bias in the dry season shown in Fig. 1 appears to be due to a severely deficient rainfall to the south of the equator in South America.

b. Topical and subtropical changes due to Amazon deforestation

Fig. 4 Air temperature at 2 m difference (K), DEFOREST minus CONTROL, averaged from 1986 to 2085. Temperatures over ocean and land have different scales. Color bar is for temperature over land.

Fig. 5 Annual mean surface wind stress on ocean (dynes/cm^2) DEFOREST minus CONTROL.

The changes in precipitation and surface temperature due to the Amazon deforestation, DEFOREST minus CONTROL, are shown in Figs. 3 (precipitation) and 4 (temperature). Deforestation locally leads to a decrease in precipitation and warming of surface air temperature in the core of the deforested region for both the wet and dry seasons as well as the annual mean. However, there is a compensating increase in rainfall and associated cooling on the southeast flank of the Amazon region.

We have conducted two sets of experiments with the GFS AGCM to separate the influences of the physical processes involved in our deforestation simulation. One set examines the changes in vegetation type, and the other the combined effects of changes in surface albedo and surface roughness. The change in vegetation reduces the resistance to evaporation, which leads by itself to enhanced rainfall and surface cooling. The increased albedo by itself would be expected to lead to decreased rainfall and surface cooling by the enhanced subsidence mechanism of Charney (1975); however, reduced cloudiness associated with the reduced rainfall acts to mitigate the surface cooling. The role of reduced roughness is less easy to anticipate, and we have not isolated the sensitivity to this process. When the changes are compared, the albedo/roughness effect turns out to be dominant.
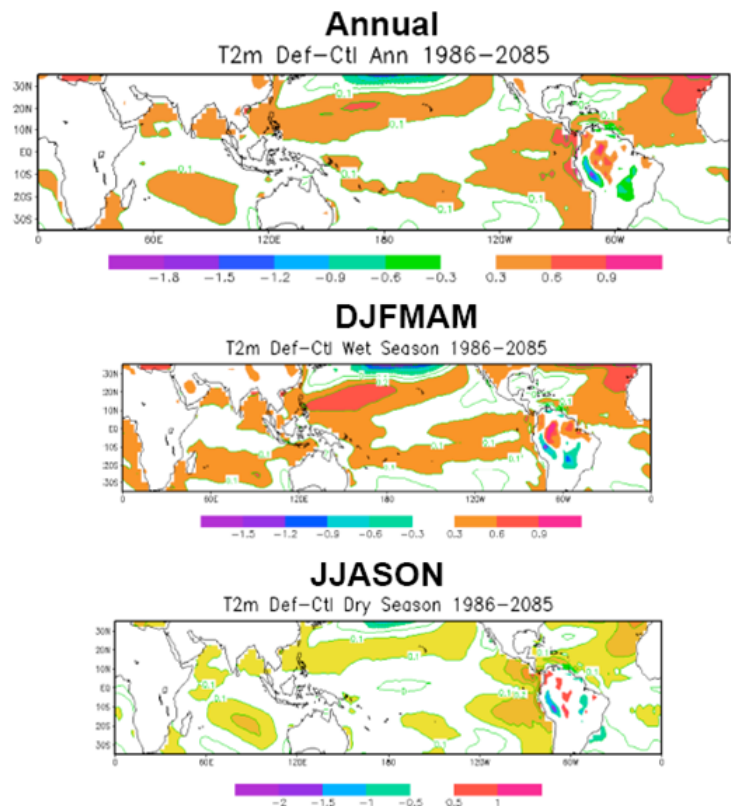
The remote effects of the deforestation include warming of the SST in the eastern tropical Pacific and tropical North Atlantic, and increased precipitation in the eastern equatorial Pacific, as well as widespread warming of SST in the northern hemisphere and southern Indian Ocean. Precipitation decreases in the near-equatorial Atlantic during the dry season, and the western tropical Pacific.

The changes in the Atlantic and Pacific can perhaps be connected to the atmospheric dynamical response to the local response over the Amazon. As shown above, the local response is: 1) reduction in magnitude and changes in distribution of the latent heat release, and 2) warmer surface temperature. Figure 5 shows that the Amazon deforestation is associated with easterly surface wind anomalies in the tropical Atlantic, with northerly/easterly anomalies north of the equator and westerlies at 5°S in the tropical Pacific east of the dateline. The precipitation and surface wind anomalies in the tropical Atlantic and eastern Pacific, and the Pacific SST are similar to those found by S06. However, the Atlantic response in S06 was a cooling to the south of the equator. The SST changes in S06 were tentatively explained there as originating from a Gill-type response (Gill 1980) to the decreased Amazon atmospheric heating competing with the response to the increased land surface temperatures. The response to surface temperature anomalies can also be viewed as a Gill-type response (Neelin 1989), but with a smaller effective depth, although this argument is not commonly applied to land surface temperatures. According to the simple model, the reduction in heating would produce surface wind directions opposite to those seen in DEFOREST minus CONTROL in the

Fig. 6 Global and annual mean 2 m air temperature for CONTROL (red) and DEFOREST (black).
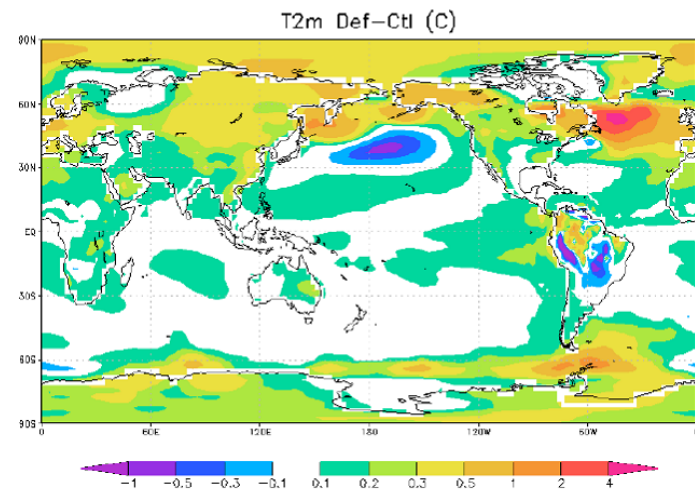
Fig. 7 Annual mean air temperature at 2 m difference (K), DEFOREST minus CONTROL, averaged from 1986 to 2085. Temperatures over ocean and land have different scales. Color bar is for temperature over ocean.
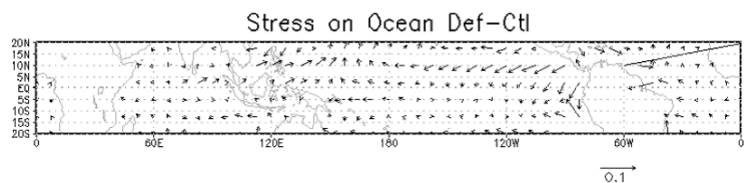
equatorial Pacific and Atlantic near South America, while the response to the warmer surface temperature would produce wind directions in agreement with those found in CGCMs. To explain the results, the land surface temperature forcing would have to be more important in producing the surface wind over the oceans near the deforested region. The response to the deep heating anomaly would be expected to be the dominant far field response. The precipitation and Pacific SST changes in the corresponding regions in VR05 also appear to be similar. In contrast to CFS, VR05 also found cooling in the tropical Atlantic.

c. Global scale changes associated with Amazon deforestation

Figure 6 shows the global mean temperature evolution for CONTROL and DEFOREST. Both simulations show evidence of a significant drift or "warming commitment," with global mean temperatures initially warming rapidly and then leveling off after about 60 years. This is the expected behavior if there is a net surface flux into the ocean (or downward top of the atmosphere heat flux) when the atmosphere is in
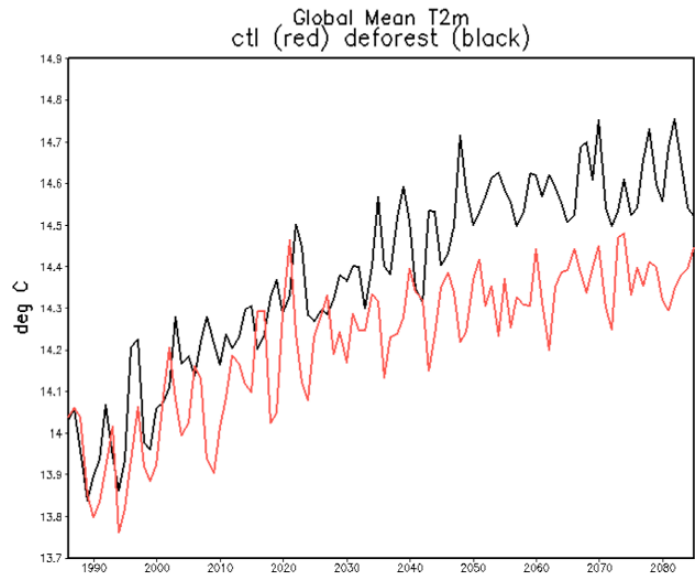
equilibrium with the ocean in the current climate configuration. A warming commitment is not an indication of biases in the model, and in fact may indicate that the model is behaving realistically, so long as the initial state is realistic (Schneider 1996). The drift is about a 0.5°C warming for CONTROL and substantially more, 0.8°C, for DEFOREST. The mechanism connecting deforestation to the larger warming is not clear; therefore we made a substantial effort to try to understand this result.

The spatial distribution of the DEFOREST minus CONTROL surface air temperature change in the global domain is shown in Fig. 7 (same data as Fig. 4 in lower latitudes). The warming in DEFOREST is greatest in high latitudes, and there is a very large warmer region in the North Atlantic. There is no obvious dynamical link between these high latitude regions and the Amazon.

The large SST anomaly in the North Atlantic suggests the possibility of involvement of the Atlantic Meridional Overturning Circulation (AMOC). The change in the AMOC is shown in Fig. 8. The surface flow in the North Atlantic shows an increase, which is of the right sign to lead to warmer SST in the North Atlantic, but the 1 Sv magnitude of the change is too weak to explain the changes shown in Fig. 7.

Further analysis shows that there may be a problem with the model physics in the vicinity of the large SST increase in Fig. 7. Figure 9 shows SST and heat flux from the ocean model output. The SST increases by order 20°C in Hudson Bay and the northern North Atlantic, while the heat flux is into the ocean in the regions of the largest SST increase and does not respond strongly to the warmer SST. Also, this strange behavior is found in a region where the specified sea ice is non-zero. If the sea ice was physically consistent with the SST, the warming SST would melt the ice. Instead, the insulating effect of the ice appears to be leading to enhanced warming of the SST. The positive heat flux into the ocean in the northern part of Hudson Bay and to the east in the North Atlantic does not seem to be physically defensible, especially in DEFOREST given the warm SST. There is also a jump in the heat flux across a line of constant latitude to the south of this region that is suspicious. It is clear that there is something amiss in the model, at least in DEFOREST,
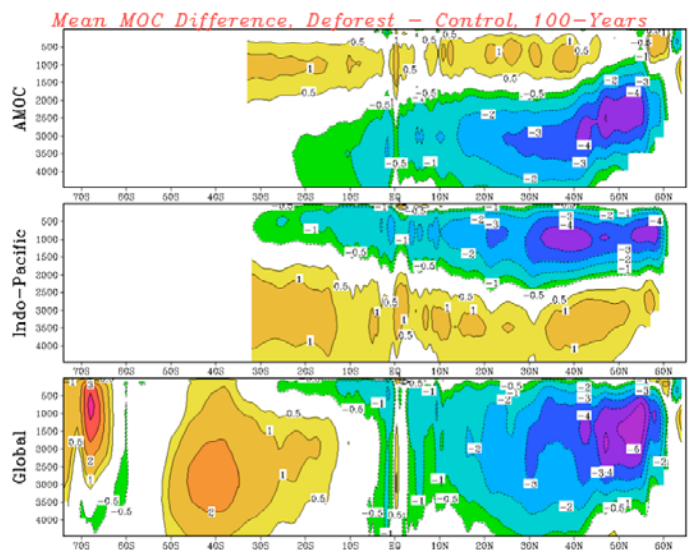

Fig. 8 Change in meridional overturning mass flux (Sv), DEFOREST minus CONTROL. Top: Atlantic; middle: Indo-Pacific; bottom: global. Positive values indicate clockwise circulation.
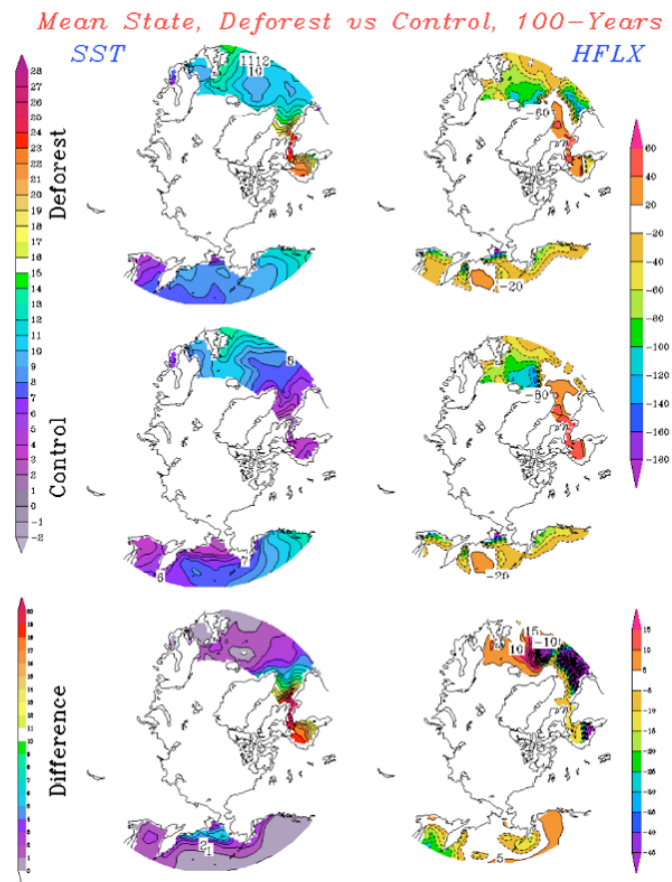

Fig. 9 SST (left) and net downward surface heat flux (right, W/m²) for DEFOREST (top), CONTROL (middle), and DEFOREST minus CONTROL (bottom).

which needs to be diagnosed, understood, and corrected before going further with this investigation. The obvious candidate processes are the sea ice and high latitude oceanic sponge layers.

4. Conclusions

We investigated the response of the coupled atmosphere-ocean-land system to a hypothetical complete Amazon deforestation. The purpose of the investigation was to identify the sensitivity of the climate system to changes in the land surface properties. The local changes were a general warming and drying in the Amazon region, primarily caused by to the changes in the surface albedo and roughness. The remote effects included warming of the eastern tropical Pacific and tropical Atlantic, increased precipitation in the eastern equatorial Pacific, and reduced precipitation in the equatorial Atlantic. The precipitation and Pacific SST changes are in agreement with other similar studies, indicating that these changes may be robust. However, the SST changes in the tropical Atlantic are opposite in sign to those in the other studies. The SST changes are apparently due to the influence on the ocean of the dynamical response of the atmosphere to the warmer Amazon surface temperatures, since the reduction in latent heat release would be expected to lead to the opposite effects on surface winds and hence tropical SST.

There was little effect on ENSO SST variability in CFS. S06 found an increase in ENSO amplitude from Amazon deforestation and attributed this to changes in the basic state (*i.e.* the changes in the ocean thermal structure climatology). However, the ENSO variability in CFS appears to be associated with stronger coupled instability and a more regular oscillation than in the COLA CGCM. Possible reasons for the lack of sensitivity of ENSO in CFS compared to the COLA CGCM then may be either the stronger coupled instability, the smaller warming in the land surface temperature and consequently smaller changes in the ocean thermal structure, or the secular warming in CFS.

The long simulations showed that CFS has a warming climate commitment of order 0.5°C for current initial states and the constant modern climate CFS CO_2 concentration. Amazon deforestation apparently enhances the global mean warming, but this effect appears to have been due to a problem in the model physics that needs to be investigated, and which may be connected to the code modifications we made.

References

Charney, J. G., 1975: Dynamics of deserts and drought in the Sahel. *Quart. J. Roy. Meteor. Soc.*, **101**, 193-202.

Dickinson, R. E., and A. Henderson-Sellers, 1988: Modelling tropical deforestation: A study of GCM land-surface parameterizations. *Quart. J. Roy. Meteor. Soc.*, **114**, 439-462.

Gill, A. E., 1980: Some simple solutions for heat-induced tropical circulation. *Quart. J. Roy. Meteor. Soc.*, **106**, 447-462.

Neelin, J. D., 1989: On the interpretation of the Gill model. *J. Atmos. Sci.*, **46**, 2466-2468.

Nobre, C. A., P. J. Sellers, and J. Shukla, 1991: Amazonian deforestation and regional climate change. *J. Climate*, **4**, 957-988.

Schneider, E. K., 1996: Flux correction and the simulation of changing climate. *Annales Geophysicae*, **14**, 336-341.

Schneider, E. K., M. Fan, B. P. Kirtman and P. A. Dirmeyer, 2006: Potential Effects of Amazon Deforestation on Tropical Climate. COLA Technical Report 226, 41 pp, (ftp://grads.iges.org/pub/ctr/ctr_226.pdf).

Voldoire, A. and J.-F. Royer, 2005: Climate sensitivity to tropical land surface changes with coupled versus prescribed SSTs. *Climate Dyn.*, **24**, 843-862. DOI 10.1007/s00382-005-0014-7.

Effects of Freshwater Flux (FWF) Forcing on Interannual Climate Variability in the Tropical Pacific

Rong-Hua Zhang and Antonio J. Busalacchi

Earth System Science Interdisciplinary Center (ESSIC)

University of Maryland, College Park, MD

ABSTRACT

The impacts of freshwater flux (FWF) forcing on interannual variability in the tropical Pacific climate system are investigated using a hybrid coupled model (HCM), constructed from an oceanic general circulation model (OGCM) and a simplified atmospheric model, whose forcing fields to the ocean consist of three components. Interannual anomalies of wind stress and precipitation minus evaporation, (P-E), are calculated respectively by their statistical feedback models that are constructed from a singular value decomposition (SVD) analysis of their historical data. Heat flux is calculated using an advective atmospheric mixed layer (AML) model. The constructed HCM can well reproduce interannual variability associated with El Niño-Southern Oscillation (ENSO) in the tropical Pacific.

HCM experiments are performed with varying strength of anomalous FWF forcing. It is demonstrated that FWF can have a significant modulating impact on interannual variability. The buoyancy flux (Q_B) field, an important parameter determining the mixing and entrainment in the equatorial Pacific, is analyzed to illustrate the compensating role played by its two contributing parts, one is related with heat flux (Q_T) and the other with freshwater flux (Q_S), respectively. A positive feedback is identified between FWF and sea surface temperature (SST) as follows. SST anomalies, generated by El Niño, non-locally induce large anomalous FWF variability over the western and central regions, which directly influences sea surface salinity (SSS) and Q_B, leading to changes in the mixed layer depth (MLD), the upper ocean stability, the mixing and the entrainment of subsurface waters. These oceanic processes act to enhance the SST anomalies, which in turn feedback to the atmosphere in a coupled ocean-atmosphere system. As a result, taking into account anomalous FWF forcing in the HCM leads to an enhanced interannual variability and ENSO cycles. It is further shown that FWF forcing is playing a different role from heat flux forcing, with the former acting to drive a change in SST, while the latter being representing a passive response to the SST change. This HCM based modeling study presents clear evidence for the role of FWF forcing in modulating interannual variability in the tropical Pacific. The significance and implications of these results are further discussed for physical understanding and model improvements of interannual variability in the tropical Pacific ocean-atmosphere system.

---

1. Introduction

The ocean is a key player in climate variability and predictability on various time-space scales. Largely driven by atmospheric forcing, the induced physical changes in the ocean can feedback to the atmosphere by which the principal oceanic quantity felt is sea surface temperature (SST). Numerous studies have identified roles of various forcings and feedbacks in the climate system, including the Bjerknes feedback (*e.g.*, Bjerknes 1969), the wind-evaporation-SST (WES) feedback (*e.g.*, Xie and Philander 1994), the SST-solar radiation feedback (*e.g.*, Waliser *et al*. 1994), and others. In the past, most studies have emphasized the forcing and feedback effects of atmospheric *wind* and *heat flux* on the coupled ocean-atmosphere system. Another less focused atmospheric forcing component to the ocean is freshwater flux (FWF) which has direct

_____

Correspondence to: Rong-Hua Zhang, Earth System Science Interdisciplinary Center (ESSIC), University of Maryland, College Park, MD; E-mail: rzhang@essic.umd.edu

effects on ocean salinity, an important variable in climate and the water cycle. While sea surface salinity (SSS) has no direct and immediate influence on the atmosphere, its variations can be forced by atmospheric FWF perturbations, which can further modify the oceanic density fields, the mixed layer depth (MLD), the mixing and entrainment, all of which can affect SST. For example, FWF forcing and its related salinity effect have been demonstrated to play an important role in climate variability in the North Atlantic, being recognized as a driving force for the thermohaline circulation and its fluctuations (*e.g.*, Schmitt *et al.* 1989).

In the tropical Pacific, a predominant role of wind forcing has been demonstrated in interannual climate variability associated with ENSO, involving a feedback loop among the SST, winds and the thermocline (*i.e.*, the Bjerknes feedback). The associated interannual changes in SST induce coherent fluctuations in the atmospheric circulation, including precipitation (P) and evaporation (E), whose interannual variabilities have been well documented in association with ENSO (*e.g.*, Xie and Arkin 1995;Yu and Weller 2007). These large variations of P and E are reflected in those of freshwater flux.

Over the tropical Pacific region, the major contribution to interannual variations in FWF comes from a net difference between P and E, with a dominance of the former over the latter. Indeed, associated with ENSO, interannual FWF variability shows a close relationship with SST in the tropical Pacific. During El Niño, SSTs are warm in the central and eastern equatorial Pacific, accompanied by an increase both in P and E in the central basin. Due to the dominance of P over E, a warming is associated with a positive FWF anomaly (an anomalous flux into the ocean). During La Niña, cold SST anomalies are accompanied by a reduction both in P and E in the central basin. The resultant FWF anomaly is negative (an anomalous loss of freshwater from the ocean). Thus, interannual variations in FWF present a non-local positive correlation with SST during ENSO cycles. This is contrasted to those in heat flux which have been demonstrated to have a negative correlation with SST (*e.g.*, Barnett *et al.* 1991; Wang and McPhaden 2001).

Recent studies indicate that FWF forcing and its directly related changes in salinity can play an active role in maintaining the Pacific climate and its low-frequency variability through their effects on the horizontal pressure gradients, the stratification, and the equatorial thermocline. Clearly, FWF forcing and its related feedback need to be taken into account in modeling studies due to its large interannual anomalies induced by ENSO.

At present, FWF forcing has not been adequately represented in simplified models. In most previous modeling studies, the effects of FWF forcing have been demonstrated mostly in forced ocean-alone experiments. For example, idealized anomalous FWF forcing fields are perpetually prescribed to examine the response of the ocean (*e.g.*, Reason 1992; Yang *et al.* 1999; Huang and Mehta 2004, 2005). Since the ocean-atmosphere is not coupled, there is no feedback from the changes in the ocean induced by FWF forcing to the atmosphere. Various coupled ocean-atmosphere models for the tropical Pacific have been developed for use in ENSO-related modeling studies, including intermediate coupled models (ICMs), hybrid coupled models (HCMs), and coupled general circulation models (CGCMs). However, FWF forcing has not been adequately represented in most state-of-the-art coupled models. For example, FWF has not been even included in most ICMs and HCMs used for simulation and prediction of ENSO (*e.g.*, Zebiak and Cane 1987; Barnett *et al.* 1993; Syu *et al.* 1995; Zhang *et al.* 2003, 2005, 2006; Zhang and Zebiak 2004). In CGCMs, the FWF forcing is included, but has not been realistically simulated. In particular, the so-called double ITCZ (the intertropical convergence zone) problem is still a big challenge to CGCM simulations in the tropical Pacific; most models tend to have excessive precipitation over the ITCZ in the tropical Pacific. This deficiency in precipitation simulation is reflected in the FWF field, resulting in large and systematic biases that affect the ocean. In addition, large uncertainties exist in observational estimates of P and E from different sources and products. Thus, FWF forcing remains a challenge to be represented realistically in diagnostic analyses and coupled modeling studies.

Indeed, previous modeling studies have mostly focused on the roles of atmospheric forcing components of *winds* and *heat flux* in the coupled ocean-atmosphere system of the tropical Pacific; FWF forcing and its related salinity effect on climate variability have not been getting much attention. In addition, its effects have been examined *mostly* in ocean-only modeling studies. In a coupled ocean-atmosphere system, changes in SST induced by FWF forcing can feedback to the atmosphere. But, these have not been clearly illustrated in

a coupled ocean-atmosphere context. Although CGCMs include the FWF forcing, its impact on interannual variability has rarely been diagnosed explicitly. Furthermore, FWF-induced feedback can also influence the strength of other forcings and feedbacks in the coupled system. For example, the changed SSTs induced by FWF forcing can modulate heat flux forcing which has been demonstrated to provide a negative feedback to interannual SST variability in the tropical Pacific. Then, what are the net effects of these related feedbacks on interannual variability? Moreover, ENSO has been observed to change significantly from one event to another. Many factors have been identified that can modulate ENSO amplitude (*e.g.*, Zhang and Busalacchi 2005; Zhang *et al.* 2008). As demonstrated in previous forced ocean-only simulations, FWF forcing can induce large changes in SST, indicating the potential for modulation of ENSO. However, the extent to which FWF forcing can play a role is not known.

In this work, a hybrid coupled modeling approach is taken to isolate the influences of anomalous FWF forcing on salinity and interannual variability in the tropical Pacific. The HCM developed at ESSIC (Zhang *et al.* 2006) consists of a layer ocean general circulation model (OGCM) and an empirical atmospheric model for interannual wind stress variability. In particular, as with wind forcing component, an additional empirical model has been developed to take into account interannual FWF variability that is explicitly related to SST anomalies. The FWF model is constructed from a SVD of the covariance matrix that is calculated from time series of monthly mean SST and (P-E) fields. Then, using this empirical (P-E) model, a FWF anomaly can be estimated from a given SST forcing, which can be included in the HCM to account for its related possible feedback. In addition, heat flux in the HCM is calculated using an advective atmospheric mixed layer (AML) model developed by Seager *et al.* (1995). Thus, the HCM has three atmospheric forcings to the ocean: wind stress, heat flux and freshwater flux, respectively. In this study, our focus is on the roles of anomalous FWF forcing in modulating interannual variability.

2. Model descriptions

Figure 1 shows a schematic of the various components of a HCM, recently developed at ESSIC (Zhang et al. 2006). The HCM consists of a layer OGCM and a simplified atmospheric representation of

Fig. 1 A schematic diagram illustrating a hybrid coupled model (HCM) for the tropical Pacific ocean-atmosphere system, consisting of an OGCM and a simplified atmospheric model, whose forcing fields to the ocean include three components: wind stress, freshwater flux and heat flux, respectively. The total wind stress (τ) consists of prescribed climatological wind stress (τ_{clim}) from observations and its interannual anomalies (τ_{inter}) associated with large scale SST anomalies (SST_{inter}); the total freshwater flux (FWF), represented by precipitation minus evaporation (P-E), consists of prescribed climatological freshwater flux [(P-E)$_{clim}$] and its interannual anomalies [(P-E)$_{inter}$]; heat flux (HF) is calculated using the Seager *et al.* (1995) advective atmospheric mixed layer (AML) model. Empirical submodels for τ_{inter} and (P-E)$_{inter}$ fields are constructed using a singular value decomposition (SVD) analysis. Buoyancy flux (Q_B) is calculated from the heat flux and freshwater flux to force a mixed layer model which is embedded in the OGCM. Climatological SST (**SST**$_{clim}$) fields are specified from a spinup run of the OGCM forced by observed climatological atmospheric fields.

three forcing fields to the ocean, including the two empirical submodels for interannual wind stress and FWF variability, respectively. For more details, see Zhang and Busalacchi (2009).
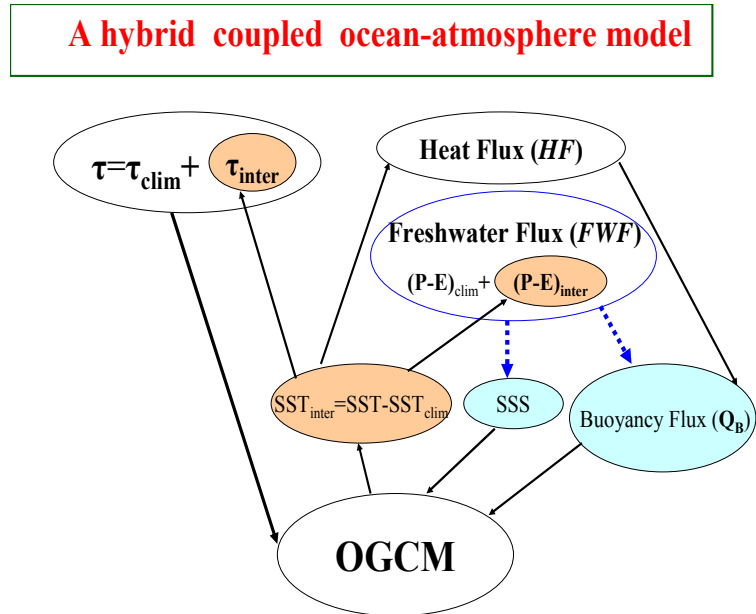
3. The effects of anomalous FWF forcing

The HCM will be used to examine the effects of FWF forcing on interannual variability in the tropical Pacific. When the FWF forcing is included in the HCM (Fig. 1), the total FWF exchange between the atmosphere and ocean can be written as: FWF = $(P-E)_{clim}$ + $\alpha_{FWF} \bullet (P-E)_{inter}$, in which its climatological part, $(P-E)_{clim}$, is specified (the P_{clim} is from observation; the E_{clim} is estimated from simualted SST_{clim} fields using the advective AML model). Its anomalous part, $(P-E)_{inter}$, is calculated using the SVD-based empirical model from interannual SST anomalies. The coefficient, α_{FWF}, represents the strength of the anomalous FWF forcing. A standard run is performed with interannual $(P-E)_{inter}$ forcing ($\alpha_{FWF} = 1$). Basically, the HCM with $\alpha_{FWF} = 1$ can quite well produce the mean ocean climatology and its interannual oscillations with about 4-year period (Fig. 2). We then perform two more HCM experiments using the identical OGCM that is coupled to the same SVD-based atmospheric wind stress and (P-E) models, but with differing α_{FWF} values to represent the strength of anomalous FWF forcing: a climatological FWF forcing run ($\alpha_{FWF} = 0.0$; the $(P-E)_{clim}$ run) and an enhanced FWF forcing run ($\alpha_{FWF} = 2.0$).
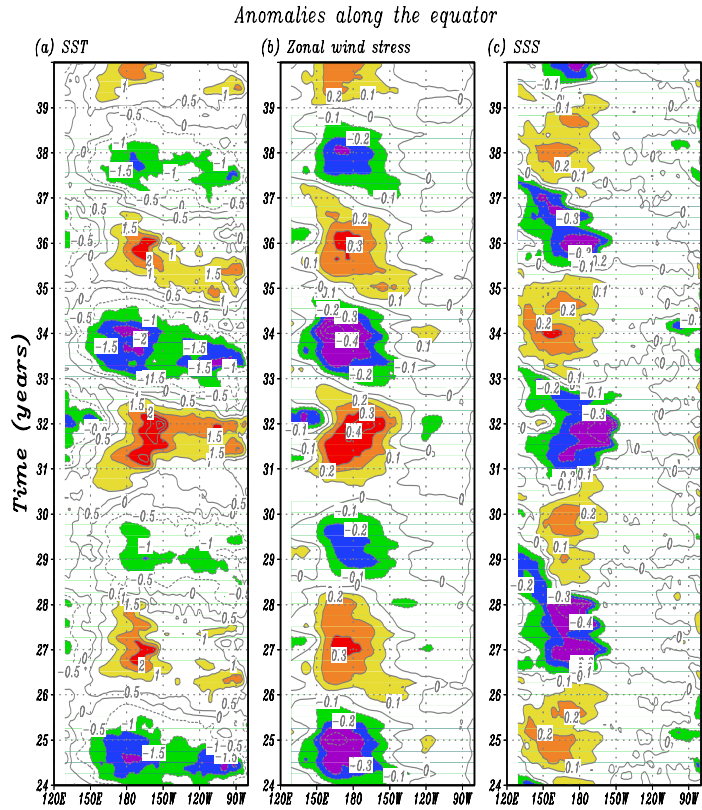
Fig. 2 Anomalies along the equator simulated from the HCM with the interannual FWF forcing: (a) SST, (b) zonal wind stress, and (c) SSS. The contour interval is 0.5 °C in (a), 0.1 dyn cm^{-2} in (b), and 0.1 psu in (c), respectively.

The effects of anomalous FWF forcing on some selected variables are quantified in Table 1. As analyzed above, the relationships among interannual anomaly fields indicate a positive effect of FWF forcing on SST during ENSO cycles. In the standard $(P-E)_{inter}$ run ($\alpha_{FWF} = 1.0$), the positive feedback between SST and FWF is included in the HCM simulation. The FWF anomalies induce additional ocean processes in such a way to reinforce the warming during El Niño and cooling during La Niña, respectively. The enhanced SST anomalies further increase interannual variability in the coupled system. When the positive SST-FWF relationship is not included as in the climatogical FWF forcing

$Nino4$ region	$\alpha_{FWF} = 0.0$ (Clim run)	$\alpha_{FWF} = 1.0$ (Standard run)	$\alpha_{FWF} = 2.0$ (Enhanced run)
SSS	0.11	0.16	0.28
SST	0.76	0.85	0.97
MLD	5.6	6.7	9.0
τ	0.16	0.19	0.23
Q_T	1.49	1.72	1.95
Q_S	0.0	0.65	1.58
Q_B	1.49	1.24	1.08
$Nino12$ SST	0.53	0.57	0.64
$Nino3$ SST	0.67	0.76	0.92

Table 1 The standard deviation of some selected anomaly fields from the HCM simulations with the climatological ($\alpha_{FWF} = 0.0$), the standard ($\alpha_{FWF} = 1.0$) and enhanced ($\alpha_{FWF} = 2.0$) FWF forcings, respectively. Shown at the Niño4 region are for SSS, SST, MLD, zonal wind stress (τ), buoyancy flux (Q_B) and its heat flux part (Q_T) and freshwater flux part (Q_S). Also shown in the last two rows are for SST at the Niño12 and Niño3 sites. The unit is: psu for SSS, °C for SST, meter for MLD, dyn cm^{-2} for τ, 10^{-6} Kg s^{-1} m^{-2} for Q_B and Q_T and Q_S, respectively.

run (α_{FWF} = 0.0), these additional oceanic effects that could be induced by the FWF forcing are disabled, and there is no positive feedback. As a result, the simulated interannual variability in the α_{FWF} = 0.0 run is weakened as compared with that in the α_{FWF} = 1.0 run. When the anomalous FWF forcing is enhanced as represented in the α_{FWF} = 2.0 run, its obvious direct effects are to increase the temporal variability of SSS and MLD in the central basin. Also, it increases the compensating effect of Q_S on Q_T, with a net reduction in Q_B variability. All these processes tend to cause more warming during El Niño and more cooling during La Niña, which acts to reinforce SST variability during ENSO cycles. As the positive feedback is exaggerated, a stronger interannual variability emerges. Clearly, the oceanic processes induced by the FWF forcing and the related feedback act in such a way to enhance the strength of ENSO cycles.

4. Concluding remarks

Most previous modeling studies have focused on the roles of atmospheric forcing components of *winds* and *heat flux*; FWF forcing and its related salinity role in coupled climate variability have not received much attention. Furthermore, the effects of FWF forcing have been examined mostly in *forced ocean-only* modeling studies. In this work, the impacts of FWF forcing on salinity and interannual variability are examined in a hybrid coupled ocean-atmosphere context in which climatological atmospheric forcing fields of wind and FWF are specified, while their anomaly parts can be added on or off separately or collectively.

We have designed various experiments using the HCM with differing strengths of anomalous FWF forcing. Three cases are considered. In a standard simulation, the climatological and anomalous FWF fields are both taken into account [i.e., FWF = (P-E)$_{clim}$+(P-E)$_{inter}$]. The constructed HCM can well reproduce interannual variability associated with ENSO in the tropical Pacific.   Two more sensitivity experiments are then performed using the HCM with the climatological forcing only [i.e., FWF = (P-E)$_{clim}$] and an enhanced FWF forcing run [i.e., FWF = (P-E)$_{clim}$+ 2.0•(P-E)$_{inter}$], respectively.

Interannual variability in the tropical Pacific is predominantly driven by wind-induced feedback; simulations with climatological FWF forcing still show large interannual variability of SSS and SST in the tropical Pacific with basic feature unchanged. This indicates that wind forcing is of primary importance for dynamics of interannual variability in the tropical Pacific climate system. However, a significant effect can be seen, arising from anomalous FWF forcing. Quantitatively, taking the (P-E)$_{inter}$ run as a standard, the SST variance at the Niño3 site can be reduced by about 12% in the climatological FWF forcing run, but enhanced by 21% in the enhanced FWF forcing run; the   variances for SSS and zonal wind stress at the Niño4 site are reduced by about 31% and 16% in the climatological FWF run, but enhanced by 16% and 75% in the enhanced FWF forcing run, respectively. Thus, anomalous FWF forcing can modulate interannual variability in a substantial way.

References

Zhang, R.-H., T. Kagimoto, and S. E. Zebiak, 2001: Subduction of decadal North Pacific thermal anomalies in an ocean GCM, *Geophys. Res. Lett.,* **28**, 2449-2452.

Zhang, R.-H., A. J. Busalacchi, and R. G. Murtugudde, 2006: Improving SST anomaly simulations in a layer ocean model with an embedded entrainment temperature submodel, *J. Climate,* **19**, 4638–4663.

Zhang, R.-H., and A. J. Busalacchi, 2009: Freshwater flux (FWF)-induced oceanic feedback in a hybrid coupled model of the tropical Pacific. *J. Climate*, **22**, 853–879.

Seasonal Prediction with CCSM:
Impact of Atmosphere and Land Surface Initialization

James L. Kinter III[1,2], Dan Paolino[1], David Straus[1,2], Ben Kirtman[3] and Dughong Min[3]

[1]*Center for Ocean-Land-Atmosphere Studies, Calverton, Maryland*
[2]*Department of Atmospheric, Ocean and Earth Sciences, George Mason University, Fairfax, Virginia*
[3]*Rosenstiel School of Marine & Atmospheric Science, University of Miami, Miami, Florida*

1. Background

As shown by Shukla and Mintz (1982) and many subsequent studies (*e.g.*, Delworth and Manabe 1988; Atlas *et al.*, 1993), the presence or absence of water at the land surface can have a profound effect on the seasonal climate. Koster *et al.* (2004) showed that that effect can vary considerably over the planet, with its most pronounced impacts occurring in the semi-arid transitional zones that often lie between humid regions and deserts. They also showed that the various coupled land-atmosphere models used to gauge this effect disagree considerably on its magnitude. These results have also been borne out in several subsequent studies, which indicate that the presence or absence of moisture in the soil enhances the predictability at intraseasonal and longer time scales (*e.g.*, Wang and Kumar, 1998; Fennessy and Shukla, 1999; Yang *et al.* 2004).

As part of a project is to explore the efficacy of using the Community Climate System Model (CCSM), in conjunction with the NOAA Climate Forecast System (CFS), to provide a multi-model ensemble of climate predictions that are superior to predictions made with either model alone, we have developed methods to initialize the land and atmosphere model components of the CCSM and CCSM with observed states. Using that process, along with the ocean initial state produced at the University of Miami, we produced a set of one-year re-forecasts. In parallel to this effort, we produced similar re-forecasts for which only the Parallel Ocean Program (POP), the ocean model component of CCSM, was initialized.

2. Experimental setup

Model: The model used in this study is the CCSM3 (Collins *et al.*, 2006), which is a coupled ice-oceanatmosphere- land climate model with state-of-the-art formulations of dynamics and subgrid-scale physical parameterizations. The atmosphere is Community Atmospheric Model (CAM3, Eulerian dynamical core) at T85 (~150 km) horizontal resolution with 26 vertical levels. Experiments have been conducted with both the CAM3.0 and CAM3.5 versions of the subgrid-scale physical parameterizations, referred to as the CCSM3.0 and CCSM3.5 experiments, respectively. The ocean model is the standard version of POP with 1° resolution, stretched to 1/3° near the equator.

Re-forecast Experiments: Retrospective forecasts cover the period 1981-2000 (1982–1998) with initial states in January (July). One set of runs was made with observed initial states for the global ocean (OCN-only; Kirtman and Min, 2009), and one set of runs was made with full initialization of the global atmosphere, ocean, and land surface (ATM-OCN-LND). Ensembles of 4 (10) and 6 (10) hindcasts were run in the OCN-only (ATM-OCN-LND) experiments for the January and July cases, respectively.

Ocean Initialization: In all experiments, the ocean initialization uses the GFDL ocean data assimilation system, based on the MOM3 global ocean model with a variational optimal interpolation scheme. The GFDL ocean initial states were interpolated (horizontally and vertically) to the POP grid using a bi-linear interpolation scheme. (Climatological data from long simulations of CCSM3 were used poleward of 65°N and 75°S.) The ocean initial state is identical for each ensemble member.

Atmosphere and Land Initialization: In the OCN-only experiments, the atmospheric and land surface initial states were taken from an extended atmosphere/land-only (CAM3) simulation with observed, prescribed SST. The atmospheric ensemble members were obtained by resetting the model calendar back one week and integrating the model forward one week with prescribed observed SST. In this way, it is possible to

generate initial conditions that are synoptically independent (separated by one week) but have the same initial date. Thus all ensemble members were initialized at the same model clock time (1 Jan or 1 July) with independent atmospheric initial conditions.

In the ATM-OCN-LND experiments, land and atmosphere were initialized for each of the 10 days preceding the date of each ocean initial state - 22-31 December for the 1 January ocean states, and 22-30 June for the 1 July ocean dates. The atmosphere was initialized by interpolating from daily reanalysis data. The land surface was initialized from daily Global Soil Wetness Project analyses (GSWP-2; 1986- 1995) and daily ERA40 (1982-1985 and 1996-1998). The observed anomalies were superimposed on a climatology for the Common Land Model (CLM), which is a component of the CCSM. The snow depth was initialized from ERA40, and the sea-ice was initialized to climatological monthly conditions based on a long simulation of CCSM3.0.

3. Results

Using the results of both sets of re-forecasts (OCN-only and ATM-OCN-LND), we performed an analysis with an eye to gauging the benefits of initializing CCSM3.0 with the observed atmospheric, land and ocean states in comparison to initializing with only the observed ocean state. We expect that a large part of the monthly to seasonal predictability in the atmosphere and land as simulated by the CCSM will be forced by temperature anomalies at the ocean surface. The working hypothesis is that a major factor in any enhanced predictability in the ATM-OCN-LND re-forecasts will be driven by long-term, large-scale anomalies of soil moisture. Therefore, we have focused on the predictability of the land surface and near surface variables.

Fig. 1 (Top row) Forecasts at 7-month lead time of the monthly mean soil moisture anomaly in the top 9 cm of soil in the contiguous U.S. (Bottom row) Analyzed values of soil moisture in the top 10 cm, based on the Global Soil Wetness Program (GSWP) analysis. In both rows, the left panels are for July 1993 and the right panels are for July 1988. All anomalies are normalized by their respective standard deviations, based on 18 years of data, at each grid point.

As an example of what can be gained by initializing the land surface, Fig. 1 shows the soil moisture anomalies at 7-month lead time for forecasts initialized in January 1993, an extreme flood year, and January 1988, an extreme drought year. In 1993, the upper Mississippi valley was well above normal soil moisture with positive anomalies to the northwest and anomalously dry conditions to the south and east. During the 1988 drought, the center of the dry anomaly was in the northeastern U.S., with dry conditions extending to the west through the upper Mississippi valley and the high plains of the northwest. Wetter than normal conditions were present in the semi-arid and desert region of west Texas and the southwest. The predicted pattern of anomalies, and to a lesser extent the predicted intensity, closely resemble the observed pattern in both cases, except the wet anomaly in the southwest.

More broadly, Fig. 2 shows the correlation of observed and first month forecast of soil moisture in the top 3 layers of the CLM, for both ATM-OCN-LND and OCN-only forecasts, and for both experiments, from the end of December and the end of June ICs. The observations are taken from the ERA-40, and represent the first layer of the TESSEL soil model, which has a depth of 7 cm. We used this same data to initialize CLM for 1981-1985 and 1996-1999. A different land surface data set from the GSWP-2, was used for 1986-1995. The forecasts with initialized soil moisture anomalies produce a much better forecast of soil moisture anomalies in
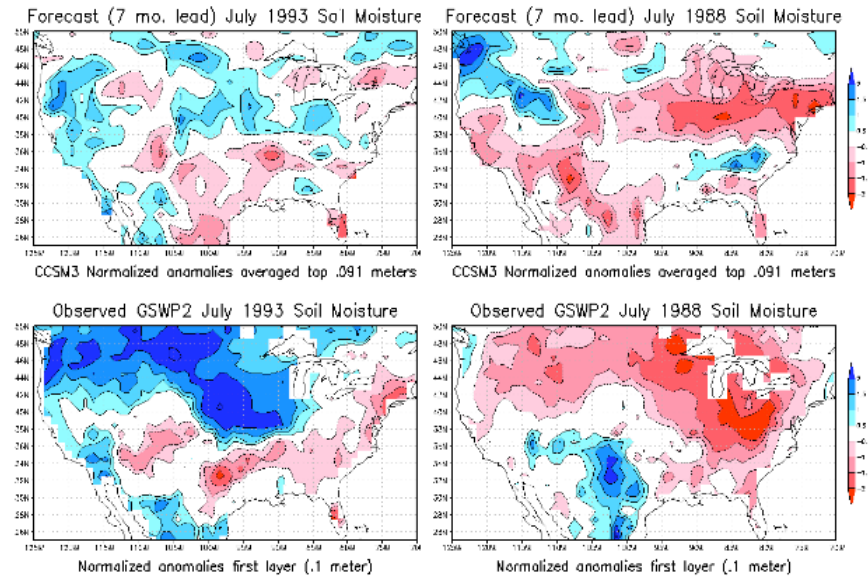
the first month. This is especially impressive, considering that these top three layers represent a total depth of only 9 cm. These maps also highlight an obvious benefit from initializing with observed snow depth. Those areas with persistent winter snow cover (and presumably frozen soil underneath) will tend to preserve their initial soil moisture anomalies. Similar correlation maps for the mid-layer (9-29 cm) CLM soil moisture (not shown) show the same results for month one, with generally higher positive correlation.

Similar, though less impressive results, are present in the forecasts of the first season (JAS, JFM) soil moisture (not shown). Longer range forecasts are adversely affected by the tendency of the simulated soil moisture in CLM3.0 to dry out over time. We hope that improvements in the CLM3.5 will reduce these systematic errors.

Figure 3 shows the correlation of the 2-meter temperature for the first month of the re-forecasts with observed surface temperature as represented by the CAMS dataset (Ropelewski *et al.*, 1985). Temperatures over the ocean have been masked out in order to focus on the land; since the two forecasts begin with the same ocean initial state, the correlations are close to identical. The simulation of the land surface temperature is

Fig. 2 A) Correlation of January monthly soil moisture in the top 9 cm from the CCSM3.5 Atm+Lnd+Ocn forecast initialized end of December versus January ERA-40 soil moisture in the top 7 cm; for Jan. 1981-1998. B) As in Fig. 2A, but for CCSM3.5 Atm+Lnd+Ocn forecast initialized end of June versus July ERA-40 soil moisture; for July 1982-1998. C) As in Fig. 2A, but for the CCSM3.5 Ocn only forecast initialized 1 January. D) As in Fig. 2A, but for the CCSM3.5 Ocn only forecast initialized 1 July. Shading indicates correlations significant at 95% and 99% levels.
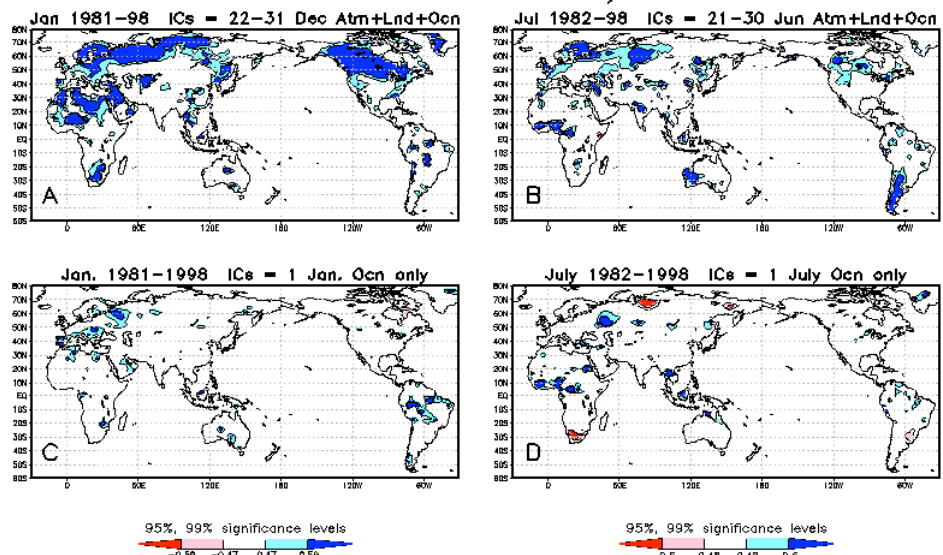
Fig. 3 A) Correlation of January monthly 2 meter temperature over land from the CCSM3.5 Atm+Lnd+Ocn forecast initialized end of December versus CAMS observed surface temperature, for Jan. 1981-1998. B) As in Fig. 3A, but for the CCSM3.5 Atm+Lnd+Ocn forecast initialized end of June; for July 1982-1998. C) As in Fig. 3A, but for the CCSM3.5 Ocn only forecast initialized 1 January. D) As in Fig. 3B, but for the CCSM3.5 Ocn only forecast initialized 1 July. Shading indicates correlations significant at 95% and 99% levels.

clearly superior for the ATM-OCN-LND case, especially for the forecasts from 1 January ICs. In general,

those areas in Fig. 2 with a good forecast of soil moisture tend also to have a good forecast of surface temperature. Figure 3 again suggests that there may be some benefit derived from the initialization of the snow depth. The correlation over snow-covered areas is generally good (*e.g.*, note southern South America in Fig. 3b.) Also note the significant correlation over northwest Europe, where the correlation of forecast and observed soil moisture was not significant.
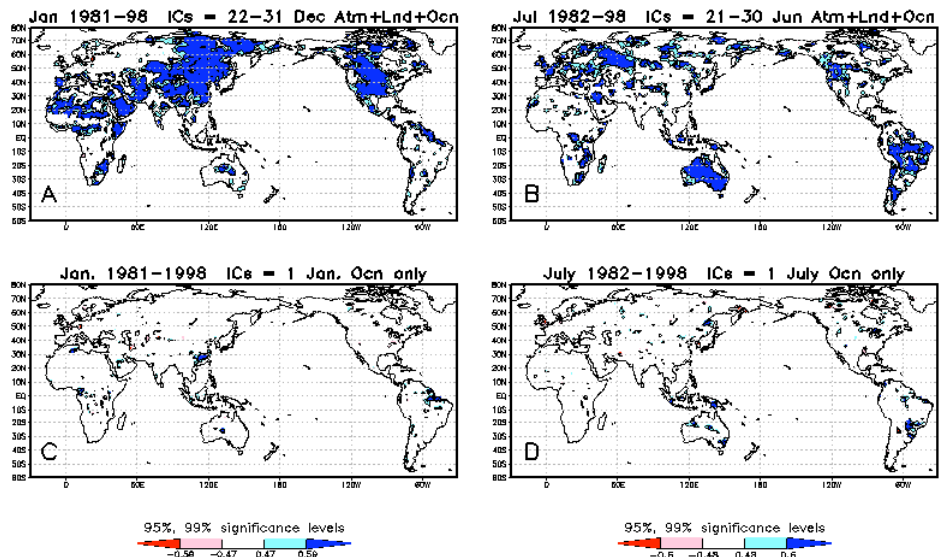

Fig. 4 A) Correlation of January monthly precipitation from the CCSM3.5 Atm+Lnd+Ocn forecast initialized end of December versus CMAP observed precipitation, for Jan. 1981-1998. B) As in Fig. 4A, but for the CCSM3.5 Atm+Lnd+Ocn forecast initialized end of June; for July 1982-1998. C) As in Fig. 4A, but for the CCSM3.5 Ocn only forecast initialized 1 January. D) As in Fig. 4B, but for the CCSM3.5 Ocn only forecast initialized 1 July. Shading indicates correlations significant at 95% and 99% levels.

Figure 4 shows the first month forecast of total precipitation versus observation as represented by CMAP (Xie and Arkin, 1997). There is little evidence that the ATM-OCN-LND initialization has provided much improvement in the forecast of precipitation over land, with the possible exceptions of Brazil and Australia in the end-of-June forecasts. Curiously, there is evidence of an improvement in the simulation of precipitation over the oceans, particularly in the extra-tropics in winter. If this improvement were to have arisen from the difference in initialization, it would seem more likely to be due to the atmospheric initialization, which might impart some skill to the first month's forecast. It might also be just an artifact of the smaller sample in the OCN-only forecasts, which had only four and six members for the 1 January cases and 1 July cases, respectively. We will investigate further.

We studied specific instances of the ability of our forecast system to simulate climate and predictability of seasonal anomalies. For example, indices of FMA rainfall over the Nordeste region of Brazil for both the ATM-OCN-LND and OCN-only forecast precipitation (Fig. 5) provide a good approximation of the variability of an identical index

Fig. 5 Time series indices of an index of February – April rainfall over the Nordeste region of Brazil for the OCN-only forecasts (blue curve), the ATM-OCN-LND forecasts (red curve), in comparison with the same index computed from the CMAP precipitation analysis (black curve).

constructed from observed rainfall. We also note that the climatology of the monsoon rainfall over India is well simulated by both sets of forecasts (not shown); although both of the forecasts do a poor job of reproducing the interannual variability.
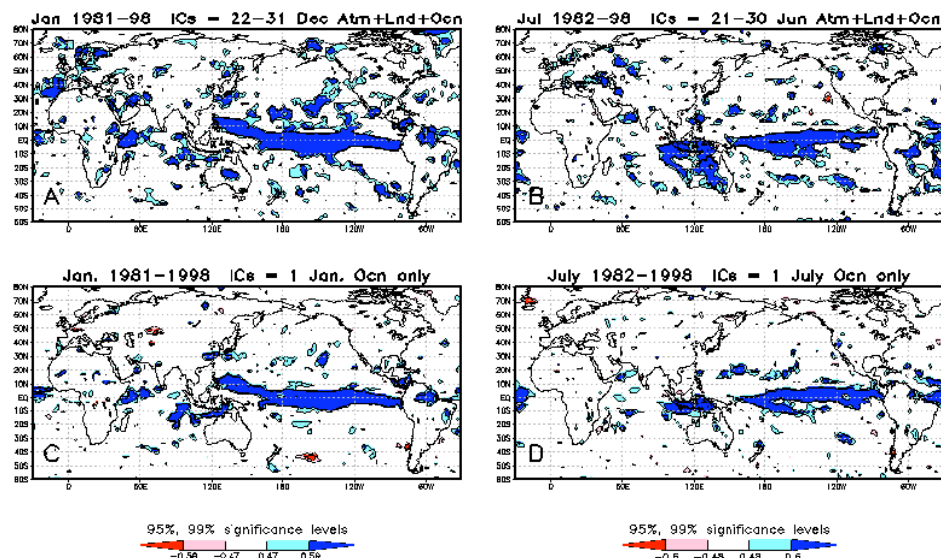
4. Future work - CCSM3.5

We have updated and improved our initialization methods for CCSM3.5, and have begun producing retrospective forecasts.

References

Atlas R., N. Wolfson, and J. Terry, 1993: The effect of SST and soil moisture anomalies on GLA model simulations of the 1998 U.S. summer drought. *J. Climate*, **6**, 2034–2048.

Collins, W.D., C.M. Bitz, M.L. Blackmon, G.B. Bonan, C.S. Bretherton, J.A. Carton, P. Chang, S.C. Doney, J.J. Hack, T.B. Henderson, J.T. Kiehl, W.G. Large, D.S. McKenna, B.D. Santer, and R.D. Smith, 2006: The Community Climate System Model Version 3 (CCSM3). *J. Climate*, **19**, 2122–2143.

Delworth T. L., and S. Manabe, 1988: The influence of potential evaporation on the variability of simulated soil wetness and climate. *J. Climate*, **1**, 523–547.

Fennessy M. J., and J. Shukla, 1999: Impact of initial soil wetness on seasonal atmospheric prediction. *J. Climate*, **12**, 3167–3180.

Kirtman B. P. and D. Min, 2009: Multi-Model Ensemble ENSO Prediction with CCSM and CFS. *Mon. Wea. Rev.*, (in press).

Koster, R. D., P. A. Dirmeyer, Z. Guo, G. Bonan, E. Chan, P. Cox, H. Davies, T. Gordon, S. Kanae, E. Kowalczyk, D. Lawrence, P. Liu, S. Lu, S. Malyshev, B. McAvaney, K. Mitchell, T. Oki, K. Oleson, A. Pitman, Y. Sud, C. Taylor, D. Verseghy, R. Vasic, Y. Xue, and T. Yamada, 2004: Regions of strong coupling between soil moisture and precipitation. *Science*, **305**, 1138-1140.

Ropelewski, C. F., J. E. Janowiak and M. F. Halpert, 1985: The analysis and display of real time surface climate data. *Mon. Wea. Rev.*, **113**, 1101-1107.

Shukla, J. and Y. Mintz, 1982: Influence of land-surface evapotranspiration on the Earth's climate. *Science*, **215**, 1498-1501.

Wang W., and A. Kumar, 1998: A GCM assessment of atmospheric seasonal predictability associated with soil moisture anomalies over North America. *J. Geophys. Res*, **103**, 28637–28646.

Xie and Arkin, 1997: Global Precipitation: A 17-Year Monthly Analysis Based on Gauge Observations, Satellite Estimates and Numerical Model Outputs. *Bull. Amer. Meteor. Soc*, **78**, 2539-2558.

Yang, F., A. Kumar, K.-M. Lau, 2004: Potential predictability of U.S. summer climate with "perfect" soil moisture. *J. Hydrometeor.*, **5**, 883-895.

How Much Do Different Land Models Matter for Climate Simulation?

Jiangfeng Wei, Paul A. Dirmeyer, Zhichang Guo, Li Zhang, and Vasubandhu Misra

Center for Ocean-Land-Atmosphere Studies, Calverton, Maryland

1. Introduction

One of the largest uncertainties in climate simulations is from the representation of land processes, because there are few observations to calibrate or constrain it. Different land surface schemes (LSSs) use quite different parameterizations to describe the complex hydrological, biogeophysical, and biogeochemical processes. Even when forced by the same atmospheric forcing and provided the same parameter settings, different LSSs can still give significantly different surface fluxes. When these LSSs are coupled to the Atmospheric General Circulation Models (AGCMs), their different behaviors will bring uncertainties into the simulated climate. As the land-atmosphere system is nonlinear, uncertainties from LSSs can be amplified or reduced during land-atmosphere interaction. This problem is systemically addressed in this study. In addition to the climatology and variability, different LSSs can lead to different coupling strength between land and atmosphere (i.e., contribution of land to prediction of atmosphere). Within the framework of Global Land-Atmosphere Coupling Experiment (GLACE), we perform GLACE-type experiments to investigate this problem.

Fig. 1 The ratio Φ (see equation 1) for 1987-2004 average JJA latent (LH; upper panel) and sensible (SH; bottom panel) heat fluxes. The regions enclosed by blue boxes are for further analysis in Figure 2.

In this study, we show results from COLA AGCM coupled to three state-of-the-art LSSs: SSiB, CLM3.5, and Noah. Two experiments are performed. In the first experiment (I), three LSSs are coupled to the AGCM individually. In the second experiment (C), the three LSSs are coupled to the AGCM in combination, i.e. the LSSs receive the same atmospheric forcing from the AGCM and the average surface fluxes from the LSSs are passed back to the AGCM at each grid point and at every time step. Experiment C is similar to three land model offline experiments with a same atmospheric forcing, but this forcing is affected by the average feedback from the LSSs.

We try to investigate the uncertainties of the three LSSs and their influence on climate simulation. We also explore the influence of land-atmosphere coupling on the simulation uncertainties. In addition, GLACE-type experiments with the COLA AGCM coupled to three land models are performed. By comparing the coupling strength of the three coupled models, we can know the impact of different land models on the coupling strength. In summary, the purpose of this study is threefold: firstly, to investigate current uncertainties in the behavior of LSSs; secondly, to investigate how much these uncertainties can influence
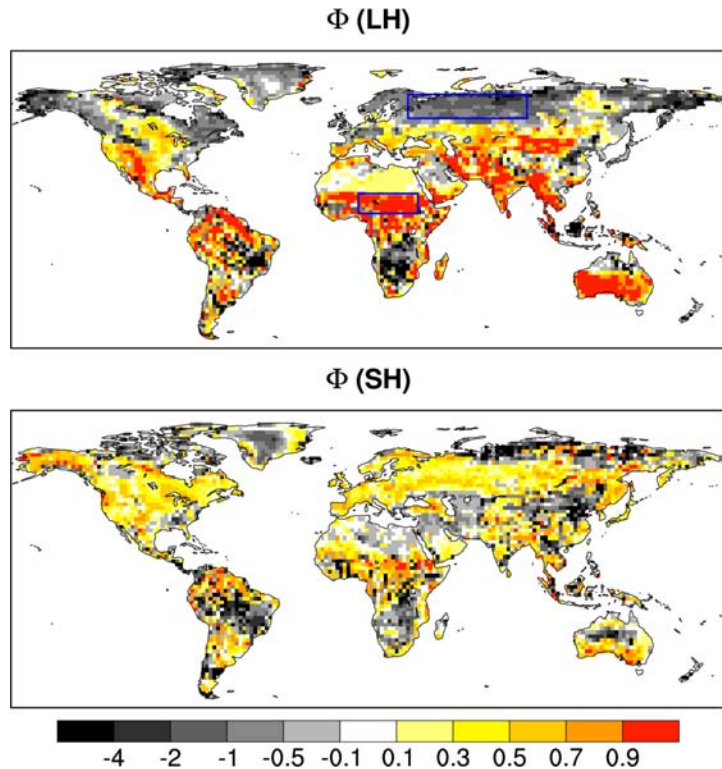
Correspondence to: Jiangfeng Wei, Center for Ocean-Land-Atmosphere Studies, 4041 Powder Mill Road, Suite 302, Calverton, MD 20705; E-mail: jianfeng@cola.iges.org

atmospheric simulation through land-atmosphere interaction; thirdly and most importantly, to have a better understanding of the mechanisms of land-atmosphere coupling.

2. Climatology

The three LSSs produce significantly different surface fluxes over most of the land, no matter whether they are coupled individually (different forcing to land) or in combination (same forcing to land). See Wei et al. (2009a) for a detailed discussion. A question is whether land-atmosphere interaction can amplify the uncertainties from LSSs if they are coupled to an AGCM.

Let $Var(I)$ and $Var(C)$ be the inter-model (3 cases) variances of fluxes from land to atmosphere in experiments I and C, respectively. Intuitively, $Var(I)$ should be larger than $Var(C)$ because the LSSs receive the same atmospheric forcing in C but not in I. Thus $Var(I)$ is the inter-model variance caused by LSS differences and land-atmosphere feedback, while $Var(C)$ is the variance caused by LSS differences only. Then the ratio

$$\Phi = \frac{Var(I) - Var(C)}{Var(I)}$$

is the percentage of inter-model variance caused by land-atmosphere feedback. If $Var(I) \geq Var(C)$, $0 \leq \Phi \leq 1$. However, if $Var(I) < Var(C)$ ($\Phi < 0$), a negative feedback between land and atmosphere is implied

Fig. 2 The simulated 1987-2004 JJA average LH, SWnet, total cloud cover, and precipitation for northern Eurasia (left column) and Sahel (right column). The areas of the two regions are marked by blue boxes in Figure 1.

and we cannot estimate the relative contributions of LSS differences and land-atmosphere interactions to the variance.

Figure 1 shows Φ averaged over JJA for sensible heat (SH) and latent heat (LH) fluxes. Over most land area, $0 \leq \Phi \leq 1$. However, there are still some areas with $\Phi < 0$. SH should have the same inter-model variance as LH if Rnet and the relatively small ground heat flux are the same for the LSSs. However, Rnet differs a lot among the models over some high latitude regions and dry regions. This is why the Φ values of SH and LH differ most over these regions (Figure 1). In order to investigate the cause of the different spread changes (positive and negative Φ), we selected the northern Eurasia and Sahel as two regions with contrasting values of Φ (blue boxes in Figure 1). Figure 2 shows the time series of LH, net shortwave radiation at surface (SW$_{net}$), total cloud cover, and precipitation over these two regions. It is evident that, compared to experiment I, the LH in experiment C strongly converge in Sahel but diverge in northern Eurasia, consistent with the value of Φ. In Sahel, the interannual time series of LH are negatively correlated with those of SW$_{net}$ but are positively correlated with those of cloud cover and precipitation. This is a semi-arid, moisture-limited area, where evapotranspiration (ET) is nominally below the potential rate, so LH is strongly controlled by the land surface states, especially soil wetness, which is largely determined by rainfall. In experiment C, each LSS experiences the same rainfall, which leads to similar soil wetness and LH. In northern Eurasia, however, the correlation
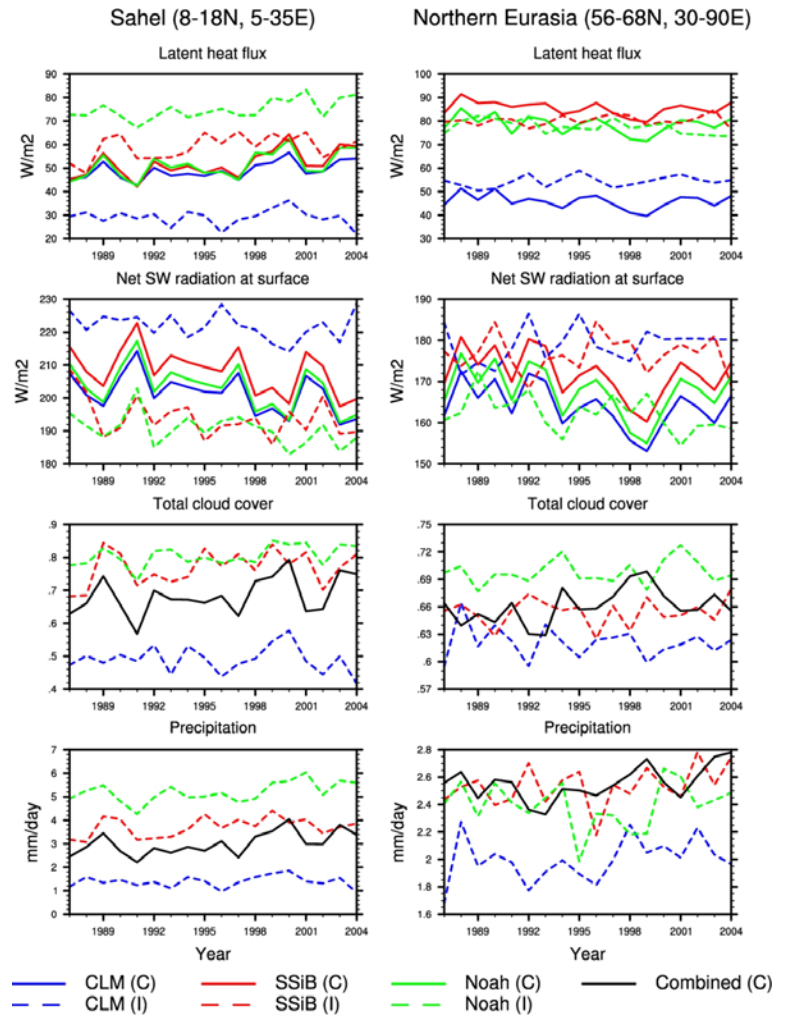
between SW$_{net}$ and LH is positive for most of the time series. The soil moisture is plentiful in this region and the control on LH is mainly the radiation at surface.

3. Variability

The memories inherent in the surface heat fluxes differ greatly among the LSSs (see Wei et al. (2009a) for a detailed discussion). It would be interesting to examine how the different memories of land surface fluxes can influence precipitation variability. Figure 3 shows the lag-2-pentad precipitation autocorrelation in JJA. This method has been used in previous studies and is based on the assumption that a wetter soil caused by a storm may last a few days and promote future storms (Koster et al. 2003). However, there is also possibility that this persistence of precipitation is caused by the internal atmospheric dynamics or some other external forcing (e.g., SST) and has nothing to do with soil moisture memory (Wei et al. 2008).

Fig. 3 The JJA lag-2-pentad autocorrelation of precipitation across 1987-2004. (a) COLA-SSiB. (b) COLA-CLM. (c) COLA-Noah. (d) Combined experiment. (e) Calculated with the average precipitation of the three individually coupled simulations. (f) From the observational dataset of GPCP (Xie et al. 2003). The model results are interpolated to the same grid as that of GPCP data (2.5°×2.5°). Values larger than 0.11 are over 95% confidence level. Seasonal cycles are not removed in this calculation; removing them can lead to results with similar patterns but smaller amplitude.

It can be seen in Figure 3 that all the model simulations show a largely similar pattern of precipitation persistence, but regional differences between models exist. The result from the combined simulation is within the range of the three individually coupled simulations. The average of the three individual simulations shows a precipitation persistence larger than any of the individual simulations because the averaging tend to suppress the short time scale precipitation variations that are inconsistent among the models. Although the memories of surface LH and SH are much lower in Noah (not shown), it does not show an overall lower precipitation persistence than the other two models in the individually coupled simulations. This suggests that the land surface heat fluxes do not play a dominant role in the global pattern of precipitation variability, but regional impacts may still exist. Compared to the observation, all the simulations here have overestimated the precipitation persistence in many areas.

4. Land-atmosphere coupling strength and its relationship to precipitation variability

Figure 4 shows the GLACE Ω values of total precipitation for ensembles W and S and their difference Ω(S)-Ω(W) (see definitions in Appendix). No matter which LSS the AGCM is coupled with, Ω show similar patterns, with largest values in the tropical rain belt, where the SST forcing has strongest influence. The patterns of W and S are very close, with large differences (Ω(S)-Ω(W)) mainly over the regions with high Ω values. This indicates that the land-atmosphere coupling strength is strongly influenced by external forcing. Globally, the COLA-SSiB has the strongest land-atmosphere coupling strength, while the coupling strength
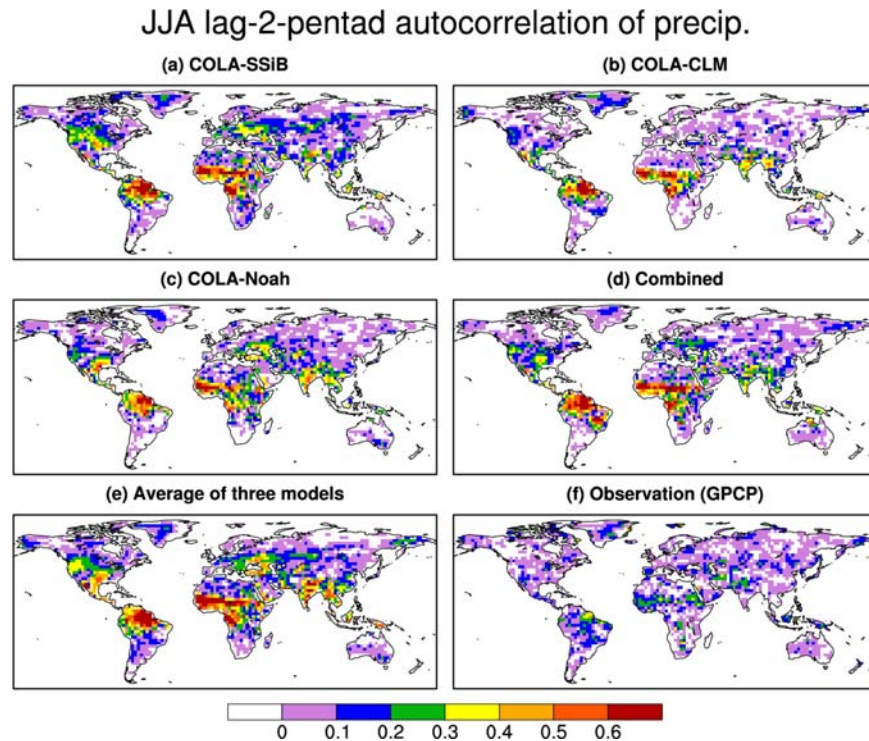
for COLA-Noah is the weakest. The difference should be mainly from the different land models because they are coupled to the same AGCM.
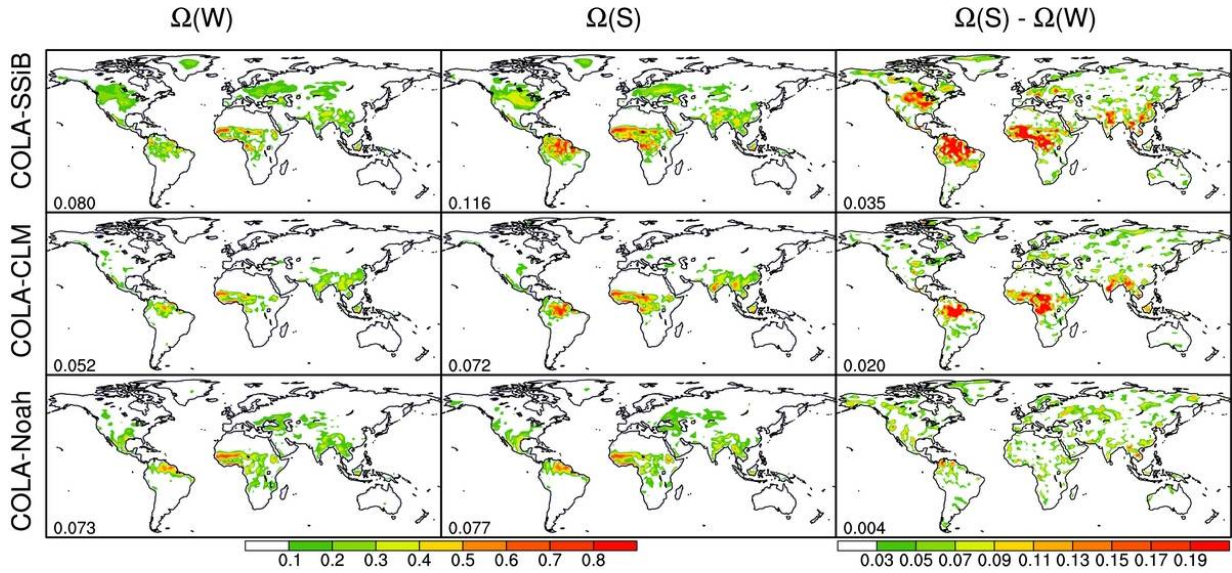

Fig. 4 The GLACE parameter Ω for precipitation from ensembles W (left column) and S (middle column), and their difference (right column). Top row: COLA-SSiB. Middle row: COLA-CLM. Bottom row: COLA-Noah. The global mean value of each panel is shown at the left corner.

It is shown in Wei et al. (2009b) that most of the precipitation predictability (Ω) and land-atmosphere coupling strength (Ω(S)-Ω(W)) are associated with the intraseasonal component of precipitation in the models, although they only account for a small percentage (~20%) of the total variance. The GLACE coupling strength can be conceptually decomposed into the impact of the slow varying external forcing (F) and the local impact of soil moisture. The external F and local soil moisture combine to determine the pattern of the coupling strength. From the output of the GLACE models, we find that most models have overestimated the low-frequency variance percentage and underestimated the high-frequency variance percentage of precipitation. It suggests that the specific mode of land-atmosphere coupling described in GLACE may be over-represented in the models. Based on the findings in this study, we adjust the land-atmosphere coupling strength estimated by GLACE. It is found that the adjusted coupling strengths are generally weaker than that from GLACE but the patterns are nearly the same.

Appendix: Global Land-Atmosphere Coupling Experiment (GLACE)

GLACE (Koster et al. 2004, 2006) is a model intercomparison study focusing on evaluation of the role of land state in numerical weather and climate prediction. It consists of three sets of 16-members ensembles of AGCM experiment: W, R and S. We only discuss W and S two sets here. Ensemble W is an ensemble of free runs with different initial land and atmosphere conditions but forced by the same SST of 1994; ensemble S is the same as ensemble W except that, at each time step, the subsurface soil moisture in the land model is replace by that from one member chosen from ensemble W. All runs cover the period of 1 June 1-31 August, 1994. A diagnostic variable Ω is defined:

$$\Omega = \frac{16\sigma^2_{<X>} - \sigma^2_X}{15\sigma^2_X}$$

where σ^2_x is the intraensemble variance of variable x, and $\sigma^2_{<x>}$ is the corresponding variance of ensemble mean time series. In calculating the variance, the first 8 days of data of each run is discarded to avoid model initial shock, and the remaining 84 days are aggregated into 14 six-day totals. Therefore, σ^2_x is a variance of 224 (16x14) six-day totals from all the ensemble members, and $\sigma^2_{<x>}$ is a variance of 14 six-day totals from the ensemble mean time series.

Theoretically, if the 16 members of an ensemble have the same time series of x, σ^2<x>will be equal to σ^2_x and Ω will be 1; if the x time series of the 16 members are completely independent, σ^2<x> will be equal to $\sigma^2_x/16$ and Ω will be 0. Without sampling error, Ω will be between 0 and 1. Ω measures the similarity (or predictability) of the time series in 16 ensemble members, and is equivalent to the percentage of variance caused by the slowly varying oceanic, radiative, and land surface processes. The difference of Ω from the two ensembles, $\Omega(S)-\Omega(W)$, is equivalent to the percentage of variance caused by the prescribed subsurface soil moisture, and is a measure of land-atmosphere coupling strength in GLACE.

References

Koster, R. D., M. J. Suarez, R. W. Higgins, and H. M. Van den Dool, 2003: Observational evidence that soil moisture variations affect precipitation, *Geophys. Res. Lett.*, **30(5)**, 1241, doi:10.1029/2002GL016571.

Koster, R. D., and Coauthors, 2006: GLACE: The Global Land-Atmosphere Coupling Experiment. Part I: Overview, *J. Hydrometeorol.*, **7**, 590–610.

Wei, J., P. A. Dirmeyer, Z. Guo, L. Zhang, and V. Misra, 2009a: How much do different land models matter for climate simulation? Part I: Climatology and variability. *COLA Tech. Rep.* 273. 35pp. Center for Ocean-Land-Atmosphere Studies, Calverton, MD. [Available online at ftp://grads.iges.org/pub/ctr/ctr273_ms.pdf]

Wei, J., P. A. Dirmeyer, and Z. Guo, 2009b: How much do different land models matter for climate simulation? Part II: A decomposed view of land-atmosphere coupling strength. *COLA Tech. Rep.* 274. 27pp. Center for Ocean-Land-Atmosphere Studies, Calverton, MD. [Available online at ftp://grads.iges.org/pub/ctr/ctr274_ms.pdf]

Wei, J., R. E. Dickinson, and H. Chen, 2008: A negative soil-moisture-precipitation relationship and its causes, *J. Hydrometeor.*, **9**, 1364–1376.

Xie, P., J.E. Janowiak, P.A. Arkin, R. Adler, A. Gruber, R. Ferraro, G.J. Huffman, and S. Curtis, 2003: GPCP Pentad Precipitation Analyses: An Experimental Dataset Based on Gauge Observations and Satellite Estimates. *J. Climate*, **16**, 2197–2214.

Methods of Multi-Model Consolidation, with Emphasis on the Recommended Three-Year-Out Cross Validation Approach

Huug van den Dool

Climate Prediction Center, NOAA/NWS/NCEP
Camp Springs, MD 20746

1. Introduction

In many contexts with limited data and no patience to wait for new and independent data, one needs to design schemes that mimic the real time forecast situation on a fixed old data set. This is done often nowadays by cross-validation (CV). The purpose of CV is to establish properties of a forecast scheme that would apply on independent future data, for instance to estimate a-priori skill. However, while CV is often a necessity, it may also itself be the source of a problem in evaluating skill. CV is not an exactly defined procedure in general, so let's focus on a situation when systematic error correction is thought to be required. Given N pairs of forecast and verification, say seasonal Nino34 forecasts for 1981-2001, we can set M pairs (M much less than N=21) aside, calculate the systematic error over the N-M cases, then apply the correction to all or some of the M cases left out. This is done exhaustively, so all data is used as (assumed independent) verification at least once. Naturally, researchers want to get away with M=0 or M=1, since it is simpler than M>1, and skill may appear higher that way. Don't we want high skill??? Yes, but not if the assessment misleads us as to the performance in real time. Dependent data generally overstate the skill level. In this write-up we want to make a strong case for M=3, i.e. keep (at least) three forecast/observed pairs out. This appears to be the right approach in the context of multi-model ensembles, where not only systematic error correction is required but also the determination of weights to be assigned to the participating models.

The procedure we recommend, used in Pena and Van den Dool (2008), is more completely named CV3RE, where CV is cross-validation, 3 means three years left out, R refers to the random choice of two of the three years left out, and E refers to an external climatology (ideally from a data set for a constant climate outside the period of experimentation.) The reason that 3 years should be taken out for the systematic error correction (SEC) is that one can show analytically that the correlation does not change upon

Mdl 4	anomaly	Obs	anomaly	year	SEC	Random Years	
25.5	.9	26.8	−.4	1981	−2.62	(1985 1989)	
25.9	1.3	28.1	.9	1982	−2.62	(2000 1989)	
23.8	−.9	27.1	−.1	1983	−2.46	(1990 1998)	
23.5	−1.3	26.7	−.5	1984	−2.44	(1993 1981)	
24.1	−.8	26.7	−.5	1985	−2.32	(1992 1995)	
26.0	1.4	27.4	.2	1986	−2.56	(1999 1987)	
26.6	2.0	28.8	1.6	1987	−2.63	(1996 1989)	←
23.6	−1.1	25.6	−1.6	1988	−2.50	(1989 1995)	
26.2	1.5	26.7	−.5	1989	−2.48	(1983 1992)	
25.8	1.1	27.3	.1	1990	−2.54	(1985 2000)	
23.5	−1.2	27.9	.7	1991	−2.42	(1990 2001)	
24.4	−.3	27.5	.4	1992	−2.49	(1996 2001)	
24.4	−.5	27.6	.4	1993	−2.32	(1985 1995)	←
23.5	−1.3	27.3	.1	1994	−2.38	(1989 1991)	
22.9	−1.8	27.0	−.2	1995	−2.48	(1986 1996)	
25.6	.9	27.1	−.1	1996	−2.45	(1991 1990)	
25.8	1.0	28.9	1.7	1997	−2.36	(1991 1990)	
23.4	−1.4	25.9	−1.2	1998	−2.37	(1991 1988)	
24.5	−.3	26.3	−.8	1999	−2.42	(2001 1995)	
25.0	.2	26.7	−.5	2000	−2.41	(2001 1991)	
25.2	.5	27.3	.1	2001	−2.50	(1998 1999)	
24.7	.0	27.2	.0	all	−2.45		

Table 1 Shown in column 1 are June temperatures for 1981-2001 (top to bottom) in the Nino34 area as predicted at a lead of 5 months by one of the Demeter models (model#4) which has its initial states in January. The observed SST is shown in column 3. The anomalies in columns 2 and 4 are wrt to the 21 year mean of model and observed data respectively. The bottom line shows 21 year averages. Column 6 shows the systematic error correction (SEC) that would be applied to the year in column 5. Columns 7&8 are two randomly selected years also withheld in calculating the recommended CV3RE SEC.

Correspondence to: Huug van den Dool, Climate Prediction Center, NOAA/NWS/NCEP, Camp Springs, MD, 20746;
E-mail: Huug.Vandendool@noaa.gov

taking out just 1 year, i.e. CV1 does not do anything. The number of elements withheld being odd (as a convenience), three would thus be the minimum. Typically that would be three successive years as a block, but here we argue that the three removed should be a) the test year, and b) two additional randomly chosen years.

2. An example of systematic error correction

Table 1 provides details of an example. Shown in column 1 are June temperatures for 1981-2001 (top to bottom) in the Nino34 area as predicted at a lead of 5 months by one of the Demeter models ("Model #4") which has its initial states in January. The observed SST is shown in column 3. The anomalies in columns 2 and 4 are wrt to the 21 year mean of model and observed data respectively. The bottom line shows 21 year averages. Column 6 shows SEC that would be applied to the year in column 5.

Fig. 1 Anomaly pattern correlation of systematic error corrected monthly SST over the tropical Pacific domain, averaged for all leads and initial months based on the 21 yr of data in the hindcasts (empty bars) and after 3 yr random cross validation (dark bars). The consolidation is done gridpoint-wise, which can be improved upon by increasing effective sample size. On the left the seven Demeter model, the CFS and CA. On the right seven entries for various MME approaches. See Pena and van den Dool (2008).

Columns 7&8 are two randomly selected years also withheld in calculating the recommended SEC.

Clearly model 4 needs systematic error correction badly, since it is about 2.5ºC too cold. This is a large error in the mean given that anomalies are rarely larger than 1.5. However, it would be wrong to assume that we know SEC = -2.45 with such certainty so as to apply it to all cases in the sample of 21 - this would be the full sample dependent data approach. If one withholds each year in turn (in the hopes of creating an independent year), plus two more years chosen at random, and calculates the difference in the mean of forecast and observation over 18 cases, one finds SEC to vary somewhat but not greatly, from -2.32 to -2.63 to be specific. Fortunately, the forecast still improves greatly as a result of applying a variable SEC, but not as much as, seemingly, when applying a constant SEC = -2.45. It is more correct to say that the dependent data case (N=21) over-estimates skill, and we have a professional duty to calculate an estimate that will hold up in true real-time. As shown in Fig.1 the skill, as measured by correlation, is around three points lower than in the dependent data result for each of the 9 models on the left in Fig.1 considered by Pena and Van den Dool (2008).

That the year for which forecast accuracy is tested should not be included in the SEC determination is easily seen in the extreme for N=1 – that would make the forecast perfect in a misguided way. But even for N=21 the test case has a noticeable impact, because of "compensation" effects that are known to affect CV. For instance, in 1987, see Table 1, the forecast and observation are 'only' -2.2ºC apart and including this case keeps the SEC at -2.45, whereas excluding it makes it -2.63. The opposite happens in 1985 and 1993, two years that feature forecast errors larger than average. Using three elements dilutes the compensation effect. In section 3 we will see a more complicated compensation effect.

In the next section we argue again that three should be taken out, but for a very different reason.

3. Degeneracy in regression

In earlier work we found highly negative correlation in CV applied to forecasts based on regression schemes, where a zero correlation would have been more reasonable. This feature was ultimately explained in Barnston and van den Dool (1993). Fig. 2, reproduced from that paper shows a synthetic data case. We
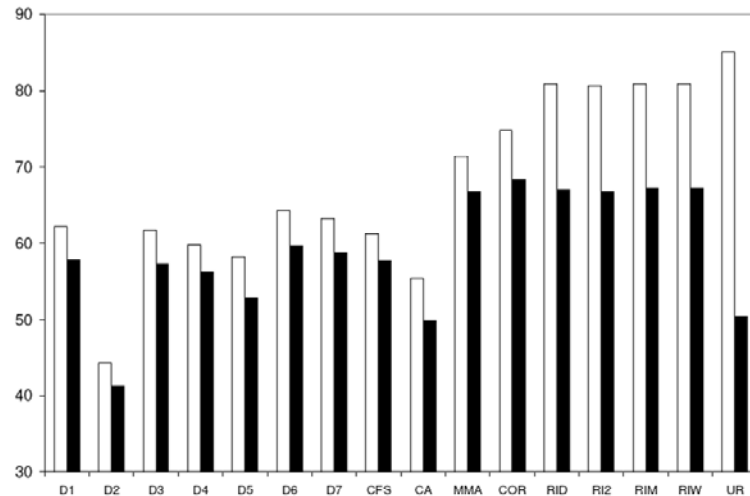
generated pairs of correlated forecast, observation data, varying the correlation along the x-axis from 0 to 1. We then did a CV-1 approach to calculate the correlation from limited data (32 pairs). When the correlation is large CV-1 functions OK and only shows some normal 'shrinkage'. But when the intended correlation is small, between 0 and 0.2 in this case, the outcome of CV-1 is a disaster. One can get a perfect -1 correlation. This happens because of compensation effects at the covariance level (in section 2 we had compensation in the mean). Suppose we have zero correlation on the full sample between forecast and observation, and thus also zero covariance. When we leave out 1 pair, which happens to co-vary by chance positively, the remaining N-1 pairs have, by necessity, a negative covariance in the mean. Thus a regression forecast based on the N-1 will be opposite to what is observed in the one case left out, thus leading to high negative correlation.

This can happen in real life regression forecasts. For instance Nino34 correlates with seasonal temperature over the US, but with opposite sign in the NW and the SE US. Along the broad band of zero and small correlation, presumably the nodal line of a teleconnection pattern, the CV-1 score of a regression forecast is highly negative. Here we get punished for our good intentions. The solution, aside from waiting forever for more years, is to take out more than 1. For instance when taking out the test year as well as two more years, the compensation effect is obviously diluted. Choosing two more years at random (as opposed to

Fig. 2 The correlation (Y axis) calculated as per CV-1 from synthetic data generated by computer with a known correlation (X – axis). For instance when generating paired data with 0.3 correlation the CV-1 procedure (applied to 32 pairs) returns an estimate for the correlation around 0.1.

a block of three, with the test year in the middle) is better because the serial correlation (caused by climate change among other things) violates the assumption of independence.

This above discussion applies to the multi-model ensemble approach because the MME is a linear combination of several forecasts, with weights derived from a limited data set as per regression. We should apply CV3RE, and we can fold the CV for SEC into the CV required for the weights (the regression aspect) into one single procedure. The seven entries on the right hand of Fig.1 are MME by different schemes subjected to CV3RE. The various ridge regression approaches fare much better under CV than an unconstrained regression (UR).

4. Conclusion

We recommend as cross-validation procedure something called CV3RE, where CV is cross-validation, 3 means three years left out, R refers to the random choice of two of the three years left out, and E refers to an external climatology (ideally from a data set for a constant climate outside the period of experimentation.). We have not laid out the case for the external climatology in this short write-up, but this aspect also helps stabilize the answers one gets. While we believe CV3RE is appropriate for the multi-model ensemble it may also be a good strategy in many other situations. However, each problem requires some deliberations of its own, and a general theory/algorithm for CV appears elusive (to me).

References

Barnston, A.G., and H.M. van den Dool, 1993: A Degeneracy in Cross-Validated Skill in Regression-based Forecasts. *J. Climate*, **6**, 963–977.

Peña, M., and H. van den Dool, 2008: Consolidation of Multimodel Forecasts by Ridge Regression: Application to Pacific Sea Surface Temperature. *J. Climate*, **21**, 6521–6538.
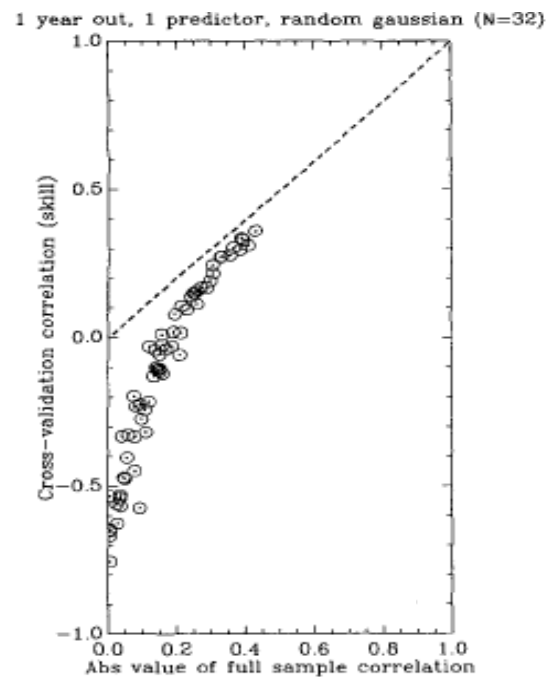
Development of Neural Network Emulations of Model Radiation for Improving the Computational Performance of the NCEP Climate Simulations and Seasonal Forecasts

V. M. Krasnopolsky[1,2], M. S. Fox-Rabinovitz[2], Y. T. Hou[1], S. J. Lord[1], and A. A. Belochitski[2]

[1]*Environmental Modeling Center, NOAA/NWS/NCEP*
[2]*Earth System Science Interdisciplinary Center (ESSIC), University of Maryland*

1. Introduction

Calculation of model physics in a GCM (General Circulation Model) usually takes a very significant part of the total model computations. Evidently, this percentage is model dependent but full model radiation is the most time-consuming component of GCMs (*e.g.*, Morcrette *et al.* 2008, Manners *et al.* 2008). In both climate modeling and NWP, the calculation of radiative transfer is necessarily a trade-off between accuracy and computational efficiency. Very accurate methods exist, such as line-by-line procedures that could be employed ideally to calculate radiative fluxes for every grid-point at every time-step. However, if the radiation transfer were to be computed for every grid point and at all time steps, it would generally require as much CPU time or more than the rest of the model components, *i.e.*, model dynamics and other physical parameterizations (Morcrette *et al.* 2008). Therefore a number of simplifications are usually made to reduce this cost to manageable levels.

For example, in the majority of modern radiative schemes, the correlated-k method (Lacis and Oinas 1991) is typically used to reduce the integration over wavelength by effectively binning wavelengths with similar absorption coefficients (k-terms). This simplification reduces greatly the number of monochromatic radiative transfer calculations required. The number of k-terms can be adjusted, which provides a trade-off between the accuracy and efficiency required for a given application. However, the correlated-k methods cannot be made sufficiently computationally efficient to allow calculations for every grid-point at every time-step.

To reduce the cost further, calculations are usually made at lower temporal and/or spatial resolutions. Quite drastic reductions in temporal resolution are often made (*e.g.*, radiation calculations are made every three hours for the climate and global forecast models at NCEP and UKMO (Manners *et al.* 2008)). Between radiative transfer calculations major changes may occur in the radiative profiles (caused primarily by two factors: changes in clouds and changes in the angle of incident solar radiation) that are not represented. A reduced horizontal resolution approach (the radiative calculations are performed on a coarser grid with a following interpolation of the results to an original finer grid) is used to speed up radiation calculations at ECMWF (Morcrette *et al.* 2007, 2008). A reduced vertical resolution approach (the full radiation is calculated at every other vertical level and interpolated on the intermediate levels) is used in the Canadian operational Global Environmental Multiscale model (*e.g.* Côté *et al.* 1998). Such approaches reduce horizontal or vertical variability of radiation fields. Thus, these approaches may reduce the accuracy of a model's radiation calculation and its spatial or/and temporal consistency with other parts of model physics and with model dynamics, which may, in turn, affect negatively the accuracy of climate simulations and weather predictions.

Such a situation is an important motivation for developing new alternative numerical algorithms that provide faster calculations of model physics while carefully preserving their accuracy. In our previous studies (Krasnopolsky *et al.* 2005, 2008) we demonstrated that the neural network (NN) emulation approach can be successfully used to speed up significantly (by one to two orders of magnitude) the calculations of model radiation while providing a sufficient accuracy of decadal (50 years) climate simulations. We also demonstrated that this approach is a generic one; namely it can be used not only for emulating any formulation of the long wave radiation (LWR) physics but also for emulating any formulation of short wave radiation (SWR) physics.

Correspondence to: V. M. Krasnopolsky, Environmental Modeling Center, NOAA/NWS/NCEP;
E-mail: Vladimir.Krasnopolsky@noaa.gov

It this study we applied the NN emulation approach to the higher complexity NCEP CFS (Climate Forecasting System), which required further development of the neural network emulation methodology. We demonstrate that the NN emulation approach for model radiation can be successfully applied to the *significantly higher resolution coupled ocean-atmosphere-land-ice model* with *time dependent CO2*. The atmospheric part of CFS has spectral T126 horizontal resolution and 64 vertical levels (T126L64); it is coupled with the 40-level interactive MOM4 ocean model, with a state-of-the-art 3D land model, and with an ice model.

In Section 2, the improved NN emulation approach and developed NN emulations for the NCEP CFS long-wave radiation (LWR) and short-wave radiation (SWR) are briefly described in terms of their design, accuracy, and computational performance. In Section 3, the results of the parallel decadal model simulations, one using both LWR and SWR NN emulations for calculation of full model radiation and the other using the original model radiation (the control run) are compared in terms of similarity of their spatial and temporal variability characteristics. Section 4 contains conclusions.

2. NN emulations for the NCEP CFS radiation

2.1 Background information on the NN emulation approach

NN emulations of model physics are based on the two following facts. First, any parameterization of model physics is a continuous or almost continuous mapping (input vector vs. output vector dependence) and can be symbolically written as:

$$Y = M(X); \quad X \in \Re^n, Y \in \Re^m \tag{1}$$

where M denotes the mapping, n is the dimensionality of the input space, and m is the dimensionality of the output space. And second, NNs (multilayer perceptrons) are generic tools for approximation of such mappings (Funahashi 1989).

NN is an analytical approximation that uses a family of functions like:

$$y_q = a_{q0} + \sum_{j=1}^{k} a_{qj} \cdot \phi(b_{j0} + \sum_{i=1}^{n} b_{ji} \cdot x_i); \quad q = 1, 2, \ldots, m \tag{2}$$

where x_i and y_q are components of the input and output vectors X and Y, respectively, a and b are fitting parameters, and $\phi(b_{j0} + \sum_{i=1}^{n} b_{ji} \cdot x_i)$ is a "neuron". The activation function Φ is usually a hyperbolic tangent, n and m are the numbers of inputs and outputs (the same n and m as in Eq. (1)), respectively, and k is the number of neurons in the hidden layer. Definitions of NN terminology can be found in many places, for example in the recent book by Bishop (2006) and in the review paper by Krasnopolsky (2007); however, eq. (2) is sufficient to understand the subject of this paper. The numerical complexity of NN (2) can be well approximated by a number of NN weights (Krasnopolsky 2007):

$$N_C = k \cdot (n + m + 1) + m \tag{3}$$

The NN numerical complexity N_C determines the time, T_{NN}, required for the estimating NN (2),

$$T_{NN} = c \cdot N_C$$

T_{NN}, is directly proportional to N_C with the coefficient of proportionality c depending mainly on a hardware and software environment of the computer used.

2.2 NN emulations for full model radiation

The LWR and SWR parameterizations together comprise the full model radiation. The LWR and SWR parameterizations or the full model radiation for the NCEP CFS have been emulated using two NNs, one for LWR and another for SWR.

The input and output vectors for NNs, emulating the LWR or SWR parameterizations, include the same parameters as those of the input and output vectors for the original LWR or SWR parameterizations, respectively. For the LWR NN emulation, these input parameters are the following nine profiles: atmospheric pressure, temperature, specific humidity, ozone mixing ratio, total cloud fraction, cloud liquid

water path, mean effective radius for liquid cloud, cloud ice water path, and mean effective radius for ice cloud. The LWR parameterization (and LWR NN emulation) output vectors consist of the profile of heating rates (HRs) and five radiation fluxes: the total sky outgoing LW radiation flux from the top layer of the model atmosphere (the outgoing LWR or OLR), the clear sky upward flux at the top of the model atmosphere, the total sky upward flux at the surface, the total sky downward flux at the surface, and the clear sky downward flux at the surface.

The NN emulation of the LWR parameterization includes all non-constant inputs of the original LWR (total 556 inputs; $n = 556$ in Eq. (1)). It has the same outputs (total 69 outputs; $m = 69$ in Eq. (1)) as the original LWR parameterization. We have developed several NNs, all of which have the same aforementioned inputs and outputs, with the number k changing from 50 to 200 in Eq. (2). Varying k, the number of terms (or neurons) in Eq. (2), allows us to demonstrate the dependence of the accuracy of approximation on this parameter as well as its convergence, and as a result, to provide a sufficient accuracy of approximation for the model (*e.g.* Krasnopolsky *et al.* 2005).

The input vectors for the SWR parameterization include 55 vertical profiles: atmospheric pressure, temperature, specific humidity, ozone, CO_2, N_2O, O_2, and CH_4 volume mixing ratios, total cloud fraction, cloud liquid water path, mean effective radius for liquid cloud, cloud ice water path, mean effective radius for ice cloud, and three profiles (optical depth, single scattering albedo, and asymmetry parameter) for each of 14 different species of aerosols. The input vectors include also the solar zenith angle, the solar constant and the surface albedo for four different bands. The SWR parameterization output vectors consist of a vertical profile of heating rates (HRs) and nine radiation fluxes: three fluxes at the top layer of the model atmosphere (the total sky outgoing SW radiation flux, the total sky downward flux, the clear sky upward flux), four radiation fluxes at the surface (the total sky upward and downward fluxes and the clear sky upward and downward fluxes), and the downward (the total and clear sky) fluxes in the UV-B spectral band.

The NN emulations of the SWR parameterization have 562 inputs and 73 outputs. We have developed several NNs, with the number k changing from 50 to 200 in Eq. (2). It is noteworthy that, as in the case of the NN emulation of LWR, the number of NN inputs is less than the number of input profiles multiplied by the number of vertical layers plus the number of relevant single level characteristics. Many input variables (*e.g.*, almost all gases) have zero or constant values for the upper vertical layers, and for some gases the entire volume mixing ratio profile is a constant (obtained from climatological data).

2.3 Generating data sets for NN training and validation

The NCEP CFS (T126L64) has been run for seventeen years to generate representative data sets. The representative data set samples adequately the atmospheric state variability, *i.e.*, it represents all possible states produced by the model as fully as possible (including the states introduced due to time dependent CO_2 concentration). All inputs and outputs of original LWR and SWR parameterizations have been saved for two days per month, *i.e.*, for one day at the beginning and one day in the middle of the month, every three hours (eight times per day) to cover the annual and diurnal cycles. From each three hour global data set three hundred events (the set of input and output profiles) have been selected. The data set was divided into three independent parts, each containing input/output vector combinations. Each part consists of about 200,000 input/output records. The first part has been used for training and the second one for tests (control of overfitting, control of NN architecture, *etc.*). The third part of the data set was used to create a validation data set independent of both the training and test data sets. The third part or the validation set was used for validation only. All approximation statistics presented in this section are calculated using this independent validation data set. The accuracy of the NN emulation, i.e., biases and rmse, are calculated against the control (the original parameterization).

It is noteworthy that along with the aforementioned requirement of representing all possible states produced by the model, the size of the training data set is limited mainly by the training time, which, in turn, is determined by the processor type and the amount of memory available. The training time is approximately proportional to the size of the training data set. In our case, the selection of about 200,000 input/output records for training is a result of an optimal choice providing a sufficient representativeness of the training set and a reasonable training time. We selected the size of the test set equal to the size of the training set

because the training and test sets are supposed to have close statistical properties. There are no serious limitations to the size of the validation set; we selected it equal to the size of the first two sets.

2.4 Bulk approximation error statistics

To ensure a high quality of representation of the LWR and SWR processes, the accuracy of their NN emulations has been carefully investigated. The NN emulations have been validated against the original NCEP CFS LWR and SWR parameterizations. To calculate the error statistics presented in Table 1; the original parameterizations and their NN emulations have been applied to the validation data set. Two sets of the corresponding HR profiles have been generated for both LWR and SWR. Total and level bias (or a mean error), total and level RMSE, profile RMSE or PRMSE, and σ_{PRMSE} have been calculated (see Krasnopolsky 2007).

Statistics Types	Statistics	LWR			SWR	
		NCAR CAM	NCEP CFS		NCAR CAM	NCEP CFS RRTMG
			RRTMG	RRTMF		
Total Error Statistics	**Bias**	$3 \cdot 10^{-4}$	$2 \cdot 10^{-3}$	$7 \cdot 10^{-4}$	$-4 \cdot 10^{-3}$	$5 \cdot 10^{-3}$
	RMSE	0.34	0.49	0.42	0.19	0.20
	PRMSE	0.28	0.39	0.30	0.15	0.16
	σ_{PRMSE}	0.2	0.31	0.30	0.12	0.12
Bottom Layer Error Statistics	**Bias**	$-2 \cdot 10^{-3}$	$-1 \cdot 10^{-2}$	$6 \cdot 10^{-3}$	$-5 \cdot 10^{-3}$	$9 \cdot 10^{-3}$
	RMSE	0.86	0.64	0.67	0.43	0.22
Top Layer Error Statistics	**Bias**	$-1 \cdot 10^{-3}$	$-9 \cdot 10^{-3}$	$2 \cdot 10^{-3}$	$2 \cdot 10^{-3}$	$1 \cdot 10^{-2}$
	RMSE	0.06	0.18	0.09	0.17	0.21
NN Complexity	N_C *See eq. (3)*	12,733	33,294	93,969	11,418	45,173
Speedup, η	*Times*	150	12	21	20	45

Table 1 Statistics estimating the accuracy of HRs (in K/day) calculations and the computational performance for NCEP CFS (T126L64) LWR and SWR using NN emulation vs. the original parameterization. For comparison, NCAR CAM (T42L26) LWR and SWR statistics are also shown. Total statistics show the bias, RMSE, PRMSE, and σ_{PRMSE} for the entire 3-D HR fields. Layer (for the top and bottom layers) statistics show the bias and RMSE for one horizontal layer (the top or bottom layer). Also, the NN complexity N_C (3) and speedup η (how many times NN emulation is faster than the original parameterization) are shown. RRTMG and RRTMF are different versions of the radiation code developed by AER Inc.

Using NN emulations simultaneously for LWR and SWR or for the full model radiation results in an overall significant, about 20 – 25% speedup of NCEP CFS climate simulations when both LWR and SWR are calculated every hour. The speedup η provided by NN emulations (see Table 1) can be also used for more frequent calculations of model radiations. For example, for calculations with higher (T382) model horizontal resolution, if full NN radiation is calculated 10 times more frequently, *i.e.*, every six minutes, at every model dynamics time step (instead of every hour), the time required for the climate simulation using full NN radiation will be still less than the time needed for the climate simulation using the original radiation with the one hour frequency.

3. Validation of parallel decadal model simulations and seasonal predictions

In this section we present comparisons between two parallel 17-year NCEP CFS model runs: one using the original LWR and SWR (the control run) and another one using their NN emulations. Both spatial and temporal characteristics of prognostic and diagnostic fields are compared for the parallel runs. To better estimate the changes introduced by NN emulations, we compare them with "background changes" between two control runs performed with the original NCEP CFS model configuration, *i.e.*, without NN emulations.

The first run was performed before and the second run after the routine changes (introduced quasi-regularly by system administrators) of the version of the FORTRAN compiler and libraries.

The results of 17-year climate simulations performed with NN emulations for both LWR and SWR, *i.e.*, for the full model radiation, have been validated against the parallel control NCEP CFS simulation using the original LWR and SWR. We analyze the differences between the parallel runs in terms of time and spatial (global) means as well as temporal characteristics.
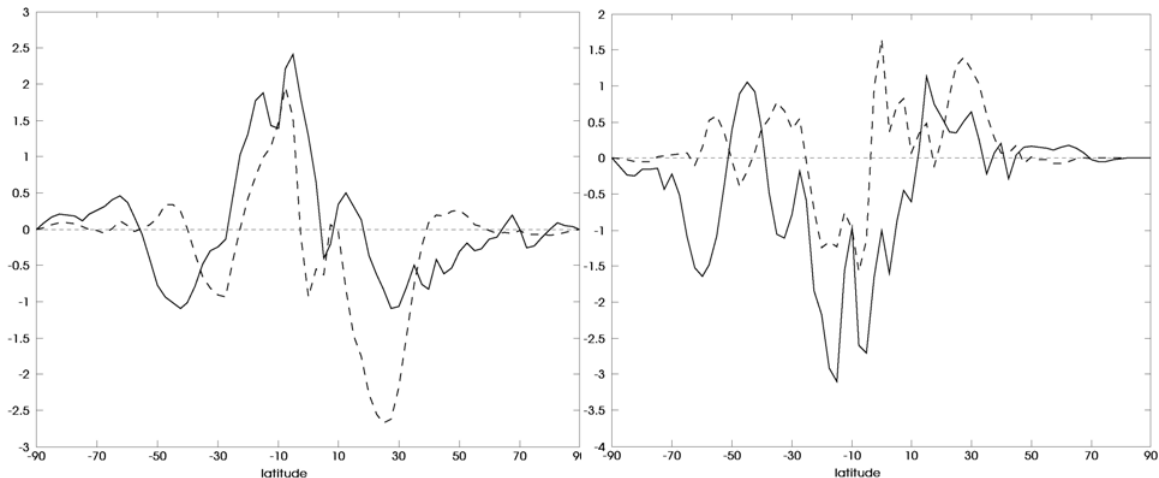

Fig. 1 Zonal and time mean Top of Atmosphere Upward Long (left panel) and Short (right panel) Wave Fluxes (in W/m^2) for the winter. The solid line – the difference (the full radiation NN run – the control (CTL)), the dash line – the background differences (the differences between two control runs) presented for comparison.   The fluxes' differences are multiplied by *cos (lat)* to equalize the areas.

Let us discuss first the differences between the parallel simulations in terms of spatial and temporal radiation characteristics. The differences between the NN radiation and control runs and the differences between two control runs for zonal and time mean LWR and SWR fluxes are presented in Fig. 1. For fluxes presented in Fig. 1 both the differences between the NN radiation and control runs and the differences between two control runs are small and similar by magnitude. They do not exceed 2-3 W/m^2 that is overall they are within observational errors and uncertainties of reanalysis (*e.g.* Kalnay *et al.* 1996).

Let us discuss now prognostic and diagnostic characteristics such as SST, precipitation, different types of clouds, and time series that are sensitive to changes in the model resulted from using NN emulations. Close similarities have also been obtained for these results of parallel runs in terms of time mean spatial fields, which are presented in Figs. 2 to 4. Figs. 2 to 4 have the same design: the upper left panel shows fields produced in the control run (CTL) and the upper right – in the full radiation NN run. The bottom left panel shows the difference (bias) between the full radiation NN and CTL runs, and the bottom right panel shows for comparison the background differences (between two control runs) described above.

The 17-year (1990-2006) time-mean SST distributions and bias/differences for the full radiation NN run vs. the control run and the background differences between two control runs are presented for summer in Fig. 2. The SST bias is very small; it is not larger than the background differences. The results for other seasons are similar.

Figure 3 shows the 17-year (1990-2006) time-mean distributions and bias/differences for total precipitation (PRATE) for the parallel full radiation NN and control runs for summer. The PRATE bias is quite limited and occurs mostly in the tropics; it is also very close by magnitude and pattern wise to the background differences. The results for other seasons are similar.

Figure 4 shows comparisons for the parallel full radiation NN and control runs for total clouds, which are very close for the above runs. The results for other seasons and for other types of clouds are similar.
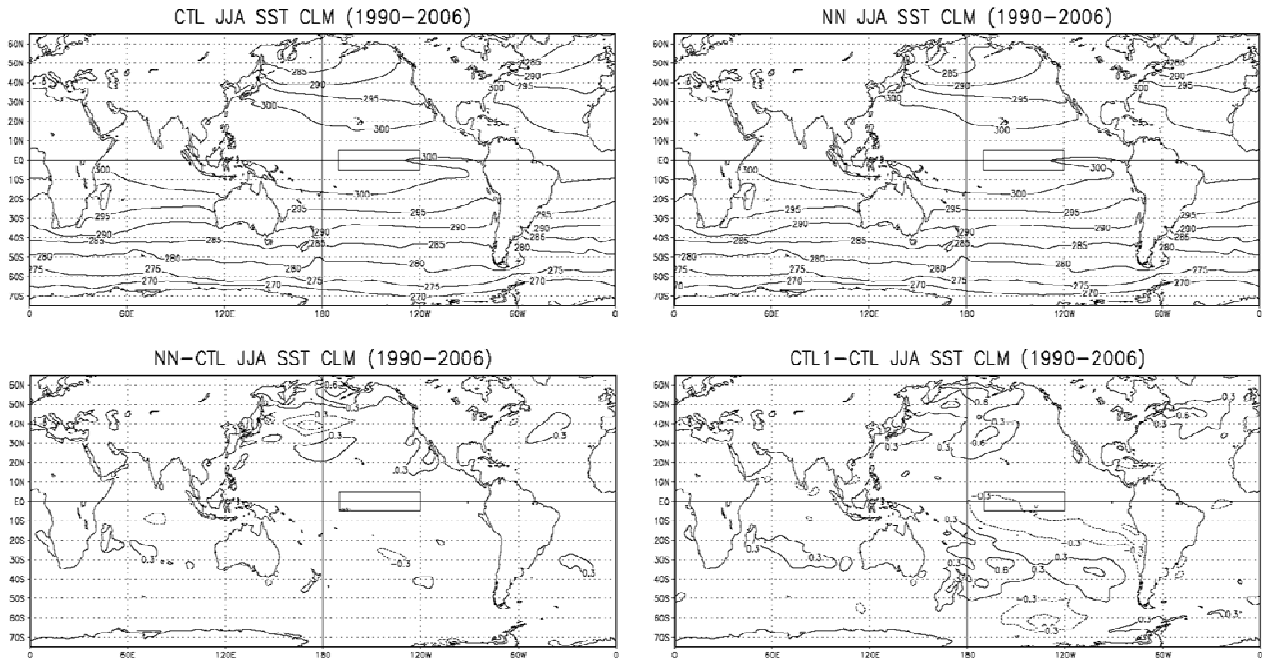
Fig. 2 The 17-year (1990-2006) time-mean SST distributions and bias/differences for summer (JJA: June-July-August) for the full radiation NN run vs. the control run. The upper row panels: left – the control (CTL) run, right – full radiation NN run. The bottom row panels: left – bias or the difference (full radiation NN run – CTL), right – the background differences between two control runs shown for comparison. The contour intervals for the SST fields are 5º K and for the SST bias and difference are 0.3º K.
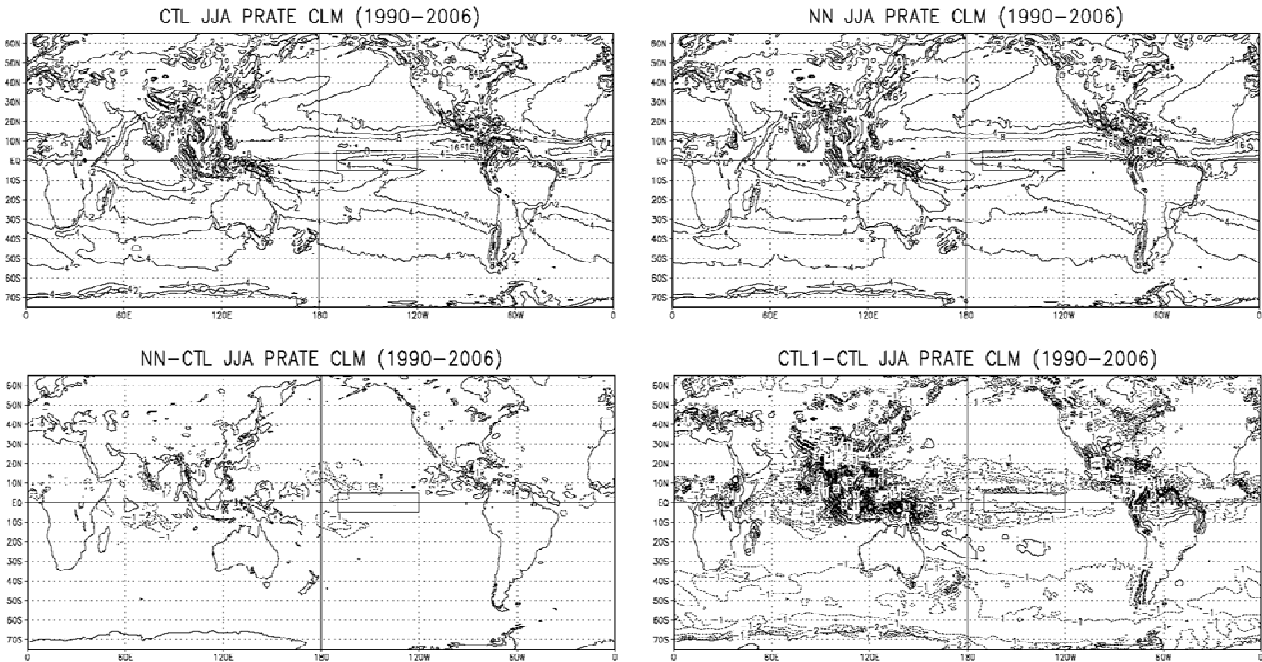

Fig. 3 The same as in Fig. 2 but for total precipitation (PRATE). The contour levels for the PRATE fields are 2, 4, 8, 16 and 32 mm/day. The contour intervals for the PRATE differences (the bottom panels) are 1 mm/day with 0 mm/day contour skipped for clarity.
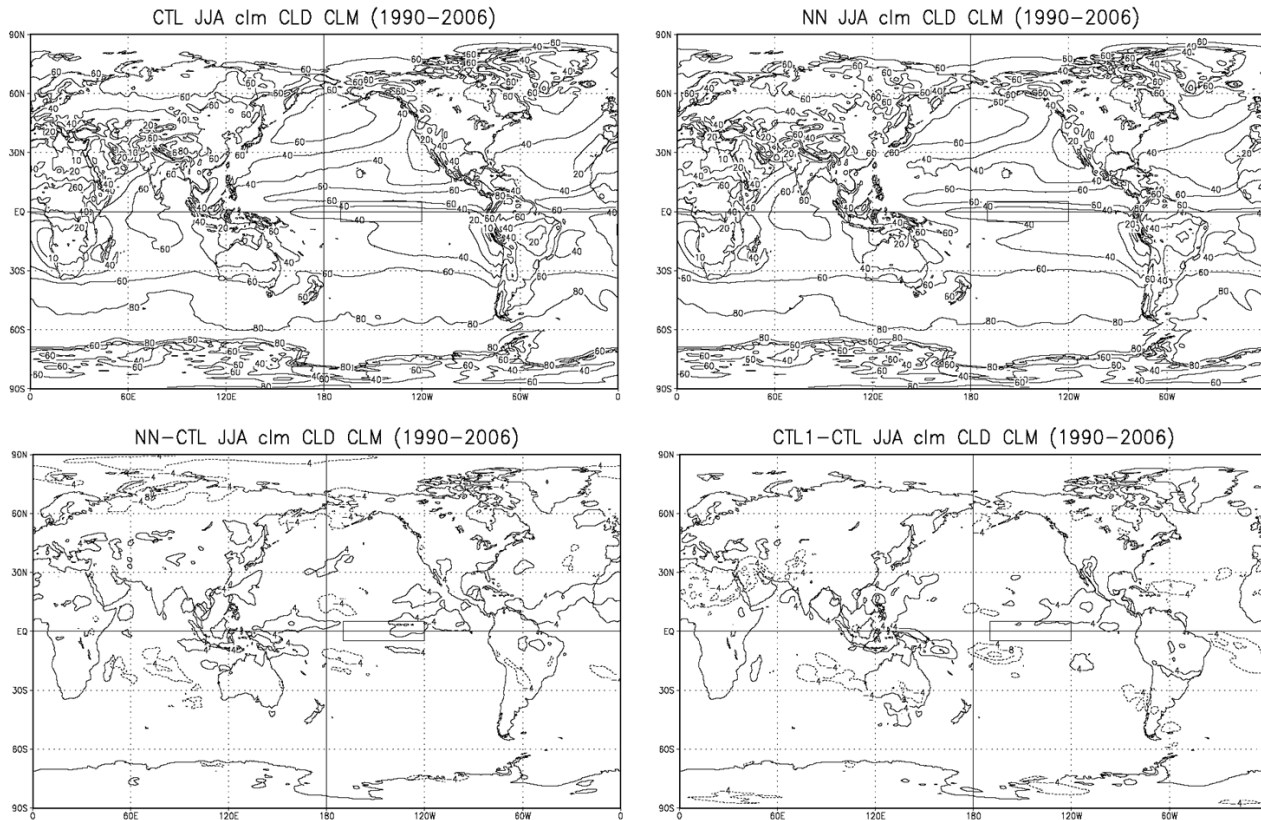
Fig. 4 The same as in Fig. 2 but for total clouds. The contour intervals for the cloud fields are 20% and for the differences 4% with 0 % contour skipped for clarity.

4. Conclusions

In this study, the NN emulation approach (Krasnopolsky *et al.* 2005, 2008) is implemented in the state-of-the-art, high resolution, coupled NCEP CFS. The developed highly accurate neural network emulations of long-wave (RRTMG-LW and RRTMF-LW) and short-wave (RRTMG-SW) radiation parameterizations are 12 to 45 times faster than the original/control long-wave and short-wave radiation parameterizations, respectively. The use of the full NN model radiation results in: (1) an overall speedup of about 20 – 25% for climate simulations and seasonal predictions, and (2) an opportunity to increase significantly the frequency of radiation calculations (for example, to calculate model radiation at every model dynamic time step) without increasing the total model calculation time.

References

Bishop, Ch. M., 2006: Pattern Recognition and Machine Learning. *Springer*, 738 pp.

Côté, J., S. Gravel, A. Méthot, A. Patoine, M. Roch, and A. Staniforth, 1998: The operational CMC-MRB global environmental multiscale (GEM) model. Part I: Design considerations and formulation, *Mon. Wea. Rev.*, **126**, 1373-1395.

Funahashi, K., 1989: On the Approximate Realization of Continuous Mappings by Neural Networks. *Neural Networks*, **2**, 183-192.

Kalnay, E., *et al.*, 1996: The NCEP/NCAR 40-Year Reanalysis Project. *Bull. Amer. Meteorol. Soc.*, **77**: 437-472.

Krasnopolsky, V.M., 2007: Neural Network Emulations for Complex Multidimensional Geophysical Mappings: Applications of Neural Network Techniques to Atmospheric and Oceanic Satellite Retrievals and Numerical Modeling, *Reviews of Geophysics*, **45**, RG3009, doi:10.1029/2006RG000200.

Krasnopolsky, V.M., M.S. Fox-Rabinovitz, and D.V. Chalikov, 2005: Fast and Accurate Neural Network Approximation of Long Wave Radiation in a Climate Model, *Mon. Wea. Rev.*, **133**, pp. 1370-1383.

Krasnopolsky, V.M., M.S. Fox-Rabinovitz, A. A. Belochitski, 2008: "Decadal Climate Simulations Using Accurate and Fast Neural Network Emulation of Full, Long- and Short Wave, Radiation.", *Mon. Wea. Rev.*, **136**, 3683-3695, doi: 10.1175/2008MWR2385.1.

Krasnopolsky, V.M., M. S. Fox-Rabinovitz, Y. T. Hou, S. J. Lord, and A. A. Belochitski, 2009: Accurate and Fast Neural Network Emulations of Model Radiation for the NCEP Coupled Climate Forecast System: Climate Simulations and Seasonal Predictions, submitted.

Lacis, A. A., V. Oinas, 1991: A description of the correlated k-distribution method for modeling non-gray gaseous absorption, thermal emission and multiple scattering in vertically inhomogeneous atmospheres. *J.Geophys. Res.* **96**: 9027-9063.

Manners, J., J.-C. Thelen, J. Petch, P. Hill & J.M. Edwards, 2008: Two fast radiative transfer methods to improve the temporal sampling of clouds in NWP and climate models, *Q. J. R. Meteorol. Soc.* **00**: 1-11.

Morcrette, J.-J., G. Mozdzynski and M. Leutbecher, 2008: A reduced radiation grid for the ECMWF Integrated Forecasting System, *Mon. Wea. Rev.*, **136**, 4760-4772, doi: 10.1175/2008MWR2590.1

Ocean Reanalyses: Prospects for Climate Studies

James A. Carton

Department of Atmospheric and Oceanic Science
University of Maryland, College Park, MD

1. Introduction

This talk reviewed progress in developing ocean reanalyses analogous to the atmospheric reanalyses, and spanning similar time periods. The questions to be addressed are: what climate signals can we detect? Where and when can we detect these signals? How large were the signals, and how large is our uncertainty? What level of diagnostic analysis is possible – for example is it possible to construct a full heat or freshwater budget? To what extent are the results contaminated by instrument and model bias (including wind bias)? What approaches can we use to identify and correct for these biases? And finally, what comes next? If this seems like a lot to cover in one talk, you are right. In fact I ended up talking mainly about the first part. If you are interested in learning more about these subjects in addition to looking at the slides you can get some up to date information and references by looking at the white papers being produced for the OceanObs'09 conference (www.oceanobs09.net). Another up-to-date source of information is the Climate Change Science Program's report (CCSP, 2008).

In order to introduce the audience, whose background is mainly in meteorology, to the results of current ocean reanalyses I present the problem of the warming of the oceans. If you, the audience member, want to evaluate the ocean's participation in global warming you can compute a volume average of the temperature of the oceans down to 700m (the well-sampled part of the ocean) and multiply by the heat capacity of seawater you can evaluate the temporal change in the volume-average heat content of the oceans (Fig. 1). Time rate of change of this quantity gives the net heat flux from the atmosphere into the ocean (a more accurate estimate, by the way, than can be evaluated from meteorological parameters).

Comparing the results from the nine reanalyses shown in Fig. 1 tells us that most of the reanalyses show

Fig. 1 Global average heat content anomalies from the individual 30-yr record means (1966-1995), integrated 0/700m and temporally smoothed with a 1-year running filter. Bold black curve shows the ensemble average of the eight no-model and sequential analyses. The linear trend of the ensemble average is 0.77×10^8 Jm^{-2}/10yr or $0.24 Wm^{-2}$, while trends of individual analyses range from 0.68 to $0.98 \times 10^8 Jm^{-2}$ /10yr (0.21-$0.31 Wm^{-2}$). Global integrated heat content can be obtained from the global average by multiplying by the surface area of the World Ocean excluding shelves, 3.4×10^{14} m^2. This figure comes from *Carton and Santorelli (2008)*.

similar rates of global warming, although they differ from each other by ~10-20%. Most of the reanalyses use sequential data assimilation. However, the one that is most different, GECCO, uses 4DVar. This immediately suggests the change from sequential approach to 4DVar will have a fundamental impact on the results. Fig. 1 is also interesting because if you look at it again you will notice that in addition to a gradual warming trend there is an anomalously rapid warming in the 1970s and corresponding cooling in the mid-
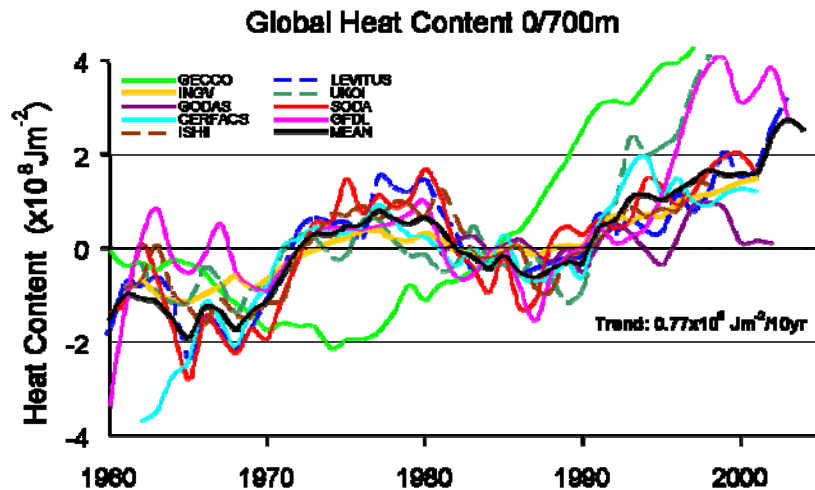
_____

1980s. This 'bump' in heat content is suspicious and, to make a long story short, turns out to be evidence of the presence of instrument bias.

2. Data and methodology

This talk describes results of a number of different data assimilation systems. For those audience members who have some idea about how data assimilation works I provide a very brief introduction to the differences among the systems I consider (see slide 9 of my presentation at ftp://ftp.cpc.ncep.noaa.gov/CTB/2008-2009/Carton-ctb.pdf). My brief introduction begins with definition of a cost function J containing weighted mean square differences between the analysis represented by the vector x (which we haven't determined yet) and the background estimate, x^b, and also the differences between the analysis and a set of observations x^o.

$$J(x) = (x - x^b)^T B^{-1} (x - x^b) + (Hx - x^0)^T R^{-1} (Hx - x^0)$$

where B denotes the background error covariance, R the observational error covariance and H the linear operator.

The data assimilation algorithms all develop from this expression and all attempt to minimize J. For most of the reanalyses considered here x is considered a function of three spatial dimensions. But for the 4DVar reanalysis (the authors prefer the term state estimate) x is additionally a function of time.

I also discuss the historical record of ocean observations. While this may seem like an esoteric subject to meteorologists, oceanography is such a data-limited field that small changes in our interpretation of the historical record can have a big impact in our understanding of ocean climate (an example is presented below).

3. Analysis of prominent results

In the introduction I mentioned the spurious 'bump' in heat content of the oceans. Recent reexaminations of the historical record have traced this bump to time-dependent errors in a particular type of instrument called an Expendable Bathythermograph (XBT). Different groups have developed corrections to the historical XBT (and earlier MBT) data which eliminate the bump. But, interestingly, they have rather different ideas about the vertical structure of this bias correction (see Fig. 2).

That means that the different bias corrections can have a rather different impact on our historical reconstruction of such variables as temperature and currents even though they may give similar estimates of heat content. And in the results presented in the talk the audience member could see the impact on data assimilation experiments using one or another of the bias corrections. Surface currents for the 1997-1998
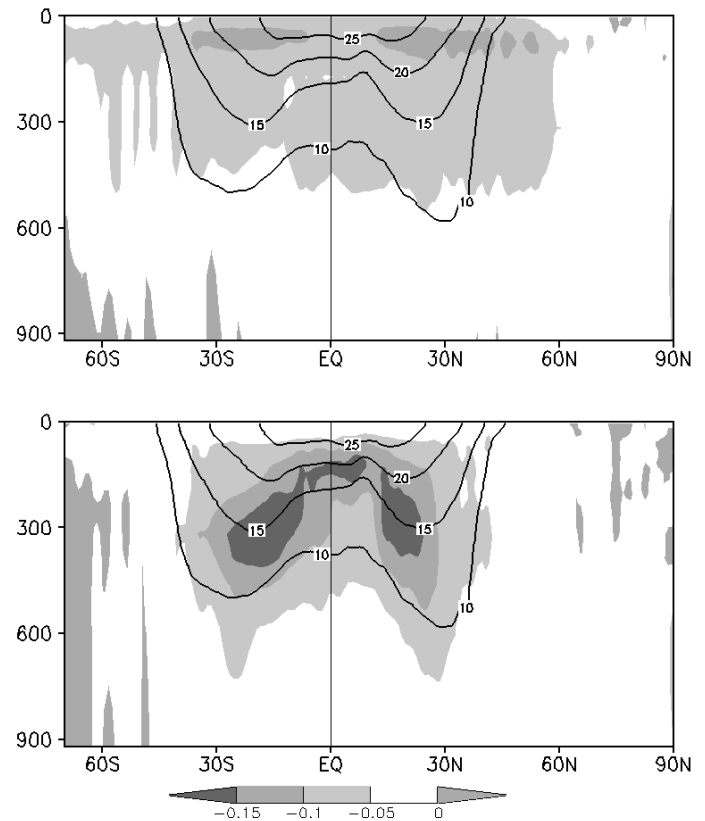

Fig. 2 Zonally and temporally (1967-2002) averaged difference in observed temperature as evaluated in different hydrographic data sets. In the upper panel the data sets are a recent version of the World Ocean Database (*Levitus et al., 2009*) which includes a new bias correction minus the older WOD05 which does not. In the lower panel the data sets are a new bias-corrected database by *Wijffels et al (2008)* minus WOD05. Mean isotherm depths are superimposed for convenience. The *Levitus et al. (2009)* bias correction is not a function of latitude, but its impact in Fig. 2 is largest in the subtropics because that's where historical data coverage is most intense.

El Nino are altered by 20% as a result of the choice of bias correction. Changes in the subtropics are smaller, but still non-negligible.

I also presented some discussion of model resolution and its impact on the ocean reanalyses. I argued that resolution of finer than 1/2° in the horizontal may well be necessary for processes involving horizontal advection, even though this resolution is much finer than the effective resolution of the historical observational network. The example I gave is the anomalous advection of freshwater in response to the Great Salinity Anomaly of the late-1960s to early 1970s. In case you are not familiar with this, a reversal of winds in the winter of 1968-1969 apparently dumped a large amount of sea ice out of the Arctic into the North Atlantic, thus introducing a pool of low salinity water. This pool gradually made its way anticyclonically around the subpolar gyre of the North Atlantic, reappearing off Norway about eight years later.

Of the nine reanalyses discussed earlier only five actually show this event in surface salinity (shown in Fig. 3). Of these, only one, SODA (Carton and Giese, 2008), actually shows the freshwater making its way around the western side of the sub-polar basin, hugging the coast as we think it should. Only this analysis has sufficient horizontal resolution (1/4°) to resolve boundary processes. The rest are too coarse (typically 1°) and as a result, too diffusive.

4. Concluding remarks

This talk has been somewhat different than

Fig. 3 Salinity anomaly from the 1962-1995 average, averaged vertically (0-250m) and in time for two 3-year periods 1968-70 and 1971-3. The two periods show early and mid stages of the 1970s Great Salinity Anomaly**.**

some of the others in this lecture series in that I do not specifically address issues related to the NOAA or NASA software suites associated with the Climate Testbed. Rather, my goal is to encourage the meteorologists to take an interest in historical reanalyses of ocean variables. I return at the end to some of the questions posed at the beginning of the talk. The most important issues for potential users of the ocean reanalyses -- what climate signals are in the historical record and how much can we trust the record – I address mainly by example, by comparison of the results among different reanalyses, and by comparison of the oceanic signals to their meteorological counterparts. I hope to have convinced audience members that there are indeed interesting, 'real' climate signals in current ocean reanalyses. For some coupled problems such as surface heat flux estimates based on the ocean reanalyses are likely more accurate than their widely discussed meteorological counterparts.

On the other hand I also expressed caution. I think it is premature to do sophisticated analyses of quantities such as relative vorticity which are sensitive to error. And we are still at the stage where the user must be on the lookout for spurious results. Finally, I discussed the potential of developments in data assimilation methodology, including ensemble methods. I discussed the prospects for extending the record back into the first half of the 20th century. And I discussed new applications such as reanalysis of ocean ecosystems based on an understanding of the changing physical properties of the oceans.
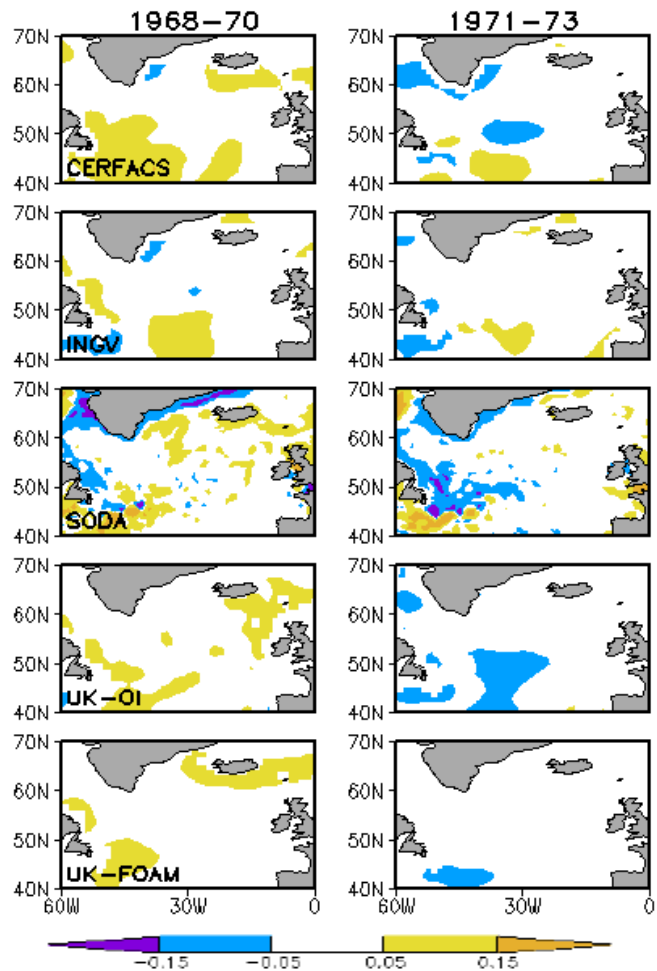
References

Carton, J.A., and B.S. Giese, 2008: A reanalysis of ocean climate using Simple Ocean Data Assimilation (SODA), *Mon. Wea. Rev.*, **136**, 2999-3017.

Carton, J.A., and A. Santorelli, 2008: Global Upper Ocean Heat Content as Viewed in Nine Analyses, *J. Clim.*, **21**, 6015–6035. DOI: 10.1175/2008JCLI2489.1.

CCSP, 2008: Review of the U.S. Climate Change Science Program's Synthesis and Assessment Product 1.3: Reanalyses of Historical Climate Data for Key Atmospheric Features: Implications for Attribution of Causes of Observed Change, *The National Academies Press*,

books.nap.edu/openbook.php?record_id=12135&page=3

Levitus, S., J.I. Antonov, T.P. Boyer, H.E. Garcia, R.A. Locarnini, and A.V. Mishonov, H.E. Garcia, 2009: Global Ocean Heat Content 1955-2007 in light of recently revealed instrumentation problems. *Geophys. Res. Lett., in press.*

Wijffels, S.E., J. Willis, C.M. Domingues, P. Barker, N.J. White, A. Gronell, K. Ridgway, and J.A. Church, 2008: Changing Expendable Bathythermograph Fall Rates and Their Impact on Estimates of Thermosteric Sea Level Rise, *J. Clim.*, **21**, 5657-5672. DOI: 10.1175/2008JCLI2290.1

Bias Correction and Forecast Skill of NCEP GFS Ensemble Week-1 and Week-2 Precipitation and Soil Moisture Forecasts

Yun Fan and Huug van den Dool

Climate Prediction Center, NOAA/NWS/NCEP

Camp Springs, MD 20746

1. Introduction

Soil moisture, the so-called land SST, has been considered important for weather and climate prediction, in particular in the warm season when land and atmosphere are more tightly coupled (Dirmeyer 2000, Kanamitsu *et al.* 2003, Koster et al 2003, Van den Dool *et al.* 2003, Zhang et al 2003, Van den Dool 2007). Soil moisture is also an important indicator for real-time drought and flood monitoring. In 1997 the NOAA Climate Prediction Center (CPC) started a soil moisture "dynamical" week1 and week2 outlook, <u>over the United States</u> only, on a daily basis, using CPC's leaky bucket (LB) land surface hydrological model (Huang *et al.* 1996, Van den Dool *et al.* 2003) forced with week1 and week2 precipitation and surface air temperature from a <u>single member</u> forecast of the NOAA National Center for Environmental Prediction (NCEP) Medium-Range Forecast (MRF), lately called the Global Forecast System (GFS). From late 2001 onward the GFS *ensemble* forecast was used to replace the single member forecast and the procedure was further improved in late 2003 to include the bias corrected GFS ensemble forecast.

The reader should understand that the LB model is kept up to date every day with observations. One can look upon this as an integration of the LB from 1931 to yesterday 12Z, and the GFS's temperature and precipitation are appended to this ongoing LB integration to jump another two weeks ahead. We do not use the GFS's soil moisture directly. We therefore avoid having to deal with the potentially very biased soil moisture states of the GFS and note the LB is integrated in an offline fashion, *i.e.* not

Fig. 1 Time series of daily spatial correlation of week-1 & week-2 observed & forecasted precipitation anomalies over North America, bias correction based on 30 days mean forecast errors on 0.5x0.5 grid. Note: the data from May to November 2007 was missing.

coupled to the atmosphere. More primitive approaches to avoid the GFS bias include considering the 2 week change in the GFS's soil moisture predicted by GFS itself, a product launched by COLA around 1995.

When we talk about research below we mean research 'on the fly' applied to products that were generated in real time, *i.e.* only a few years worth of data has been saved and nothing was rerun.

In mid-2007, the CPC initiated its monitoring and prediction of the variability of the Global (African, Asian, Australian, and American) Monsoons Systems, to collaborate with the international community on improving monsoon monitoring and providing timely and useful weather and climate information for different users and decision makers worldwide. With releasing the CPC gauge based Global Unified Land
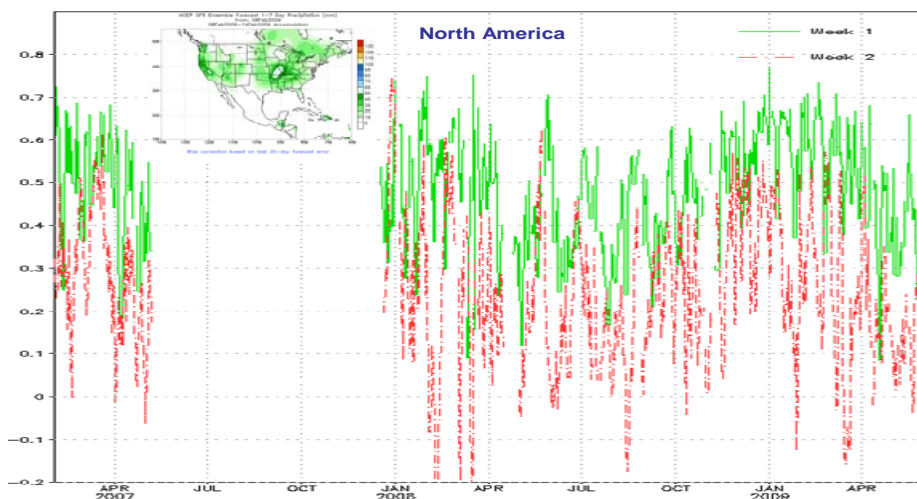
Surface Precipitation Analysis in late 2007, the daily bias corrected GFS ensemble week1 and week2 precipitation forecasts have been *expanded to the global land surface*.

The NCEP GFS is not a frozen system but has been upgraded frequently in terms of dynamical core and physics package in the past years. In the early stage of CPC's soil moisture "dynamical" outlook, both bad and good comments were received. In recent years, more and more good comments were gathered from different users. So it is time to verify and quantify the daily bias corrected GFS ensemble week1 and week2 precipitation and soil moisture forecast thereof. The first part of this work is to assess the GFS ensemble week 1 and week 2 precipitation forecasts over the global land. The main attention is on the skill of the bias corrected GFS ensemble precipitation forecasts over the

	Week 1		Week 2	
	Bias Correction	No Bias Correction	Bias Correction	No Bias Correction
North America	0.49	0.48	0.24	0.26
South America	0.45	0.25	0.31	0.18
Asia Australia	0.47	0.40	0.29	0.26
Africa	0.40	0.24	0.25	0.13

Increased by 67%   *Increased by 80%*

The effectiveness of bias correction is mainly space dependent.
Bias correction can correct spatial distribution of P_f & reduce its error.

Similarity of P_f & P_o

Table 1 Averaged (May 1, 2008 – June 7, 2009) spatial correlations of observed and GFS forecasted precipitation anomalies over different monsoon regions.

	Week 1		Week 2	
	Bias Correction	No Bias Correction	Bias Correction	No Bias Correction
North America	19.18	22.82	21.61	23.58
South America	29.55	41.06	32.27	41.72
Asia Australia	22.65	27.62	25.24	29.15
Africa	17.06	19.47	17.66	19.33

Reduced by 28%   *Reduced by 23%*

Distance of P_f & P_o

Table 2 Averaged (May 1, 2008 – June 7, 2009) RMSE of GSF forecasted precipitation anomalies over different monsoon regions (unit: mm/week).

North American, South American, Asia-Australian and African monsoon regions. Detailed analysis is conducted on the spatial-temporal distribution of the bias, in order to address questions like: what does the bias look like and is it removable? Does bias correction improve GFS forecast skill? The second part of this research focuses on the predictability of the land surface, but over the US only. Since the predictability of soil moisture critically depends on the quality of the GFS ensemble predicted precipitation, further analysis is done on the temporal-spatial features of the GFS driven soil moisture forecast skills, i.e. when, where and to what extent the soil moisture can be predicted on week1 and week2 time scales beyond the skill of a persistence forecast.

2. Methodology

Every day the week-1 and week-2 GFS precipitation ensemble forecasts have been corrected with the past N days mean forecast errors, defined as follows:

$$Bias1 = 1/N \, \Sigma \, [\, Pf \, (week1) - Po \, (week1) \,] \hspace{3cm} (1)$$

$$Bias2 = 1/N \, \Sigma \, [\, Pf \, (week2) - Po \, (week2) \,] \hspace{3cm} (2)$$

where Pf is the NCEP GFS ensemble week-1 and week-2 precipitation forecasts, Po is the observed week-1 and week-2 precipitation from CPC daily US and Global Unified Precipitation Analysis. N is number of days (*e.g.* 30 or 7 days, these being the only choices being maintained in real time). The choice of N is a little bit subjective. In general, the mean forecast errors calculated from larger N (*e.g.* 30 days) are more robust then those from the smaller N (*e.g.* 7 days or 1 day). Of cause, one can calculate the mean forecast errors for the bias correction with more complicated methods, such as non-equal weighting (giving larger weights to more recent days and reducing weights with the time of past days increasing) or use probability density function (PDF) adjustment based on the forecasted and observed precipitation in the past days.
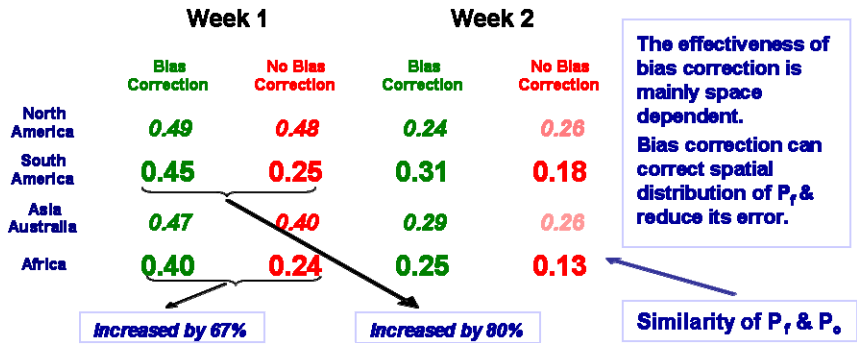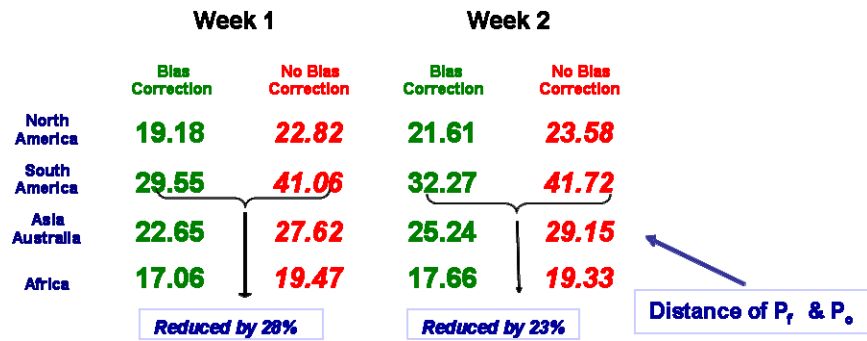
The very same bias correction is also applied every day to the week-1 and week-2 GFS ensemble 2 meter surface air temperature (T2m) forecasts, but over the US only. Results for T2m are not shown in this paper.
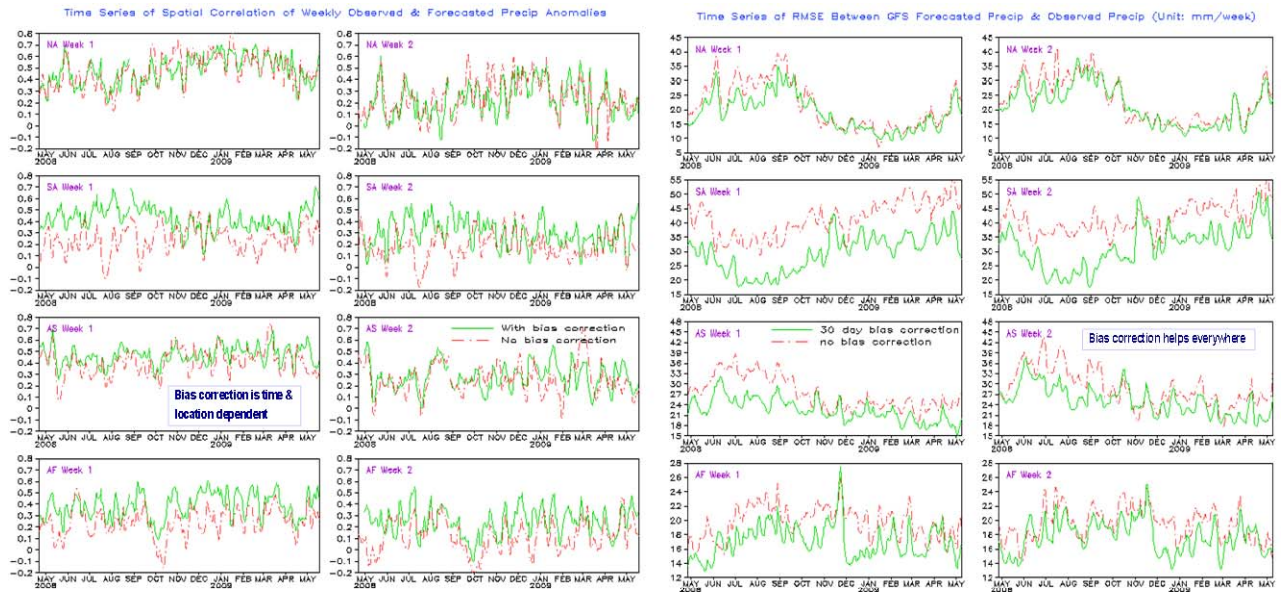

Fig. 2 Time series of 5-day running mean spatial correlation & RMSE of week-1 & week-2 observed & forecasted precipitation anomalies over North America (NA), South America (SA), Asia-Australia (AS) & Africa (AF), bias correction based on 30 days mean forecast errors.
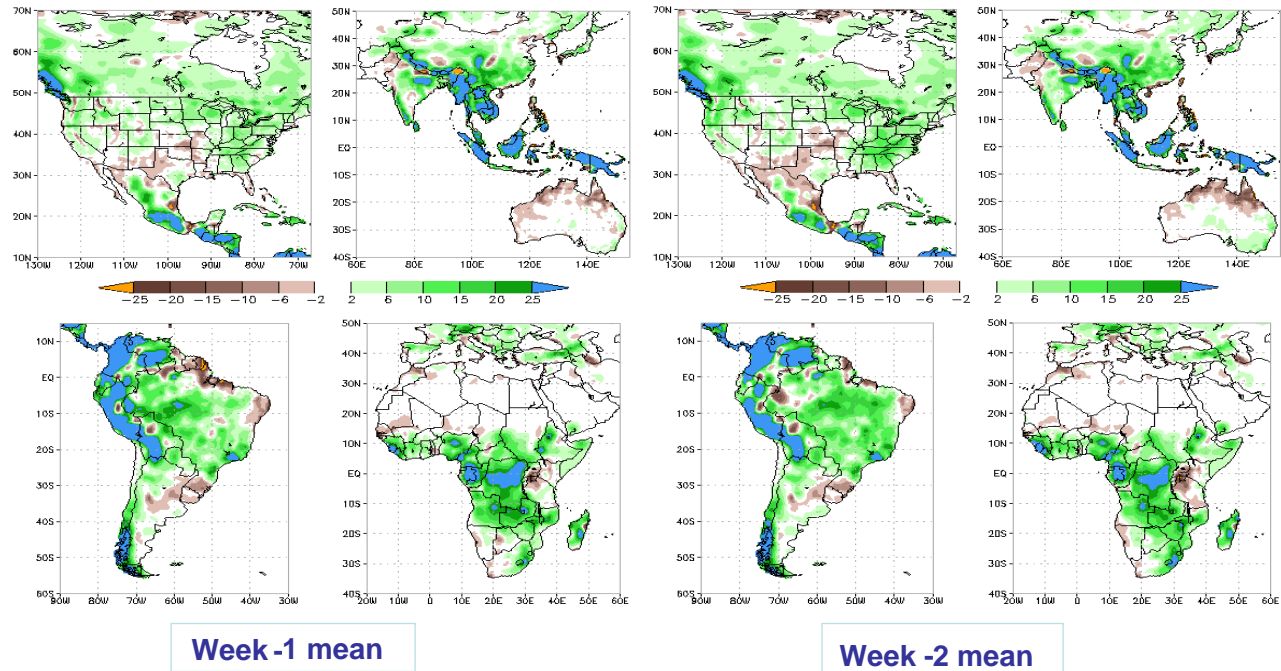

Week -1 mean   **Week -2 mean**

Fig. 3 Annual mean of week-1 (left) & week-2 (right) forecasted precipitation errors over North America, South America, Asia-Australia & Africa (unit: mm/week).

3. Performance of NCEP GFS week-1 and week-2 ensemble precipitation forecasts

Since the above bias corrections (with both 30 days and 7 days mean forecast errors) are performed every day, the data sets are archived on a daily basis for verification and research. Figure 1 shows the time evolution of daily spatial correlation of the week-1 and week-2 observed precipitation anomalies and GFS forecasted

precipitation anomalies over North America, corrected with 30 days mean forecast errors. The dominant features are a large day to day fluctuation and a clearly seasonal cycle in the GFS precipitation forecast skill, with the relative higher skill in the cold season and lower skill in warm season. In general, the annual mean of spatial correlation skill for the week-1 GFS precipitation forecasts is around 0.49 and 0.24 for the week-2 GFS precipitation forecasts.

Similar features for the bias corrected GFS ensemble precipitation forecasts are found in other regions, such as in South America, Asia-Australia and Africa monsoon regions, but with somewhat different forecast skills for week-1 and week-2 time scales (See Table 1 and Table 2 for more details).

Because the resolution of the GFS forecasts used here is on a 2.5x2.5 degree grid and the observed CPC daily Unified Global Precipitation Analysis is on a 0.5x0.5 degree grid, one can do the verification on either grid. A test has been conducted on both grids and the results show that the skill assessment does not depend much on the

Fig. 4 First 4 EOF patterns and their PCs of week-1 (left two columns) & week-2 (right two columns) ensemble forecasted precipitation errors over North America, South America, bias correction based on 30 days mean forecast errors.

grids, despite some higher resolution information may be lost when working on 2.5x2.5 grid. Some comparisons also have been done on the forecast skills from bias corrections based on 30 and 7 days mean forecast errors. The results show that the 30 days mean forecast errors are more robust than the 7 days mean forecast errors. In general, the forecast skills from bias correction based on 30 days mean forecast errors are slightly better than those from bias correction based on 7 days mean forecast errors.

Here one of major question is: Can bias correction improve GFS forecast skill? The results (Figure 2 and Table 1 & 2) show that in terms of spatial anomaly correlation the bias correction offers very little help in
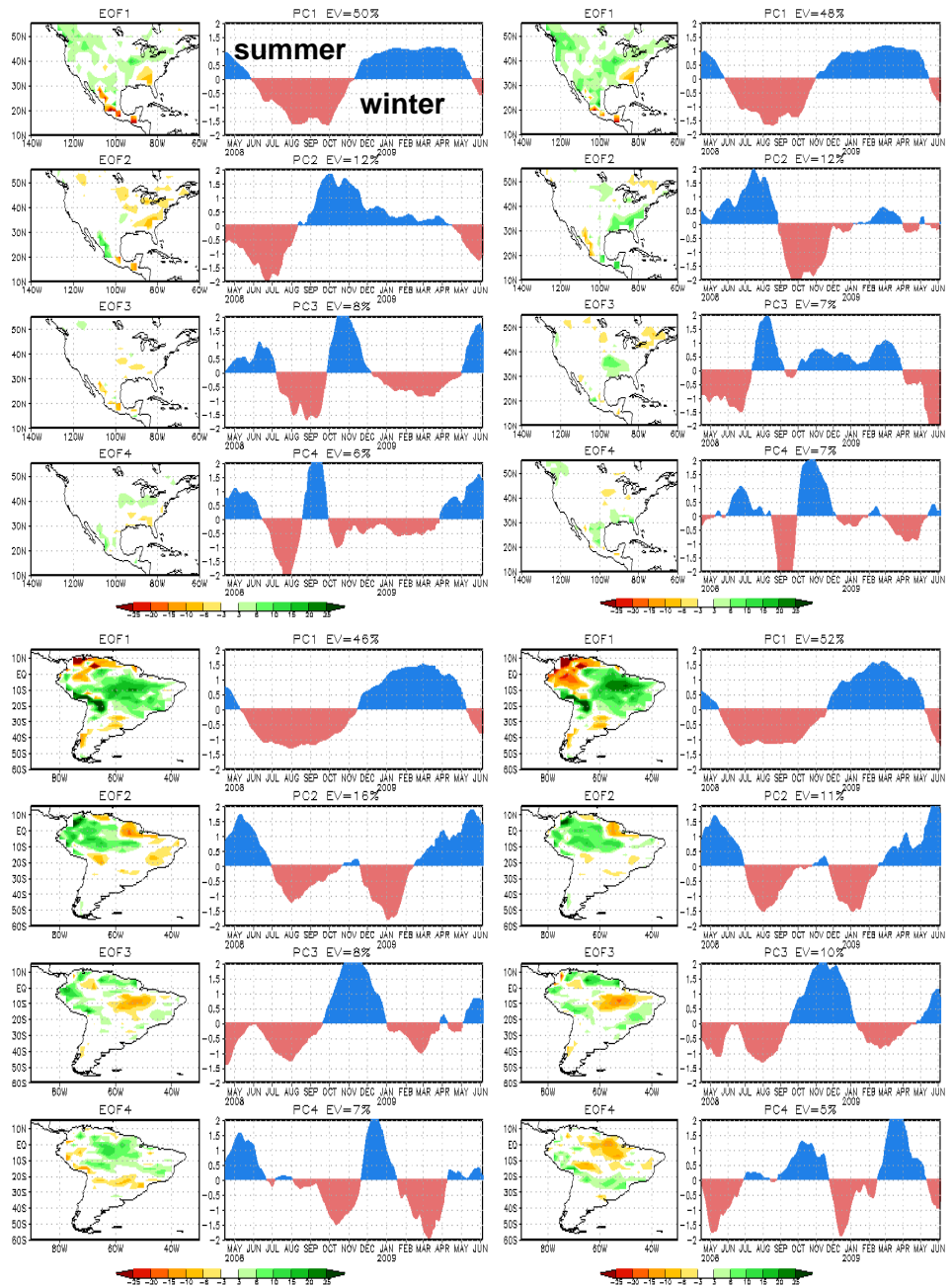
North America, considerable help in South America and Africa, and some help in Asia-Australia monsoon regions. In terms of root mean square error (RMSE), bias correction helps everywhere!!
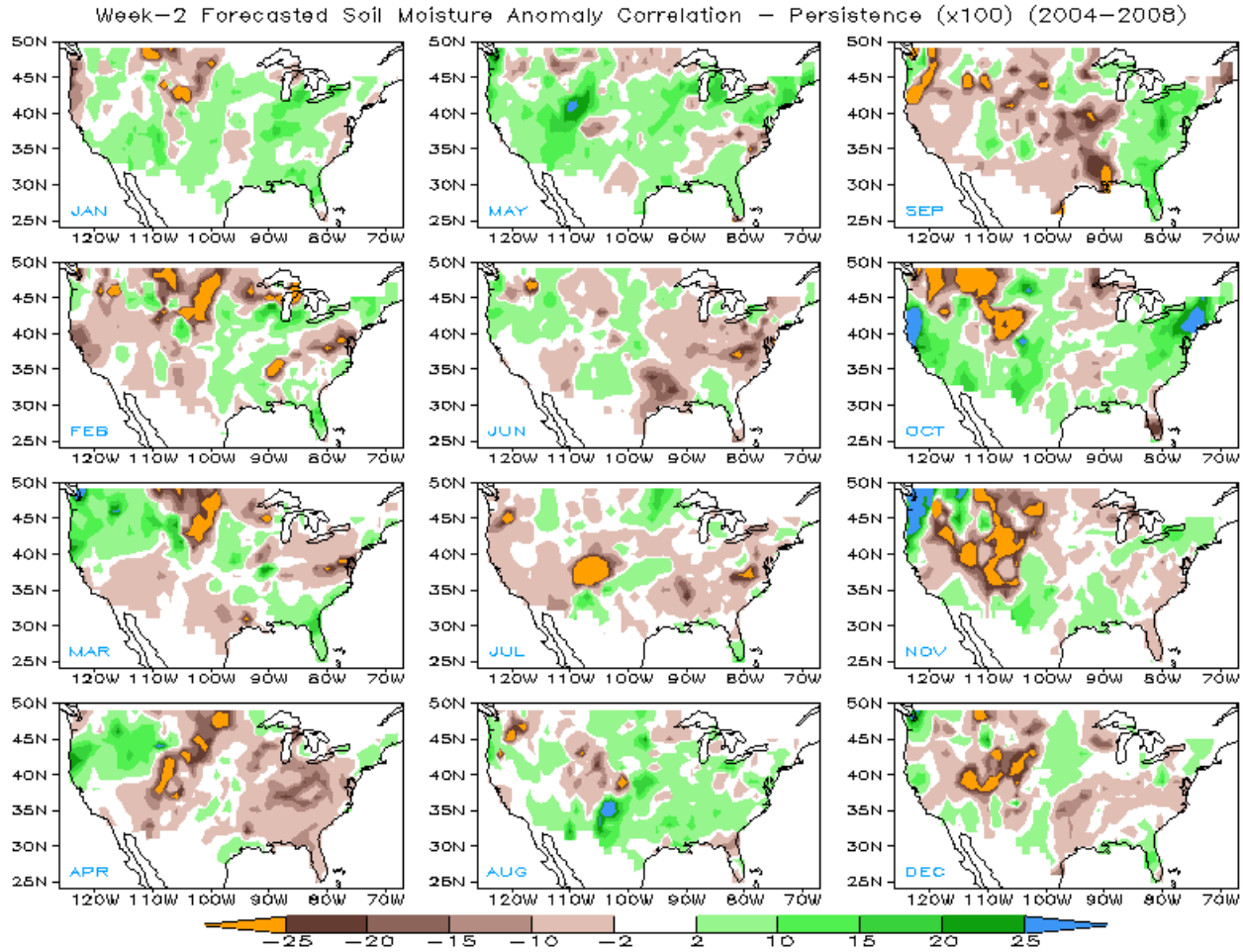

Fig. 5 The temporal anomaly correlation of the daily GFS week-2 soil moisture forecast minus its persistence.

4. Analysis of week-1 and week-2 forecast errors

In order to understand why bias correction works while it varies in space and time, some detailed analysis on the spatial-temporal structure of the mean forecast errors has been conducted. In general, the GFS forecast errors can be separated into two parts, i.e. the annual mean forecast error and its variation part around the annual mean, which was further decomposed by using EOF analysis (see equation 3).

$$Bias1,2(s,t) = Mean + \sum_{m=1}^{M} PC_m(t) * EOF_m(s) \qquad (3)$$

The annual mean of the GFS week-1 and week-2 ensemble precipitation forecast errors shows that the GFS tend to produce too much rainfall in most regions (Figure 3). The pattern and amplitude of the week-1 and week-2 forecast errors are very similar, indicating the GFS forecast errors are nearly saturated in week-1 period. The variation part (against annual mean) of the GFS week-1 and week-2 ensemble precipitation (30 day mean) forecast errors is displayed in Figure 4. The unexpected and most prominent features are that the GFS forecast errors are relative large-scale and low-frequency (annual and semi-annual cycles). The first two EOF modes of the GFS week-1 and week-2 ensemble forecast errors explains about 60% of the total variances. The above features exist almost everywhere (Asia-Australia and Africa are not shown here). The Bias correction shows a very large part of the annual mean forecast errors can be removed and some part of the variable forecast error can also be removed, especially in the cold season.

5. Application of the GFS ensemble forecast: soil moisture outlook

The bias corrected week-1 and week-2 GFS ensemble precipitation and T2m forecasts are used to drive the CPC leaky bucket land surface hydrological model forward up to two weeks over the US only. Because there is very little ground truth can be used, all land surface initial conditions and verification datasets are from the CPC leaky bucket model forced with daily observed precipitation and T2m.

Since the sea surface temperature and land surface soil moisture are the two important lower boundary variables and both of them have high persistence (or memory), so one interesting question (and an old "standard" in meteorology) is: can the soil moisture "dynamical" outlook (forced with GFS week-1 and week-2 ensemble forecasts) beat its persistence (*i.e.* provide more useful information than persistence)? For most land surface models, the land surface hydrological budget can be represented as:

$$dW/dt = P - E - R = F \qquad\qquad\qquad (4)$$

or $$W(t+1) = W(t) + F \qquad\qquad\qquad (5)$$

It is clear that if the F does not have sufficient skill, the GFS "dynamical" forecasts will lose against persistence (*i.e.* F=0). Figure 5 displays the spatial-temporal distribution of daily GFS forecasted week-2 soil moisture anomaly correlation minus its persistence in different 12 months for periods of Jan.1, 2004 to Dec. 31, 2008. In general, the GFS shows some useful skill over the west coast region, south east US and Texas, but constantly (except May) loses against persistence over the Rocky Mountain regions, which seriously degenerates the US overall performance of the GFS. Figure 6 depicts time evolution of the forecast skill and its persistence of week-1 and week-2 soil moisture anomalies averaged over the U.S. In general, both forecast and persistence reach their lowest values (most unpredictable time) around September, when soil moisture is in its driest season climatologically in the year. Overall, in terms of spatial correlation, the GFS dynamical forecast hardly beats persistence only by a very small number in week-1 and loses to the persistence in week-2. In terms of RMSE, the GFS dynamical forecast loses to persistence in both week-1 and week-2.

6. Summary

The above results show the bias corrected forecast skill of the NCEP GFS week-1 and week-2 ensemble precipitation presents large day to day fluctuation with a clear seasonal cycle. The overall week-1 and week-2 precipitation forecast skill is moderate. The GFS 30 day mean forecast errors are dominated by low-frequency (annual and semi-annual cycles) and relatively large-scale error patterns. Part of the forecast errors is removable. The effectiveness of the bias correction is time and space dependent.

Fig. 6 The time series of spatial correlation and RMSE from GFS week-1 (top) & week-2 (bottom) forecasted soil moisture anomaly and its persistence over the US. 30-day running mean is applied. The numbers in the plots are the means averaged over the whole periods (Nov. 1, 2003 to June 20, 2009).

The dynamical soil moisture forecast (*i.e.* land model forced with the bias corrected GFS week-1 and week-2 ensemble precipitation and 2 meter surface air temperature) has very high skill, but indicates that in general the current GFS is not good enough to beat soil moisture persistence (which is very high also) over the US. The inability to outperform the persistence relates to the skill of forecasted
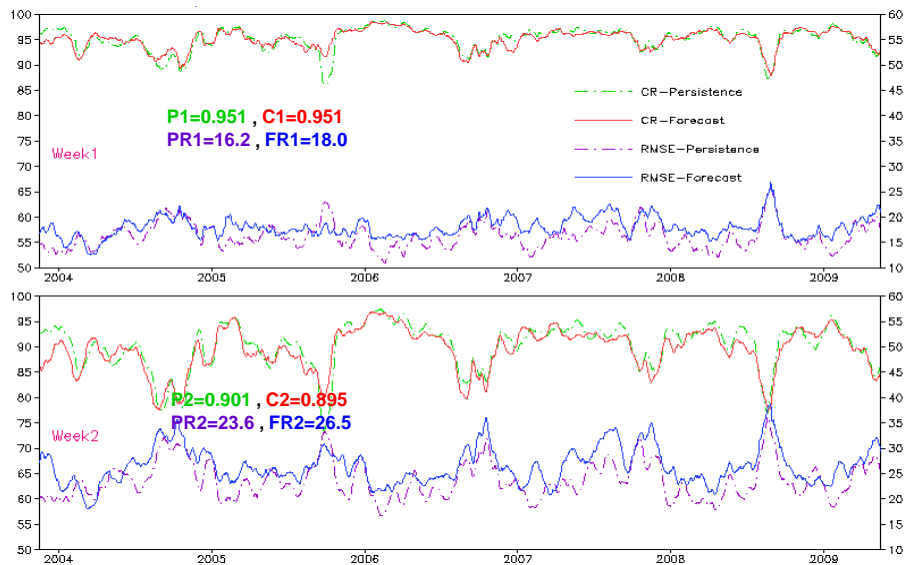
week-1 and week-2 precipitation not being above the threshold (*i.e.* anomaly correlation (AC) > 0.5 is required).

References

Dirmeyer, P., 2000: Using a global soil wetness data set to improve seasonal climate simulation. *J. Climate*, **13**, 2900-2922.

Huang, J., H.M. van den Dool, and K.P. Georgakakos, 1996: Analysis of model-calculated soil moisture over the United States (1931-1993) and applications to long-range temperature forecasts. *J. Climate*, **9**, 1350-1362.

Kanamitsu, M., C. Lu, J. Schemm, W. Ebisuzaki, 2003: The Predictability of Soil Moisture and Near-Surface Temperature in Hindcasts of the NCEP Seasonal Forecast Model. *J. Climate*, **16**, 510–521.

Koster, R. D., M. J. Suarez, R. W. Higgins, and H. M. Van den Dool, 2003: Observational evidence that soil moisture variations affect precipitation. *Geophys. Res. Lett.*, **30(5)**, 1241, doi:10.1029/2002GL016571, 2003

Van den Dool, H. M., Jin Huang and Yun Fan, 2003: Performance and Analysis of the Constructed Analogue Method Applied to US Soil Moisture over 1981-2001. *J. Geophys. Res.*, **108(D16)**, 8617, doi:10.1029/2002JD003114.

Van den Dool, H., 2007: Empirical Methods in Short-Term Climate Prediction. *Oxford University Press*, 215 pages.

Zhang, H. and C.S. Frederikson, 2003: Local and nonlocal impacts of soil moisture initialization on AGCM seasonal forecasts: A model sensitivity study. *J. Climate*, **16**, 2117-2137.

Office of Science and Technology
National Weather Service
National Oceanic and Atmospheric Administration
U.S. Department of Commerce

http://www.nws.noaa.gov/ost/climate/STIP/fy09jsctb.htm