Assessing the idealized predictability of precipitation and temperature in the NCEP Climate Forecast System

Lifeng Luo[1,2] and Eric F. Wood[2]

Received 28 November 2005; revised 6 January 2006; accepted 13 January 2006; published 23 February 2006.

[1] The 24-year retrospective forecast data set from the NCEP Climate Forecast System (CFS) is analyzed to study its idealized predictability of precipitation and temperature under its current configuration. The analysis approach assumes the forecasting model and system to be predicted share exactly the same physics so that the idealized predictability is calculated and serves as the upper limit of the predictive skill in practical forecasts. The analysis shows that CFS is not capable of predicting itself over much of the mid-latitudes land areas for precipitation and temperature anomalies having small temporal (monthly) and spatial ($2.5° \times 2.5°$ grid) scales at lead-times longer than a month. Anomalies become more predictable with the increase in temporal and spatial scales and with the decrease in lead-times, as illustrated with results from the central US region. The results imply that additional care should be taken when using climate model seasonal forecast products.
Citation: Luo, L., and E. F. Wood (2006), Assessing the idealized predictability of precipitation and temperature in the NCEP Climate Forecast System, *Geophys. Res. Lett.*, *33*, L04708, doi:10.1029/2005GL025292.

1. Introduction

[2] An accurate weather forecast beyond approximately two weeks is impossible due to the chaotic nature of the climate system [*Lorenz*, 1963] and the inability to obtain perfectly accurate initial conditions. A seasonal prediction of the climate system is thought possible because there are slowly varying components in the coupled ocean-atmosphere-land system. If the slowly varying components of the climate system, such as the sea surface temperature, can be predicted several months in advance, then there is a good chance that the mean state of the climate system at longer lead-times can be described with reasonable accuracy. General circulation models (climate models) are very useful tools for making such seasonal predictions. However, the predictive skill of current climate models is still limited in seasonal predictions, especially for variables that are mostly relevant to our daily life, i.e., precipitation and near surface air temperature. To improve the predictive skill, it is essential to understand the predictability of precipitation and temperature along with many other climate model variables. The question that we are trying to answer is, what are the predictability limits of our forecast models?

[3] With the increased understanding about the interaction among different components of the climate system and improved numerical models, as well as the dramatic increase in computing power, long-lead seasonal forecasting with coupled ocean-atmosphere-land models has become feasible. In August 2004, the Climate Forecast System (CFS), a fully coupled model representing the interaction between the Earth's oceans, land and atmosphere, became the operational system for seasonal forecasting at the National Centers for Environmental Prediction (NCEP). For the first time in U.S. operational seasonal prediction, a dynamical modeling system has demonstrated a level of skill in forecasting U.S. surface temperature and precipitation that is comparable to the skill of the statistical methods used by the NCEP Climate Prediction Center [*Saha et al.*, 2006]. To facilitate real-time seasonal forecasts, a set of fully coupled retrospective forecasts covering a 24-year period (1981–2004) were produced with CFS. These multi-member ensemble retrospective forecasts provide meaningful information to measure the forecast skill of the system. This data set also provides the necessary information to study the idealized predictability of precipitation and near surface air temperature at seasonal time scales within the system.

[4] The predictive skill of precipitation or temperature is a multidimensional variable. It should vary geographically with location (x, y), lead-time (τ), season (t), and with temporal (T) and spatial scales (L). If predictability is defined as the possible maximum predicative skill that a forecast system can achieve, then predictability is also a 6-dimensional (6-D) variable. Higher predictability means an event is more predictable and a high predictive skill is potentially achievable. Understanding the predictability of a system helps us to concentrate on improving predictions of the predictable components, and to prevent us from spending time trying to predict the unpredictable.

[5] *Koster et al.* [2000, 2004] studied the seasonal predictability of precipitation with the NSIPP (NASA Seasonal-and-Interannual Prediction Project) model, and they provided a metric to measure predictability within a specific modeling framework. As CFS is now the operational system for US seasonal forecasting, similar studies with CFS are needed that will assess its predictability and provide guidelines for future model improvements and thus forecasts. In this study, we use the 24-year retrospective forecast data set to assess the idealized predictability of warm-season precipitation and temperature in CFS.

2. Data and Methodology

[6] The 24-year (1981–2004) nine-month multi-member ensemble retrospective forecast data set from CFS is pro-

[1]Program in Atmospheric and Oceanic Sciences, Princeton University, Princeton, New Jersey, USA.
[2]Environmental Engineering and Water Resource, Department of Civil and Environmental Engineering, Princeton University, Princeton, New Jersey, USA.

0094-8276/06/2005GL025292$05.00

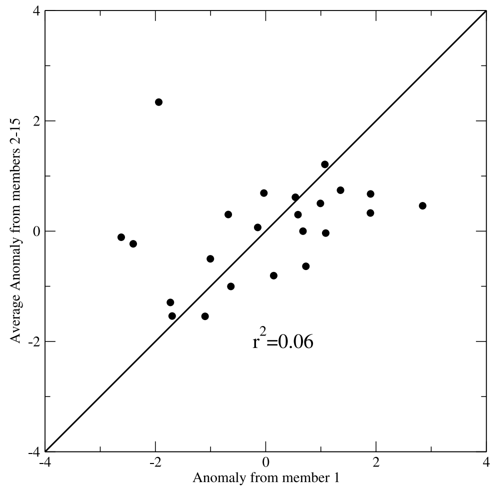

Figure 1. Scatter plot for the idealized predictability analysis, showing the degree to which the Climate Forecast System can "predict itself" at a central U.S. grid cell (37.5°N, 97.5°W). The lead-time for this prediction is 16 days, and the anomaly is from a 32-day running mean. The x-axis represents the precipitation anomaly generated by the first members of all May forecasts, and the y-axis represents the forecasted precipitation anomaly averaged over the remaining fourteen ensemble members from the same ensemble set. Twenty-four points are from the twenty-four years (1981–2004). The solid line is the 1:1 line.

vided by NCEP. For each calendar month, fifteen runs were made from different initial conditions that were carefully selected to span the evolution of both the atmosphere and ocean in a continuous fashion [*Saha et al.*, 2006]. Each run is a full nine-month integration in addition to the first partial month. A total of 4320 runs were performed and the entire data set is equivalent to a 3240-year model integration. The technical details about CFS and the production of this retrospective forecast data set are given by *Saha et al.* [2006] and the NCEP web site (http://cfs.ncep.noaa.gov/). Both monthly and sub-daily (12 hourly) fields of precipitation, 2-m air temperature and other variables are available. The spatial resolution of this data set is 2.5° × 2.5°, and can be aggregated to larger spatial scales. The data can also be temporally averaged to produce output at larger temporal scales. This data set offers a unique opportunity to determine the 6-D variation of predictability in CFS.

[7] The estimation of the CFS idealized predictability follows the method of *Koster et al.* [2004], where one member of the ensemble is assumed to be the "truth" while the rest of the ensemble are considered as model forecasts. Because observed climate is in fact one of many possible realizations of the climate system, this approach will produce an analog to the real world where the forecast model and the underlying "climate system" share exactly the same physics. Ideally, we would like our climate models to share the same physics as the real world, but this is unlikely because the real climate system is much more complex to be captured with finite grids, simplifications for resolved processes and parameterizations for unresolved sub-grid processes in current climate models.

[8] With the pair formed by the "truth" (from one member of the ensemble) and the prediction (from the remainder of the ensemble), the predictive skill is expressed by the square of the correlation coefficient (r^2) and can be calculated at each grid (x, y) for a given lead-time (τ) and given season (t), when similar forecasts from all years are used. Figure 1 shows an example where the first members of May forecasts of all years are considered as truth and the averages of the remaining 14 members are taken as the best predictors. Each of the 24 points on the scatter plot represents a pair of truth and predictor of the monthly mean precipitation of May with a lead-time of 16 days at a 2.5° × 2.5° grid centered at (37.5°N, 97.5°W). A value of r^2 equal to 0.06 is obtained at this grid for the given T, L, t and τ. A higher correlation would indicate that the averages of the 14 ensemble members are good predictors of the first member that is taken as "truth". The same analysis is performed repeatedly by letting the 2nd, 3rd, …, 15th member as the truth respectively, and in total 15 different values of r^2 are calculated. These 15 values of r^2 are then averaged as the final estimate of the predictive skill of the system in predicting itself, which is also referred as idealized predictability. Since the correlation is a 6-D variable, we can calculate r^2 for each grid, lead-time, season, and with different temporal and spatial averaging lengths. The array of r^2 then can be analyzed with respect to each of the six dimensions. Without using observations, this approach can assess how well CFS predicts itself with uncertain initial conditions, and the idealized predictability is the upper limit of the predictability and predictive skill that CFS can achieve in operational forecast.

3. Results

[9] In this study, the averaged r^2 values are calculated using all the retrospective forecasts (15 members × 24 years) initialized between April and May, so the 9-month target period is May to January (of the following year). The climatology of precipitation and temperature is defined as the average of the 360 runs with a 7-day smoothing. The anomalies are then spatially averaged over the regions indicated by the boxes in Figure 2, and all the boxes are

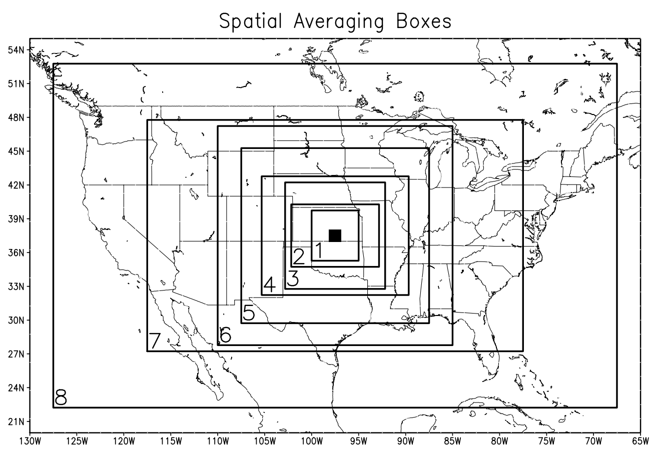

Figure 2. The regions used in spatial averaging. All the grids inside a box including the ones passed through by the outlines are used in the averaging. The solid black square indicates the original grid for which we want to make a forecast.

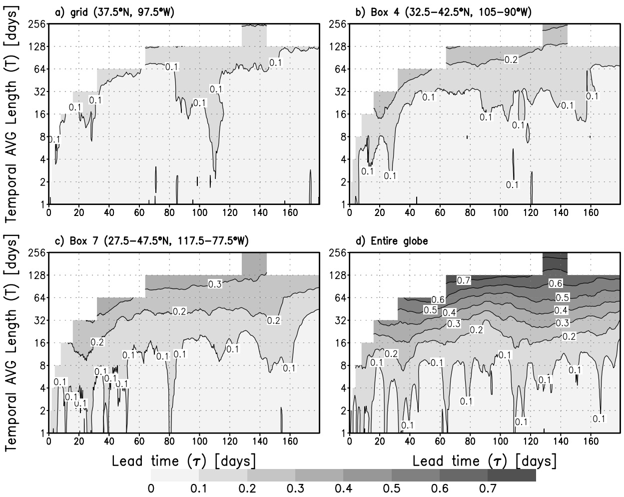

Figure 3. (a–d) The variation of precipitation predictability, expressed as the averaged square of correlation coefficient (r^2), with lead-time (x axis), temporal averaging length (y axis) and spatial averaging scale.

centered at the grid (37.5°N, 97.5°W) in the central US. The temporal running averaging uses a window size of 1, 2, 4, 8, 16, 32, 64, 128, and 256 days, respectively. Figures 3 and 4 show the results for precipitation and 2-m air temperature respectively, where the averaged r^2 is plotted against lead-time (τ), temporal averaging scale (T), and at four different spatial averaging scales (L). Here the lead-time is defined as the time between the target date and the first day when all 15 ensemble members are available. Because of the specific way these ensemble members were generated, the atmospheric states in each member at lead-time zero are already quite different. Therefore, r^2 does not appear to be close to 1 when τ is close to 0.

[10] Over the selected grid (Figure 3a), r^2 hardly reaches 0.1 when the temporal averaging length is shorter than 32 days (approximately monthly averaging). If we use $r^2 = 0.1$ as the threshold to determine the predictability, then Figure 3a suggests that monthly mean precipitation over a 2.5° × 2.5° grid in the central US is almost unpredictable, even within the idealized system. But at the seasonal scale (T > 90 days), anomalies can be predicted at a skill level above 0.1. Comparison among all the panels in Figures 3 and 4 suggest that an anomaly is more predictable when it is at a larger spatial scale, or at a longer temporal scale and a shorter lead-time. For example, at a lead-time of two months, the anomaly in monthly mean precipitation over the entire US (box 7) is more predictable than the anomaly over one single grid or a smaller region. It is not surprising that temperature is more predictable than precipitation in general. In fact, under the same condition, the predictive skill for temperature can be significantly higher. The global mean surface temperature is highly predictable with CFS, but it may not have much practical usefulness.

[11] Figure 5 shows the global distribution of averaged r^2 for monthly mean precipitation forecasts at the CFS grid scale with lead-time of 16 days. Similar maps for different lead-times and spatial scales can be easily produced but are not shown. It is obvious that precipitation anomaly over the tropical ocean is much more predictable than that over mid-latitude land. In fact, there are not many land areas where monthly mean precipitation shows any significant predictability at the CFS grid scale and 16-day lead-time. The monthly mean precipitation at longer lead-time shows even less predictability virtually everywhere.

4. Concluding Remarks

[12] The 24-year retrospective forecasts from the NCEP climate forecast system (CFS) are used to assess its idealized predictability—the capability of CFS to predict itself. Shown in this paper is the idealized predictability in the *current* CFS configuration, which will always be an upper limit of practical predictability and predictive skill because its physics are simpler than the real climate system. However, when the initialization procedures and parameterizations are improved, the idealized predictability will change and needs to be re-assessed. Additionally, it is possible that a new configuration will have improved predictive skill that exceeds the idealized predictability of earlier configurations.

[13] For real applications of a forecast system, the actual predictive skill will be less than its idealized predictability. This implies that knowledge of the idealized predictability for a forecast system provides necessary knowledge of the upper limit, but is insufficient for assessing actual predictive skill of the forecast system. The results presented here suggest that there is a significant limit to the idealized predictability of the current CFS. This suggests that there are significant limitations in the predictive skill in the CFS forecasts of precipitation and temperature at seasonal timescales. This further suggests that seasonal prediction of the climate system may be limited at this time and ultimately may be limited in nature. It is worth mentioning that the current results are obviously model-dependent. A different climate modeling system may have higher idealized predictability within the same framework, as shown for the NSIPP model used by *Koster et al.* [2004]. The results from the Global Land-Atmosphere Coupling Experiment (GLACE) [*Koster et al.*, 2006; *Guo et al.*, 2006] show that the NCEP Global Forecast System (GFS) (a higher resolution

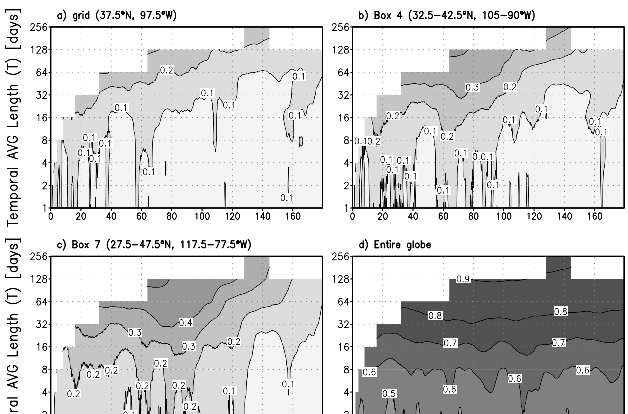

Figure 4. (a–d) Same as Figure 3, but for 2-m air temperature.

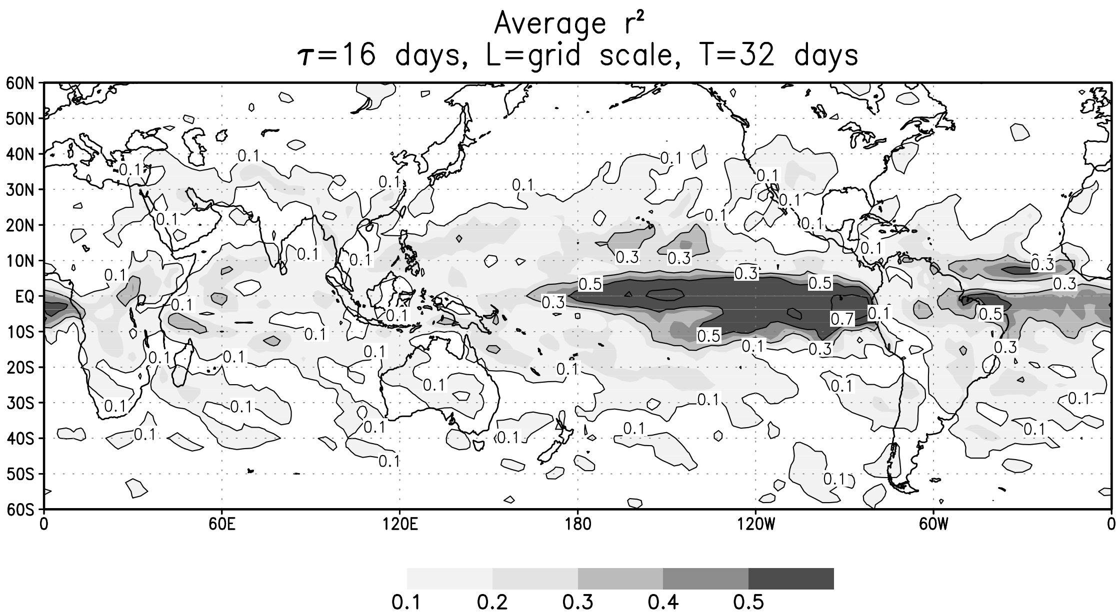

Figure 5. Global distribution of averaged r^2 for monthly mean (32-day running average) precipitation at grid scale ($2.5° \times 2.5°$) with lead-time of 16 days.

of CFS atmosphere model) coupled with the OSU land model, has the weakest land-atmosphere coupling strength among many climate models. Whether this explains the low idealized predictability seen in the current CFS needs further investigation, but these studies should help the climate modeling community move forward in making more useful seasonal predictions.

[14] The work presented here was for the May-to-January forecasts. Not shown is the variability in the idealized predictability with seasons, or with different climatic conditions such as ENSO versus non-ENSO periods. We recognize that the idealized predictability may change under different climatic conditions, but the limited length of the available retrospective forecast data set is too short to stratify by climatic categories and still obtain statistically meaningful results.

[15] Although the computed predictability is for an idealized system (i.e., predictions of the modeled climate system), the results have important implications for actual seasonal forecasts. For example, in a hydrological forecast system using seasonal climate predictions of precipitation and temperature, the hydrological predictions of snow, soil moisture and streamflow will be impacted by the underlying predictive skill of the precipitation and temperature predictions. This implies that the spatial and temporal scales, and forecast lead time, at which the climate model forecast shows skill and is meaningful must be considered in designing a hydrologic forecast system.

[16] **Acknowledgments.** This work was supported by NOAA grants NA17RJ2612 and NA03O3AR4310001. We thank Randy Koster and another anonymous reviewer for their helpful review and comments.

References

Guo, Z., et al. (2006), GLACE: The Global Land-Atmosphere Coupling Experiment. 2. Analysis, *J. Hydrometeorol.*, in press.

Koster, R. D., M. J. Suarez, and M. Heiser (2000), Variance and predictability of precipitation at seasonal-to-interannual timescales, *J. Hydrometeorol.*, *1*, 26–46.

Koster, R. D., M. J. Suarez, P. Liu, U. Jambor, A. Berg, M. Kistler, R. Reichle, M. Rodell, and J. Famiglietti (2004), Realistic initialization of land surface states: Impact on subseasonal forecast skill, *J. Hydrometeorol.*, *5*, 1049–1063, doi:10.1175/JHM-387.1.

Koster, R. D., et al. (2006), GLACE: The Global Land-Atmosphere Coupling Experiment. 1. Overview, *J. Hydrometeorol.*, in press.

Lorenz, E. N. (1963), Deterministic nonperiodic flow, *J. Atmos. Sci.*, *20*, 130–141.

Saha, S., et al. (2006), The NCEP Climate Forecast System, *J. Clim.*, in press.

L. Luo and E. F. Wood, Environmental Engineering and Water Resource, Department of Civil and Environmental Engineering, Princeton University, Princeton, NJ 08544, USA. (lluo@princeton.edu)