ECPC's Weekly to Seasonal Global Forecasts

John O. Roads,* Shyh-Chin Chen,* and Francis Fujioka+

ABSTRACT

The Scripps Experimental Climate Prediction Center (ECPC) has been making experimental, near-real-time seasonal global forecasts since 26 September 1997 with the NCEP global spectral model used for the reanalysis. Images of these forecasts, at daily to seasonal timescales, are provided on the World Wide Web and digital forecast products are provided on the ECPC anonymous FTP site to interested researchers. These forecasts are increasingly being used to drive regional models at the ECPC and elsewhere as well as various application models. The purpose of this paper is to describe the forecast and analysis system, various biases and errors in the forecasts, as well as the significant skill of the forecasts. Forecast near-surface meteorological parameters, including temperature, precipitation, soil moisture, relative humidity, wind speed, and a fire weather index (a nonlinear combination of temperature, wind speed, and relative humidity) are skillful at weekly to seasonal timescales over much of the United States and for many global regions. These experimental results suggest there is substantial forecast skill, out to at least a season, to be realized from current dynamical models.

1. Introduction

The Scripps Experimental Climate Prediction Center has been making experimental, near-real-time, long-range global dynamical forecasts since 27 September 1997. Images from these forecasts are regularly shown on its World Wide Web (WWW) site (http://ecpc.ucsd.edu/; see Roads et al. 2000b). The global model used for these forecasts is the National Centers for Environmental Prediction's (NCEP's) global spectral model (GSM; Kalnay et al. 1996; Roads et al. 1998, 1999) used for the NCEP–National Center for Atmospheric Research (NCAR) reanalysis. The initial conditions and SST boundary conditions for these experimental global forecasts come from the NCEP Global Data Assimilation (GDAS) 0000 UTC operational analysis, which is available nearly every day in near–real time on NCEP rotating disk archives, to interested researchers. Transforming NCEP's higher-resolution operational analyses to lower- (vertical and horizontal) resolution initial conditions for the global model, 7-day global forecasts are made every day and every weekend, these global forecasts are extended to 12 weeks. A sufficient number of these experimental forecasts have now been made (104; once a week for the period Oct 1997–Oct 1999) to begin estimating the forecast skill.

Since these experimental long-range forecasts extend well beyond traditional atmospheric predictability limits (e.g., Lorenz 1965, 1982), it is important to stress that the foundation for this type of long-range prediction is probably based upon the interaction of the chaotic atmosphere with slowly varying boundary conditions. Persistent boundary conditions include anomalous sea surface temperatures (SSTAs), soil moisture (see, e.g., Huang et al. 1996; Fennessy and Shukla 1999), and snow (see Marshall et al., 2000, manuscript submitted to *Ann. Glaciol.*). In that regard, it should be noted that slowly varying boundary conditions are externally provided to the global forecast

*Experimental Climate Prediction Center, Scripps Institution of Oceanography, University of California, San Diego, La Jolla, California.
+Riverside Fire Laboratory, U.S. Forest Service, Riverside, California.
Corresponding author address: Dr. John O. Roads, Experimental Climate Prediction Center, Scripps Institution of Oceanography, 0224, University of California, San Diego, La Jolla, CA 92093.
E-mail: jroads@ucsd.edu
In final form 12 September 2000.

model by persisting the initial SSTAs throughout the entire forecast. Internally, the model's persistent soil moisture, which provides an intrinsic low-frequency land forcing, especially during the summer when the surface fluxes are positive, may also be influential. The model's forecast snowfields may also be important, especially during springtime, since their presence or absence can have a large impact upon atmospheric fields (although it will be shown later that the lowest skill tends to occur during spring). In brief, it is this persistent surface forcing that dynamical as well as statistical models are presumably exploiting for making skillful long-range predictions.

Systematic dynamical long-range forecast experiments, like those evaluated here, began many years ago as part of the dynamical extended-range forecast experiments by NCEP (see, e.g., Roads 1989; Tracton et al. 1989; Barker and Horel 1989). Due to the computer limitations at the time, only 30-day forecasts were made for a limited time period (90 days). With the advent of newer and better forecast models and faster computers, dynamical seasonal forecasts have subsequently become more commonplace and several long-range forecast experiments, mainly focused upon Northern Hemisphere winter, have begun to be reported on in the literature (see, e.g., Kumar et al. 1996; Anderson et al. 1999; Shukla et al. 2000; Brankovic and Palmer 2000, manuscript submitted to *Quart. J. Roy. Meteor. Soc.*).

The purpose here is to evaluate the forecast skill, errors, and biases of two complete years of weekly to seasonal forecasts. Now, it is sometimes argued that since long-range experiments provide only limited and probabilistic forecast skill, an ensemble approach is needed to provide increased significance for specific forecasts. Although we do not have the computer capacity to make ensemble forecasts on a regular basis, it should be noted that our skill evaluation is made on an ensemble of 104 forecasts. This large ensemble is more than sufficient to provide an estimate of the climatological average and distribution of typical forecasts for at least the available forecasts. As will be shown, there is substantial skill in these long-range predictions. However, a major caveat is that our long-range forecasts during the evaluation period may be overly predictable because of a strong El Niño–Southern Oscillation (ENSO) cycle. That is, there may be strong and persistent tropical forcing associated with the large tropical sea surface temperature anomalies of the 1997–98 El Niño and 1998–99 La Niña. In fact, a number of papers have commented upon the increased predictability during previous ENSO events (see Barnston 1994; Brankovic et al. 1994; Kumar and Hoerling 1998). Other caveats about our evaluation will be discussed later.

In section 2, we briefly describe the global spectral model, as well as the initial conditions and validating analysis. Section 3 discusses our error and skill measures. Section 4 describes geographic characteristics of seasonal forecasts. Section 5 describes temporal characteristics of seasonal forecasts. Section 6 describes the ensemble forecast skill of weekly to seasonal forecasts. Conclusions are provided in section 7.

2. Tools and methods

a. *Global spectral model*

The GSM used for this study is based upon the Medium-Range Forecast (MRF) Model used at NCEP for making the four-times-daily GDAS analysis and for making the long-range (6–14 day) predictions. This GSM, which has undergone steady improvement for a number of years (see Caplan et al. 1997), became on 10 January 1995 the basic global model used for the NCEP–NCAR reanalysis (hereafter referred to as NCEPR; see Kalnay et al. 1996).

The GSM uses a primitive equation or hydrostatic system of virtual temperature, humidity, surface pressure, mass continuity, vorticity, and divergence prognostic equations on terrain-following sigma (sigma is defined as the ratio of the ambient pressure to surface pressure) coordinates. Our particular version of the GSM uses spherical harmonics with a triangular truncation of T62 and 18 irregularly spaced vertical levels (L18T62). These levels are concentrated near the lower boundary and tropopause.

As discussed by Hong and Leetma (1999; see also Hong and Pan 1996), the physics package for the GSM includes longwave and shortwave radiation interactions between cloud and radiation; boundary layer processes, such as shallow clouds and convection; large-scale condensation; gravity wave drag; and enhanced topography. Vertical transfer throughout the troposphere, including the boundary layer, is related to eddy diffusion coefficients dependent upon a Richardson number–dependent diffusion process (Kanamitsu 1989). For large-scale condensation, the GSM has a vertical iteration starting from the top sigma level that checks for supersaturation at each level. Supersaturated layers are set to a saturated state and excess water is precipitated. Lower layers evaporate or augment this precipitation until it reaches the surface in the form of rain or snow (depending upon whether the lowest sigma layer is below freezing.

A key parameterization development for the GSM was the innovative land surface parameterization (LSP), which underwent a radical change from the former bucket model. The new LSP (see Pan 1990) consists of two soil layers in which soil moisture and temperature as well as moisture present in the vegetation and snow are carried as dependent variables. Exchange between the two soil layers is modeled as a diffusion process. Evaporation occurs from bare soil, leaf canopy, as well as transpiration through leaf stomata. Chen et al. (1996) made further improvements (see also Betts et al. 1997) to NCEP's associated Eta Model, and these changes are being incorporated in more recent versions of the GSM.

Another key GSM Development effort has been concerned with the cumulus convection parameterization, which currently uses a simplified Arakawa–Schubert (SAS; Arakawa and Schubert 1974) parameterization. Pan and Wu (1995), following the Grell (1993) simplification, developed this for the NCEPR. SAS removes large-scale instabilities by relaxing temperature and moisture profiles toward prescribed equilibrium values on a prescribed timescale. The convection scheme also allows entrainment into the updraft and detrainment from the downdraft between the level from which the updraft air originates and the level of free convection (LFC). The level of maximum moist static energy between the surface and 400 hPa is used as the level from which the updraft air originates. Convection is suppressed when the distance between the updraft air originating level and the LFC exceeds a certain threshold (150 hPa). Cloud top is determined as the first neutral level above the cloud base.

It should be noted that a number of more recent improvements have been implemented in NCEP models, which may ultimately prove useful in increasing the forecast skill (see, e.g., Hong and Leetma 1999; M. Kanamitsu 1999, personal communication). In that regard, it is our intention to eventually transition our forecast system to a more recent version of the NCEP model and to reexamine the skill in the new system. It should also be noted that a regional spectral model (see Juang and Kanamitsu 1994; Chen et al. 1999; Anderson et al. 2000; Roads and Chen 2000), with the same basic parameterizations as the GSM, is also being used to make higher-resolution forecasts for specific regions. The forecast skill of the higher-resolution forecast model will eventually be compared to the forecast skill of this global model as soon as we can obtain a similar number of forecasts.

b. Initial conditions

The initial conditions for the GSM forecasts come from the NCEP GDAS operational analysis (L28T126), which is posted in a timely fashion on a rotating disk archive at NCEP. We have managed to access these initial conditions almost every day for the 0000 UTC analysis from 27 September 1997 to present (28LT126 analyses were available from 0000 UTC 27 September 1997 to 0000 UTC 16 March 2000; thereafter 42LT170 became available). These higher-resolution analyses are then transformed to lower-resolution initial conditions (18LT62) by linearly interpolating between vertical sigma levels, spectrally truncating the spectral components, and bilinearly interpolating the higher-resolution surface grids to our lower-resolution grids (and land mask). We have also tried to access initial conditions for the three other analysis times (0600, 1200, 1800 UTC) but have been slightly less successful. Also, on 27 September 1999, there was a fire at NCEP that severely impacted the computer system and since then only twice-daily analysis times (0000, 1200 UTC) were made available for initialization and validation.

The GSM LSP is capable of changing the surface land conditions (snow and soil moisture), and thus only initial conditions are needed over the land regions. However, ocean conditions must be specified during the course of a prediction. We change the SST climatological component continuously throughout the integration, and persist the initial SSTA throughout the forecast integration. The sea–ice distribution is only changed climatologically. Obviously it is of interest to eventually use coupled ocean–atmosphere models as well as coupled land–atmosphere models to make seasonal predictions. Unfortunately, no coupled ocean–atmosphere model has yet been shown superior to simply persisting the SSTA for at least the first 12 weeks.

c. Validating analyses

Although the operational GDAS analyses are sufficient to start our GSM forecasts, they are not sufficient to evaluate the desired forecast variables. For example, only atmospheric state variables such as temperature, humidity, winds, surface pressure, and surface state variables, such as soil moisture and snow, are available in the GDAS sigma files and surface files. Another file, the so-called flux file, developed from 6-h forecasts with the MRF models contains near-surface information such as maximum and minimum 2-m temperature; humidity; 10-m winds; and surface latent, sensible, radiative, and top-of-the-atmosphere radiation fluxes; and precipitation. These GDAS flux files were more difficult to access initially and it was not until 15 March 1998 that we were successful in getting the four-times-daily flux files to evaluate our forecasts. These analysis files then formed our basic validation dataset until the NCEP fire (27 September 1999) at which point only twice-daily GDAS flux files became available and adversely affected the daily averages we were making from the four-times-daily forecasts. Although we could have also used the NCEP reanalysis files to validate the model (and we did use these to develop preliminary climatologies before we had the GDAS files), we were never able to access these files in as timely a manner as the GDAS files.

To extend backward the validation forecast period to the time when we first started archiving initial states, to extend a consistent validation beyond the NCEP fire, and to have available in near–real time validating observations, we ultimately decided to develop our own flux files. Therefore, for our main validation effort, we now use 1-day forecasts made every day from 0000 UTC analysis initial conditions. There are only three missing 0000 UTC initial states in more than a 2-yr period and for these periods we used a previous 2-day forecast to generate the associated daily flux files. These 1-day forecasts are not exactly the same as the operational analysis, which is based upon four-times-daily 6-h forecasts from the latest high-resolution global model, or the reanalysis, which is based upon four-times-daily 6-h forecasts with the model we use for our forecasts, but they do form a useful approximation that can at least be used to estimate forecast skill upper bounds. As we shall see, forecasts validated against these 1-day forecasts are slightly more skillful than forecasts validated against NCEP's operational analysis. For future reference, the validating 1-day forecast analyses will be referred to as the V1 or 1-day forecast analysis and the more-limited GDAS operational analysis will be referred to as the VO or operational analysis.

Although the 1-day forecast analysis validation set is certainly useful in the absence of other validating datasets, at least there are better approximations to the global precipitation. In particular, Xie and Arkin (1997) have developed a global precipitation dataset that not only extends back to 1979 monthly means but also provides higher temporal resolution pentad means mostly from satellite and gauge measurements (numerical weather prediction estimates are mainly used only in high latitudes to fill in missing grid points). These pentads were interpolated into weekly means (by first interpolating them into daily means and then subsequently accumulating the daily values into weekly means). This analysis will be referred to as the VX or Xie and Arkin analysis.

d. Forecasts

Seven-day GSM forecasts are made and displayed on the WWW every day and digital files are transferred to an anonymous FTP site in order to provide general information to interested researchers as well as to develop the basic 1-day validating analysis. These daily forecasts are also used to drive associated regional models, which will be discussed in another paper. Here we examine the 12-week GSM forecasts, which are made only once a week (every weekend when the greatest computer capacity is available). These 12-week forecasts are then archived into weekly averages, which can be further averaged into three monthly (4 week) averages and a seasonal (12 week) average. Because of limited archive capacity, we decided not to evaluate timescales of less than a week, at least initially.

3. Skill evaluation

A basic evaluation of forecast skill is concerned with the average difference or bias of the forecasts:

$$Bias = [F - O] = [F] - [O]$$

where [. . .] indicates a climatological average, which is time (monthly and forecast lag) dependent; F is the forecast average (weekly, monthly, seasonal); and O is the corresponding validating observation or analyses with the same time averages. The ensemble average can also be a spatially (cosine) weighted average:

$$\{Bias\} = \{[F]\} - \{[O]\},$$

where {. . .} indicates the spatially (cosine) weighted average.

We develop climatological averages of weekly to seasonal forecasts. In that regard, biases can have a seasonal dependence. As will be shown, for spatial averages over the U.S. land area, there is a tendency for seasonal temperature forecasts to have a negative bias during the fall and a positive bias during the spring, suggesting a shifted GSM seasonal cycle. Biases can also have a geographic dependence. As will be shown, the northern tier of states has a warm bias and the southern tier of states has a cold bias. Also, biases depend upon forecast lag. Forecast biases for the 1st week are different from the forecast biases for the 12th week. For example, initial GSM soil moisture and relative humidity are wet but by the 12th week become dry. Biases thus depend upon averaging length; seasonal biases are likely to be different from weekly biases.

Another skill evaluation is concerned with how well the model forecasts anomalies from the mean climatological state. It is usually quite difficult for dynamical models to exactly reproduce the mean state, even though they can still forecast the variations about the mean state. To evaluate the skill of these anomaly forecasts, we remove the associated (forecast or observed) climatological state for each grid point, for each initial state month, and for each forecast lag time. Averaging each monthly average with a three-point Shapiro temporal filter removes noisy two-gridpoint temporal variations in the monthly climatologies. The anomalies are then defined separately for the forecast, F, and validating analysis, O, by

$$F' = F - [F]$$

and

$$O' = O - [O].$$

Now one useful measure of anomalous skill is an rms measure:

$$\mathrm{rms} = [(F' - O')^2]^{1/2},$$

where O' is the observed deviation. However, unless one knows what the background rms is in the absence of no skill, $[F'^2 + O'^2]^{1/2}$, an rms value does not immediately indicate if the forecasts are good or bad.

In that respect, correlations provide normalized values as well as a body of standard literature about them that can be used to evaluate skill significance:

$$\mathrm{COR} = \frac{[F'O'] - [F'][O']}{\left([F'^2] - [F']^2\right)^{1/2}\left([O'^2] - [O']^2\right)^{1/2}}.$$

Now a correlation can be calculated for each grid point to give spatial maps of the ensemble correlations for a complete annual cycle. For 104 (once a week from Oct 1997 to Oct 1999) independent forecasts, a correlation above 0.16 is considered significant at the 95% level (see von Storch and Zwiers 1999). Allowing for some degree of dependence between the weekly forecasts of seasonal means, we only plot correlations above 0.2 in the geographic maps (Figs. 1–6) of seasonal forecast mean correlation.

Correlations can also be calculated for various domain averages to provide an instantaneous average value, which can also be used to estimate how significant ensemble averages are. In this case the time-varying correlations consist of areally weighted grid points for each forecast:

$$\mathrm{COR}(t) = \frac{\{F'O'\} - \{F'\}\{O'\}}{\left(\{F'^2\} - \{F'\}^2\right)^{1/2}\left(\{O'^2\} - \{O'\}^2\right)^{1/2}}.$$

By calculating the temporal variations of areally averaged correlations, one can estimate the standard deviations of each areally weighted ensemble forecast correlation. That is, the standard deviation of the spatial correlations provides a measure of the climatological uncertainty for each forecast variable. This correlation can then be further transformed to a normal variable (see Roads 1989) although this makes only a small difference in practice.

Unfortunately, this measure of uncertainty breaks down for geographically small areas where the observed and forecast variations can be of one sign; that is, for small areas, $\{F'\}$ and $\{O'\}$ can be nonzero instantaneously, even though $[\{F'\}]$ and $[\{O'\}]$ are zero. Although removing the spatial average presumably affects both the numerator and denominator, in practice it appears to reduce the numerator more. Thus

$[\mathrm{COR}(t)]$ *can* be noticeably smaller than

$$\frac{[\{F'O'\}] - [\{F'\}][\{O'\}]}{\left([\{F'^2\}] - [\{F'\}]^2\right)^{1/2}\left([\{O'^2\}] - [\{O'\}]^2\right)^{1/2}}.$$

Because of this areal mean problem in calculating instantaneous correlations for small areas, the normalized covariance was used to estimate the skill significance for individual times (see Figs. 7–12). That is,

$$\mathrm{COV}(t) = \frac{\{F'O'\}}{\left(\{F'^2\}\right)^{1/2}\left(\{O'^2\}\right)^{1/2}},$$

and then because $[F']$ and $[O']$ are identically zero for the climatological mean,

$$[\mathrm{COV}] \cong \frac{[\{F'O'\}]-[\{F'\}][\{O'\}]}{\left([\{F'^2\}]-[\{F'\}]^2\right)^{1/2}\left([\{O'^2\}]-[\{O'\}]^2\right)^{1/2}}.$$

Indeed, time (actually ensemble) averages of the temporally varying normalized covariance were much closer to time- and space-averaged correlations, which validates the use of instantaneous normalized covariance to estimate the significance of the ensemble correlations. It should be further noted, however, that it is mainly the time average that is affected by the choice of normalized covariance or correlation. Temporal variations in normalized covariance are actually quite comparable to temporal variations in correlation.

Anyway, our measure of uncertainty here for an individual spatially averaged forecast will be two standard deviations of the normalized covariance. That is,

$$V = 2[(\mathrm{COV} - [\mathrm{COV}])^2]^{1/2}.$$

This provides significance at the 95% level. It should be further noted that the standard deviation, U, for an ensemble of N independent forecasts is

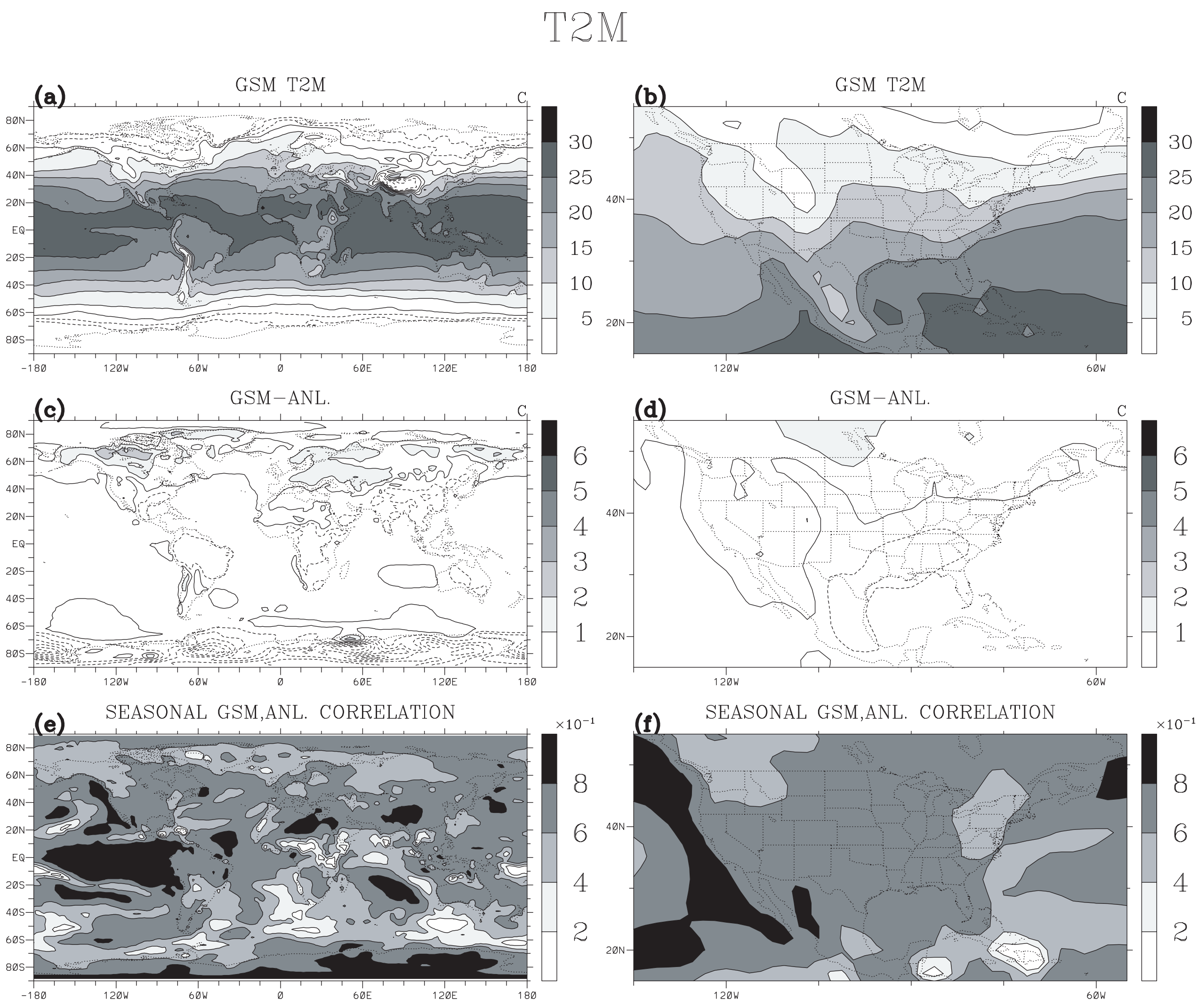

FIG. 1. The 2-m temperature seasonal predictions (Oct 1997–Oct 1999; 104 forecasts): (a) GSM seasonal mean, °C; (b) U.S. focus; (c) GSM–V1 analysis mean, °C; (d) U.S. focus; (e) global correlation (GSM, V1 analysis); and (f) U.S. focus.

$$U = V/N^{1/2}.$$

It should be noted that the amount of serial dependence depends upon forecast variable as well as average, lag, etc. We assume this serial independence mainly for simplicity, recognizing that we could therefore overestimate the significance. In summary, to evaluate the significance of the areally and temporally (ensemble) averaged forecasts, we assume that the skill is significantly different from zero if the correlation minus U, which is an estimate of two standard deviations of the ensemble mean, is greater than zero. As will be shown, almost all forecasts thus appear to be skillful, although some are obviously more skillful than others.

4. Geographic characteristics of seasonal forecasts

In this section we describe global characteristics of the seasonal mean (12-week average) forecast fields and their biases and correlations with respect to V1 and VX. Both global and focused U.S. views are shown. We examine six variables that we have regularly displayed on the WWW: 2-m temperature, 2-m relative humidity, 10-m wind speed, soil moisture (vertically integrated soil moisture in the upper 2 m of soil), precipitation, and a fire weather index. The basis for the fire weather index (FWI) was described by Roads et al. (1991, 1997). Suffice it to say that the FWI represents a simplified index for fire potential that has proven useful in several cases (including the recent, May 2000, Los Alamos fires), although even more compre-

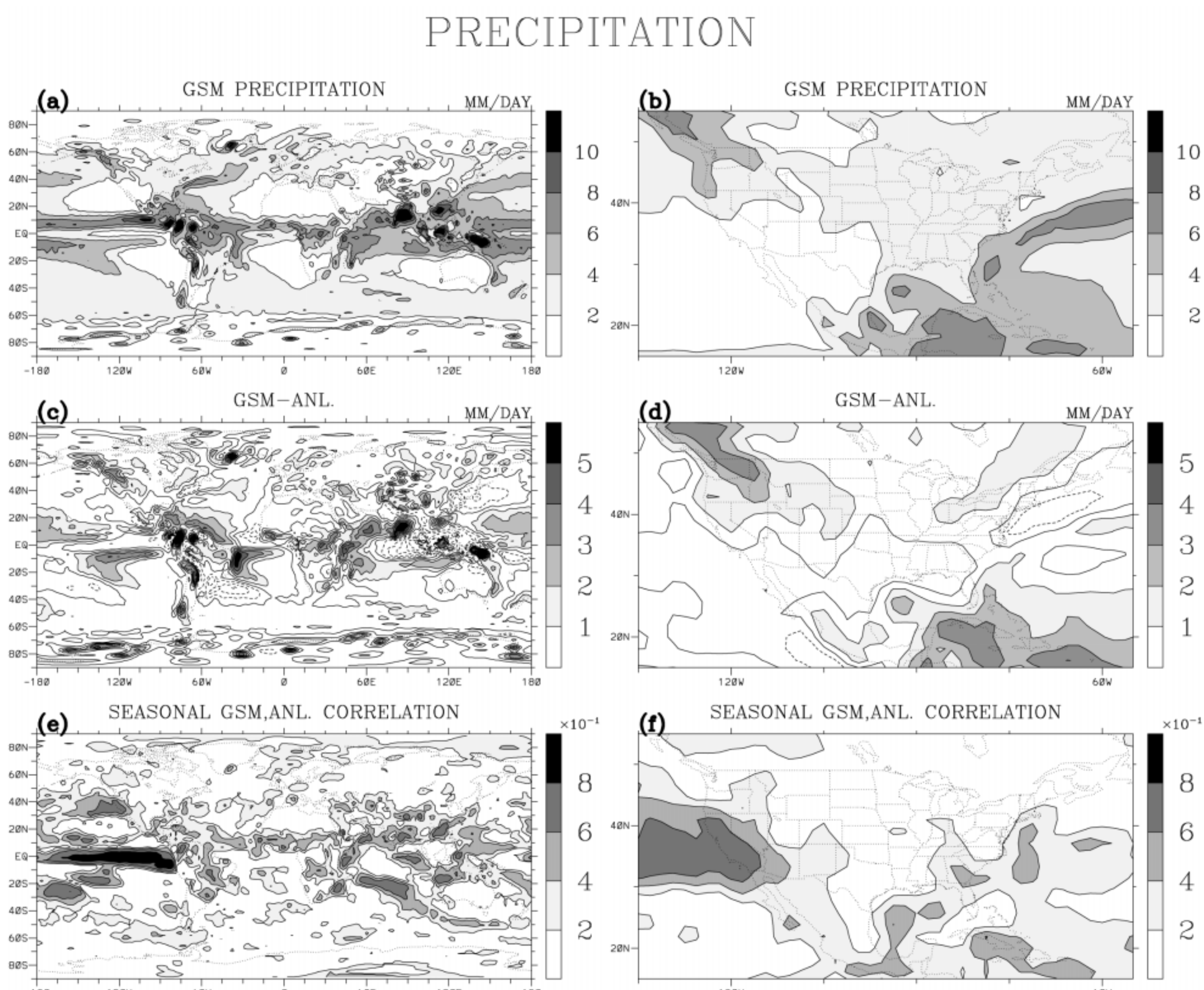

FIG. 2. Precipitation seasonal predictions (Oct 1997–Oct 1999; 104 forecasts): (a) GSM seasonal mean, mm day^{-1}; (b) U.S. focus; (c) GSM–VX analysis mean, mm day^{-1}; (d) U.S. focus; (e) global correlation (GSM, VX analysis); and (f) U.S. focus.

hensive measures, which include vegetation characteristics, are currently used by the U.S. Forest Service (USFS) to forecast daily fire danger.

As shown in Figs. 1a,b, GSM seasonal forecasts for the daily averaged 2-m surface temperature provide a typical 2-m-temperature field, high in the Tropics and decreasing toward high latitudes and elevations. The continents modify the relative zonal symmetry; coastal temperatures tend to be lower on the western boundaries and higher on the eastern boundaries. Figures 1c,d show there is a tendency for the forecasts to have a small cold bias over middle, low, and southern latitudes and a warm bias over high northern latitudes. Over the United States, the bias is small. The correlation of seasonal mean forecasts (Figs. 1e,f) is quite high just about everywhere. Globally, the highest skill is found in the eastern equatorial Pacific, Atlantic, and southern Indian Oceans, as well as a few land regions. The lowest skill appears over equatorial Africa and in the roaring 50°s of the Southern Hemisphere. Over the United States the forecasts are most skillful along the west coast (0.8) and lowest along the east coast (0.4). Especially low skill is found in the Caribbean over Haiti. In general the seasonal forecast skill of temperature is quite high and significant. Since the forecast skill of this field is related to persistent SSTAs, it would be of interest to eventually compare these forecasts with forecasts using observed and climatological SSTAs to better understand current and future (coupled ocean–atmosphere) temperature prediction capability.

GSM seasonal forecasts for the daily averaged precipitation have a characteristic global precipitation

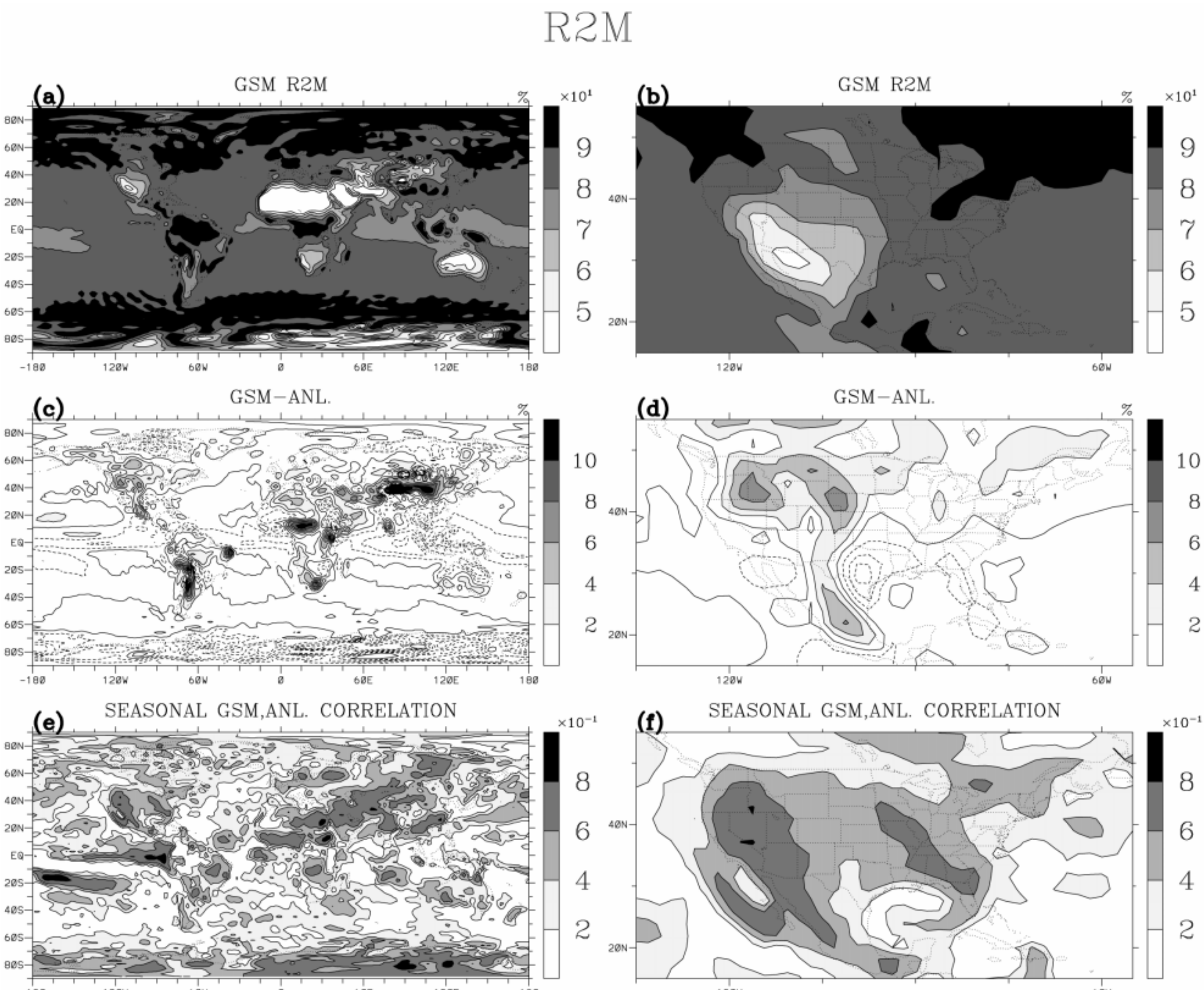

FIG. 3. Relative humidity seasonal predictions (Oct 1997–Oct 1999; 104 forecasts): (a) GSM seasonal mean, %; (b) U.S. focus; (c) GSM–V1 analysis mean, %; (d) U.S. focus; (e) global correlation (GSM, V1 analysis); and (f) U.S. focus.

field. Maximum precipitation occurs over the intertropical convergence zones as well as over continental coastal regions (Fig. 2a). Over the United States (Fig. 2b), the precipitation is especially strong over the northwest and southeast. This seasonal forecast field has considerable bias with respect to the validating analysis. Seasonal precipitation forecasts diminish in amplitude, especially over the western tropical oceans (Fig. 2c), while eastern tropical oceans have a wet bias, as do midlatitude regions. Over the United States, the difference is positive over the northwest and negative over the southwest (Fig. 2d). In high latitudes, the spectral noise of the reanalysis model (Roads et al. 1999) is noticeable. It should be noted again, however, that the bias is taken into account before evaluating the correlations. That is, as discussed previously, each forecast lead has an associated (monthly and forecast lead time) climatology removed before evaluating the correlations. The seasonal mean forecast correlations are statistically significant in several global regions—mainly those associated with the ENSO phenomenon in the Pacific and Indian Ocean (Fig. 2e). Over the United States, the correlations are especially strong along the west coast, extending across the southern tier of states, and on up the east coast (Fig. 2f). The least skillful seasonal forecasts occur over the central part of the country.

GSM seasonal forecasts of the daily averaged 2-m relative humidity field have characteristic high uniform values over the oceans (Fig. 3a), with the lowest values in the tropical oceans and the higher values in the polar oceans. Land regions, especially subtropical regions, have relatively low relative humidity, reflect-

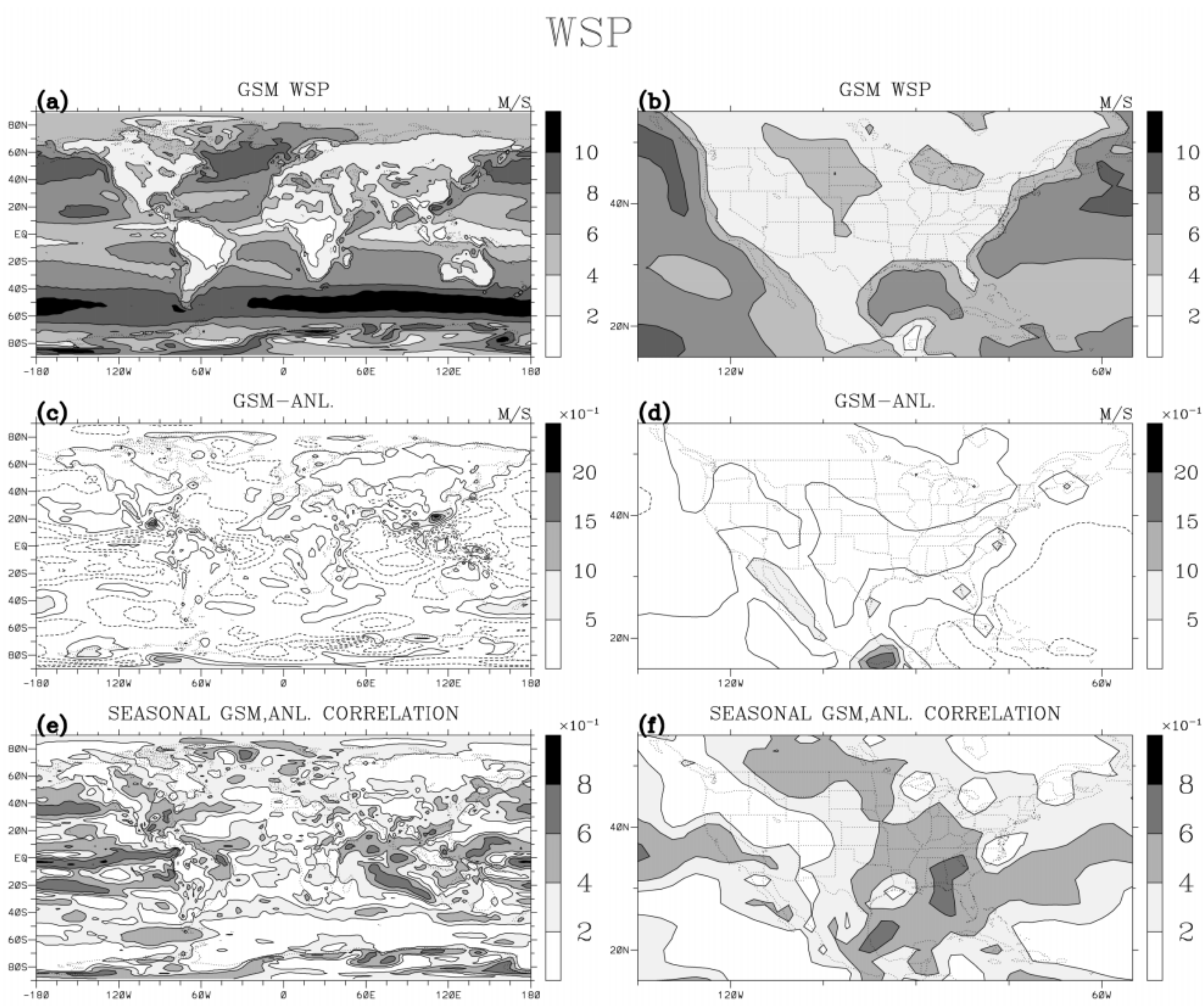

FIG. 4. Wind speed seasonal predictions (Oct 1997–Oct 1999; 104 forecasts): (a) GSM seasonal mean, m s^{-1}; (b) U.S. focus; (c) GSM–V1 analysis mean, m s^{-1}; (d) U.S. focus; (e) global correlation (GSM, V1 analysis); and (f) U.S. focus.

ing, presumably, the relatively low evaporation from the drier land surface. Over the United States (Fig. 3b) the relative humidity is lowest in the southwest. As shown in Figs. 3c,d, there is a tendency for the forecasts to have a positive bias over middle latitudes and a negative bias over tropical regions. Over the United States, the bias is positive to the north and negative toward the south, except for the tongue of high-elevation bias extending to Mexico. Again, the high-latitude spectral noise is noticeable. Again, this bias is removed before calculating the correlation. The correlation of seasonal mean forecasts (Figs. 3e,f) is high mostly over land regions; although over the United States the Rocky Mountains show a smaller correlation. Extremely low skill is found in the Caribbean, consistent with the low skill in temperature forecasts there. The lowest skill appears over ocean areas with the exception of the tropical oceans, which have high skill in the eastern tropical Pacific, presumably due to the strong SST signal. Note, however those correlations are low over the equatorial regions, which does not correspond to the high correlation found for precipitation forecasts in this region. Getting both the precipitation and relative humidity forecast accurately appears to be a critical problem with this model. It should also be noted that areas with large biases do not necessarily have low skill and areas with low biases do not necessarily have high skill. Whether variations are easy or difficult to predict is not necessarily dependent upon getting the climatology right.

The GSM seasonal forecast wind speed is a maximum over the oceans (Fig. 4a), especially in the midlatitude westerlies. It is weaker in the subtropics (horse latitudes) and near the equator (doldrums). Over the

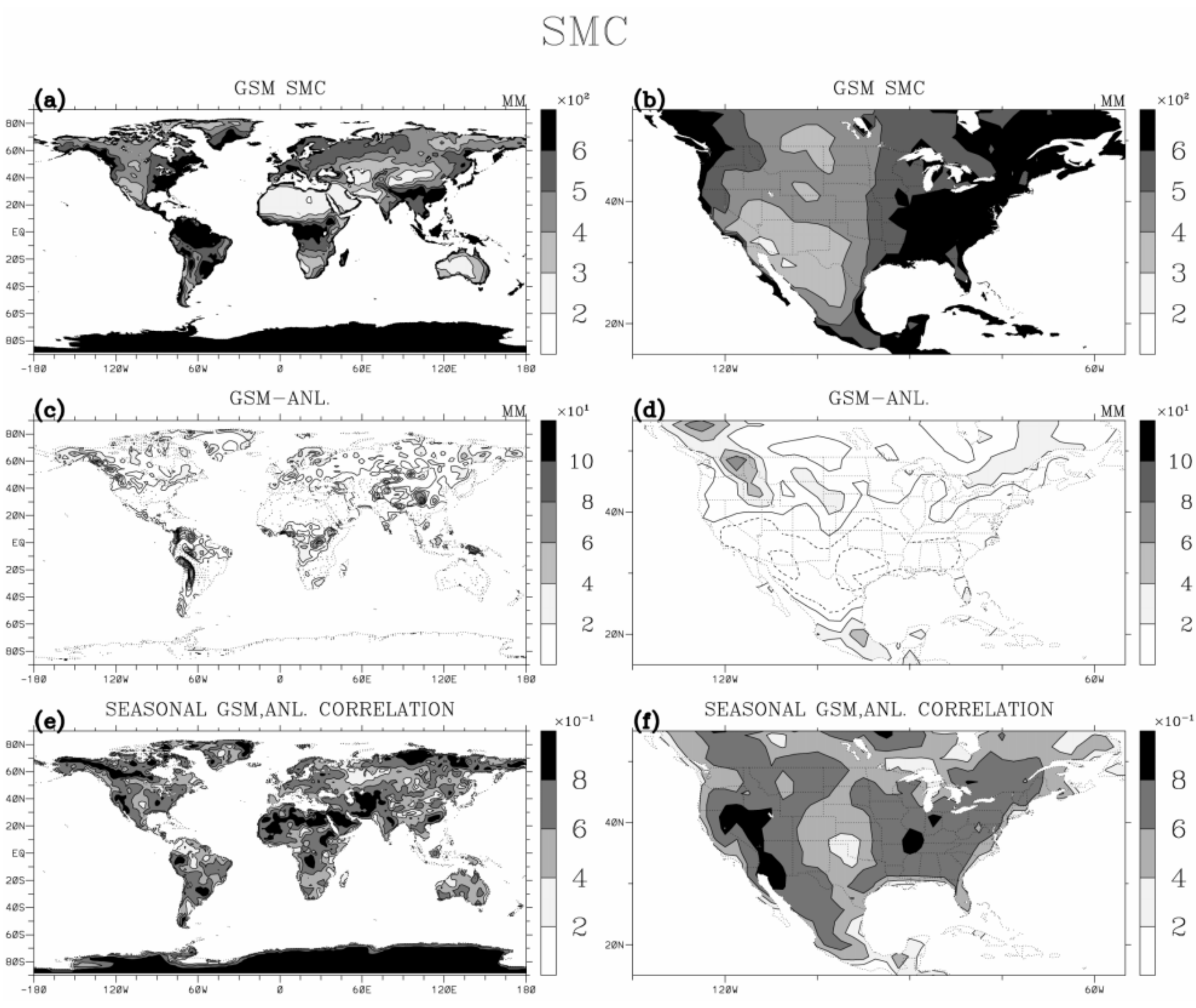

FIG. 5. Soil moisture seasonal predictions (Oct 1997–Oct 1999; 104 forecasts): (a) GSM seasonal mean, mm; (b) U.S. focus; (c) GSM–V1 analysis mean, mm; (d) U.S. focus; (e) global correlation (GSM, V1 analysis); and (f) U.S. focus.

United States (Fig. 4b), there is a relative maximum in the north-central part of the country. As shown in Figs. 4c,d there is a tendency for the forecasts to have relatively weak winds just about everywhere, especially over the tropical oceans. It is critical to take this relative weakness into account when using this model to drive ocean models [Auad et al. 2000, mauscript submitted to *J. Geophys. Res. (Oceans)*]. Fortunately, the correlation of seasonal mean forecasts (Figs. 4e,f) is quite high in limited regions. Over the United States, the correlation is strong over the eastern part of the country and along the central California coast. In the West, the wind speed correlations are low, especially in the Northwest and farther down, at the tip of Baja. This indicates that removing the wind speed bias will be critical before using this model to force ocean and fire danger models.

The GSM seasonal forecast soil moisture (Figs. 5a,b) is similar to the relative humidity field over land in that the subtropical deserts, especially the Sahara Desert, have relatively low values of soil moisture. Higher values occur in high latitudes and on the eastern portions of the continents. Over the United States, the soil moisture is high in the northwest and east and relatively dry over much of the west. Figures 5c,d show the biases of seasonal forecasts with respect to the validating analysis, and basically there is a tendency for the forecasts to have a low bias over middle latitudes and a positive bias over high northern latitudes and tropical regions. Globally and over the United States, the bias is noticeable and is related to the biases in the forecast precipitation field. For example, the previously mentioned high-latitude forecast spectral noise shows up especially strongly in the soil moisture bi-

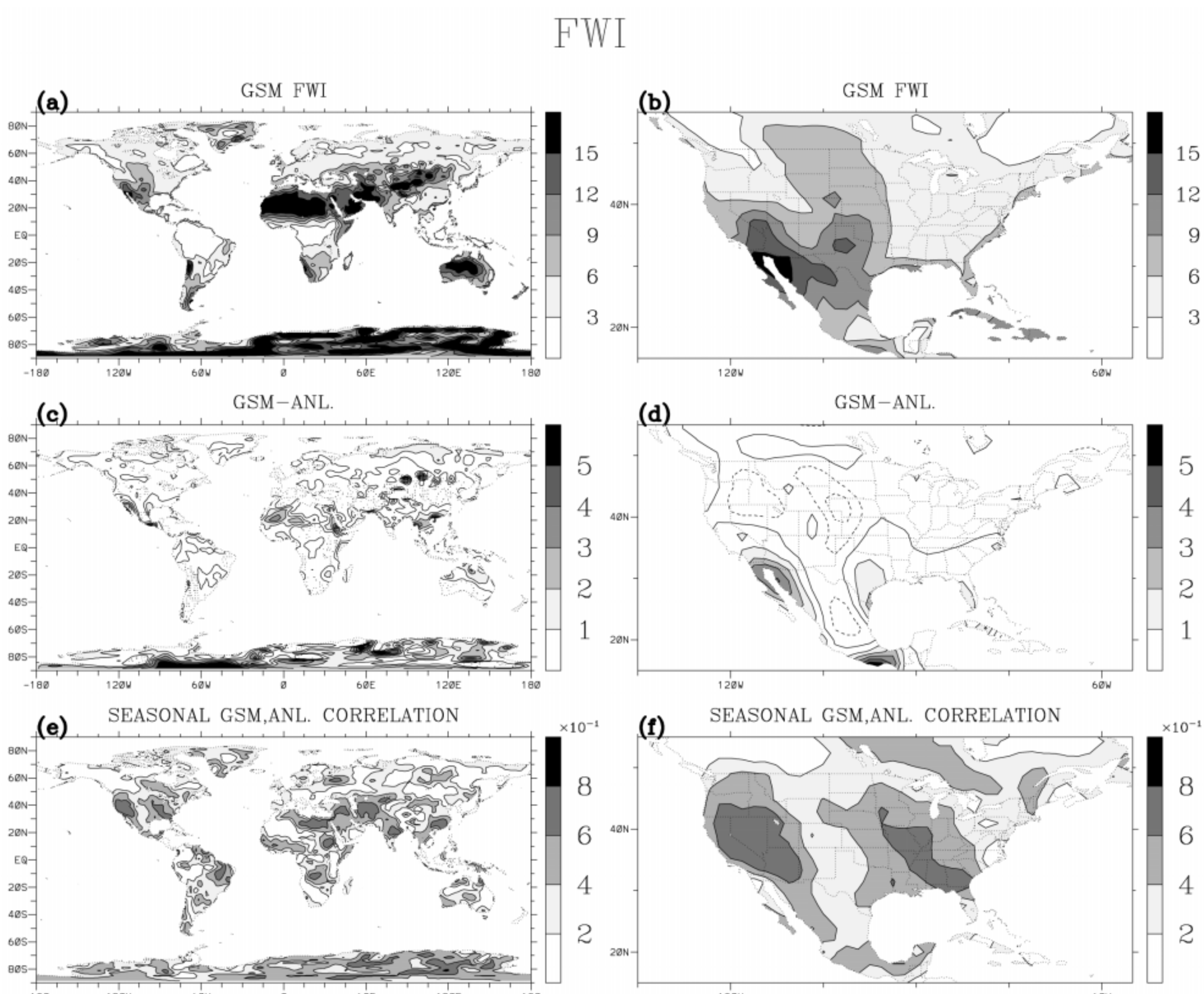

FIG. 6. FWI seasonal predictions (Oct 1997–Oct 1999; 104 forecasts): (a) GSM seasonal mean, (b) U.S. focus, (c) GSM–V1 analysis mean, (d) U.S. focus, (e) global correlation (GSM, V1 analysis), and (f) U.S. focus.

ases. The seasonal forecasts show considerable skill, comparable to the skill previously shown for the temperature field. Globally the correlations are fairly high just about everywhere (Fig. 5e). Over the United States, there are relatively low correlations along the Front Range for the relative humidity (Fig. 5f). Everywhere else the correlations are especially strong, ranging from 0.6 to 0.8. As shall be shown later, all of this forecast skill is presumably related to the low-frequency variability of this field. In that regard it appears to be more important to obtain an accurate soil moisture analysis for initializing the model than it is to make the actual forecast.

The GSM seasonal forecast FWI, which basically reflects wind speed and relative humidity, is the inverse of the relative humidity and soil moisture and is relatively high in those regions of low soil moisture and relative humidity (Figs. 6a,b). Although variations in the FWI are also reflected by the wind speed, it does not include vegetation stress, which must somehow be related to soil moisture, and which is better incorporated in standard fire danger indices (see, e.g., Roads et al. 2000a). There is a tendency for the forecasts to have a negative bias (Figs. 6c,d), which can be traced to the tendency for the model to have relatively high relative humidity over the land regions and a negative wind bias. Still, seasonal forecast correlations (Figs. 6e,f) are high over much of the United States, except for the Front Range of the Rocky Mountains. Globally the highest forecast correlations are found over most land regions with the major exceptions being the northwestern United States, Africa, and South America regions. The correlation pattern resembles more the relative humidity correlation pattern (instead of the wind speed correlation pattern), indicating that it is the relatively accurate forecasts of relative humidity, more than wind speed, that provide some skill for the forecast FWI at long (seasonal) timescales.

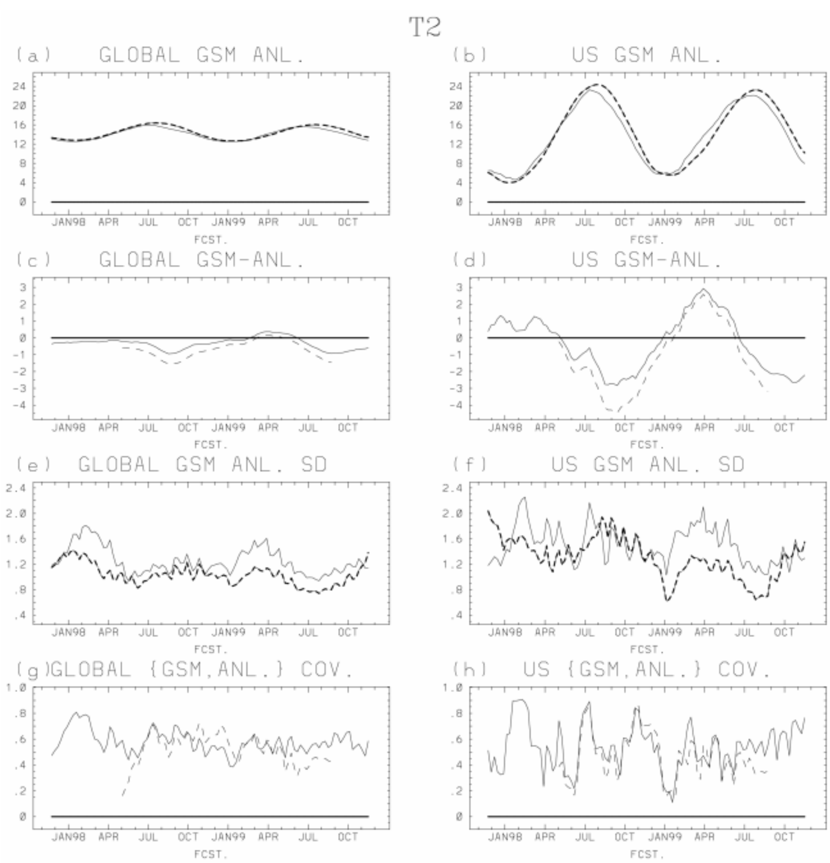

FIG. 7. The 2-m temperature seasonal forecast temporal variations (Oct 1997–Oct 1999; 104 forecasts; smoothed by five-forecast running mean): (a) global, GSM (solid), V1 (dashed), °C; (b) United States; (c) global GSM–V1 (solid), GSM–VO (dashed), °C; (d) United States; (e) global GSM (solid), V1 (dashed) rms, °C; (f) United States; (g) global (GSM, V1) (solid), GSM, VO (dashed), normalized seasonal covariance; and (h) United States.

5. Temporal characteristics of seasonal forecasts

Let us now consider the temporal characteristics of forecast averages for global (only land regions with latitudes less than 60° are included for soil moisture and the FWI) and U.S. land regions. All seasonal temporal variations (Figs. 7–12) are plotted with respect to the validating time. In the case of seasonal forecasts, we choose the initial time (which is how all the forecasts are tracked) plus 6 weeks for this validation time. Again, our main validating analyses is the 1-day (V1) forecasts and the Xie and Arkin precipitation (VX), although we also compare forecasts to the VO analysis. It should be noted that V1, VO, and VX analyses have a distinct correlation with each other that is always higher than any forecast validation correlation (not shown). That being said, there is a smaller correlation between the forecasts and the VO or VX analyses than between the forecasts and V1. This is not unexpected even though it should be stressed that evaluations using VO occur over a more limited period.

Temporal characteristics of global and U.S. (land only) temperature variations with respect to two sets of validating analyses, V1 and VO, are shown in Fig. 7. Note that the seasonal forecast temperature fields are quite consistent with the validating analysis (Figs. 7a,b). In comparison to the strong seasonal variations differences between the validating analysis and forecasts cannot be easily seen. The differences only show up in the biases (Figs. 7c,d), which show a tendency for the seasonal forecasts, especially over the United States, to have a cold bias during the spring–summer and a warm bias during the fall–winter. Similar features occur for the global means. Note that the model has the coldest bias with respect to the GDAS operational analysis. Figures 7e,f show that the seasonal forecast variances are slightly larger than the analysis (V1) variations, but still faithfully represent the intraseasonal variations. Figures 7g,h show the temporal variations in the global and U.S. normalized covariance. Note that the normalized covariance calculated from the validating analysis datasets (V1) and the GSM seasonal forecasts are similar. Also note that skill has fairly strong temporal variations but does not demonstrate, especially over the United States, any significant seasonal variation. The reasons for the strong intraseasonal variations in skill are currently unknown.

Temporal characteristics of global and U.S. (land only) precipitation variations with respect to two sets of validating analyses (VX and VO) are shown in Fig. 8. It should be noted that the seasonal forecast precipitation fields show strong biases with the validating analysis, which in turn are somewhat different (Figs. 8a,b). As shown in Figs. 8c,d, the model is biased high with respect to the operational analysis (VO) but not with respect to the VX analysis. Clearly even operational analyses can have large biases and one has to be careful to not get too concerned with absolute values. Despite this strong bias, the variations in the seasonal forecast are comparable to the analysis variations (Figs. 8e,f). Figures 8g,h show that skill has a seasonal pattern, with the lowest U.S. skill occurring during the summer to fall seasons. These seasonal forecasts have small but significant skill in many places.

Temporal characteristics of global and U.S. (land only) relative humidity seasonal forecast variations are shown in Fig. 9. Seasonally, the highest relative humidity occurs in the winter, especially over the United States (Figs. 9a,b). Note also the strong interannual signal superimposed upon the seasonal U.S. cycle. The relative humidity was much higher during the beginning of the period when the El Niño was affecting the United States and then became drier as the La Niña began. Generally, the seasonal forecasts show little bias (Figs. 9c,d) with respect to the 1-day forecast V1 analysis but a large positive bias with respect to the operational VO analyses. However, the same general

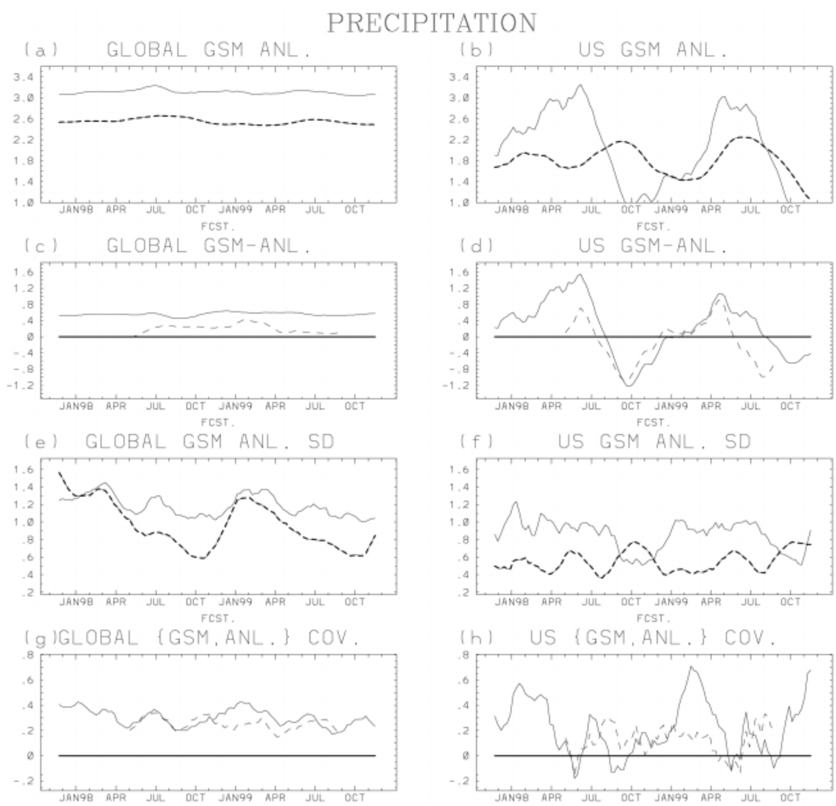

FIG. 8. Precipitation seasonal forecast temporal variations (Oct 1997–Oct 1999; 104 forecasts; smoothed by five-forecast running mean): (a) global, GSM (solid), VX (dashed), mm day^{-1}; (b) United States; (c) global GSM–VX (solid), GSM–VO (dashed), mm day^{-1}; (d) United States; (e) global GSM (solid), VX (dashed); rms, mm day^{-1}; (f) United States; and (g) global (GSM, VX) (solid), GSM, VO (dashed), normalized seasonal covariance; and (h) United States.

characteristics of relatively low relative humidity during the winter and relatively high relative humidity during the summer still seem to prevail over the United States, indicating a reduced seasonal cycle in the seasonal forecast. The forecast variations (Figs. 9e,f) show that the forecasts have slightly larger variations. Seasonal forecast normalized covariance (Figs. 9g,h) is relatively higher with respect to the 1-day forecast analysis, especially over the United States. The weakest skill occurs during the summertime. Except for some strong intraseasonal variations, the skill (normalized covariance) of these forecasts is moderately high and significant. Note the strong disagreement between the forecast skill using the operational and 1-day forecasts, especially in early 1999. In that regard we note that the early 1999 period relative humidity forecast skill with respect to the V1 analysis is consistent with the high forecast skill of the precipitation with respect to the VX analysis. Further investigation is needed to determine why these two analyses (V1 and VO) indicate such different skill.

Temporal characteristics of global and U.S. (land only) wind speed variations are shown in Fig. 10. Globally, the wind speed is fairly constant, but over the United States, the wind speed is much stronger during the spring (Figs. 10a,b). The bias in the forecast wind speed is fairly noticeable with the forecast wind speed having a 10% low bias in wind speed. The bias is a bit larger if the operational analysis is used. Over the United States, the wind speed was actually somewhat larger than the V1 analysis during the early part of the period and lower during the latter part of the period. If the operational analysis is used, the bias is almost always negative, although still most strongly negative during the latter part of the period (Figs. 10c,d). By contrast, the forecast variations are much larger everywhere by a noticeable amount (Figs. 10e,f). Despite the various biases, the model at least provides a significant forecast at seasonal timescales. The seasonal normalized covariance is on the order of 0.3 for a seasonal forecast (Figs. 10g,h). The normalized covariance is only slightly lower, with respect to the operational analysis. Seasonal variations are relatively small, except that the transition seasons appear to be most difficult to accurately forecast.

Temporal characteristics of global (land only and latitudes less than 60°) and U.S. (land only) soil moisture variations with respect to two sets of validating analysis are shown in Fig. 11. Like the relative humidity (and to a lesser extent the precipitation) the characteristic soil moisture is wet during the first part of the period and then dries considerably during the latter part. These variations reflect the ENSO cycle variations, especially over the United States (Figs. 11a,b). It should also be noted that the seasonal forecast soil moisture fields (Figs. 11c,d) show weak dry biases with respect to the validating analyses, indicating a tendency for the model to be mostly dry, especially over the United States during

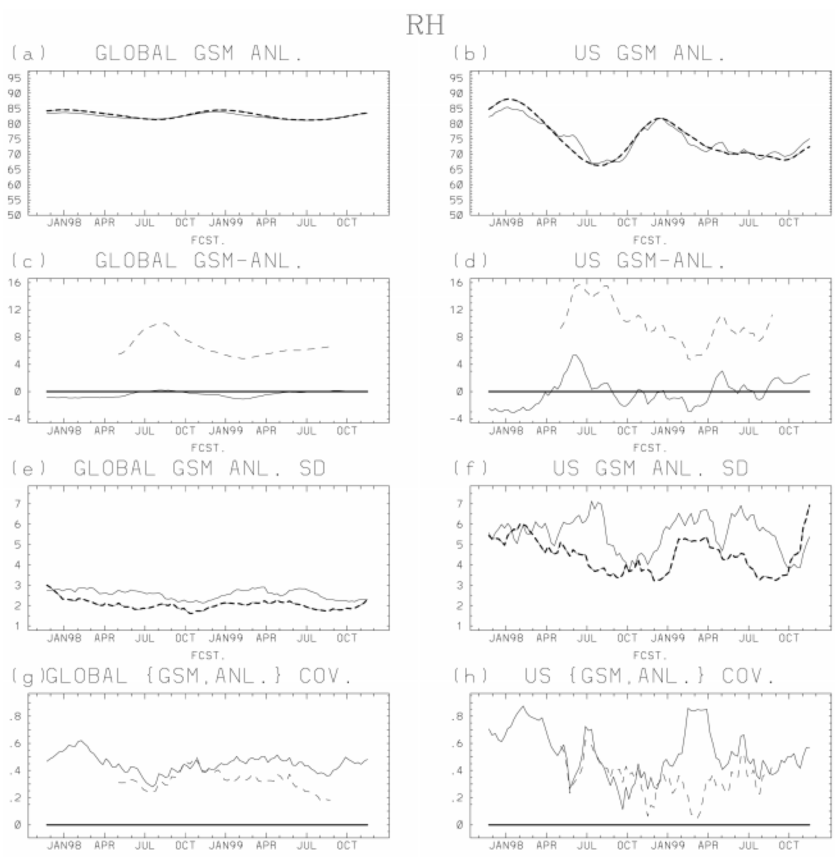

FIG. 9. Relative humidity seasonal forecast temporal variations (Oct 1997–Oct 1999; 104 forecasts; smoothed by five-forecast running mean): (a) global, GSM (solid), V1 (dashed), %; (b) United States; (c) global GSM–V1 (solid), GSM–VO (dashed), %; (d) United States; (e) global GSM (solid), V1 (dashed) rms, %; (f) United States; (g) global (GSM, V1) (solid), GSM, VO (dashed), normalized seasonal covariance; and (h) United States.

the summer. Figures 11e,f show that the analysis and forecast GSM variations are comparable albeit somewhat higher in the global and U.S. forecasts, like all the other forecast fields. Figures 11g,h show that skill, as measured by normalized covariance, has a seasonal pattern, with the highest skill occurring during the wintertime, especially over the United States. With respect to the operational analysis, the seasonal variation is less noticeable and the skill is noticeably smaller, but still significant.

Figure 12 shows temporal characteristics of global and U.S. (land only) FWI variations. Like the soil moisture, there is a distinct interannual variation with lower FWI during the first part of the period and higher FWI during the latter part of the period (Figs. 12a,b). This variation is notable, despite there being a substantial bias in the FWI, especially with regard to the operational analysis (Figs. 12c,d). This bias is due mainly to the substantial bias in the relative humidity although the model's weaker wind speed also contributes. By contrast, the forecast model standard deviations are substantially stronger than the analysis standard deviations especially over the United States during the springtime (Figs. 12e,f). The normalized covariance is fairly significant and shows little seasonal variation (Figs. 12g,h) but strong intraseasonal variation, especially over the United States and especially when using the 1-day forecasts to validate the seasonal forecasts.

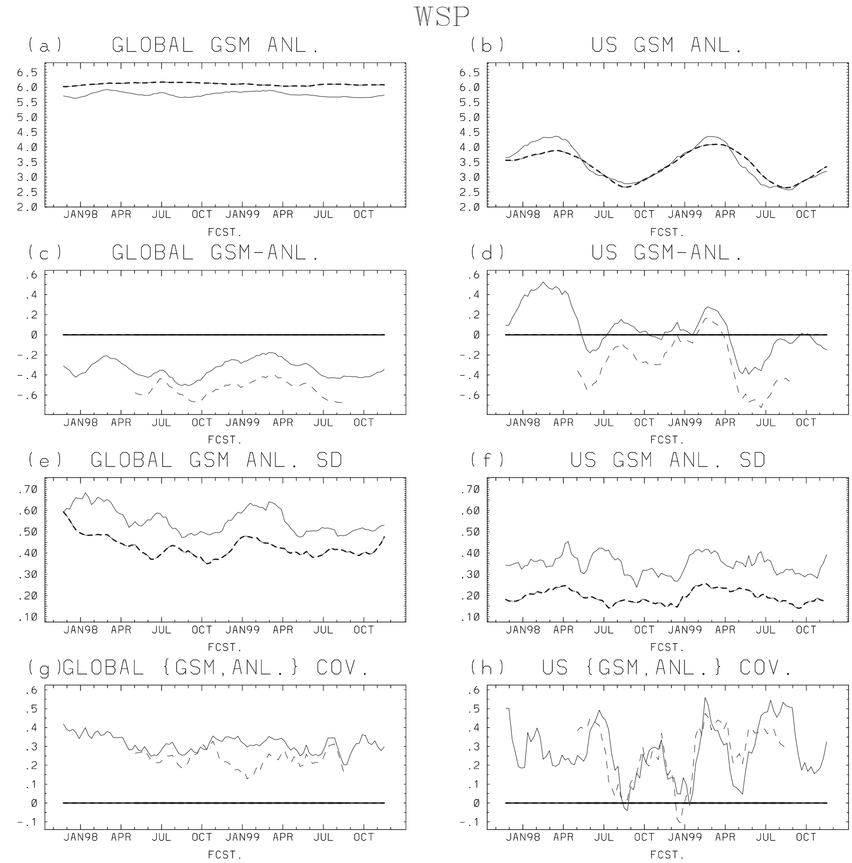

FIG. 10. Wind speed seasonal forecast temporal variations (Oct 1997–Oct 1999; 104 forecasts; smoothed by five-forecast running mean): (a) global, GSM (solid), V1 (dashed), m s^{-1}; (b) United States; (c) global GSM–V1 (solid), GSM–VO (dashed), m s^{-1}; (d) United States; (e) global GSM (solid), V1 (dashed) rms, m s^{-1}; (f) United States; (g) global (GSM, V1) (solid), GSM, VO (dashed), normalized seasonal covariance; and (h) United States.

6. Ensemble forecast skill and systematic bias of weekly to seasonal forecasts

Figure 13 shows the ensemble (104 forecasts) U.S. average skill, as measured by spatially averaged correlations for weekly GSM forecasts (thin solid lines) versus weekly persistence forecasts V1 (thin dashed line):

$$[\mathrm{COR}] = \frac{[\{F'O'\}] - [\{F'\}][\{O'\}]}{\left([\{F'^2\}] - [\{F'\}]^2\right)^{1/2}\left([\{O'^2\}] - [\{O'\}]^2\right)^{1/2}}.$$

Error bars ($\pm U$) indicate an estimate for two standard deviations in the ensemble forecast skill. These estimates are made by calculating the standard deviations of the normalized covariance of individual forecasts (the normalized seasonal covariance was shown previously in Figs. 7–12). Also shown are monthly forecast (O) and seasonal forecast skill (*) along with their respective monthly and seasonal ensemble error bars.

There is significant skill at weekly to seasonal timescales that is distinctly larger for the forecast model than persistence, for all variables except soil moisture (Fig. 13e). Here the forecast model has less skill than the initial state of the analysis, which is not too dissimilar to what has been found from efforts to forecast midlatitude sea surface temperature (see, e.g.,

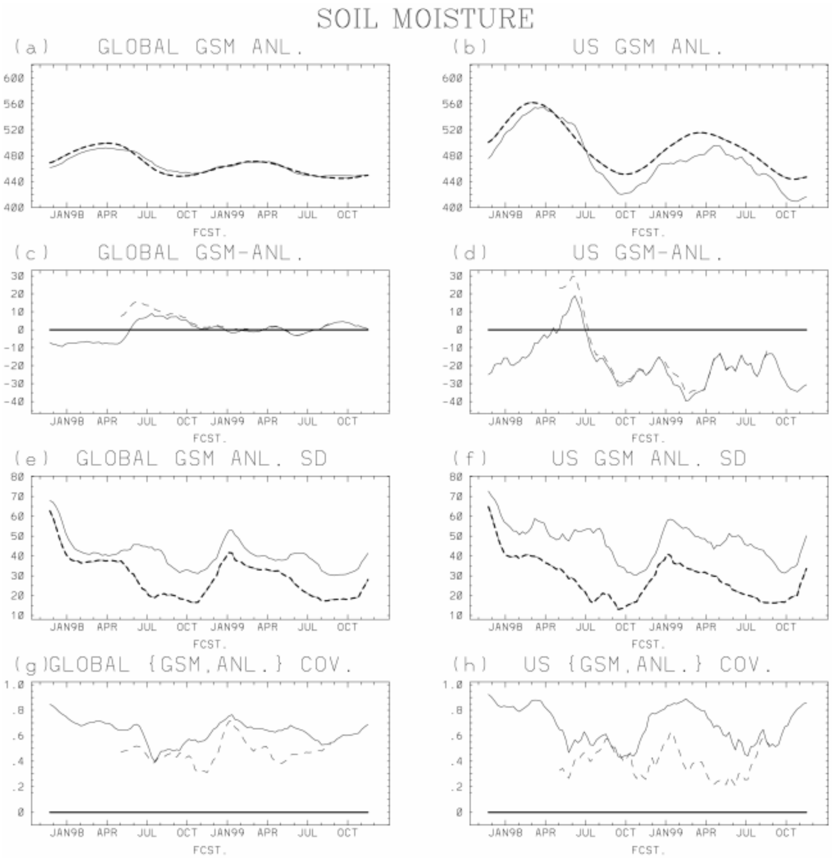

FIG. 11. Soil moisture seasonal forecast temporal variations (Oct 1997–Oct 1999; 104 forecasts; smoothed by five-forecast running mean): (a) global, GSM (solid), V1 (dashed), mm; (b) United States; (c) global GSM–V1 (solid), GSM–VO (dashed), mm; (d) United States; (e) global GSM (solid), V1 (dashed) rms, mm; (f) United States; (g) global (GSM V1) solid), GSM, VO (dashed), normalized seasonal covariance; and (h) United States.

Miller and Roads 1990). For other variables, the skill is significant out to 12 weeks, especially if 4-week averages (monthly means) or 12-week averages (seasonal means) are used. Presumably this time-average skill arises because of the persistent boundary conditions. The skill is especially high for the temperature (Fig. 13b) and soil moisture (Fig. 13e). Precipitation (Fig. 13a) appears to be the most difficult variable to forecast although wind speed (Fig. 13d) is also difficult. Relative humidity forecasts (Fig. 13c) show relatively high skill, although given the strong influence of temperature on relative humidity, this should perhaps not be too surprising. Fortunately the FWI (Fig. 13f) depends upon relative humidity as well as wind speed. For all variables, forecast skill over the United States is slightly greater than the skill over the corresponding global (or global land) domain (not shown). There are likely various reasons for this increased U.S. skill, including intrinsic skill. Because of the larger domain, and the greater number of independent grid points, variations in individual global forecasts are much smaller and the significance of the skill is therefore much larger, despite the lower overall skill.

The model soil moisture provides similar but still somewhat lower skill than the persisted soil moisture anomaly. These attempts to predict soil moisture are similar to previous efforts to predict SST from coupled models. Coupled models have so far only succeeded in making better SST predictions in the Tropics. Globally, SST and soil moisture are still predicted better by persistence, although, again, it is this persistence that presumably helps to provide the significant forecast skill demonstrated in other variables. In that regard, it should be noted that the U.S. Forest Service fire danger index (not evaluated in this paper) includes a number of vegetation characteristics that presumably act like soil moisture. Thus fire danger has potential long-range forecast capability that may be achieved in part by the slowly varying initial state as well as the faster varying and more difficult to predict near-surface wind speed and relative humidity.

Except for soil moisture, all variables show the positive influence of time averaging, indicating that forecasts of time averages, even for lags of 2 months, still show significant skill. In the previous experiments discussed by Roads (1989) time averaging did not have a strong impact, presumably because the forecast time (up to a month) was too short to be influenced by slowly varying boundary conditions. Now that we have the capability to make longer forecasts, time averaging does seem to help. Note that seasonal forecast skill at zero lag is also almost as good as forecasts of 1 month at zero lag, indicating the usefulness of dynamical models (in a time-averaged sense) at this time range.

Figure 14 shows the average U.S. biases that develop during the forecasts. As was previously shown, there is a significant amount of precipitation spindown (Fig. 14a). Temperature (Fig. 14b) has a cold bias although not as cold a bias if the model was being compared against the VO analysis (see Fig. 7). Relative humidity has a small positive bias during the first part of the forecast period, which then changes to a negative bias for the longest lags (Fig. 14c). Since V1 RH is so much higher than VO RH, however, the RH would have a strong positive bias if compared to VO. Wind speed shows an initial small positive and then a negative bias, decreasing to a minimum during the third week and then increasing slightly (Fig. 14d). Soil moisture has a positive bias initially and then begins a strong decrease that is still not finished at the end of the 12th week (Fig. 14e). Because of the initial relative humidity bias the FWI is too weak initially but then becomes positive with increasing forecast times (Fig. 14f). This change mimics the RH variations although clearly the wind speed bias is also important. Similar biases occur over the global domain, with perhaps the most noticeable difference being that the biases in the precipitation and wind speed are set up more quickly and the global wind speed is, on the average, less than the analysis wind speed for all forecast times. In short, it is important to recognize the inherent biases in the forecasts, for each geographic region (not shown), for each season (not shown), and for each time lag, but to still recognize that despite these biases, skillful seasonal predictions of the anomalous variations can still be made, so long as proper accounting for the biases are made.

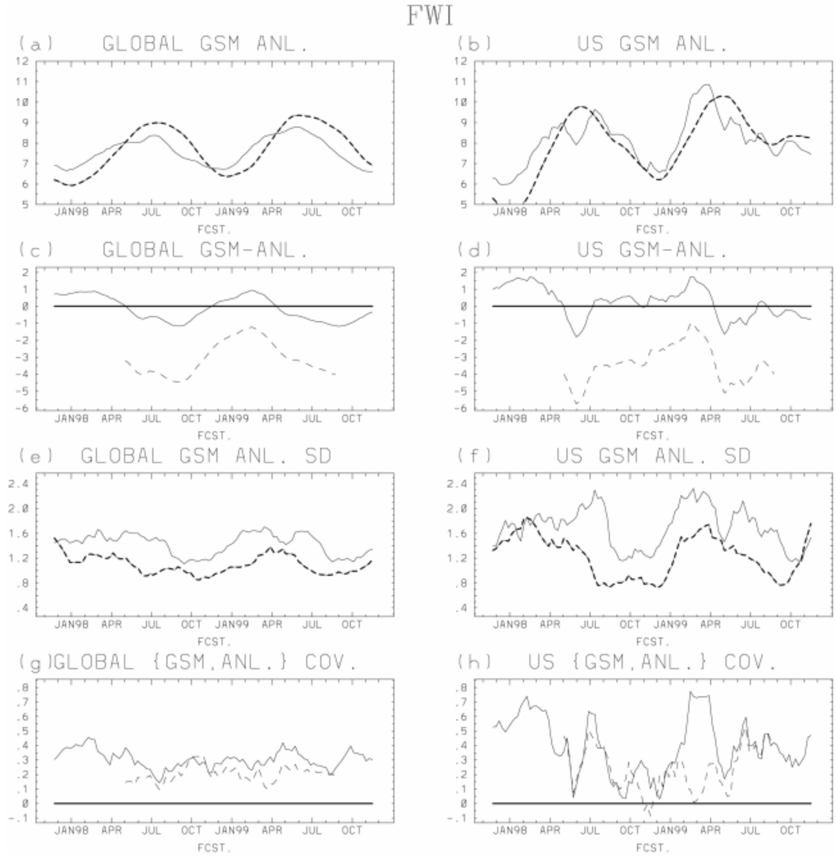

FIG. 12. FWI seasonal forecast temporal variations (Oct 1997–Oct 1999; 104 forecasts; smoothed by five-forecast running mean): (a) global, GSM (solid), V1 (dashed); (b) United States; (c) global GSM–V1 (solid), GSM–VO (dashed); (d) United States; (e) global GSM (solid), V1 (dashed) rms; (f) United States; (g) global (GSM, V1) (solid), GSM, VO (dashed), normalized seasonal covariance; and (h) United States.

7. Summary

The Scripps Experimental Climate Prediction Center has been making experimental, near-real-time global forecasts since 27 September 1997. The global forecast model is based upon the NCEP GSM (Kalnay et al. 1996). The initial conditions and SST boundary conditions come from the NCEP operational global analysis at 0000 UTC. Seven-day forecasts are made everyday, and every weekend these forecasts are extended to 12 weeks. Although the SST climatology varies throughout the forecast, the initial SSTA is artificially persisted throughout the forecast. Over 104 12-week forecasts have now been made (a 12-week forecast is made once a week; on 26 September 2000 we will have 156 forecasts) and the purpose of this paper was to describe the forecast system as well as the various biases and errors. In brief, many relevant near-surface meteorological parameters (including temperature, precipitation, soil moisture, relative humidity, wind speed, and a fire weather index, which is a nonlinear combination of weather variables) are skillful at weekly to seasonal timescales over much of the United States and in many global regions.

Surface temperature forecasts are the most skillful, with ensemble seasonal forecast correlations of 0.7 for the United States and 0.62 for the globe. Precipitation has much lower forecast skill, 0.3 over the United States and 0.24 globally. Relative humidity is a bit more skillful at seasonal timescales with correlations over the United States of 0.5 and 0.3 globally. Wind speed forecasts are more problematic with seasonal forecast skill of 0.3 over the United States and 0.27 globally. FWI, which is a nonlinear combination of wind speed and relative humidity, has higher forecast skill, which presumably arises from contribution of the relative humidity as well as wind speed to the FWI. Finally, soil moisture forecasts are skillful but show no skill beyond what is available from simply persisting the initial state. Nonetheless, the strong persistence of the soil moisture is presumably one of the controlling features on the ability of the model to make skillful seasonal forecasts of temperature and other variables. It should in fact be noted that for all other variables the model forecast skill is generally greater than persistence. It should also be noted that forecast skill almost always increases with averaging length. This is due in part to the inclusion of skillful initial forecasts; however, even monthly forecasts with 2-month lags are more skillful than corresponding weekly forecasts with similar lags, indicating the positive influence of time averaging here.

Even better weekly to seasonal forecasts can probably be made. Now that we have examined the forecast skill of this model, we intend to upgrade it to the latest operational model and repeat the comparisons. A number of noticeable improvements have been made since this model was frozen for the first NCEP reanalysis and many of these improvements were incorporated in the NCEP reanalysis II model (see Kanamitsu et al. 2000). Many regional extensions to ECPC's experimental global forecast system are also under development. In particular, the

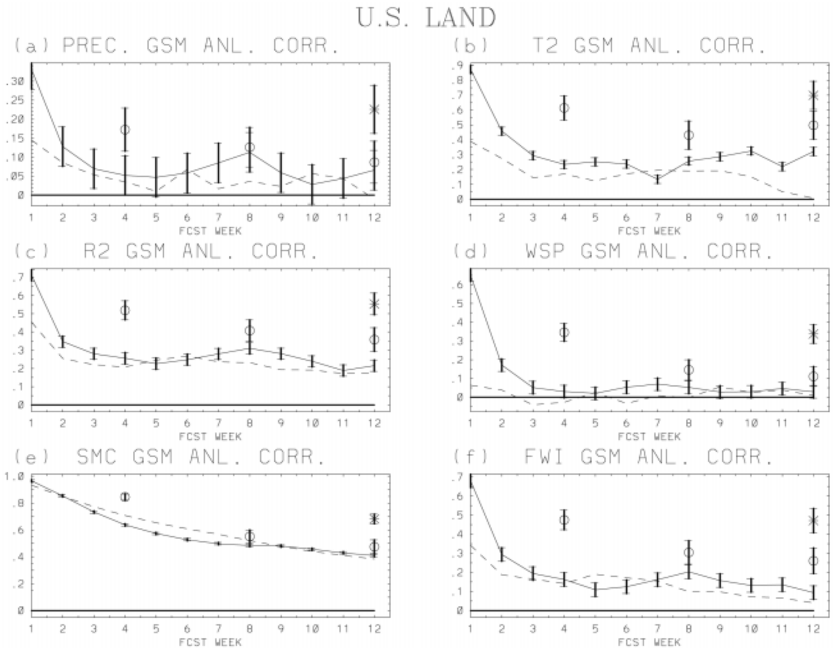

FIG. 13. The U.S. ensemble correlations (GSM, V1) for weekly GSM forecasts (solid) vs persistence forecasts (dashed). Correlations (GSM, V1) for monthly means ( ) and seasonal means (*). Estimates of the ensemble forecast variations (discussed in the text) are shown by the error bars for the weekly, monthly, and seasonal (GSM, V1) means: (a) precipitation, (b) 2-m temperature, (c) relative humidity, (d) wind speed, (e) soil moisture, and (f) FWI.

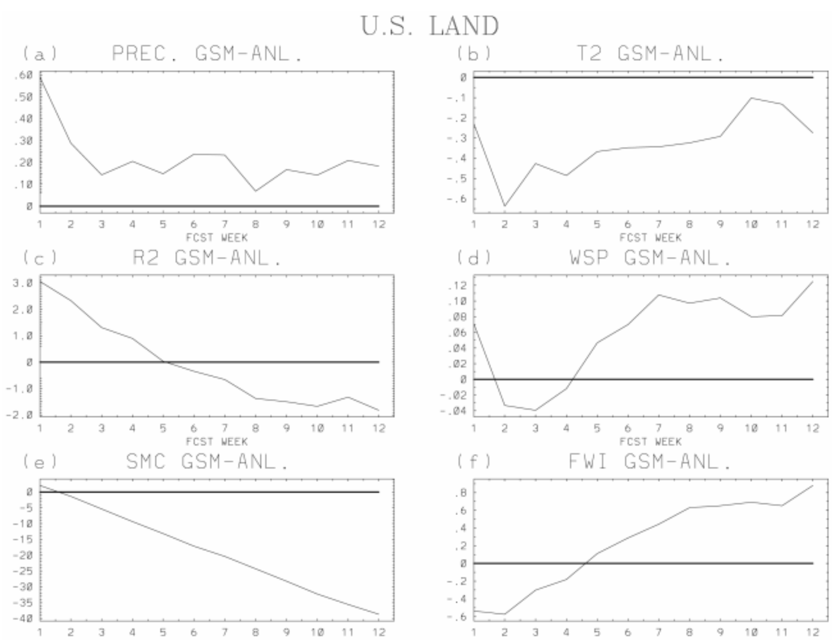

FIG. 14. The U.S. ensemble forecast biases (GSM–V1) for weekly GSM forecasts: (a) precipitation, mm day^{-1}; (b) 2-m temperature, K; (c) relative humidity, %; (d) wind speed, m s^{-1}; (e) soil moisture, mm; and (f) FWI.

GSM currently forces a regional spectral model (RSM; Juang and Kanamitsu 1994; Juang et al. 1997) in order to gain increased spatial resolution (50–25-km resolution) for shorter timescales (four times daily, 4 weeks) for several limited regions (see also Chen et al. 1999; Roads et al. 2000b). The GSM output will also eventually be used to force regional models at other application centers (see, e.g., Stevens et al. 2000). The GSM also forces a single-column model (S. Iacobellis and R. Somerville 2000, personal communication), as well as an ocean model [Auad et al. 2000, manuscript submitted to *J. Geophys. Res. (Oceans)*], which will eventually be coupled to the atmospheric model. Limited output products from the GSM and RSM are also being used for a wide variety of experimental applications (Roads et al. 2000a). These regional extensions and applications will be discussed in future papers.

Acknowledgments. This research was funded by a cooperative agreement from NOAA (NA77RJ0453) and U.S. Forest Service Cooperative Agreement USDA PSW-99-0017CA and NASA NAG54775. The views expressed herein are those of the authors and do not necessarily reflect the views of NOAA or the U.S. Forest Service. Jack Ritchie has performed the Herculean task of maintaining this quasi-operational system since 1998. We are also grateful to M. Kanamitsu and H. Juang of NCEP for providing us the model as well as many useful answers to many questions. We thank Mr. Thomas Reichler for his careful review of the manuscript. We also thank the anonymous reviewers for their careful reviews.

References

Anderson, B. T., J. Roads, S. Chen, and H.-M. Juang, 2000: Regional modeling of the low-level monsoon winds over the Gulf of California and southwest United States: Simulation and validation, *J. Geophys. Res. (Atmos.),* **105**(D14), 17 955–17 969.

Anderson, J., H. Van den Dool, A. Barnston, W. Chen, W. Stern, and J. Ploshay, 1999: Present-day capabilities of numerical and statistical models for atmospheric extratropical seasonal simulation and prediction. *Bull. Amer. Meteor. Soc.,* **80,** 1349–1361.

Arakawa, A., and W. Schubert, 1974: Interaction of a cumulus ensemble with the large-scale environment. Part I. *J. Atmos. Sci.,* **31,** 674–704.

Barker, T. W., and J. D. Horel, 1989: The impact of climatology and systematic errors upon the skill of DERF forecasts. *Mon. Wea. Rev.,* **117,** 2835–2842.

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

Betts, A. K., F. Chen, K. Mitchell, and Z. Janjic, 1997: Assessment of land-surface and boundary layer models in two operational versions of the Eta Model using FIFE data. *Mon. Wea. Rev.,* **125,** 2896–2915.

Brankovic, C., and L. Ferranti, 1994: Predictability of seasonal atmospheric variations. *J. Climate,* **7,** 217–237.

Caplan, P., J. Derber, W. Gemmil, S.-Y. Hong, H.-L. Pan, and D. Parrish, 1997: Changes to the 1995 NCEP operation medium-range forecast model analysis–forecast system. *Wea. Forecasting,* **12,** 581–594.

Chen, F., K. Mitchell, J. Schaake, Y. Xue, H. Pan, V. Koren, Q. Duan, and A. Betts, 1996: Modeling of land-surface evaporation by four schemes and comparison with FIFE observations. *J. Geophys. Res.,* **101,** 7251–7268.

Chen, S., J. Roads, H.-M. H. Juang, and M. Kanamitsu, 1999: Global to regional simulation of California's wintertime precipitation. *J. Geophys. Res.,* **104** (ND24), 31 517–31 532.

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

Grell, G., 1993: Prognostic evaluation of assumptions used by cumulus parameterization. *Mon. Wea. Rev.,* **121,** 764–787.

Hong, S.-Y., and H.-L. Pan, 1996: Nonlocal boundary layer vertical diffusion in a medium-range forecast model. *Mon. Wea. Rev.,* **124,** 2322–2339.

——, and A. Leetmaa, 1999: An evaluation of the NCEP RSM for regional climate modeling. *J. Climate,* **12,** 592–609.

Juang, H.-M. H., and M. Kanamitsu, 1994: The NMC nested regional spectral model. *Mon. Wea. Rev.,* **122,** 3–26.

——, S.-Y. Hong, and M. Kanamitsu, 1997: The NCEP regional spectral model: An update. *Bull. Amer. Meteor. Soc.,* **78,** 2125–2143.

Kalnay, E., and Coauthors, 1996: The NCEP/NCAR 40-Year Reanalysis Project. *Bull. Amer. Meteor. Soc.,* **77,** 437–471.

Kanamitsu, M., 1989: Description of the NMC global data assimilation and forecast system. *Wea. Forecasting,* **4,** 335–342.

——, W. Ebisuzaki, J. Woolen, J. Potter, and M. Fiorino, 2000: Overview of NCEP/DoE Reanalysis-2. *Proc. Second WCRP International Conf. on Reanalyses,* WCRP-109, WMO/TD-No. 985, 1–4. [Available from WCRP Joint Planning Staff, c/o World Meteorological Organization, 7 bis, Avenue de la Paix, Case Postale 2300, 1211 Geneva 2, Switzerland.]

Kumar, A., and M. Hoerling, 1998: Annual cycle of Pacific–North American seasonal predictability associated with different phases of ENSO. *J. Climate,* **11,** 3295–3308.

——, ——, M. Ji, A. Leetmaa, and P. Sardeshmukh, 1996: Assessing a GCM's suitability for making seasonal predictions. *J. Climate,* **9,** 115–129.

Lorenz, E., 1965: A study of the predictability of a 28-variable atmospheric model. *Tellus,* **17,** 321–333.

——, 1982: Atmospheric predictability experiments with a large numerical model. *Tellus,* **34,** 505–513.

Miller, A. J., and J. Roads, 1990: A simplified coupled model of extended range predictability. *J. Climate,* **3,** 523–542.

Pan, H.-L., 1990: A simple parameterization of evapotranspiration over land for the NMC medium-range forecast model. *Mon. Wea. Rev.,* **118,** 2500–2512.

——, and W. Wu, 1995: Implementing a mass flux convective parameterization package for the NMC medium-range forecast model. NMC Office Note 409, NCEP/EMC, Camp Springs, MD, 40 pp.

Roads, J., 1989: Dynamical extended range forecasts of the lower tropospheric thickness. *Mon. Wea. Rev.,* **117,** 3–28.

——, and S. Chen, 2000: Surface water and energy budgets in the NCEP Regional Spectral Model. *J. Geophys. Res.,* **105**(D24), 29 539–29 550.

——, K. Ueyoshi, S. Chen, J. Alpert, and F. Fujioka, 1991: Medium-range fire weather forecasts. *Int. J. Wildland Fire,* **1,** 159–176.
——, S. Chen, F. M. Fujioka, H. Juang, and M. Kanamitsu, 1997: Global to regional fire weather forecasts. *Int. Forecast Fire News,* No. 17, 33–37.
——, ——, M. Kanamitsu, and H. Juang, 1998: Vertical structure of humidity and temperature budget residuals over the Mississippi River basin. *J. Geophys. Res.,* **103,** 3741–3759.
——, ——, ——, and ——, 1999: Surface water characteristics in NCEP global spectral model reanalysis. *J. Geophys. Res.,* **104,** 19 307–19 327.
——, ——, F. Fujioka, and R. Burgan, 2000a: Development of a seasonal fire weather forecast for the contiguous United States. Preprints, *Third Symp. on Fire and Forest Meteorology,* Long Beach, CA, Amer. Meteor. Soc., 99–102.
——, ——, and J. Ritchie, 2000b: ECPC's September 2000 forecasts. *ELLFB Quarterly Bulletin,* **9** (3), 6 pp.
Shukla, J., and Coauthors, 2000: Dynamical seasonal prediction. *Bull. Amer. Meteor. Soc.,* **81,** 2598–2606.
Stevens, D., D. Funayama, J. Roads, S. Chen, W. Smith, C. McCord, H. Juang, and F. Fujioka, 2000: Experimental short-term weather forecasts for Hawaii. MHPCC Application Briefs, 2 pp. [Available from MHPCC, Kihei, Maui, HI 96753.]
Tracton, M. S., K. Mo, and W. Chen, 1989: Dynamical extended range forecasting (DERF) at the National Meteorological Center. *Mon. Wea. Rev.,* **117,** 1604–1635.
von Storch, H., and F. W. Zwiers, 1999: *Statistical Analysis in Climate Research.* Cambridge University Press, 216 pp.
Xie, P., and P. A. Arkin, 1997: Global precipitation: A 17-year monthly analysis based on gauge observations, satellite estimates, and numerical model outputs. *Bull. Amer. Meteor. Soc.,* **78,** 2539–2558.