

Using Observed Spatial Correlation Structures to Increase the Skill of Subseasonal Forecasts

RANDAL D. KOSTER

Global Modeling and Assimilation Office, NASA Goddard Space Flight Center, Greenbelt, Maryland

THOMAS L. BELL

Climate and Radiation Branch, Atmospheric Sciences Division, NASA Goddard Space Flight Center, Greenbelt, Maryland

ROLF H. REICHLER

Global Modeling and Assimilation Office, NASA Goddard Space Flight Center, Greenbelt, and Goddard Earth Sciences and Technology Center, University of Maryland, Baltimore County, Baltimore, Maryland

MAX J. SUAREZ AND SIEGFRIED D. SCHUBERT

Global Modeling and Assimilation Office, NASA Goddard Space Flight Center, Greenbelt, Maryland

(Manuscript received 9 May 2007, in final form 27 September 2007)

ABSTRACT

Model deficiencies limit a subseasonal or seasonal forecast system's ability to produce accurate predictions. In this paper, an approach for transforming the output of a forecast system into a revised forecast is presented; it is designed to correct for some of the deficiencies in the system (particularly those associated with the spatial correlation structures of the forecasted fields) and thereby increase forecast skill. The approach, based on the joint consideration of the correlation structures present in the observational record and the inherent potential predictability of the model, is tested on a preexisting subseasonal forecast experiment. It is shown to produce modest but significant increases in the accuracy of forecasted precipitation and near-surface air temperature at monthly time scales.

1. Introduction

The strategy employed for forecasting meteorological variables such as precipitation and air temperature in large part relies on the lead time of the forecast. For medium-range weather forecasts, extending out to about a week, forecast skill relies mostly on the correct initialization of atmospheric states. For seasonal-to-interannual (SI) forecasts, however, atmospheric initialization has almost no impact on skill. SI forecasts must instead rely on the slower-moving components of the coupled earth system, most notably the state of the ocean. If the responses of precipitation and air temperature to predictable ocean states are known, then

precipitation and air temperature are themselves predictable at SI leads.

At subseasonal time scales (up to about a month), soil moisture, another slower-moving component of the climate system, can also be an important source of predictability, particularly over the continents. Observational analyses (Vinnikov and Yeserkepova 1991; Entin et al. 2000) show that in most places soil moisture anomalies have time scales of persistence that exceed a month, and climate system models typically capture these time scales successfully (Seneviratne et al., 2006). Some statistical studies (e.g., Huang et al., 1996) show a connection between soil moisture and subsequent temperature, even at leads beyond a month. Some modeling studies suggest that, for midlatitude continental regions during summer, the contribution of land initialization (and atmospheric initialization, to the extent that it contributes to the evolution of soil moisture in the first week or so of the forecast) to subseasonal pre-

Corresponding author address: Randal Koster, Global Modeling and Assimilation Office, Code 610.1, NASA/GSFC, Greenbelt, MD 20771.

E-mail: randal.d.koster@nasa.gov

precipitation and air temperature prediction skill should strongly outweigh that of ocean initialization (Koster and Suarez 1995; Koster et al. 2000). Indeed, a number of modeling studies have examined the impact of soil moisture on precipitation (e.g., Delworth and Manabe 1988, 1989; Atlas et al. 1993; Hong and Kalnay 2000; Dirmeyer 2000, 2001; Douville et al. 2001; Koster et al. 2006; Guo et al. 2006; among others). These studies generally show that the impact is strong—simulated precipitation does tend to respond in a predictable way to variations in soil moisture.

True initialization studies, however, in which forecast simulations are initialized with realistic land moisture states and the resulting precipitation forecasts are compared to observations, are few and far between. Those that exist (e.g., Beljaars et al. 1996; Fennessy and Shukla 1999; Douville and Chauvin 2000, Viterbo and Betts 1999) find some cause for encouragement—suggestions that the initialization does improve the simulation of precipitation and/or temperature. Koster et al. (2004, hereinafter K04) provide perhaps the first precipitation forecast study that both utilizes realistic soil moisture initial conditions and examines enough independent forecasts (75 in all) to tease out a statistical characterization of soil moisture impacts on precipitation forecast skill. The results, some of which are shown in section 5 below, show a small but statistically significant contribution of land initialization to forecast skill in the center of the United States, the only large-scale region on the globe in which the model's inherent potential predictability is relatively large (see section 2 below) and precipitation measurements are adequately dense.

The skill levels quantified by K04 fall far below potential predictability levels. Possible reasons for the modest skill levels include the following: 1) the model has errors in its representation of physical processes; 2) the data used to initialize the model and validate its performance have errors; and 3) the size of the forecast ensemble (nine members) is insufficient to average out all of the background noise in the signal. One obvious way to increase the skill levels is to redo the forecast experiment after significantly improving the model and the quality of the input datasets. Such improvements are a continuing goal at modeling and data centers.

It may be possible, however, to apply statistical techniques to increase the skill levels now, without waiting for model and data improvements. We examine this possibility in the present paper. A simple statistical approach is applied that corrects for known biases in the forecast model's climatological statistics (particularly in the spatial correlation structures of the forecasted variables), leading to improvements in its levels of forecast

skill. The approach should be transferrable to any forecast system.

We begin in sections 2 and 3 with a discussion of potential predictability and the spatial correlation structures seen in models and observations. In section 4 we then describe a simple statistical approach that uses both of these pieces of information to improve skill. In section 5, we use the approach to transform the forecasts of K04, and we quantify the improvement in the skill of the forecasts. A mathematically formal description of the approach is provided in the appendix.

2. Potential predictability: An example

In this paper, we define “potential predictability” as the contribution of the land surface and/or sea surface initialization to the “signal” contained in an ensemble forecast. That is, it represents the degree to which the precipitation and air temperature anomalies produced by the different ensemble members of a forecast are similar because of the land and/or ocean initialization, even in the face of disparate atmospheric initialization (representing, in an extreme way, internal atmospheric noise, or “chaos”). The potential predictability is an inherent property of a given forecast model. Here we describe an example of its calculation, using results from an existing forecast experiment.

a. Forecast system

The forecast experiments performed by K04 utilized the atmosphere and land components of the seasonal prediction system of the former National Aeronautics and Space Administration (NASA) Seasonal-to-Interannual Prediction Project (NSIPP), which is now a part of the NASA Global Modeling and Assimilation Office (GMAO). The atmospheric general circulation model (AGCM) is a state-of-the-art, finite-difference model run at 2° latitude \times 2.5° longitude resolution. It uses the relaxed Arakawa–Schubert scheme (Moorthi and Suarez 1992) for convection, sophisticated codes for shortwave and longwave radiation (Chou and Suarez 1994), and fourth-order advection of vorticity and all scalars in the modeled dynamics. The Mosaic land surface model (LSM) of Koster and Suarez (1996) is used, which is a soil–vegetation–atmosphere transfer (SVAT) model that uses tiling to account for subgrid vegetation distributions. The behavior of the coupled land–atmosphere system relative to observations is well documented (Bacmeister et al. 2000; Koster et al. 2000). Although the model is far from perfect, it does successfully reproduce the broad features of observed precipitation statistics across the globe.

b. Forecast experiment

The K04 experiment consisted of two series of 1-month ensemble forecasts. A brief description of these two series is provided here; the reader is referred to K04 for a full description. Each series employed 75 forecast start dates—5 start dates (the first days of May, June, July, August, and September) for each year in the 1979–93 period. Each 1-month forecast utilized nine ensemble members. The forecast variables examined were precipitation and near-surface air temperature.

Series 1 [“Atmospheric Model Intercomparison Project (AMIP) mode”] served as the control. Sea surface temperatures (SSTs) were prescribed throughout the forecast period to observed, time-varying states, taken from Rayner et al. (1996) for 1979–82 and Reynolds and Smith (1994) for 1983–93. Land states and atmospheric states were *not* initialized to realistic values; instead, the initial states were taken from values generated for the start date in question by parallel AMIP-style (Gates 1992) simulations, that is, long-term atmospheric simulations with SSTs prescribed to observations. The land and atmosphere initial conditions were therefore not similar between the nine ensemble members, though any given set was consistent with the given year’s SST distribution. [Note that the AMIP mode simulations are described here in terms of initialized forecasts to contrast them with the Land Data Assimilation System (LDAS) mode forecasts discussed next. In actuality, the AMIP mode forecasts are simply subsets of existing long-term AMIP integrations.]

Series 2 (“LDAS mode”) focused on land initialization. To generate realistic land states for the initialization of the forecasts, the land model in the forecast system was driven offline with realistic meteorological forcing for the 15-yr period, as provided by Berg et al. (2003) and processed through the Global Land Data Assimilation System (GLDAS; Rodell et al. 2004). Land states were scaled for consistency with the forecast system’s climatology (see K04). All other aspects of the LDAS mode simulations were identical to those of the AMIP mode simulations. (Note that the resulting inconsistency between the land and atmospheric initial conditions may have resulted in a “shock” at the beginning of the LDAS forecasts, a problem that could only hinder forecast skill levels.) In short, the LDAS mode forecasts differed from the AMIP forecasts only in their use of a realistic initialization of land surface states.

Neither the LDAS nor AMIP series can rigorously be called a true forecast experiment, because for both series observed SSTs were prescribed throughout the forecast period. The strategy employed by K04 was to

isolate the impacts of land initialization through a comparison of the LDAS and AMIP experiments; if forecast skill was higher for the LDAS than AMIP series, then the skill increase could be attributed to the land initialization, because, again, that was the only difference between the two series. Even so, given that SSTs typically persist over time scales much greater than a month, the AMIP series can give a first-order indication of how predicted SSTs affect 1-month forecasts, and thus the true potential predictability associated with ocean initialization.

c. Quantification of potential predictability

A measure of a model’s potential predictability can be derived from an extensive series of independent ensemble forecasts. Various approaches for quantifying predictability can be found in the literature (e.g., Zwiers 1996). The approach used here is fairly simple. One ensemble member in each forecast is assumed as “nature” and the average of the remaining ensemble members make up the “forecast.” The square of the correlation coefficient (r^2) between the multiple independent nature–forecast pairs is then determined. The process is repeated several times, each time using a different ensemble member as nature. The resulting average r^2 is the desired measure of potential predictability; given the construct of the K04 experiment, it measures the degree to which the model’s weather is controlled by the prescribed SSTs and the land surface initialization (for the LDAS forecasts) rather than by noise associated with variations in the atmosphere’s initial conditions.

Such an analysis was used by K04 to determine the potential predictability for boreal warm season months (May through September) using the 15-yr forecast experiment described in section 2b. Results are shown in Fig. 1. We focus on the North American region in this paper because, as explained by K04, it is the only large-scale region that is both rich in data and home (in this model) to a significant area of potential precipitation predictability.

Results are shown for two meteorological variables—precipitation and near-surface air temperature—and for two combinations of boundary conditions. The panels on the left were derived from the AMIP forecasts; they show the potential predictability that stems from knowledge of the SST boundary. Those on the right show results from the LDAS forecasts, that is, the predictability stemming from knowledge of both SSTs and land surface initial conditions. The land initialization leads to a significant increase in potential predictability, particularly in the center of the continent, where the r^2 increases to more than 0.2 for precipitation and more

Potential Predictability (r^2)

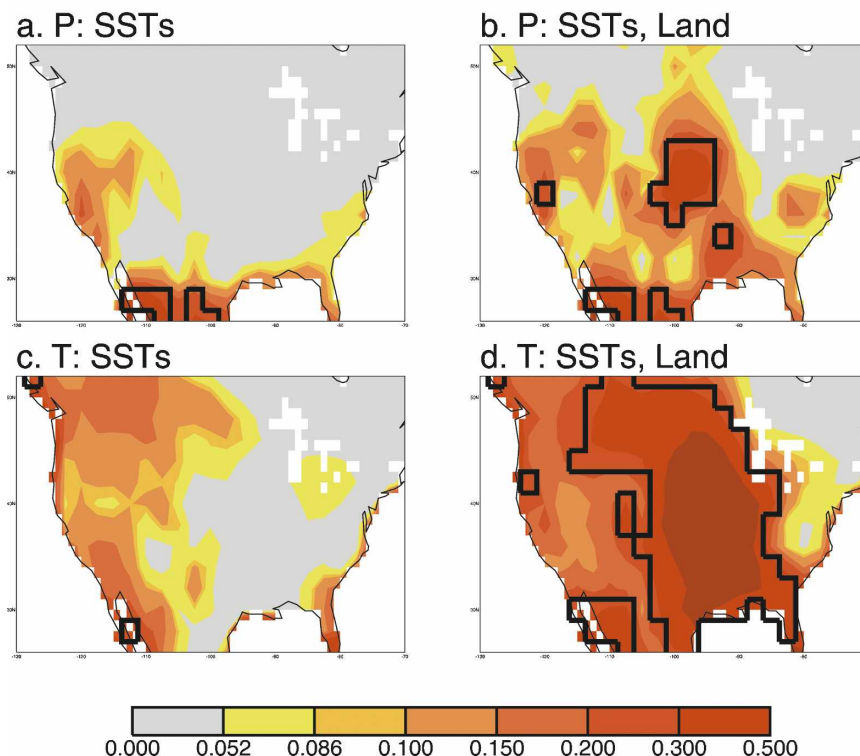


FIG. 1. (a) Potential predictability (r^2 from idealized analysis) of 1-month precipitation totals when SSTs are prescribed. (b) Same as (a), but when SSTs are prescribed and all ensemble members use the same land initialization. (c) Same as (a), but for air temperature. (d) Same as (b), but for air temperature. The dark lines indicate regions for which the potential predictability exceeds 0.25. All values above 0.052 are significantly different from zero at the 95% level; those above 0.086 are significant at the 99% level.

than 0.3 for air temperature. In this forecast system, knowledge of the land state at the start of the forecast can potentially lead to important increases in forecast skill, beyond that which can be obtained from knowledge of SSTs alone. Note, however, that even with both SST prescription and land surface initialization, the potential predictability of precipitation, and even that of temperature, is either negligible or modest in many parts of the continent.

3. Spatial correlation structures in models and observations

The second piece of information used in the proposed approach for transforming forecasts is the spatial correlation structure inherent in observed meteorological fields. In this section, we illustrate the errors that can occur in the simulation of these spatial structures. For the model-based spatial correlations examined here, we analyzed 612 values of monthly (June) pre-

cipitation and near-surface air temperature collected from nine parallel AMIP-style simulations with the model described in section 2a. The contributing simulations covered (roughly) the latter two-thirds of the twentieth century.

Although the observational spatial correlations were based on a much smaller dataset, the sample sizes were still large enough to capture the first-order structure of the fields. For precipitation, we used the multidecadal (1948–97) daily precipitation reanalysis of Higgins et al. (2000). The input data for the reanalysis was a Unified Rain gauge Database (URD) for the United States, which consists of daily rain gauge reports from multiple sources in the United States, including the River Forecast Centers (about 7000 sites per day), the National Climatic Data Center (NCDC) daily cooperative network (about 6000 sites per day), and the NCDC Hourly Precipitation Network (about 2500 sites per day; aggregated into daily accumulations). Several quality control tests were applied to the daily gauge data (Higgins et al.

2000). The daily precipitation data were gridded at a horizontal resolution of $\frac{1}{4}^\circ$ latitude \times $\frac{1}{4}^\circ$ longitude over the domain 20° – 60° N, 140° – 60° W using a Cressman (1959) scheme with modifications (Glahn et al. 1985; Charba et al. 1992). For the present analysis, to maintain consistency with the AGCM data, the $\frac{1}{4}^\circ \times \frac{1}{4}^\circ$ daily dataset was aggregated in space and time into a $2^\circ \times 2\frac{1}{2}^\circ$ monthly dataset.

For near-surface air temperature, we use a gridded version (M. Fennessy 2005, personal communication) of the Climate Anomaly Monitoring System (CAMS) temperature dataset maintained by the Climate Prediction Center of the National Centers for Environmental Prediction. The dataset, constructed from station data (Ropelewski et al. 1985), covers 58 yr (1946–2003) with full coverage over the continental United States. Because of some problems with the data during the first 3 yr, we use only the 1949–2003 period for the statistics.

All data were converted to standard normal deviates (i.e., all precipitation rates and air temperatures were converted to anomalies relative to their long-term means and were then normalized by their local standard deviation) prior to performing any statistical analysis. The standardization was performed independently for simulated and observed data, and it used (here and throughout this paper) monthly means and standard deviations rather than annual values, so as to remove the imprint of the seasonal cycle on the anomalies.

The top three panels of Fig. 2 show, for three representative points in North America (both inside and outside the region of high potential predictability, from Fig. 1), the correlations between total June precipitation at the indicated point and total June precipitation elsewhere on the continent, as simulated by the model. The second row of panels shows the corresponding correlations as derived from the observations. Notice that in all cases, the observations show a larger correlation structure for monthly precipitation anomalies. Stated another way, the spatial extent of a given monthly anomaly tends to be larger in the observations. For example, for the observations, wet or dry conditions in the central United States tend to coexist with similar conditions in the west (center panel of Fig. 2b), whereas in the model, conditions in the central United States are uncorrelated with conditions elsewhere (center panel of Fig. 2a). {Sampling error must, of course, be considered here, particularly for the observations. With the observational sample size of 50, an underlying zero correlation may give a false correlation of 0.2 or higher with a probability of 0.08, and may give a false correlation of 0.4 or higher with a probability of 0.002, with the same

probabilities for the corresponding negative correlations. With these probabilities, the orange [0.4] contours in particular indicate that the observations and the model have fundamentally different underlying correlation patterns.}

The bottom two rows show the corresponding correlations for average June near-surface air temperature. The observed and modeled correlations look more similar for air temperature than they do for precipitation, though the simulated spatial length scales for air temperature appear slightly too large in the western and eastern United States and somewhat too small in the center of the continent. The observations also show a distinct negative correlation between temperature in the eastern United States and temperature in the northwest, a feature that is absent in the model.

4. Transformation of GCM forecasts

We now describe a simple statistical approach that uses observations-based spatial correlations along with an estimate of the model's potential predictability to improve forecast skill for precipitation and near-surface air temperature.

a. Overview of approach

The shaded contours in Fig. 3 show the correlation patterns for rainfall in the central United States. (The map is copied from the middle panel of Fig. 2b.) Overlaid on the map are heavy lines showing where the model demonstrates a high degree of potential predictability for rainfall (from Fig. 1b). Notice that the model's forecast skill is potentially high at the grid cell holding the asterisk, but it is much lower at the grid cell marked by the white circle, which lies outside the heavy lines. The figure suggests the potential for improving a forecast: if precipitation is known at the asterisk, then according to the observed correlations, it is known (to some degree) at the white circle. By using the observed spatial correlation structure, we may be able to "translate" the forecast at the asterisk spatially to a forecast at the white circle, regardless of the model's inherent skill level at the latter location. We may be able to compensate, to a degree, for errors in the model's simulated spatial correlation structure.

Described mathematically (and more generally, i.e., not constrained to only a single cell contributing forecast information), we can produce a transformed forecast vector $\tilde{\mathbf{x}}$ (of length N) from a set of "predictor forecasts" \mathbf{x} (also of length N):

$$\tilde{\mathbf{x}} = \mathbf{A}\mathbf{x}, \quad (1)$$

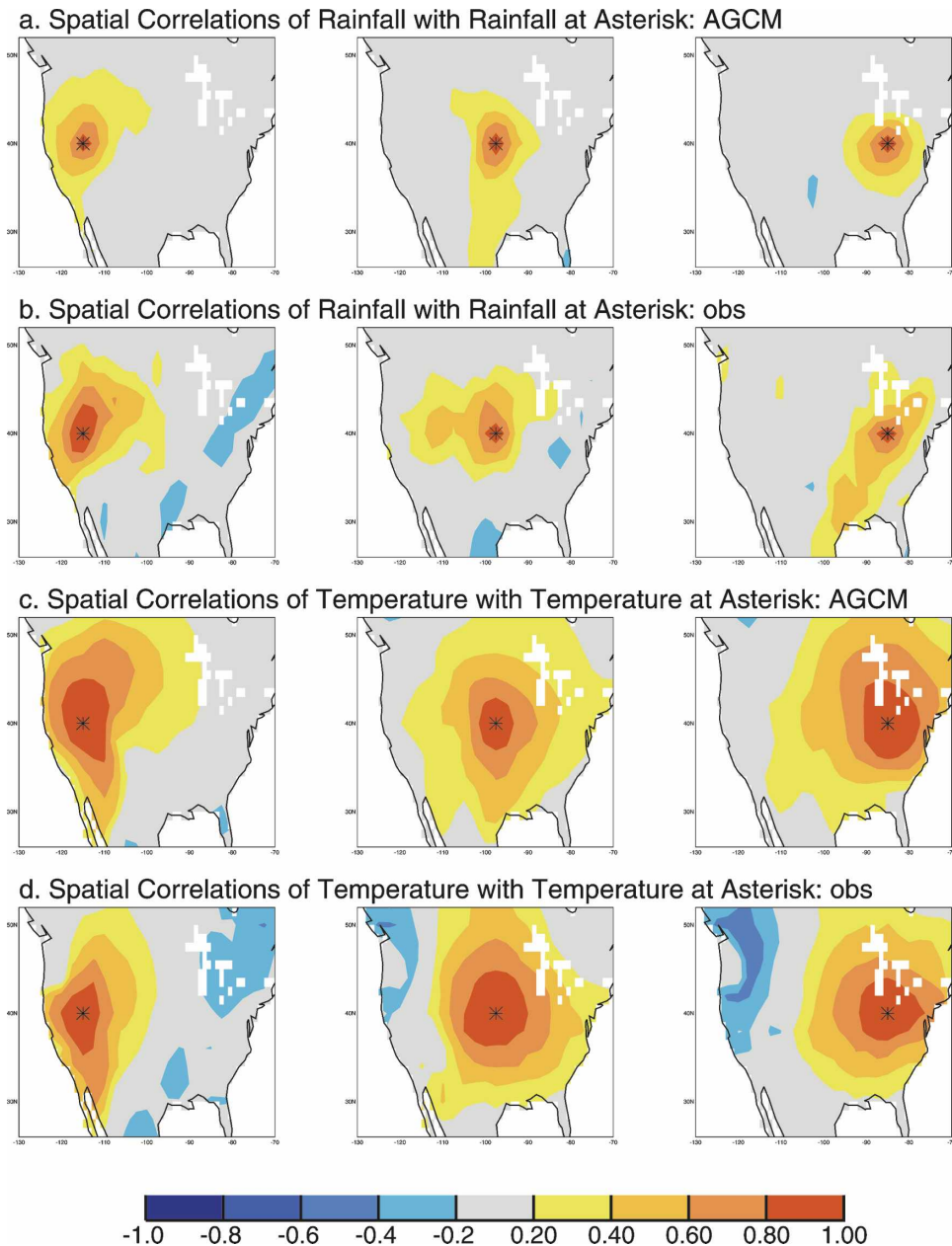


FIG. 2. (a) Correlations between monthly June rainfall in three representative points (marked with asterisks) in North America and that elsewhere on the continent, as derived from 600+ yr of AGCM simulation data. (b) As in (a), but based on 50 yr of observational data. (c) Correlations between June surface air temperature in three representative points (marked with asterisks) in North America and that elsewhere on the continent, as derived from 600+ yr of AGCM simulation data. (d) As in (c), but based on 55 yr of observational data.

where \mathbf{A} is an $N \times N$ transformation matrix and \mathbf{x} is a vector holding the original set of model forecasts in the region considered. Statistically optimal approaches for determining \mathbf{A} include the reprocessing of data generated in a forecast experiment (such as that of K04); taking \mathbf{x}_{obs} to be the vector of *observed* precipitation rates, we could compute

$$\mathbf{A} = \langle \mathbf{x}_{\text{obs}} \mathbf{x}^T \rangle \langle \mathbf{x} \mathbf{x}^T \rangle^{-1}, \quad (2)$$

where the “ $\langle \rangle$ ” symbols denote time averages and T denotes the vector or matrix transpose. In essence, $\langle \mathbf{x}_{\text{obs}} \mathbf{x}^T \rangle$ is the cross-correlation matrix between model predictions and observations, and $\langle \mathbf{x} \mathbf{x}^T \rangle$ is the spatial correlation matrix of the model forecasts. Unfortunately,

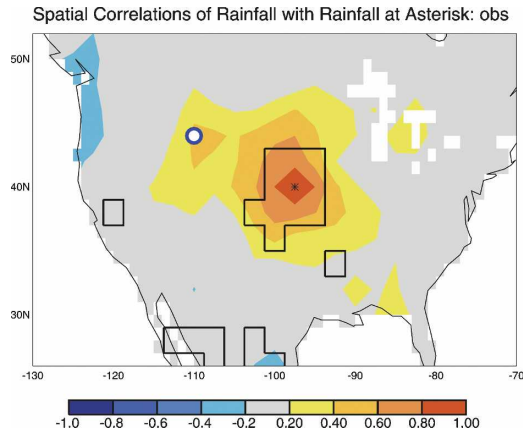


FIG. 3. Correlations between monthly June rainfall in the central United States (marked with an asterisk) and that elsewhere on the continent, from observations. Overlaid on the plot are black lines showing the area of high potential predictability for precipitation ($r^2 > 0.25$) in the model.

while in theory (2) is an optimal fit (in a least squares sense) to the training observations, in practice it comes with a cost: it requires an extensive “training period” of forecasts, that is, a large enough sample space for computing accurate values for the elements of $\langle \mathbf{x}_{\text{obs}} \mathbf{x}^T \rangle$. The results of an extensive analysis, not presented here, reveals that the 15-yr forecast experiment of K04 (e.g.) is inadequate for an effective determination of $\langle \mathbf{x}_{\text{obs}} \mathbf{x}^T \rangle$, apparently because the signals we seek are too subtle.

We therefore design an alternative approach for generating \mathbf{A} , an approach that *does not require* an extensive forecast training period. The alternative approach instead relies on auxiliary information characterizing model behavior and observations. It recognizes that for a forecast at grid cell m to improve the forecast of a quantity at a remote cell n , the following two conditions must be met: (a) the model must have some predictability for the quantity at cell m , and (b) in the observations, the quantity at cell m must be correlated with that at cell n . The approach thus relies solely on two important and (relatively) robustly determined characteristics of the model and nature—the potential predictability of the model (as plotted in Fig. 1) and the spatial correlation structure of the observations (as described in section 3). The observed spatial correlation structure is accurately estimated from multidecadal datasets (50+ yr), as in Fig. 2. The potential predictability, an internal model characteristic, is far more robustly determined from the 15-yr forecast experiment than the term $\langle \mathbf{x}_{\text{obs}} \mathbf{x}^T \rangle$ in (2), particularly when the predictability is low. Note that while in this paper we use the forecast experiment to establish the potential predictability, it could just as easily be derived from idealized model

prediction experiments that do not use an observations-based initialization. These idealized experiments would instead initialize a series of forecasts with states produced by the modeling system in free-running climate mode, a much cheaper and easier way of generating initial states than relying on an observations-based analysis system. In other words, with the approach described here, we could in principle generate the matrix \mathbf{A} prior to performing a single forecast experiment utilizing observations.

In the simpler context of Fig. 3, we know that a prediction at the asterisk may have some correlation to the observed truth at the white circle, and that with a long enough training period, we could get at that correlation through (2). Here, we are instead attempting to estimate a serviceable version of \mathbf{A} through two more easily (and more cheaply) obtained relationships: that describing the model’s predictability at the asterisk (a surrogate for the relationship between the model prediction and the observational truth at the asterisk) and that between the observational truth at the asterisk and the observational truth at the white circle. Given the high cost of training a forecast system, and the potential difficulty in generating the multidecadal observational data needed to perform the training in the first place, the alternative approach, if it works, has a distinct advantage over the direct application of (2).

b. Details of approach

For a given region of interest (such as the continental United States), N is set to the total number of grid cells in the region. To generate the $N \times N$ matrix \mathbf{A} , we perform the following algorithm at each target (predictand) grid cell n independently [If cell n corresponds to the n th element of vector $\tilde{\mathbf{x}}$ in (1), then the algorithm computes the n th row of \mathbf{A} .]:

- The k “predictor grid cells” for target cell n (i.e., the k elements that have nonzero values in the n th row of \mathbf{A}) are identified.
- Multiple regression is performed on the observational dataset to determine how observed precipitation data (*not* forecasts) in the k predictor cells, suitably degraded with noise based on the forecast model’s internal predictability, can best be used to forecast the observed precipitation in the target cell n . The regression coefficients produced in this exercise provide a potential n th row of the matrix \mathbf{A} .
- Steps a and b are repeated for different trial values of k . Each k value produces a different potential n th row of \mathbf{A} . The k value that best allows the observations to reproduce themselves at cell n determines which of the potential rows is chosen.

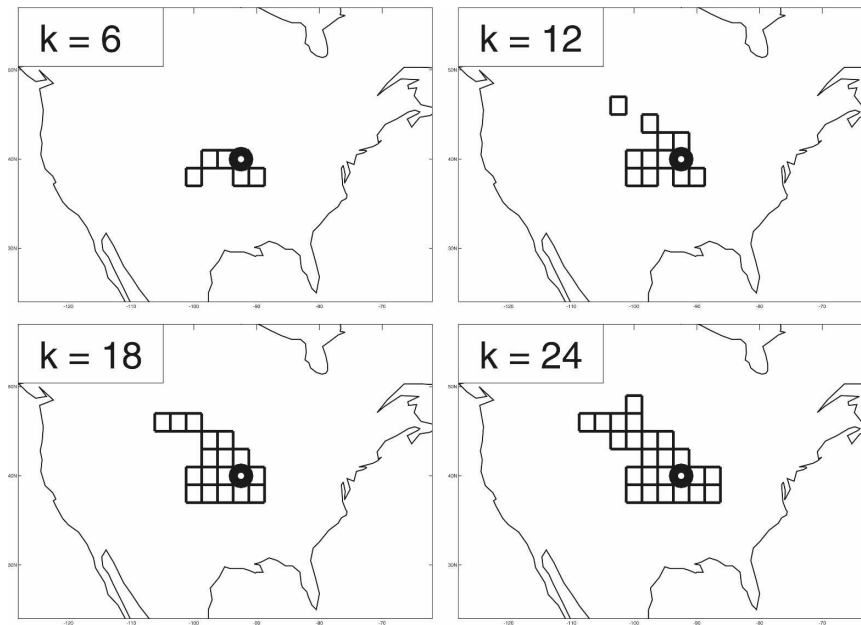


FIG. 4. Predictor grid cells for the grid cell indicated by the heavy black circle, for four different values of k .

These steps are now described in more detail. Discussion focuses on precipitation, but the same approach can be applied to other meteorological fields, such as near-surface air temperature.

1) IDENTIFICATION OF PREDICTOR CELLS

The first step of the approach is simple—for each target grid cell n , we identify the k grid cells ($m = 1, 2, \dots, k$) for which the product of $r_{\text{pot}}^2(P_m)$ and $\text{Corr}^2(P_m, P_n)$ is largest. Here, $r_{\text{pot}}^2(P_m)$ is the potential predictability of the model for precipitation at grid cell m (as in Fig. 1, but for the specific month considered), and $\text{Corr}^2(P_m, P_n)$ is the square of the correlation, from the observations, between precipitation at cells m and n (e.g., from Fig. 2 for June) for the month in question. This selection criterion recognizes the fact that a cell m will be a useful contributor to a prediction at cell n only if the following two conditions are satisfied: (i) the model shows some predictability for precipitation at m (as represented by the first factor), and (ii) the observations show that precipitation is correlated at the two locations (as represented by the second factor). To illustrate, Fig. 4 shows, for four different values of k , the identified predictor cells for precipitation forecasts at a specific grid cell in the central United States. In each case, the target cell itself is one of the predictor cells; the subsequent regression analysis (see next section) is made nontrivial by the addition of noise to the data at this cell. To some extent, the robustness of the idealized

predictability and observed correlation estimates implies some robustness for the selection of contributor cells.

2) MULTIPLE REGRESSION ANALYSIS

The ability of a set of k predictor cells ($m = 1, 2, \dots, k$) to contribute to a forecast at target cell n is established through multiple regression analysis on the observational data (not the forecasts). To explain how this is done, we first describe a simple but inadequate approach. A given year t of the T years of observational record is chosen as the “target year.” Multiple regression is then performed on the $T - 1$ remaining years of observations (i.e., not using the data for year t); the $T - 1$ observed precipitation values for the target grid cell n are regressed against the $T - 1$ sets of k observed values at the predictor cells to produce a set of k regression coefficients. Using these coefficients in a predictor equation, we can then “predict” the precipitation at n in year t from the observed precipitation values at cells m ($m = 1, 2, \dots, k$) in year t . We can then repeat the process T times, taking each year in sequence as the target year. This produces T “predictions” of the precipitation at cell n , which can be directly compared with the T actual (observed) precipitations at n . A linear regression of the predictions against the actual values gives a measure of predictive skill using this approach. The k coefficients from the multiple regression would be the nonzero elements of the matrix \mathbf{A} for the given tested year.

The flaw in this strategy is that in a forecast environment, the precipitation values at the k predictor cells in year t will themselves be inaccurate, largely because forecasted precipitation rates are subject to chaotic noise. Indeed, this issue underlies the potential predictability map shown in Fig. 1. By design, the potential predictability (the model’s predictive skill under the assumption of perfect model physics and perfect data) is less than one solely due to chaotic dynamics in the modeling system. Consider two predictor cells that show the same observed correlation with target cell n . If the potential predictability in the model at the first predictor cell is smaller than that at the second, we want the first cell to contribute correspondingly less to the transformed forecast at n . The approach outlined above does not allow this.

Thus, rather than regressing the $T - 1$ precipitation values for the target grid cell n against the $T - 1$ sets of k values at the predictor cells, we instead regress them against “degraded” versions of the k values. The degradation is keyed to the potential predictability in the model—the lower the $r_{\text{pot}}^2(P_m)$ value at a predictor cell is, the more the observational values there are degraded, and thus the less likely they are to contribute to the predictor equation. The degradation is performed as follows:

$$P_{m,\text{degraded}}(t) = r_{\text{pot}}(P_m)P_m(t) + [1 - r_{\text{pot}}^2(P_m)]^{1/2}\zeta(t), \tag{3}$$

where $\zeta(t)$ is a random variable with zero mean and unit variance at the given predictor cell. Because the meteorological variables we consider here are already converted to standard normal deviates, the coefficients utilized in (3) ensure that the degraded time series also has zero mean and unit variance.

The multiple regression on the degraded dataset provides a set of k regression coefficients that constitute the nonzero elements of the n th row of matrix \mathbf{A} for the given tested year. A technical note: to extract the true inherent signal in the data in the face of the noise added through (3), the multiple regression is actually performed on a greatly extended set of time series. First, the original $T - 1$ precipitation values at the target cell n are replicated 50 times, and the 50 time series are concatenated into a single time series of $50(T - 1)$ elements. Then, the k time series (one for each predictor cell) of $T - 1$ degraded values are constructed 50 separate times, using different sets of random numbers each time; these time series are concatenated into k time series of $50(T - 1)$ values. The time series of $50(T - 1)$ values at the target cell are regressed on the k time series of $50(T - 1)$ values at the predictor cells.

3) OPTIMIZING THE NUMBER OF CONTRIBUTOR CELLS

We still must determine how many contributor cells should contribute to a transformed forecast. We allow k to vary as a function of month and location. In essence, then, we construct five different \mathbf{A} matrices (one for each month from May to August), and the rows of a given matrix have different numbers of nonzero elements, corresponding to different values of k for different locations.

To choose k for a given location (target cell n), we proceed as follows. First, for each trial k , we choose contributor cells (step a) and perform the aforementioned multiple regression (step b). We then apply the resulting regression coefficients to the observations themselves (rather than to model forecasts) after degrading the observations with (3). In essence, for a set of regression coefficients derived from observational data spanning every year but year t , we determine the ability of these coefficients to reproduce the observational data value at cell n in year t . The k that produces the best estimates of the observations there (as measured with the r^2 skill score, computed over the length of the observational record) is assumed to be the “optimal” k for grid cell n . The n th row of the \mathbf{A} matrix will have that number of nonzero elements.

At first glance, it may seem surprising that the optimal k can be less than the maximum k attempted. Recall, however, that we are performing the regression on the years *outside* the year being tested. A k that is too large will pin the regression coefficients to noise in the $T - 1$ year that is irrelevant to what is happening in year t , hindering the success of the regression in year t . In practice, we have found that examining k values from 1 to 30 is sufficient; the selected value of k is generally less than 30, and the computationally demanding exercise of examining k values higher than 30 does not add or detract from the results.

c. Additional remarks

The above approach may appear ad hoc but is nevertheless built on the following two reasonable assumptions: (i) a model forecast at a location with strong potential predictability (in the model) can contribute to a forecast at a remote location with weak potential predictability if the observations show a significant correlation between the forecasted fields at the two locations, and (ii) the contribution of the forecast to the remote location can be determined through statistical analysis of the observational record, suitably degraded to reflect the impact of noise on a forecast. The first assumption, in particular, recognizes the presence of

strong model errors (relative to observations) in the simulation of the statistical structures of meteorological fields (Fig. 2).

In essence, our approach assumes that some predictability exists in nature both at the circle and at the asterisk in Fig. 3, but that the forecast system, because of various biases (e.g., excessive atmospheric noise), misses the predictability at the circle. The approach effectively uses the observed correlations to correct this assumed deficiency. Of course, the model may also be wrong about the predictability at the asterisk, and, if so, using the predictability there to transform a signal could cause problems. Still, the raw, untransformed model forecasts are every bit as limited by biases in model predictability; predictability biases help define the limitations of a forecast model, and they are a problem whether the forecast is transformed or not. Our approach effectively relies on the implicit assumption that the predictability in the model, where it does exist, has some connection to reality. The validity of this assumption will be demonstrated if the approach does indeed increase forecast skill (see section 5).

For the interested reader, a mathematically formal discussion of the approach, with the underlying statistical assumptions clearly outlined, is provided in the appendix. The appendix shows that the only truly ad hoc feature of the approach is the use of the product of $r_{\text{pot}}^2(P_m)$ and $\text{Corr}^2(P_m, P_n)$ to determine the contributor cells, and, again, this feature of the approach is built on reasonable assumptions regarding which cells should indeed be able to contribute to the revised forecast.

The potential predictability of a model [$r_{\text{pot}}^2(P_m)$] can, in principle, be determined with idealized forecast experiments that do not rely on real-world initialization and verification data, using instead, for example, snapshots of model conditions from a free-running model experiment as initial conditions. As a result, the construction of the transformation matrix \mathbf{A} does not rely on the traditional “training” of a system through joint analysis of real forecasts with verification data. The multiple regressions underlying \mathbf{A} require only the long-term observational record and the $r_{\text{pot}}^2(P_m)$ estimates. In other words, the transformation matrix \mathbf{A} can be constructed prior to performing any real forecast with the system, a distinct advantage to the approach, given the difficulty of producing long records of realistic initialization and verification data for the training exercise.

We note the possibility that alternative methods (e.g., based on canonical correlation analysis) could improve the construction of the transformation matrix \mathbf{A} even further using information on idealized predictability and observed correlations. Such developments are

left for future study. In any case, the effectiveness of the approach as presented here will be established in the next section through improvements in the model forecasts. We consider our tests to be fair because the \mathbf{A} matrices used to transform a forecast in a given year are derived solely from information outside that year.

5. Application to a forecast experiment

Figures 5a,b show the original (untransformed) skill of the forecast model in predicting monthly precipitation under AMIP and LDAS modes, as documented by K04. Skill in this analysis is shown here as the correlation coefficient between the model forecast and the observations. (More advanced measures of skill, such as the Brier skill score, were avoided given the limited sample size in our study. Future work will address alternative skill assessments, including those that look at the variance and structure of the forecasted variables, rather than at just their means.) Again, the forecasts of K04 did not make use of realistic atmospheric initialization—all skill was derived from either the prescribed SST boundary condition (under AMIP mode) or the combination of the prescribed SSTs and the initialization of the land states (under LDAS mode). The LDAS skill levels are significantly larger than the AMIP skill levels, particularly where the model has strong potential predictability (Fig. 1). (Differences of 0.27 and 0.38 between the LDAS and AMIP skill levels are significant at the 95% and 99% confidence levels, respectively.)

We emphasize again that our tests of the forecast transformation approach involve careful cross validation. In generating a transformed forecast for a given month of year t , we use only information outside that year to standardize the data, compute the observations-based correlations, compute the model’s idealized predictability, establish the contributor cells, and perform the multiple regressions that establish the \mathbf{A} transformation matrix. In effect, for the tests below, the \mathbf{A} matrix differs for each tested month of each tested year.

Figures 5c,d show the skill levels for the transformed precipitation forecasts, that is, the forecasts obtained through the use of (1), with \mathbf{A} computed as described in section 4. While the skill levels remain modest, both modes of the forecast, especially the LDAS mode, show significant skill increases. The improvement is highlighted in the histogram plot in Fig. 5e. In LDAS mode, the average precipitation forecast skill across the continental United States has increased by over 60%.

Figure 6 shows the corresponding results for air temperature. Again, for both AMIP and LDAS modes, using (1) to transform the forecasts improves the skill of the forecasts almost everywhere. As seen in the histo-

Forecast Skill: Precipitation

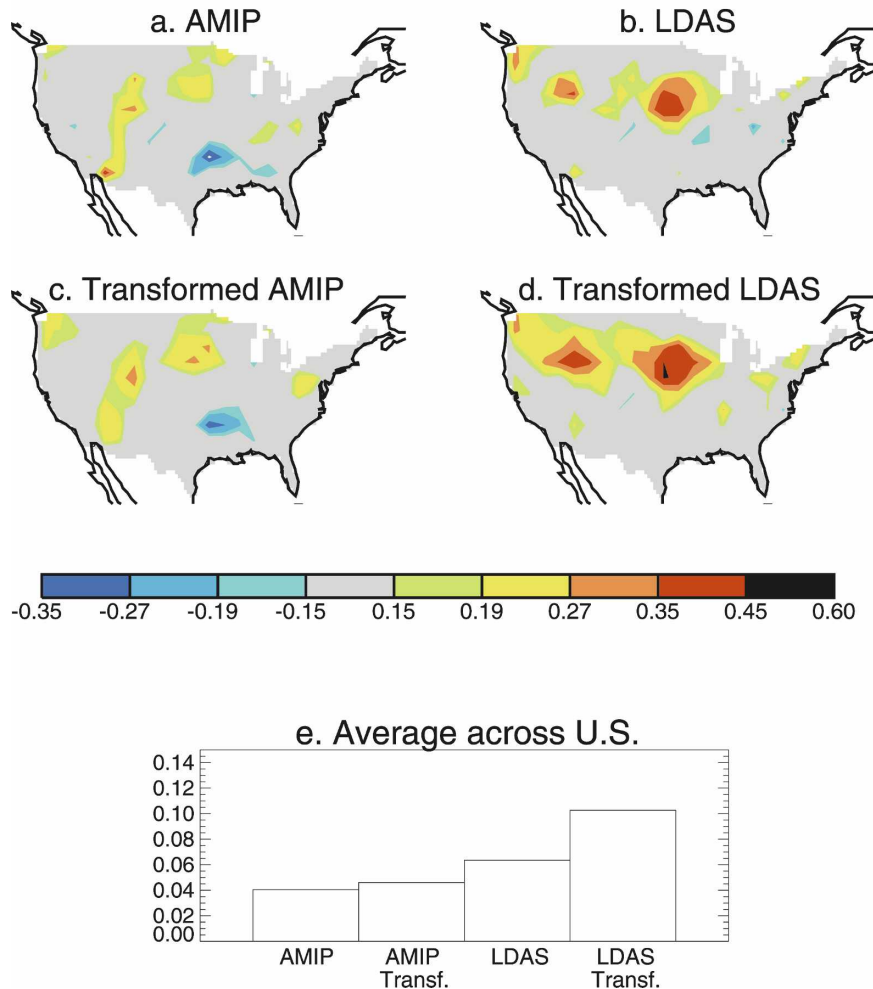


FIG. 5. (a) Forecast skill (measured as the correlation coefficient between model forecasts and observations) for precipitation in AMIP mode, that is, given prescribed SSTs. (b) As in (a), but in the LDAS mode, that is, given prescribed SSTs and a realistic initialization of land surface states. (c) Forecast skill in AMIP mode when the forecasts are transformed with the matrix **A**. (d) As in (c), but for forecast skill in LDAS mode. (e) Averages of the skill levels over the continental United States. For (a)–(d), values of 0.15, 0.19, 0.27, and 0.35 are statistically different from zero at the 90%, 95%, 99%, and 99.9% levels, respectively.

gram, skill in the AMIP mode has increased by about 35%, while that in the LDAS mode (which started off being much higher than that in AMIP mode) has increased by about 15%.

Note that the standardized ensemble mean forecasts and the standardized transformed forecasts of a quantity X (precipitation or air temperature) can be rebuilt into absolute predictions of X using the observed mean value (\bar{X}_{obs}) and observed standard deviation ($\sigma_{X,\text{obs}}$):

$$X_{\text{fcst,abs}} = X_{\text{fcst,standardized}}\sigma_{X,\text{obs}} + \bar{X}_{\text{obs}}. \quad (4)$$

When this is done, the resulting root-mean-square errors (RMSEs, not shown) relative to observations are, as expected from the correlation results, generally smaller for the transformed forecasts.

The plots in Figs. 5 and 6 are a fair test of the effectiveness of the forecast transformation procedure itself, because the transformation matrices used for a given year were not calibrated with observational data from that year. As mentioned before, though, the skill levels shown are not for “rigorously proper” forecasts, because the SST fields used for both the AMIP and

Forecast Skill: Temperature

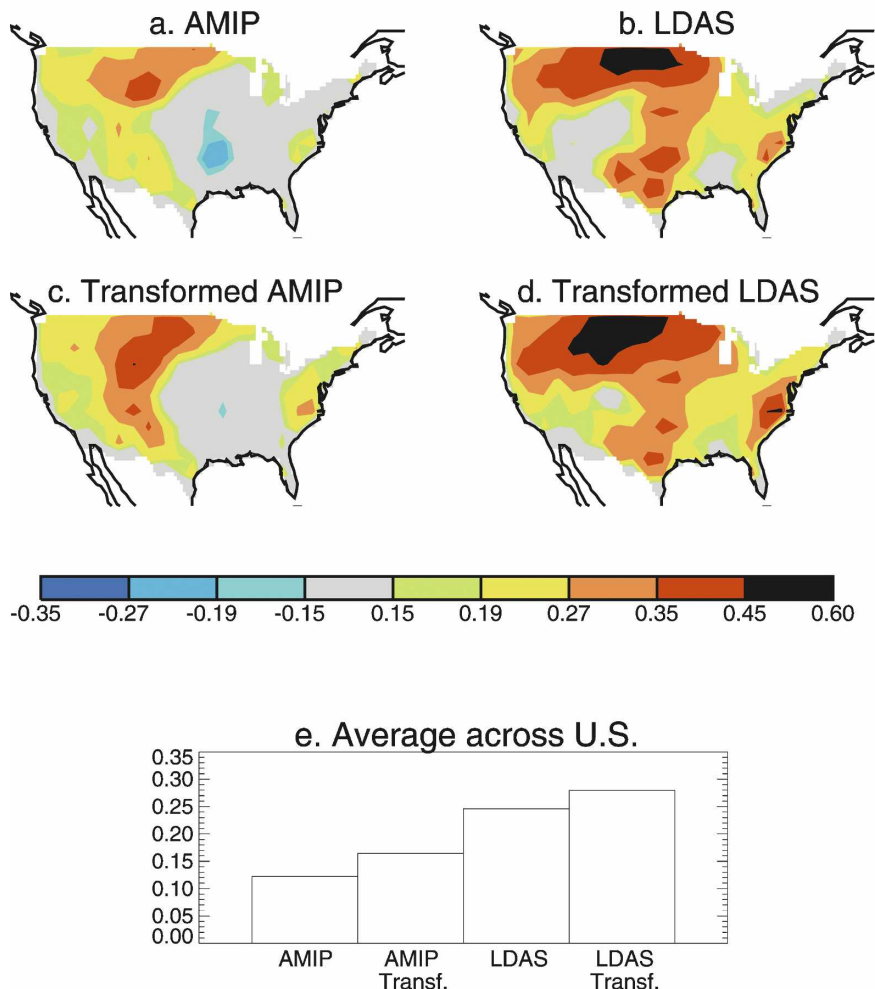


FIG. 6. As in Fig. 5, but for near-surface air temperature.

LDAS modes were prescribed to observations rather than predicted. Again, because these are 1-month forecasts and SSTs have much longer autocorrelation time scales, corresponding plots obtained with a system using a full ocean model or persisted SST anomalies would presumably be very similar. In any case, the strategy behind the K04 study was to isolate the increase in forecast skill associated with land moisture initialization by subtracting the skill obtained through the AMIP mode from that obtained through the LDAS mode. In this way the impact of the SSTs used in the forecasts, in particular, whether they are prescribed or predicted, becomes less important.

A similar strategy can be employed here. Figure 7 shows the difference between the skill obtained under the AMIP mode from that obtained under the LDAS

mode, for both the original forecasts and for the forecasts transformed with (1). Here, the skill for each mode is represented as the *square* of the correlation coefficient prior to the subtraction; in this way we quantify the degree to which the land initialization “explains” the observed variance. Results for the original forecasts are essentially those shown in K04, though with contour levels reflecting a slightly improved estimation of significance levels.

For precipitation, the “benefit” of land moisture initialization to a forecast, as measured by the plotted difference in forecast skill, is larger for the transformed forecasts. In other words, by this measure, K04 underestimated the positive impact of land moisture initialization on a forecast. For air temperature, however, the apparent benefit of land moisture initialization is

Benefits Due to Soil Moisture Initialization (LDAS - AMIP)

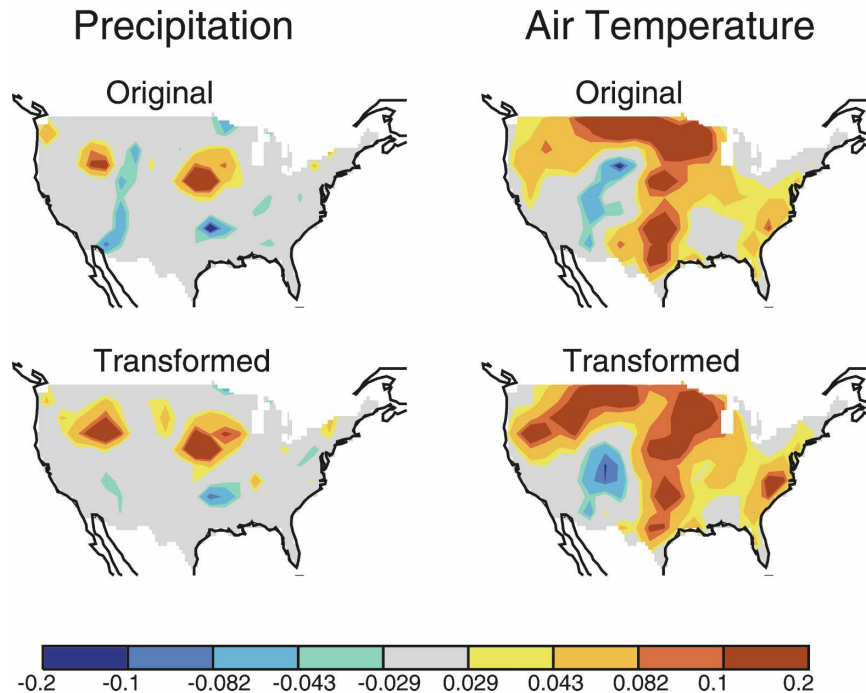


FIG. 7. Forecast skill (expressed for this figure as the square of the correlation coefficient between the forecasts and observations) in the LDAS mode minus that in the AMIP mode, a measure of the contribution of land state initialization to forecast skill. Results are shown for (right) precipitation and (left) near-surface air temperature; (top) the results for the original K04 study and (bottom) the corresponding results for the transformed forecasts are shown. Differences of 0.029, 0.043, and 0.082 are statistically different from zero at the 90%, 95%, and 99% levels, respectively.

roughly the same for the transformed forecasts as it was for the original forecasts. For air temperature, a reduction in benefit is seen in parts of the southern United States, while the northwestern part of the country sees some increase in benefit.

Regardless of whether or not the land initialization benefit appears to increase, we can conclude two things. First, realistic soil moisture initialization does increase skill, both for the original and for the transformed forecasts. Second, the transformation of the forecasts through (1) increases skill for both the AMIP and LDAS modes. In other words, the greatest total skill, for both precipitation and air temperature prediction, stems from the use of both land initialization and forecast transformation (Figs. 5e and 6e).

6. Summary

We propose an approach for jointly utilizing two disparate pieces of information—the statistical structure

of observed meteorological fields and the potential predictability within a forecast system—to develop a transformation matrix that improves the subseasonal forecasts generated by the system. In essence, the approach is based on the idea that a forecast can be spatially “translated” from a location where the model has predictive skill to a location where it does not, provided that the observations show the two locations are linked. (The linkage may be absent in the forecast system because of biases in the modeled climate.) The statistical structure of the observations can be robustly determined from existing multidecadal datasets. The potential predictability, an inherent property of a given model, can be established through actual or idealized prediction experiments. By using idealized prediction experiments, in which soil moistures are initialized to specific, model-consistent values that may have no connection to observations, the transformation matrix can be derived before any true forecast is attempted with the system, which is an advantage if “training” a trans-

formation matrix through extensive forecast experiments is difficult due to computational constraints or observational data limitations.

The effectiveness of the approach is tested by deriving the transformation matrices for the forecast system utilized by K04 and then applying them to the subseasonal (1 month) forecasts analyzed in that study. The increases in skill obtained through the transformation are modest but significant. On average, across the United States, the transformation matrix increases the subseasonal forecast skill stemming from SST prescription and land initialization by about 60% for precipitation and by about 15% for near-surface air temperature.

Skillful subseasonal (and seasonal to interannual) forecasts have obvious societal benefits. As models, data collection, and data assimilation strategies improve, the forecasts should also improve. Each increment of forecast system improvement, particularly those increments unrelated to the spatial statistical structures of simulated meteorological fields, perhaps can be enhanced further through the application of the transformation strategy outlined in this paper.

Acknowledgments. The AGCM runs were funded by the Earth Science Enterprise of NASA Headquarters through the EOS Interdisciplinary Science Program and the NASA Global Modeling and Assimilation Office (GMAO), with computational resources provided by the NASA Center for Computational Sciences. Ping Liu and Sarith Mahanama assisted with the data processing.

APPENDIX

Transformation Approach: Mathematical Basis

The approach to improving subseasonal forecasts described above can be put into a more conventional statistical context if we formulate it as a problem in devising a statistical model that can help us find an optimal choice for \mathbf{A} in (1), based on the combined database of observations, the forecast model runs, and the predictability study results. The statistical model should be as consistent as possible with what we know, but we are limited in how flexible we can make the model because of the finite amount of data available to us for establishing model parameters.

To simplify notation, we shall use ξ here to denote what is notated as \mathbf{x}_{obs} in the main text. In other words, ξ is the “true,” observed state of the atmosphere when \mathbf{x} is the ensemble mean forecast. Equation (1) is then written

$$\xi = \mathbf{A}\mathbf{x} + \boldsymbol{\eta}, \quad (\text{A1})$$

where we have made the prediction error $\boldsymbol{\eta}$ relative to the hoped-for improved forecast $\mathbf{A}\mathbf{x}$ explicit. The vectors ξ , \mathbf{x} , and $\boldsymbol{\eta}$ are of length N . As in the main text, the components of the vectors ξ and \mathbf{x} are standardized, $\langle x_m^2 \rangle = \langle \xi_m^2 \rangle = 1$, and $\langle x_m \rangle = \langle \xi_m \rangle = 0$.

If we assume that the prediction errors are approximately normally distributed, then the optimal choice for \mathbf{A} is the standard linear regression result

$$\mathbf{A} = \langle \xi \mathbf{x}^T \rangle \langle \mathbf{x} \mathbf{x}^T \rangle^{-1}, \quad (\text{A2})$$

which is just (2) in the new notation, where T indicates matrix transpose. Because we lack a sufficient number of pairs of observations and forecasts $\{\xi, \mathbf{x}\}$ from the forecast runs, however, the inverse covariance matrix $\langle \mathbf{x} \mathbf{x}^T \rangle^{-1}$ would be singular if we used only these data to estimate it. We therefore try to obtain an estimate of (A2) by bringing in the predictability studies and observational data in our database.

The matrix \mathbf{A} in (A2) could be determined if we knew the joint probability distribution $P(\xi, \mathbf{x})$. The approach described here produces a model for the distribution. To show that the approach creates a statistically consistent description of $P(\xi, \mathbf{x})$, it is helpful to decompose it into the product of two distributions using Bayes’s theorem:

$$P(\xi, \mathbf{x}) = P(\mathbf{x}|\xi)P(\xi). \quad (\text{A3})$$

The conditional probability $P(\mathbf{x}|\xi)$ is the probability that the forecast is \mathbf{x} when the actual state of the atmosphere is ξ . (Forecasts vary unpredictably depending on details of the initialization and boundary conditions.) The probability distribution $P(\xi)$ on the right of (A3) describes the distribution of possible states of the atmosphere, of which the observational data provide a sample. Looking ahead, the methods we describe next produce a model for $P(\mathbf{x}|\xi)$ that is entirely based on the model predictability studies, while $P(\xi)$ is based entirely on observational data. The joint distribution $P(\mathbf{x}, \xi)$ is thus a realizable distribution for ξ and \mathbf{x} .

We try to construct a statistical model for $P(\mathbf{x}|\xi)$, keeping in mind the limitations of our datasets. We therefore attempt to represent the statistics of \mathbf{x} by a simple linear model

$$\mathbf{x} = \mathbf{R}\xi + \boldsymbol{\epsilon}, \quad (\text{A4})$$

which estimates what the forecast would be knowing the actual state of the atmosphere ξ , where $\boldsymbol{\epsilon}$ is the unpredictable error in this estimate of \mathbf{x} . Note that this model is the reverse of (A1) and should be viewed as a

best effort at representing the statistics of \mathbf{x} given ξ , in the “least squares” sense, whereas (A1) is a best effort at representing the statistics of ξ given \mathbf{x} , also in the least squares sense.

We could in principle find an optimal matrix \mathbf{R} in (A4) from the forecast runs, but, again, the actual number of runs available is insufficient for this. It is here that we bring in what we have learned from the predictability studies. Recall that the predictability studies employ ensembles of multiple runs where each member of the ensemble is successively treated as the actual state of the atmosphere, which we can denote as ξ_{mod} , and the remainder are used to make the ensemble mean forecast \mathbf{x} . We estimate \mathbf{R} and $\langle \epsilon \epsilon^T \rangle$ for (A4) by replacing ξ with ξ_{mod} and using the predictability study runs as data. The standard linear regression estimate of \mathbf{R} is then

$$\mathbf{R} = \langle \mathbf{x} \xi_{\text{mod}}^T \rangle \langle \xi_{\text{mod}} \xi_{\text{mod}}^T \rangle^{-1}, \quad (\text{A5})$$

where the angular brackets indicate an average over the ensemble runs and choices ξ_{mod} . In replacing ξ with ξ_{mod} in estimating \mathbf{R} , we are in effect making the assumption that the model’s potential predictability provides a useful approximation to \mathbf{R} , which in reality should be calculated from the model’s error in predicting the true atmosphere’s behavior. The covariance statistics of ϵ are given by

$$\begin{aligned} \Sigma_\epsilon \equiv \langle \epsilon \epsilon^T \rangle &= \langle (\mathbf{x} - \mathbf{R} \xi_{\text{mod}})(\mathbf{x} - \mathbf{R} \xi_{\text{mod}})^T \rangle \\ &= \langle \mathbf{x} \mathbf{x}^T \rangle - \mathbf{R} \langle \xi_{\text{mod}} \xi_{\text{mod}}^T \rangle \mathbf{R}^T, \end{aligned} \quad (\text{A6})$$

using (A5). We have thus used the predictability study part of our database to specify the parameters in the linear model (A4).

With the parameters \mathbf{R} and Σ_ϵ of the linear model (A4) in hand, we could in principle obtain \mathbf{A} in (A2) using (A4) and the covariance of the observational data

$$\Sigma_\xi \equiv \langle \xi \xi^T \rangle, \quad (\text{A7})$$

because $\langle \xi \mathbf{x}^T \rangle = \Sigma_\xi \mathbf{R}^T$ and $\langle \mathbf{x} \mathbf{x}^T \rangle = \Sigma_\epsilon + \mathbf{R} \Sigma_\xi \mathbf{R}^T$. This would give us \mathbf{A} as

$$\mathbf{A} = \Sigma_\xi \mathbf{R}^T (\Sigma_\epsilon + \mathbf{R} \Sigma_\xi \mathbf{R}^T)^{-1}. \quad (\text{A8})$$

Recall that our original problem was that the limited number of forecast experiments available for estimating \mathbf{A} using (A2) caused the matrix $\langle \mathbf{x} \mathbf{x}^T \rangle$ to be singular and the estimate (A2) to be unusable. The estimate of \mathbf{A} given by (A8) might be finite even when our estimate of Σ_ξ is singular (because of insufficient observational data), because of the addition of the symmetric matrix Σ_ϵ to the “denominator” of (A8). The incorporation of the predictability study information seems, in effect, to

have stabilized our estimate of (A2) in a manner analogous to singular value decomposition methods for solving linear equations with an insufficient number of equations (e.g., Press et al. 1992). Because there are fewer years of observational data than the number N of grid points, however, the stability characteristics of (A8) are unclear to us. We have therefore chosen to follow a more conservative route to estimating \mathbf{A} than the straightforward evaluation of (A8), and we introduce further simplifying assumptions.

We first reduce the number of parameters in (A4) considerably by ignoring off-diagonal terms in \mathbf{R} and Σ_ϵ , assuming

$$(\mathbf{R})_{mm'} = r_m \delta_{mm'} \quad \text{and} \quad (\text{A9})$$

$$(\Sigma_\epsilon)_{mm'} = \sigma_m^2 \delta_{mm'}, \quad (\text{A10})$$

where $\delta_{mm'}$ is the Kronecker delta. These assumptions might eventually be relaxed if more predictability runs become available, but we feel they are necessary for now in order to make progress. The variables \mathbf{x} and ξ are standardized, $\langle x_m^2 \rangle = \langle \xi_m^2 \rangle = 1$. Equation (A6) therefore implies that

$$\sigma_m^2 = 1 - r_m^2. \quad (\text{A11})$$

The definition of potential predictability introduced in the main text allows us to identify the diagonal elements r_m and σ_m^2 in (A9) and (A10), using (A5) and (A11), as

$$r_m = \langle x_m(\xi_{\text{mod}})_m \rangle \equiv r_{\text{pot},m} \quad \text{and} \quad (\text{A12})$$

$$\sigma_m^2 = \langle \epsilon_m^2 \rangle = 1 - r_{\text{pot},m}^2. \quad (\text{A13})$$

The modeling approach described here can be related directly to the description of the transformation approach described in section 4. Recall that in the transformation approach, the observations in a given year are degraded using (3), and a multiple regression is performed to determine how the degraded observations relate to an observation at a target cell. Performing the multiple regression separately for each target cell fills out the matrix \mathbf{A} . The degradation in (3) corresponds exactly to (A4) in the formalism of this appendix, with \mathbf{R} taken to be the diagonal matrix holding the idealized predictability estimates, as indicated by (A12), and ϵ being a vector of uncorrelated random unit normal deviates scaled by the factor $(1 - r_{\text{pot},m}^2)^{1/2}$, as indicated by (A13).

Said another way, (A4) provides a framework for generating a very large number of artificial “model forecasts” \mathbf{x} that reflect the model’s own inherent predictability characteristics, and (A4) can be used easily given the assumptions leading to (A12) and (A13), which make (A4) equivalent to (3). By concatenating multiple repeat time series of the observations, as described in the main text, we have enough data for a reliable multiple regression [as in (A2), but cell by cell] to fill out the matrix \mathbf{A} . When combined with the objective method described in the main text of reducing the number of grid cells in the predictor equations as much as possible to minimize the prediction error, the modeling approach outlined here becomes identical to the procedure described in the main text.

The multiple regressions on the degraded time series in section 4 should approach (A8) (with the diagonal forms of \mathbf{R} and Σ_e) in the limit of a very large number of simulations. The use of (A8) in its more general form may be a profitable direction for future exploration of the method.

REFERENCES

- Atlas, R., N. Wolfson, and J. Terry, 1993: The effect of SST and soil moisture anomalies on GLA model simulations of the 1988 U.S. summer drought. *J. Climate*, **6**, 2034–2048.
- Bacmeister, J., P. J. Pegion, S. D. Schubert, and M. J. Suarez, 2000: Atlas of seasonal means simulated by the NSIPP 1 atmospheric GCM. NASA Tech. Memo. 2000-104606, Vol. 17, 194 pp.
- Beljaars, A. C. M., P. Viterbo, M. J. Miller, and A. K. Betts, 1996: The anomalous rainfall over the United States during July 1993: Sensitivity to land surface parameterization and soil moisture anomalies. *Mon. Wea. Rev.*, **124**, 362–383.
- Berg, A. A., J. S. Famiglietti, J. P. Walker, and P. R. Houser, 2003: Impact of bias correction to reanalysis products on simulations of North American soil moisture and hydrological fluxes. *J. Geophys. Res.*, **108**, 4490, doi:10.1029/2002JD003334.
- Charba, J. P., A. W. Harrell III, and A. C. Lackner III, 1992: A monthly precipitation amount climatology derived from published atlas maps: Development of a digital data base. NOAA/U.S. Department of Commerce, TDL Office Note 92-7, 20 pp.
- Chou, M.-D., and M. Suarez, 1994: An efficient thermal infrared radiation parameterization for use in general circulation models. NASA Tech. Memo. 104606, Vol. 3, 85 pp.
- Cressman, G. P., 1959: An operational objective analysis system. *Mon. Wea. Rev.*, **87**, 367–374.
- Delworth, T. L., and S. Manabe, 1988: The influence of potential evaporation on the variabilities of simulated soil wetness and climate. *J. Climate*, **1**, 523–547.
- , and —, 1989: The influence of soil wetness on near-surface atmospheric variability. *J. Climate*, **2**, 1447–1462.
- Dirmeyer, P. A., 2000: Using a global soil wetness dataset to improve seasonal climate simulation. *J. Climate*, **13**, 2900–2922.
- , 2001: An evaluation of the strength of land–atmosphere coupling. *J. Hydrometeorol.*, **2**, 329–344.
- Douville, H., and F. Chauvin, 2000: Relevance of soil moisture for seasonal climate predictions: A preliminary study. *Climate Dyn.*, **16**, 719–736.
- , —, and H. Broqua, 2001: Influence of soil moisture on the Asian and African monsoons. Part I: Mean monsoon and daily precipitation. *J. Climate*, **14**, 2381–2403.
- Entin, J. K., A. Robock, K. Y. Vinnikov, S. E. Hollinger, S. Liu, and A. Namkhai, 2000: Temporal and spatial scales of observed soil moisture variations in the extratropics. *J. Geophys. Res.*, **105**, 11 865–11 877.
- Fennessy, M. J., and J. Shukla, 1999: Impact of initial soil wetness on seasonal atmospheric prediction. *J. Climate*, **12**, 3167–3180.
- Gates, W. L., 1992: AMIP: The Atmospheric Model Intercomparison Project. *Bull. Amer. Meteor. Soc.*, **73**, 1962–1970.
- Glahn, H. R., T. L. Chambers, W. S. Richardson, and H. P. Perrotti, 1985: Objective map analysis for the local AFOS MOS Program. NOAA/U.S. Department of Commerce, NOAA Tech. Memo. NWS TDL 75, 34 pp.
- Guo, Z., and Coauthors, 2006: GLACE: The Global Land–Atmosphere Coupling Experiment. Part II: Analysis. *J. Hydrometeorol.*, **7**, 611–625.
- Higgins, R. W., W. Shi, E. Yarosh, and R. Joyce, 2000: Improved United States precipitation quality control system and analysis. NCEP/Climate Prediction Center ATLAS No. 7. [Available online at http://www.cpc.ncep.noaa.gov/research_papers/ncep_cpc_atlas/7/index.html.]
- Hong, S., and E. Kalnay, 2000: Role of sea surface temperature and soil-moisture feedback in the 1998 Oklahoma-Texas drought. *Nature*, **408**, 842–844.
- 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.
- Koster, R. D., and M. J. Suarez, 1995: The relative contributions of land and ocean processes to precipitation variability. *J. Geophys. Res.*, **100**, 13 775–13 790.
- , and —, 1996: Energy and water balance calculations in the Mosaic LSM. NASA Tech. Memo. 104606, Vol. 9, 59 pp.
- , —, and M. Heiser, 2000: Variance and predictability of precipitation at seasonal-to-interannual timescales. *J. Hydrometeorol.*, **1**, 26–46.
- , and Coauthors, 2004: Realistic initialization of land surface states: Impacts on subseasonal forecast skill. *J. Hydrometeorol.*, **5**, 1049–1063.
- , and Coauthors, 2006: GLACE: The Global Land–Atmosphere Coupling Experiment. Part I: Overview. *J. Hydrometeorol.*, **7**, 590–610.
- Moorthi, S., and M. J. Suarez, 1992: Relaxed Arakawa–Schubert: A parameterization of moist convection for general circulation models. *Mon. Wea. Rev.*, **120**, 978–1002.
- Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, 1992: *Numerical Recipes in C*. Cambridge University Press, 1021 pp.
- Rayner, N. A., E. B. Horton, D. E. Parker, C. K. Folland, and R. B. Hackett, 1996: Version 2.2 of the global sea-ice and sea surface temperature data set, 1903–1994. Met Office Climate Research Tech. Note 74, 43 pp. [Available from Met Office, London Road, Bracknell, Berkshire RG12 2SY, United Kingdom.]

- Reynolds, R. W., and T. M. Smith, 1994: Improved global sea-surface temperature analyses using optimum interpolation. *J. Climate*, **7**, 929–948.
- Rodell, M., and Coauthors, 2004: The Global Land Data Assimilation System. *Bull. Amer. Meteor. Soc.*, **85**, 381–394.
- Ropelewski, C. F., J. E. Janowiak, and M. S. Halpert, 1985: The analysis and display of real time surface climate data. *Mon. Wea. Rev.*, **113**, 1101–1106.
- Seneviratne, S. I., and Coauthors, 2006: Soil moisture memory in AGCM simulations: Analysis of Global Land–Atmosphere Coupling Experiment (GLACE) data. *J. Hydrometeor.*, **7**, 1090–1112.
- Vinnikov, K. Y., and I. B. Yeserkepova, 1991: Soil moisture, empirical data and model results. *J. Climate*, **4**, 66–79.
- Viterbo, P., and A. K. Betts, 1999: Impact of the ECMWF re-analysis soil water on forecasts of the July 1993 Mississippi flood. *J. Geophys. Res.*, **104**, 19 361–19 366.
- Zwiers, F. W., 1996: Interannual variability and predictability in an ensemble of AMIP climate simulations conducted with the CCC GCM2. *Climate Dyn.*, **12**, 825–847.