Prediction of Monthly-Mean Temperature: The Roles of Atmospheric and Land Initial Conditions and Sea Surface Temperature

MINGYUE CHEN, WANQIU WANG, AND ARUN KUMAR
Climate Prediction Center, National Centers for Environmental Prediction, Camp Springs, Maryland

(Manuscript received 12 February 2009, in final form 21 August 2009)

ABSTRACT
Using the retrospective forecasts from the National Centers for Environmental Prediction (NCEP) coupled atmosphere–ocean Climate Forecast System (CFS) and the Atmospheric Model Intercomparison Project (AMIP) simulations from its uncoupled atmospheric component, the NCEP Global Forecast System (GFS), the relative roles of atmospheric and land initial conditions and the lower boundary condition of sea surface temperatures (SSTs) for the prediction of monthly-mean temperature are investigated. The analysis focuses on the lead-time dependence of monthly-mean prediction skill and its asymptotic value for longer lead times, which could be attributed the atmospheric response to the slowly varying SST. The results show that the observed atmospheric and land initial conditions improve the skill of monthly-mean prediction in the extratropics but have little influence in the tropics. However, the influence of initial atmospheric and land conditions in the extratropics decays rapidly. For 30-day-lead predictions, the global-mean forecast skill of monthly means is found to reach an asymptotic value that is primarily determined by the SST anomalies. The lead time at which initial conditions lose their influence varies spatially. In addition, the initial atmospheric and land conditions are found to have longer impacts in northern winter and spring than in summer and fall. The relevance of the results for constructing lagged ensemble forecasts is discussed.

1. Introduction
The short- to medium-range numerical prediction of day-to-day weather relies on atmospheric initialization: the accurate specification of atmospheric pressures, temperatures, winds, and humidity at the beginning of the forecast. For the time scales of seasonal or longer, on the other hand, a primary source of atmospheric prediction skill is the lower boundary conditions such as the variability associated with the tropical Pacific sea surface temperatures (SSTs) related to El Niño–Southern Oscillation (ENSO). However, prediction of the variability on time scales between individual weather events and seasonal-mean climate presents a challenging problem, because it depends both on the atmospheric initial conditions and predictability resulting from slowly varying boundary conditions (Walsh and Ross 1988; Phelps et al. 2004; Vitart 2004; Reichler and Roads 2005a,b), and it is the focus of this paper.

During the past three decades, the predictability of seasonal climate associated with surface boundary conditions has been studied extensively with atmospheric general circulation model (AGCM) simulations. Anomalous SST associated with ENSO are found to be the primary source of predictability of seasonal-mean circulation (Shukla and Wallace 1983; Brankovic et al. 1994; Livezey et al. 1996; Barnett et al. 1997; Kumar and Hoerling 1998; Anderson et al. 1999; Kumar et al. 2003). The atmospheric response (i.e., the predictive signal) to the anomalous SST forcing is characterized by large geographical variations with largest signal-to-noise ratio in the tropics. The extratropical response generally has a smaller signal-to-noise ratio, partially because of larger noise associated with midlatitude synoptic-scale systems. Over certain regions, however, extratropical atmospheric seasonal variability is documented to have a pronounced response, such as over the Pacific North America (PNA) region in Northern Hemisphere winter, to interannual variability in tropical SST (Kumar and Hoerling 1995, 1998; Shukla 1998; Shukla et al. 2000). Compared to the seasonal means, the response of monthly mean to ocean surface forcing, however, has not been well documented, although it is expected that, because of larger internal
variability of the atmospheric monthly means, the predictability would be smaller.

It is also expected that initial conditions will have a large influence on the prediction skill of monthly means. The influence of initial conditions on the predictability of monthly means, however, has been difficult to quantify. Shukla (1981) performed a study on the dynamical predictability of monthly means with prescribed boundary condition and concluded that the extratropical atmospheric response is sensitively dependent upon the initial conditions and that the atmospheric monthly means may be predictable up to a one-month lead time. Recently, Phelps et al. (2004) investigated the role of atmospheric initial condition in the predictability of both monthly and seasonal means and found that the atmospheric initial conditions have little influence on the monthly and seasonal means for a lead time of one month or longer. Phelps et al. (2004) showed that the model forecast skills are high over the tropics and the PNA region, especially for the period of strong ENSO events, and predictability is mainly determined by the interannual variations in the SST anomalies. However, their analysis did not consider the influence of initial condition with a lead time shorter than one month. Further, their study was based on January and January–March (JFM) retrospective forecast from an AGCM and used observed SST. For operational forecast systems, SST itself is predicted and the results may differ compared to an analysis where the observed SST is specified.

Reichler and Roads (2005a,b) analyzed the relative roles of initial atmospheric and land surface conditions and oceanic surface conditions in the forecast of tropical monthly-mean fields. The forecasts were based on perfect-model experiments with different combinations of initial and prescribed ocean boundary conditions. It was found that initial conditions dominate the forecast for the first three weeks, after which the oceanic boundary forcing becomes the main contributor to tropical prediction skill. Their study focused on the tropical region in boreal winter and the analysis was performed through the comparison of perturbation runs with a control run from the same model (rather than with observations).

In this study, we extend the previous analysis on predictability and prediction skill of monthly means to investigate initialized predictions that incorporate the influence of observed atmospheric and land initial conditions and slowly varying ocean surface conditions. Our analysis is based on retrospective forecasts from the current National Centers for Environmental Prediction (NCEP) operational coupled Climate Forecast System (CFS) and the Atmospheric Model Intercomparison Project (AMIP) simulations from its atmospheric component, the NCEP atmospheric Global Forecast System (GFS). Building upon the previous studies that used prescribed observations of SSTs, the present analysis is based on prediction of SST by the CFS; further, a comparison with the AMIP simulations allows an assessment of the influence of atmospheric and land initial conditions with the forecast lead time. The foci of our analysis include the following: 1) How does prediction skill of monthly means evolve with lead time, and what is the time scale that the influence of atmospheric and land initial conditions lasts? 2) What is the origin for the asymptotic value of prediction skill for longer lead-time initialized predictions? Our analysis is also relevant to constructing forecasts for monthly means based on lagged ensemble. Although the use of lagged members increases ensemble size, the forecast skill may be degraded if the lead time of the forecasts included in the ensemble is too long. An analysis of lead-time dependence of monthly-mean forecast will help determine the optimal construction of lagged ensemble forecasts of monthly means.

A brief description of the models, model forecasts and simulations, observations, and methodology used in the analysis is provided in section 2. Analysis and results are given in section 3. Summary and discussion are presented in section 4.

2. Models and data

a. Models

The model used in this study is the NCEP coupled CFS. It is the current NCEP operational system for dynamical seasonal forecasts. The atmospheric component of the CFS is the 2003 version of NCEP atmospheric GFS model, with a spectral truncation of 62 waves (T62) in the horizontal (equivalent to nearly a 200-km grid) and 64 layers in vertical. The oceanic component of the CFS is the Geophysical Fluid Dynamics Laboratory Modular Ocean Model V.3 (MOM3; Pacanowski and Griffies 1998) with a zonal resolution of 1° and a meridional resolution of 1/8° between 10°S and 10°N, gradually increasing through the tropics until becoming fixed at 1° poleward of 30°S and 30°N. The vertical resolution of MOM3 is 10 m from the surface to 240-m depth, gradually increasing to about 511 m in the bottom layer. The atmospheric and oceanic components are coupled without any flux adjustment. The two components exchange daily averaged quantities once a day. The CFS is an initialized dynamical prediction system. The oceanic initial conditions are from the NCEP Global Ocean and Data Assimilation System (GODAS), and the atmospheric and land initial conditions are from the NCEP/Department of Energy (DOE) reanalysis-2 (R2) (Kanamitsu et al.
More details of the CFS and its simulations can be found in Wang et al. (2005) and Saha et al. (2006).

b. Data

Monthly-mean retrospective forecasts (hindcasts) from the CFS for 1981–2006 are used in this study. For each month, CFS produced three sets of forecast runs from 0000 UTC initial conditions around the 11th, 21st, and 1st day of the month, each set consisting of five forecast runs from an identical oceanic initial state and five observed atmospheric and land initial conditions that are one day apart. Each forecast run is a full nine-month integration. More details about the design of the CFS retrospective forecasts can be found in Saha et al. (2006).

In this study, monthly-mean forecasts with lead times ranging from 0 days to 2 months are used to analyze the lead-time dependence of forecast skills. For simplicity of presentation, we assume two adjacent sets of forecasts are 10 days apart, although the exact gap between adjacent sets varies from 8 days (e.g., between 21 February and 1 March) to 11 days (e.g., between 21 November and 1 December). The definitions of the lead time are further summarized in Fig. 1. For a specific target month, the 0-day-lead monthly-mean forecast is defined from an ensemble mean of five forecast runs from the initial condition set centered on the 1st day of the target month, the 10-day lead from the 21st day of the month before the target month, the 20-day lead from the 11th day of the month before the target month, and so on, up to a 60-day lead.

The AMIP integration data used in the analysis are simulations from the NCEP atmospheric model GFS, which is the atmospheric component of the coupled CFS model. The simulations are forced with observed SST (Reynolds et al. 2002) and cover the same time period from 1981 to 2006 and include five ensemble members. Different AMIP simulations start with slightly different atmospheric initial conditions but are forced with the same observed SSTs. Further, even though the SST forcing for different AMIP simulations is the same, the evolution of monthly means, because of atmospheric internal variability, can differ considerably from each other (Phelps et al. 2004). As for the CFS forecasts, an ensemble mean of five AMIP simulations is used to assess the simulation skill and to compare with the forecast skill in the CFS with different lead times.

Our analysis is focused on monthly-mean 2-m temperature, one of the key variables in the operational forecast (O’Lenic et al. 2008). The observed monthly-mean 2-m temperature data are from the Climate Anomaly Monitoring Systems (CAMS) constructed at the National Oceanic and Atmospheric Administration (NOAA)/NCEP/Climate Prediction Center (CPC; Ropelewski et al. 1985).

c. Methodology

The analyses in this study are based on retrospective forecasts from the CFS and AMIP simulations. Because the CFS is a fully coupled ocean–land–atmosphere initialized dynamical seasonal prediction system, its predictions
incorporate both the influence of rapidly evolving observed initial conditions (e.g., atmosphere) and of slowly varying boundary conditions (that are themselves initialized and predicted: e.g., SST and land). On the other hand, the source of the predictability in the AMIP simulations is the observed SST alone. Although SSTs are predicted by the CFS for the short-lead forecasts discussed in this paper (less than one month), because SST forecast skill is very high, the influence of differences in SSTs between AMIP simulations and CFS forecasts on the atmosphere are likely to be small. Differences between AMIP simulations and CFS forecasts, therefore, are primarily due to atmospheric and land initial conditions. Improvements in the ocean observing system and better specification of ocean initial conditions could result in marginal improvement in the SST skill for the short-lead CFS forecasts and would lend further credibility to our assumption.

Availability of the CFS retrospective forecasts and AMIP simulation together allow us to examine the lead-time dependence of prediction skill. The prediction skill and its lead-time dependence are assessed by analyzing the monthly-mean forecasts with different lead times from the CFS. For longer lead times, when the influence of initial conditions is expected to diminish, the skill in the CFS forecasts is compared with the simulation skill from the AMIP integrations to ascertain the influence of the SST initial condition. The asymptotic value of prediction skill resulting from the boundary condition of SSTs is therefore examined by comparing the prediction skills from the CFS at different lead times with those from the AMIP simulation.

3. Results

The measure of prediction skills is defined here in terms of monthly anomaly correlation (AC) between forecasts and observations. The monthly anomalies of the CFS forecasts, AMIP simulations, and observational temperature are computed with respective climatologies for the time period from 1981 to 2006. The analysis is summarized for land areas only.

a. Lead-time dependence forecast skill

We begin the analysis with spatial maps of monthly-mean forecast skills. Figure 2 shows the spatial structure of temporal anomaly correlation skill for the monthly-mean temperature forecast over the global land and for forecast lead times from 0 to 40 days. For each calendar month, the anomaly correlation for the monthly means is computed over 1981–2006 and results are averaged over all calendar months. Areas with AC less than 0.2 are masked with light gray color, and white areas indicate that observational data for verification are not available.

The correlation for 0-day lead is greater than 0.3 over most of the land and is above 0.5 over some regions, such as North America, Europe, western and northern Asia, northern Africa, southeast and northwest of Australia, and part of eastern Brazil. The forecast skill decreases rapidly from 0-day lead to 10-day lead over most of the land areas, especially over high latitudes. For a large part of the regions with AC > 0.5 at the 0-day lead, the AC values drop below 0.2 at 10-day lead. As the lead-time further increases, the AC continues to decrease in the extratropics but stays largely unchanged in the tropics. For lead times of 30 days or longer, the correlation higher than 0.3 is mostly confined to a few tropical areas (equatorial Africa, part of the Maritime Continent, and northern South America). The distribution of AC at lead times beyond 40 days is nearly the same as that at a 40-day lead (not shown).

The lead-time dependence of the correlation averaged over the entire global land is summarized in Fig. 3. The forecast skill is above 0.5 at shortest 0-day lead. As the lead-time increases to 10 days, the AC decreases rapidly from above 0.5 to below 0.3. Afterward, the decrease of the AC with increasing lead time becomes slower. For lead times longer than 40 days, the global-mean value of AC becomes less than 0.1. These features are consistent with the spatial distribution in Fig. 2. The black line in Fig. 3 is the corresponding AC from the AMIP simulation and will be discussed later.

The global distribution of the correlation in Fig. 2 shows large latitudinal variations. To further analyze the changes of the lead-time dependence with latitude, zonal mean of forecast skill is shown in Fig. 4. Each curve in Fig. 4 represents forecast skill for a specific lead time from 0 to 50 days. For the 0-day-lead forecast, the skill is larger in high latitudes of both hemispheres than that for the tropics, with largest values near 60°N. Overall, the skill in the Northern Hemisphere is higher than that in the Southern Hemisphere. At lead times of 10 days and longer, the skill is much smaller than that at 0-day lead and the skill in the extratropics becomes smaller than that in the tropics.

In the extratropics, the skill reduces significantly compared to the skill at 0-day lead and decreases monotonically with lead time. Beyond a 30-day lead time, the skill becomes less changeable. In the tropics, the change in skill with lead time is very small with correlation around 0.34. The faster decrease of AC in the extratropics compared to that in the tropics indicates that the initial condition may have a stronger influence on the monthly-mean forecast in the extratropical regions than in the tropical region; furthermore, most of predictability...
resulting from the influence from the initial condition is lost within the lead time of 10 days. A much smaller change in AC values for different lead times over the tropics suggests that the prediction skill is likely related to other factors (e.g., slow variations related to SSTs).

In previous studies of the initial condition effect on the monthly-mean forecast (e.g., Phelps et al. 2004; Reichler and Roads 2005a,b), atmospheric circulation response was analyzed only for the winter season. The availability of forecasts and simulations for all calendar months from 1981 to 2006 allows us to analyze the seasonality in the lead-time dependence of the forecast skill.

In general, the meridional variation in AC skill for the four seasons [December–February (DJF), March–May (MAM), June–August (JJA), and September–November (SON)] with lead time is similar to that for the entire year (Fig. 4). The AC values at different forecast lead times for each season over northern extratropical (20°–60°N) and

![Figure 2](image-url)
tropical (20°S–20°N) regions are summarized in Table 1. The skill is highest in extratropical regions at the shortest lead times but decreases rapidly and becomes lower than in tropics as lead time increases. The AC values in the Northern Hemisphere are higher in winter and spring (DJF and MAM) than in summer and fall (JJA and SON). There are two possible reasons for this seasonal variation of forecast skill. One possibility is the impact associated with initial snow cover in winter season and soil moisture in spring season. The other possibility is the extratropical impacts of dynamical modes, such as the Madden–Julian oscillation (MJO), which are stronger in winter and spring seasons (Madden and Julian 1994), and El Niño–Southern Oscillation (ENSO), which induces stronger atmospheric response in the extratropics in the winter and spring seasons (Barnston and Smith 1996; Kumar and Hoerling 1998).

b. The asymptotic value of the forecast skill and the role of boundary SST

The results shown in section 3a suggest that the initial atmospheric and land conditions contribute to the predictability of monthly mean, especially for extratropical high-latitude regions. The latter is consistent with the influence of land initial conditions on the short-lead-time evolution of near-surface atmospheric fields; for example, surface temperatures (Huang et al. 1996; Wang and Kumar 1998; Mahanama et al. 2008).

The prediction skill in the extratropics drops rapidly within 10 days of lead time and decays more slowly afterward. In this subsection, we examine the source of the skill associated with the ocean surface conditions and address two questions: First, is the skill at longer lead times due to the initial condition, SST boundary condition, or both? Second, is the asymptotic value of the CFS prediction skill due to atmospheric response to slowly varying boundary SSTs? The analysis is based on a comparison between the retrospective forecasts from CFS, which incorporate the influence of initial atmospheric and land conditions and SST conditions, with AMIP simulations from the GFS forced with observed SST. Figure 5 shows spatial distribution of the correlation for monthly-mean temperature in the AMIP-type simulation.
from the GFS. The atmospheric response to the slowly varying boundary SSTs is largely confined to tropical regions and is consistent with the strong influence of SST variability on tropical land temperatures (Kumar et al. 2003). The major difference in forecast skill between the CFS forecast and GFS AMIP simulation is that the CFS forecast skill at 0-day lead is significantly higher than the AMIP simulation skill over most of the globe, with the largest differences in the high latitudes of the Northern Hemisphere. Similar differences with smaller amplitudes also exist between the 10-day lead forecast and the AMIP simulation. At longer leads, the skill of the GFS simulation is similar to that of the CFS forecast. Comparison of global-mean correlation shows that the AMIP skill is comparable to the CFS forecast skill at 30–40-day lead (Fig. 3). This suggests that, at a lead time of 30–40 days, the influence of initial conditions on the predictability of CFS essentially vanishes and the predictability of monthly mean in the coupled CFS reaches an asymptotic value primarily because of anomalous SSTs (as implied by the skill for the AMIP simulations).

As the lead time increases further, the CFS forecast skills become even lower than the AMIP skills, suggesting additional degradation of forecast skill resulting from imperfect forecast of SST in the CFS.

The AMIP simulation and the CFS forecast skill is further examined for zonal mean of the correlation to investigate latitudinal variations (Fig. 4 and Table 1). For longer lead times, the CFS forecast skill approaches that for the AMIP simulation. In the tropics (10°S–10°N), the skill of the AMIP simulation is close to that of the CFS forecast at 0-day lead, indicating that the skill in tropical forecast of monthly means is primarily due to SST. In higher latitudes, the CFS forecast skill drops to skill values of the AMIP simulation at various lead times from 20 to 40 days, except for south of 40°S, where the AMIP simulation is better than the CFS 10-day-lead forecast.

The skill differences between the CFS forecasts and GFS AMIP simulations for individual seasons (Table 1) are similar to annual mean, except that the CFS skill in the extratropics becomes comparable to GFS at lead

<table>
<thead>
<tr>
<th>Lat range</th>
<th>AMIP</th>
<th>0-day lead</th>
<th>10-day lead</th>
<th>20-day lead</th>
<th>30-day lead</th>
<th>40-day lead</th>
</tr>
</thead>
<tbody>
<tr>
<td>DJF</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20°–60°N</td>
<td>0.12</td>
<td>0.55</td>
<td>0.29</td>
<td>0.16</td>
<td>0.12</td>
<td>0.13</td>
</tr>
<tr>
<td>20°S–20°N</td>
<td>0.33</td>
<td>0.45</td>
<td>0.35</td>
<td>0.31</td>
<td>0.34</td>
<td>0.30</td>
</tr>
<tr>
<td>MAM</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20°–60°N</td>
<td>0.13</td>
<td>0.58</td>
<td>0.28</td>
<td>0.21</td>
<td>0.19</td>
<td>0.13</td>
</tr>
<tr>
<td>20°S–20°N</td>
<td>0.29</td>
<td>0.38</td>
<td>0.33</td>
<td>0.29</td>
<td>0.28</td>
<td>0.25</td>
</tr>
<tr>
<td>JIA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20°–60°N</td>
<td>0.15</td>
<td>0.42</td>
<td>0.19</td>
<td>0.11</td>
<td>0.04</td>
<td>0.05</td>
</tr>
<tr>
<td>20°S–20°N</td>
<td>0.31</td>
<td>0.32</td>
<td>0.24</td>
<td>0.25</td>
<td>0.24</td>
<td>0.24</td>
</tr>
<tr>
<td>SON</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20°–60°N</td>
<td>0.15</td>
<td>0.43</td>
<td>0.20</td>
<td>0.14</td>
<td>0.07</td>
<td>0.10</td>
</tr>
<tr>
<td>20°S–20°N</td>
<td>0.29</td>
<td>0.35</td>
<td>0.25</td>
<td>0.24</td>
<td>0.20</td>
<td>0.21</td>
</tr>
</tbody>
</table>

Fig. 5. The spatial distribution of the GFS AMIP simulation skills (AC) in monthly-mean temperature. The areas with AC < 0.2 are masked with light gray color, and the areas without observational data are not colored. The contour interval is 0.1.
times as short as 10 days for JJA and SON. It is also noticeable that, in the tropics, the CFS skill for MAM is better than that of the AMIP simulation for lead times up to 30 days.

4. Summary and discussion

The roles of initial atmospheric and land conditions and the lower boundary condition of SST on the prediction skill of monthly-mean temperature were analyzed using retrospective forecasts from the NCEP coupled CFS and AMIP-type simulation from the uncoupled atmospheric model GFS. Results show that the CFS monthly-mean prediction skill is highest at 0-day lead with higher values in the extratropics than in the tropics. As lead time increases, the CFS prediction skill changes little in the tropics but decreases rapidly in the extratropics, with the largest drop from 0-day lead to 10-day lead, resulting in lower skill in the extratropics than in the tropics.

The CFS monthly forecast was also compared with GFS AMIP simulation to understand the evolution of skill for longer lead times. Based on the similarity of skill scores for the AMIP simulations and that for the CFS forecasts, we conclude that for the tropical latitudes the predictability of monthly means is primarily due to the anomalous SST, and atmospheric and land initial conditions play a minor role. On the other hand, for the extratropical latitudes, specification of atmospheric and land initial conditions contributes positively to the forecast skill of monthly means. This influence, however, decays rapidly with time; for forecasts with lead times longer than 20–30 days, the skill of monthly-mean forecasts is largely consistent with the atmospheric response to anomalous SST (Kumar and Hoerling 1998; Anderson et al. 1999; Kumar et al. 2003). The experimental design of CFS forecasts, where both atmospheric and land conditions are initialized, does not allow separation of the influence of atmospheric versus land boundary conditions. However, the fact that the combined influence of initial conditions on the monthly forecasts decays with a lead time of 20–30 days does provide an upper limit as to the influence of only the land boundary conditions.

This study indicates that extratropical monthly forecast requires the best possible initialization of the land and atmosphere. Because of the rapid decay of forecast skill from 0-day lead to 10-day lead, the construction of monthly-mean forecasts should also rely on the latest possible initial conditions. For monthly prediction using a lagged ensemble, the inclusion of forecasts beyond 10-day lead in an attempt to increase the ensemble size and adequately sample the spread in monthly means may not lead to any significant improvement in prediction skill.

However, the present analysis is based on assigning equal weights to forecasts from initial conditions with different lead times, and more sophisticated approaches to generate lagged ensembles (Raftery et al. 2005) may allow the use of longer-lead initial conditions, possibly leading to some improvements in prediction skill.

Acknowledgments. Constructive comments by two anonymous reviewers and by Drs. Philip Pegion and Song Yang led to significant improvements in the final version of the manuscript.

REFERENCES


