

Sources of Bias in the Monthly CFSv2 Forecast Climatology

MICHAEL K. TIPPETT

Department of Applied Physics and Applied Mathematics, Columbia University, New York, New York, and Center of Excellence for Climate Change Research, Department of Meteorology, King Abdulaziz University, Jeddah, Saudi Arabia

LAURIE TRENARY, TIMOTHY DELSOLE, AND KATHLEEN PEGION

George Mason University, and Center for Ocean–Land–Atmosphere Studies, Fairfax, Virginia

MICHELLE L. L'HEUREUX

Climate Prediction Center, National Weather Service, National Centers for Environmental Prediction, National Oceanic and Atmospheric Administration, College Park, Maryland

(Manuscript received 18 October 2017, in final form 5 February 2018)

ABSTRACT

Forecast climatologies are used to remove systematic errors from forecasts and to express forecasts as departures from normal. Forecast climatologies are computed from hindcasts by various averaging, smoothing, and interpolation procedures. Here the Climate Forecast System, version 2 (CFSv2), monthly forecast climatology provided by the NCEP Environmental Modeling Center (EMC) is shown to be biased in the sense of systematically differing from the hindcasts that are used to compute it. These biases, which are unexpected, are primarily due to fitting harmonics to hindcast data that have been organized in a particular format, which on careful inspection is seen to introduce discontinuities. Biases in the monthly near-surface temperature forecast climatology reach 2°C over North America for March targets and start times at the end of January. Biases in the monthly Niño-3.4 forecast climatology are also largest for start times near calendar-month boundaries. A further undesirable consequence of this fitting procedure is that the EMC forecast climatology varies discontinuously with lead time for fixed target month. Two alternative methods for computing the forecast climatology are proposed and illustrated. The proposed methods more accurately fit the hindcast data and provide a clearer representation of the CFSv2 model climate drift toward lower Niño-3.4 values for starts in March and April and toward higher Niño-3.4 values for starts in June, July, and August.

1. Introduction


“Forecast climatologies” are used in weather and climate prediction to correct systematic model errors and to express forecasts as anomalies. A forecast climatology is the expected (average) forecast value for a specified start time, lead time, and target period. The calculation of a forecast climatology is similar in many ways to that of an observational climatology, except that a forecast climatology can depend on lead time as

well as target period and is computed from historical forecasts or hindcasts rather than observations.

Forecast climatologies are especially important for seasonal-prediction systems that are based on coupled ocean–atmosphere models whose climatology may differ substantially from observations and may drift with lead time (Kumar et al. 2012). In seasonal- and subseasonal-prediction systems, forecasts are expressed as anomalies with respect to a forecast climatology. For instance, a forecast f for a specified start time, lead time, and target period can be written as the sum of the forecast climatology μ_f and an anomaly f_a ,

$$f = \mu_f + f_a.$$

Given an estimate $\hat{\mu}_f$ of the forecast climatology, the corresponding estimate \hat{f}_a of the forecast anomaly is

 Denotes content that is immediately available upon publication as open access.

Corresponding author: Michael K. Tippett, mkt14@columbia.edu

DOI: 10.1175/JAMC-D-17-0299.1

© 2018 American Meteorological Society. For information regarding reuse of this content and general copyright information, consult the [AMS Copyright Policy \(www.ametsoc.org/PUBSReuseLicenses\)](https://www.ametsoc.org/PUBSReuseLicenses).

$$\hat{f}_a = f - \hat{\mu}_f = (f - \mu_f) + (\mu_f - \hat{\mu}_f) = f_a + (\mu_f - \hat{\mu}_f).$$

Error in the estimation of the forecast climatology results in estimated forecast anomalies that are biased, and this bias increases mean-square error.

A simple (naive) method of estimating a forecast climatology for a specified start time, lead time, and target period is to average hindcasts with the same start time, lead time, and target period from different years. That is to say, the naive estimate of the forecast climatology from N years of hindcast data is

$$\hat{\mu}_f = \frac{1}{N} \sum_{i=1}^N f_i, \quad (1)$$

where f_i are hindcasts whose start time, lead time, and target period match those of the forecast climatology being estimated. No observational data are involved. The accuracy of the naive estimate depends on the number of years in the hindcast as well as the variability of the quantity being estimated. In particular, the variance of the naive estimate is σ_a^2/N , where σ_a^2 is the variance of the forecast anomaly. The naive approach may be problematic when a forecast climatology is required for start times that are not present in the hindcasts (e.g., when the operational forecast schedule differs from that of the hindcasts) or when the number of hindcasts is relatively small in comparison with the forecast anomaly variance. Curve (or surface) fitting methods are an alternative to the naive method. Fitting methods can estimate forecast climatologies for start times that are not in the hindcast (interpolation) and can reduce the errors due to sampling variability (smoothing).

Here we examine issues that lead to biases in the monthly forecast climatology of the NCEP Climate Forecast System, version 2 (CFSv2; Saha et al. 2014), provided by the NCEP Environmental Modeling Center (EMC). By bias, we mean the difference between the forecast climatology and averages of the corresponding hindcasts. The naive estimate in Eq. (1) of the forecast climatology is equal to the hindcast average and therefore has no bias. A fitting method is required to compute the CFSv2 forecast climatology because the naive estimate cannot be used for initialization times that are not available in the hindcasts. Details of the CFSv2 hindcasts, forecasts, and forecast climatology are provided in section 2. A standard method for computing a forecast climatology is to fit the hindcast data to some specified dependence on start time, forecast target, and lead time. Section 3 describes least squares estimation of forecast climatologies with linear, locally linear, and periodic dependence. The EMC forecast climatology appears to assume periodic

dependence on start time. Section 4 examines bias in the EMC monthly climatologies of near-surface temperature and the Niño-3.4 index and relates those biases to the fitting procedure. We propose two alternative methods for computing the CFSv2 monthly forecast climatology that depend on fewer parameters and that better fit the hindcast data. A summary and conclusions are given in section 5.

2. Data

The CFSv2 variables examined here are 2-m temperature and the Niño-3.4 index computed from sea surface temperature (SST). We use 2-m temperature hindcast data from 1982–2010 and SST hindcast data from 1999–2010. CFSv2 seasonal hindcasts have initializations on every fifth day starting from 1 January (not counting 29 February in leap years) at 6-hourly intervals (0000, 0600, 1200, and 1800 UTC; all start times here are UTC, and we omit the time zone). Operational CFSv2 seasonal forecasts began in early 2011 and are initialized every day at 6-hourly intervals, which means that there are no hindcasts with matching start times for most operational CFSv2 seasonal forecasts (e.g., there are no seasonal hindcasts initialized on 2–5 January). Operational forecasts of 2-m temperature that were initialized in January and February of 2017 are used to illustrate the effect of the EMC forecast climatology bias on forecast anomalies.

NCEP's EMC provides a CFSv2 monthly forecast climatology that matches the start times of operational seasonal forecasts. The forecast targets are the monthly averages of the nine calendar months that follow the start time. We use the index k to label these nine forecasts, $k = 1, \dots, 9$. That is, for each start time, we refer to the first of the nine monthly forecasts as the $k = 1$ forecast, the second as the $k = 2$ forecast, and so on up to the $k = 9$ forecast. (An alternative terminology, which we will not use here, is to refer to the $k = 1$ forecast as the lead-month-1 forecast.) For instance, for a 0000 6 January start time, forecast climatologies are provided for the $k = 1$ forecast, which is the February average, through the $k = 9$ forecast, which is the October average. Although *partial* calendar-month averages (e.g., the 6–31 January average from a 0000 6 January start) are included in operational products, these targets are not examined here and do not appear in the forecast climatologies provided by EMC. Also, by convention, the $k = 1$ forecast of a forecast initialized at 0000 1 January is the February average. Target months are plotted as a function of start time for $k = 1, k = 3$, and $k = 5$ forecasts in the upper panel of Fig. 1. We define the lead time L of a forecast to be the time (in days) from the forecast

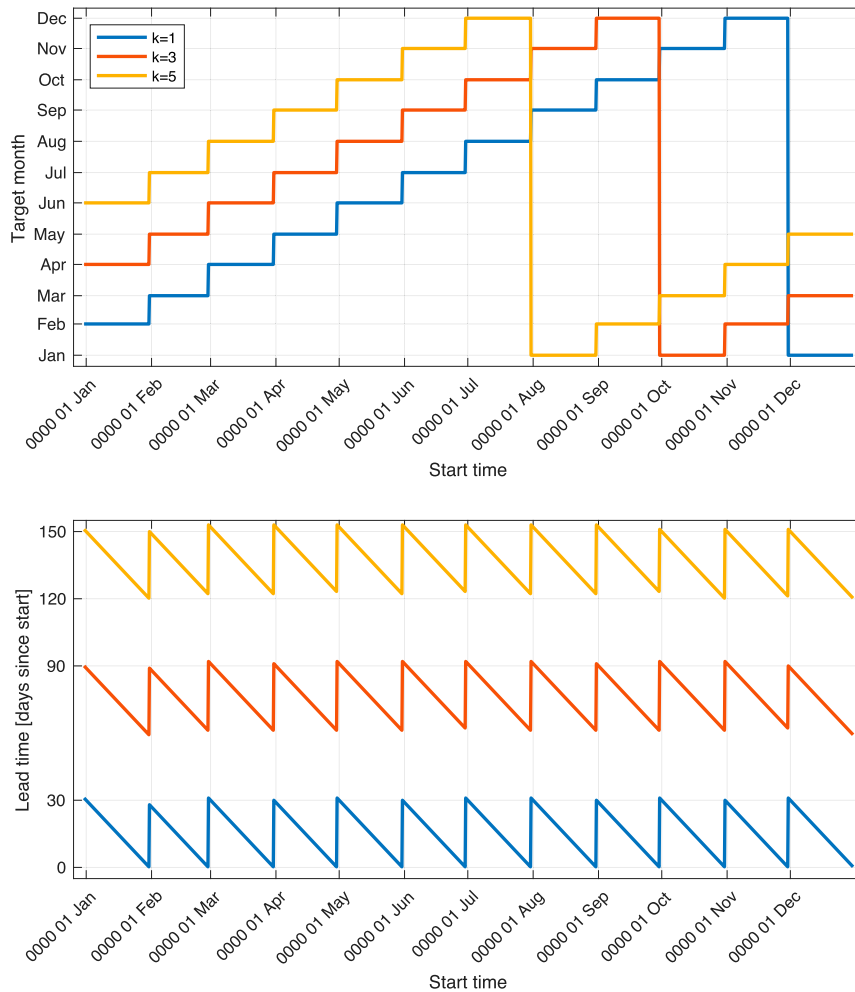


FIG. 1. Schematic showing the dependence of the (top) target month and (bottom) lead time on start time for $k = 1$, $k = 3$, and $k = 5$ forecasts during a year with 365 days.

start time S to the beginning of the target period. For instance, the lead time of a forecast starting at 0000 6 January with February target ($k = 1$) is 26 days. Lead times are plotted as a function of start time for $k = 1$, $k = 3$, and $k = 5$ forecasts in the lower panel of Fig. 1.

Two sets of CFSv2 forecast climatologies are available from EMC: one for the full hindcast period (1982–2010) and one for the last 12 years (1999–2010) of the hindcast period. Saha et al. (2014) recommends using the latter in the tropics for SST and precipitation over ocean because of a time-varying systematic error related to model initialization (Kumar et al. 2012). The skill of Niño-3.4 forecasts tends to be higher when forecast anomalies are computed using the 1999–2010 forecast climatology (Barnston and Tippett 2013; Barnston et al. 2018). Here we use the 1982–2010 EMC forecast climatology for 2-m temperature and the 1999–2010 EMC forecast climatology for SST.

CFSv2 hindcasts, operational forecasts, and EMC forecast climatologies are available for download from the International Research Institute for Climate and Society (IRI) Data Library (IRI 2011). EMC forecast climatologies are available from other locations as well (http://cfs.ncep.noaa.gov/pub/raid1/cfsv2/climo_9mon_mmtser; https://nomads.ncdc.noaa.gov/modeldata/cfs_reforecast_calclim_mm_9mon_fixf_1999-2010/).

3. Methods

a. Linear and local linear regression

Linear regression is a simple and powerful method for estimating a forecast climatology from hindcast data. For simplicity, consider a forecast for given target and spatial location so that the forecast climatology μ_f is a function of the lead time L alone, where lead time is measured from start time to the beginning of the target

period. Over some limited range of lead times, it may be reasonable to assume that the forecast climatology depends linearly on lead time. In that case, we write

$$\mu_f(L) = p_0 + p_1 L, \tag{2}$$

where the slope p_1 and intercept p_0 are unknown parameters to be estimated from the hindcast data. After the slope and intercept are calculated, the forecast climatology can be computed at any lead time whether or not it appears in the hindcast data. The slope and intercept are chosen to minimize

$$\sum_{i=1}^N [f(L_i) - \mu_f(L_i)]^2 = \sum_{i=1}^N [f(L_i) - (p_0 + p_1 L_i)]^2, \tag{3}$$

where $L_i, i = 1, \dots, N$ are lead times at which hindcasts $f(L_i)$ for the given target and spatial location are available. Here we have applied the principle that the climatology minimizes the sum of squared error.

The linear dependence of the forecast climatology on the unknown parameters p_0 and p_1 allows us to write Eqs. (2) and (3) in matrix form as

$$\begin{aligned} \begin{bmatrix} \mu_f(L_1) \\ \vdots \\ \mu_f(L_i) \\ \vdots \\ \mu_f(L_N) \end{bmatrix} &= \begin{bmatrix} 1 & L_1 \\ \vdots & \vdots \\ 1 & L_i \\ \vdots & \vdots \\ 1 & L_N \end{bmatrix} \begin{bmatrix} p_0 \\ p_1 \end{bmatrix} \quad \text{and} \\ \boldsymbol{\mu}_f &= \mathbf{X} \mathbf{p} \\ \sum_{i=1}^N [f(L_i) - (p_0 + p_1 L_i)]^2 &= \|\mathbf{f} - \mathbf{X}\mathbf{p}\|^2, \end{aligned} \tag{4}$$

respectively, where \mathbf{f} is the vector whose N elements are $f(L_1), \dots, f(L_N)$. The least squares estimate of the slope and intercept is given by $\hat{\mathbf{p}} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{f}$. The same least squares formalism can be used to estimate other functional forms of the forecast climatology as long as the forecast climatology is a linear function of the unknown parameters. For instance, quadratic dependence of the forecast climatology on lead time could be included by adding a column to the matrix \mathbf{X} that contains the squares of the lead times.

The assumption that the forecast climatology for given target and spatial location depends linearly on lead time may be unrealistic over a sufficiently wide range of lead times. In that case, *local linear regression* is a generalization of linear regression that can capture nonlinear dependence. For a particular lead time L_* , local linear regression finds the slope and intercept that

best fit (in a least squares sense) the hindcast data that have lead times near L_* (Hastie et al. 2009; Tippett and DelSole 2013). This selective fitting of the hindcast data is accomplished by giving more weight to the terms in the sum of squares in Eq. (4) that have values of L_i that are close to L_* . To be specific, local linear regression finds the slope and intercept that minimize the weighted sum of squares

$$\sum_{i=1}^N K(L_*, L_i) [f(L_i) - (p_0 + p_1 L_i)]^2 = \|\mathbf{W}_*^{1/2} (\mathbf{f} - \mathbf{X}\mathbf{p})\|^2, \tag{5}$$

where the weight matrix \mathbf{W}_* is an $N \times N$ diagonal matrix whose i th diagonal element is $K(L_*, L_i)$, with K being a kernel function. The kernel function measures the similarity of its two arguments and is large when they are similar and small when they are not. The least squares estimate of the local slope and intercept that minimizes Eq. (5) is $\hat{\mathbf{p}}_* = (\mathbf{X}^T \mathbf{W}_* \mathbf{X})^{-1} \mathbf{X}^T \mathbf{W}_* \mathbf{f}$ and depends on L_* through the weight matrix \mathbf{W}_* . Local linear regression fits the hindcast data more closely than linear regression because the slope and intercept are allowed to vary smoothly with L_* .

Here we choose a Gaussian kernel function,

$$K(L_*, L_i) = \frac{1}{\sqrt{2\pi}} \exp \left[-\frac{(L_* - L_i)^2}{\Delta^2} \right],$$

which depends on a bandwidth Δ . All of the data are weighted equally in the limit of large bandwidth, and the linear regression solution is recovered. Here we take $\Delta = 15$ days.

b. “EMC” fitting method

The linear regression method described in section 3a can also be used to fit periodic (annual or diurnal) behavior in the forecast climatology. The EMC CFSv2 monthly forecast climatology appears to have been computed by fitting the hindcast data to annual and diurnal harmonics in start time. The online description (<http://cfs.ncep.noaa.gov/cfsv2.info/CFSv2.Calibration.Data.doc>) of the EMC CFSv2 forecast climatology refers to the documentation of the CFS (version 1; <http://cfs.ncep.noaa.gov/cfs.daily.climatology.doc>) daily forecast climatology, which was computed by fitting four annual harmonics in start time to each lead of the hindcast data (Epstein 1988).

We were able to match the EMC forecast climatology of the Niño-3.4 index very well (see section 4) when we fit an intercept, 10 annual harmonics, a diurnal harmonic, and their interaction terms separately

to the $k = 1$ hindcasts, the $k = 2$ hindcasts, and so on. We refer to this model for the forecast climatology as *periodic in S*, where S is the time at which the forecast

is initialized. The periodic-in- S forecast climatology for start time S_i (in yeardays) and forecast k has the form

$$\begin{aligned} \mu_f(S_i, k) = & C_k + A_k \cos(2\pi\nu_{\text{day}}S_i) + B_k \sin(2\pi\nu_{\text{day}}S_i) + \sum_{n=1}^{10} a_{nk} \cos(2\pi\nu S_i) + b_{nk} \sin(2\pi\nu S_i) \\ & + \sum_{n=1}^{10} c_{nk} \cos(2\pi\nu_{\text{day}}S_i) \cos(2\pi\nu S_i) + d_{nk} \cos(2\pi\nu_{\text{day}}S_i) \sin(2\pi\nu S_i) \\ & + \sum_{n=1}^{10} e_{nk} \sin \sin(2\pi\nu_{\text{day}}S_i) \cos(2\pi\nu S_i) + f_{nk} \sin(2\pi\nu_{\text{day}}S_i) \sin(2\pi\nu S_i), \end{aligned} \quad (6)$$

for $i = 1, \dots, N$ and $k = 1, \dots, 9$, where $\nu_{\text{day}} = 1/(1 \text{ day})$, $\nu = 1/(365.25 \text{ days})$, and N is the number of lead- k hindcasts. For each value of k , the forecast climatology in Eq. (6) is a linear function of 63 parameters: $C_k, A_k, B_k, a_{nk}, b_{nk}, c_{nk}, d_{nk}$, and e_{nk} , $n = 1, 2, \dots, 9$, for a total of 567 parameters. The unknown parameters are estimated from the hindcast data by the linear regression method that was described in section 3a. Alternatively, two diurnal harmonics, equivalent to fitting 0000, 0600, 1200 and 1800 starts separately, should match the EMC forecast climatology (H. van den Dool, personal communication). The form of Eq. (6) means that the EMC forecast climatology is estimated independently for different values of k .

4. Results

There are two fundamental problems with the method used to compute the EMC forecast climatology. First, since the forecast climatology is fit separately for each value of k [Eq. (6) in section 3b], the method does not constrain the forecast climatology to vary smoothly as a function of lead time. Each value of k corresponds to a set of lead times with a range of about 30 days (Fig. 1, lower panel). Therefore, fitting the forecast climatology for each value of k separately is equivalent to fitting each range of lead times separately. For instance, fitting the $k = 1$ forecast climatology is equivalent to fitting lead times from 6 h to 31 days, whereas fitting the $k = 2$ forecast climatology fits lead times from 28.25 to 62 days. The EMC forecast climatology is constrained to be a smooth function of lead time only over lead-time ranges that correspond to the same value of k . Discontinuities in the EMC forecast climatology are possible across lead times that correspond to different values of k because the forecast climatology values are fit separately.

The second fundamental problem with the EMC forecast climatology is that harmonics in start time do not accurately fit the hindcast data. For each value of k , the target periods of the hindcast data vary discontinuously

with start time (Fig. 1, upper panel). These jumps in target period occur when start times cross calendar-month boundaries. For instance, February is the target of $k = 1$ forecasts that have start times during January, but, as the start time changes from 1800 31 January to 0000 1 February, the target of the $k = 1$ forecast changes from February to March. Therefore, by construction, the hindcast data for fixed values of k are necessarily discontinuous with respect to start time because their target periods change at calendar-month boundaries. An accurate forecast climatology should match these discontinuities, but using harmonics in start time smooths out the discontinuities and can introduce spurious oscillations (Gibbs phenomena).

The consequences of these two issues are apparent in the EMC forecast climatology of 2-m temperature over the contiguous United States. First, there are discontinuities with respect to lead time for fixed target months. For instance, March is the $k = 1$ target of a forecast starting at 0000 1 February, with lead time of 28 days (assuming no leap year). The $k = 2$ target of a forecast starting at 1800 31 January is also March, and the lead time of this forecast is 28.25 days. We expect the forecast climatologies associated with these two forecasts to be similar since their targets are the same and their lead times differ by only 6 h. In fact, they are dramatically different in the EMC forecast climatology (Figs. 2a,b). The EMC forecast climatology from the 0000 1 February start is substantially cooler than the one from the 1800 31 January start, by more than 5°C in some locations (see the difference map in Fig. 2c). Moreover, neither EMC forecast climatology matches the average of the March target hindcasts that start on 31 January. The EMC forecast climatology with 1800 31 January start is up to 2° warmer (Fig. 2d), and the one with 0000 1 February start is up to 5° cooler (Fig. 2e).

The second issue with the EMC method—that the target period is a discontinuous function of start time for fixed k —provides an explanation for why the forecast climatologies in Figs. 2a and 2b are so different despite

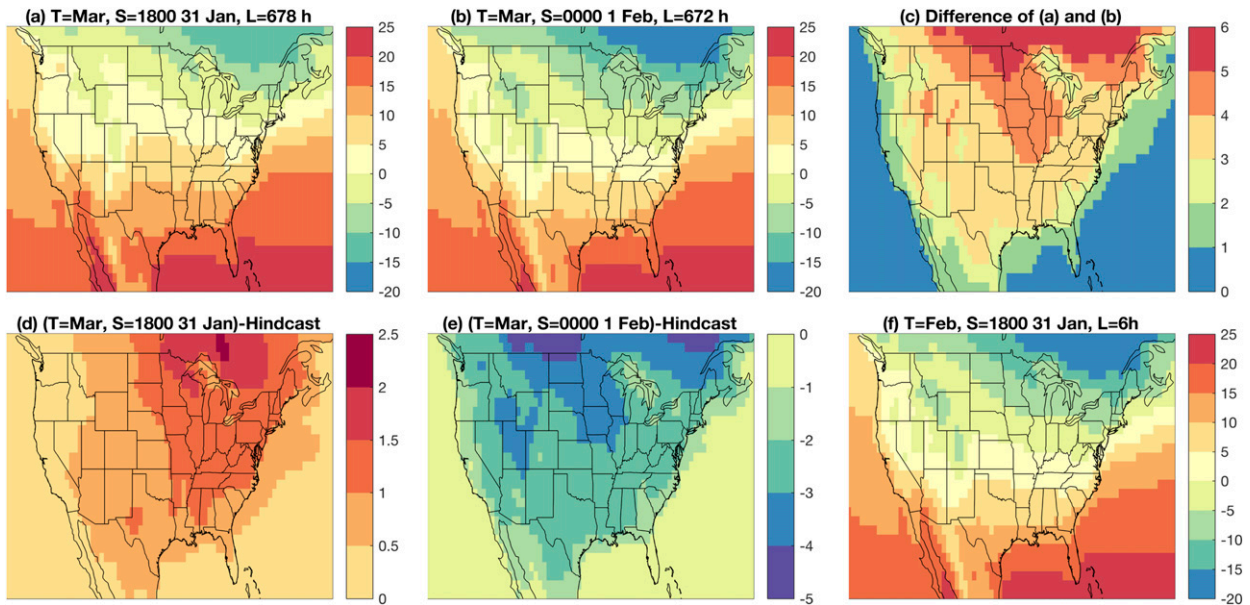


FIG. 2. March values of 2-m temperature from the EMC monthly forecast climatology for (a) starts at 1800 31 Jan and (b) starts at 0000 1 Feb, and (c) their difference. Also shown is the difference of the EMC forecast climatology for March targets and (d) 1800 31 Jan and (e) 0000 1 Feb starts with the 1982–2010 average of hindcasts with March targets starting on 31 Jan. The EMC (f) February forecast climatology for $k = 1$ starts at 1800 31 Jan. The T, S, and L labels denote target month, start time, and lead time, respectively.

their target months being the same and their lead time differing by only 6 h. The fact that the target period is a discontinuous function of start time for any value of k means that, even if a large number of harmonics in start time are used, the forecast climatology will be overly smooth and biased for start times that are near the discontinuities because harmonic approximations of discontinuous functions converge slowly with the addition of higher harmonics. In short, the spurious lead-time discontinuity in the EMC 2-m temperature forecast climatology (Figs. 2a,b) is possible because hindcast data are fit separately for each value of k , and the discontinuity is large because harmonics in start time are being fit to discontinuous hindcast data that result from the target period changing at calendar-month boundaries. The nearly indistinguishable change in the $k = 1$ climatology as the start time changes from 1800 31 January to 0000 1 February and the target changes from February to March (Fig. 2f) is also unrealistic given the strong seasonality in near-surface temperature and is further evidence of excessive smoothing in the EMC forecast climatology.

The EMC forecast climatology for the Niño-3.4 index provides a more detailed picture of the problems that arise from fitting the hindcasts to annual and diurnal harmonics in start time. Data for $k = 1$, $k = 2$, and $k = 7$ are shown in Fig. 3; other values are not shown. First we note that the EMC Niño-3.4 forecast climatology is nearly

indistinguishable from the periodic-in- S forecast climatology [Eq. (6) in section 3b; the blue lines in Figs. 3 and 4], even at the level of diurnal variation, which is particularly large for $k = 7$ (Fig. 4). The 1999–2010 hindcast averages (circles in Fig. 3) clearly show the jumps that are expected when start times cross calendar boundaries and the target month changes. Despite the large number of harmonics, the EMC forecast climatology (red lines in Fig. 3) is overly smooth and does not match the hindcast averages. The jumps in the $k = 1$ hindcast averages at target transitions are larger when seasonality is strong (from late winter through late summer) and are smaller during autumn and early winter. Dependence of the model climatology on lead time (model climate drift) is visible in hindcast averages whose starts are in the same calendar month and thus have the same target month but varying lead time. Hindcast averages for starts in June, July, and August and $k = 1$ decrease by more than 0.5°C for fixed target periods as the lead time increases from 6 h to 1 month (positive slope). The hindcast averages show weaker dependence on lead time for starts in winter and at longer leads.

Because the dependence of the hindcast averages on lead time within calendar months appears approximately linear in the lead time, we propose an alternative forecast climatology model that is linear in the lead time L and periodic (four annual harmonics) in target period, namely,

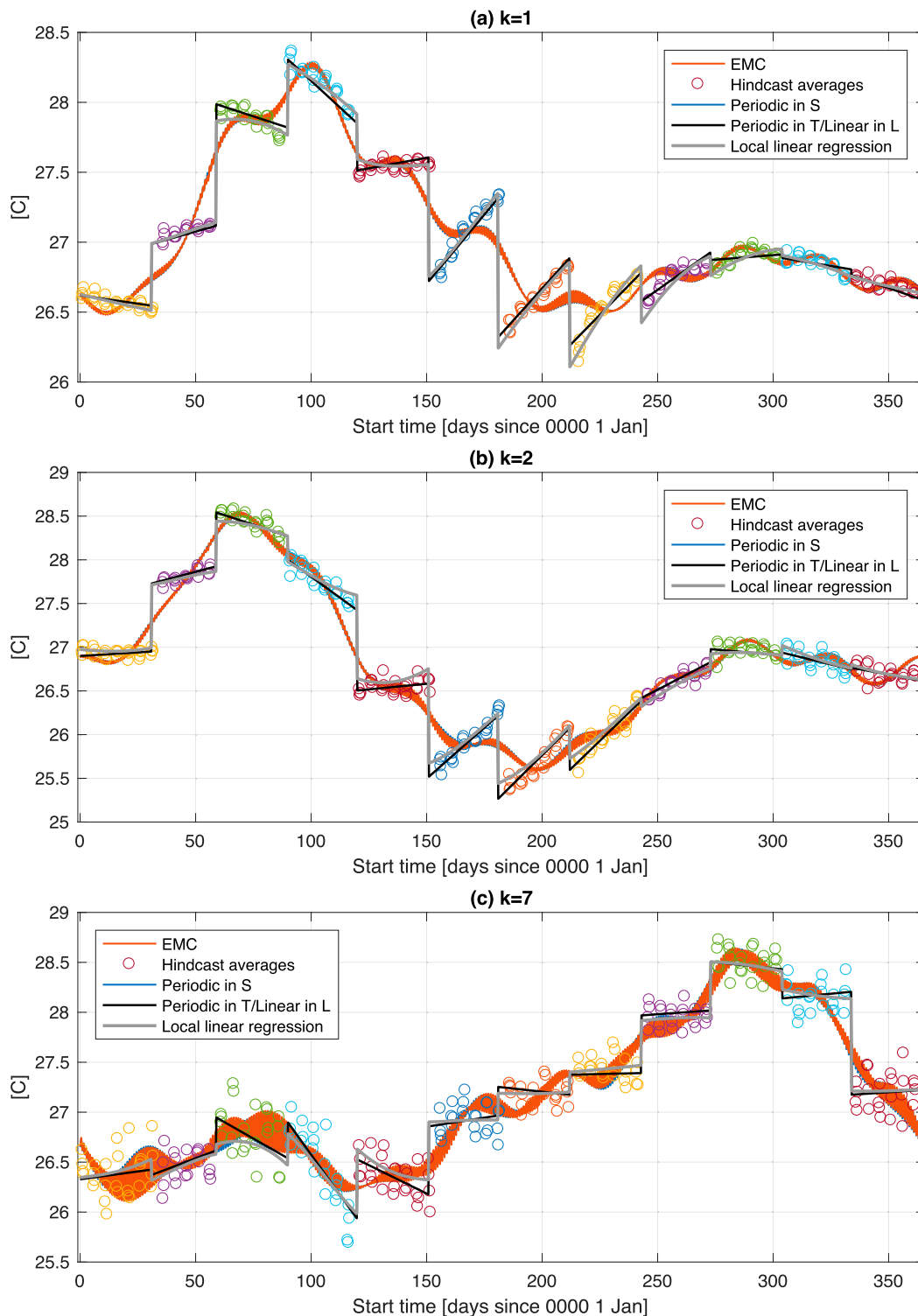


FIG. 3. The CFSv2 Niño-3.4 forecast climatology as a function of start time for (a) $k = 1$, (b) $k = 2$, and (c) $k = 3$. The EMC (red lines) and periodic-in-start-time- S (blue lines) forecast climatologies are nearly indistinguishable. The periodic-in-target-month- T /linear-in-lead-time- L (black lines) and local linear regression (gray lines) forecast climatologies match better the 1999–2010 hindcast averages (circles with color that depends on start month). The EMC forecast climatology is shifted by 4 days to match the hindcast averages.

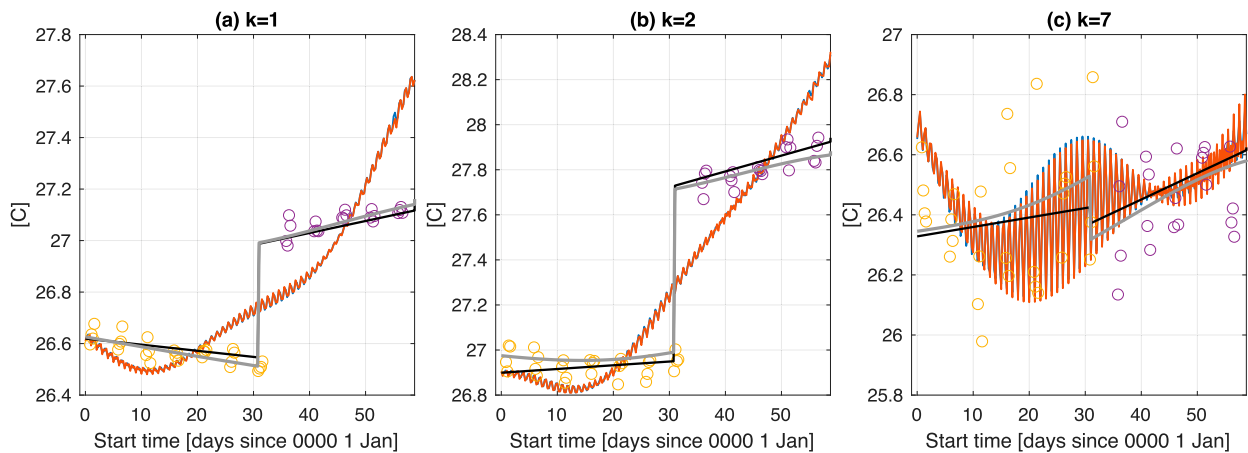


FIG. 4. As in Fig. 3, but restricted to January and February starts.

$$\begin{aligned} \mu_f(S_i, L_i, k) = & C_k + \sum_{n=1}^4 a_{nk} \cos[2\pi\nu(S_i + L_i)] + b_{nk} \sin[2\pi\nu(S_i + L_i)] + \sum_{n=0}^4 c_{nk} L_j \cos[2\pi\nu(S_i + L_i)] \\ & + d_{nk} L_j \sin[2\pi\nu(S_i + L_i)], \end{aligned} \quad (7)$$

for $i = 1, \dots, N$ and $k = 1, \dots, 9$. We include periodic dependence on target month through the quantity $S_i + L_i$, which is the beginning time (yearday) of the target period. We refer to this model for the forecast climatology as *periodic in T/linear in L*. The model in Eq. (7) contains 18 parameters for each value of k , for a total of 162 parameters. As in the EMC method, the parameters are estimated from hindcasts by the least squares method separately for each k . The choice of four annual harmonics is arbitrary but reasonable (e.g., van den Dool et al. 1997) and matches the number of harmonics used in the CFS (version 1) daily forecast climatology. Statistical significance testing and cross validation offer methods for objectively choosing the number of harmonics (Epstein 1991; Narapusetty et al. 2009). The periodic-in- T /linear-in- L forecast climatology fits lines through the hindcast data for each value of k , and the slope and intercept of the lines vary periodically with target month. The periodic-in- T /linear-in- L forecast climatology (black lines in Fig. 3) fits the hindcast averages very well (circles in Fig. 3). Although the periodic-in- T /linear-in- L forecast climatology is a smooth function of start time S and lead time L , it has the correct jumps as start times cross calendar boundaries because of the discontinuous dependence of L on start time (Fig. 1, lower panel). There is no clear indication that including a diurnal cycle in lead time would better fit the hindcast data. The slopes estimated in Eq. (7) vary with target and reflect dependence on lead time. In the absence of model climate drift, the slopes would

be zero, and the forecast climatology would be a function of target month alone (Fig. 1, upper panel).

Examination of the linear-in- L /periodic-in- T forecast climatologies for $k = 1$, $k = 2$, and $k = 7$ shows that Niño-3.4 forecasts starting in the same month often have roughly the same model climate drift (slope) regardless of target (Fig. 3). For instance, starts during June–August have hindcast averages that decrease with lead time (positive slope). This feature is clear in the hindcast data and the linear-in- L /periodic-in- T forecast climatologies but is less so in the EMC forecast climatology, which shows the opposite model climate drift for $k = 1$ and $k = 2$ forecasts starting in June. Starts during March and April have hindcast averages that increase with lead time (negative slope). Less consistency between start month and lead-time dependence is apparent in other months.

Although the periodic-in- T /linear-in- L model from Eq. (7) fits the hindcast data fairly well for fixed k , it does not constrain the resulting forecast climatology to be a smooth function of lead time for a given target month. Fitting Eq. (7) to hindcasts separately only guarantees smooth dependence on lead time (linear dependence, in fact) for start times that are in the same month. To examine this aspect of the forecast climatologies and hindcast data, we fix the target month and allow the lead time to range from 6 h to 270 days (~ 9 months). With this arrangement, the EMC forecast climatology, which is an excessively smooth function of start time for fixed values of k , displays unrealistic jumps as a function of lead time (October–March targets in Fig. 5; other targets

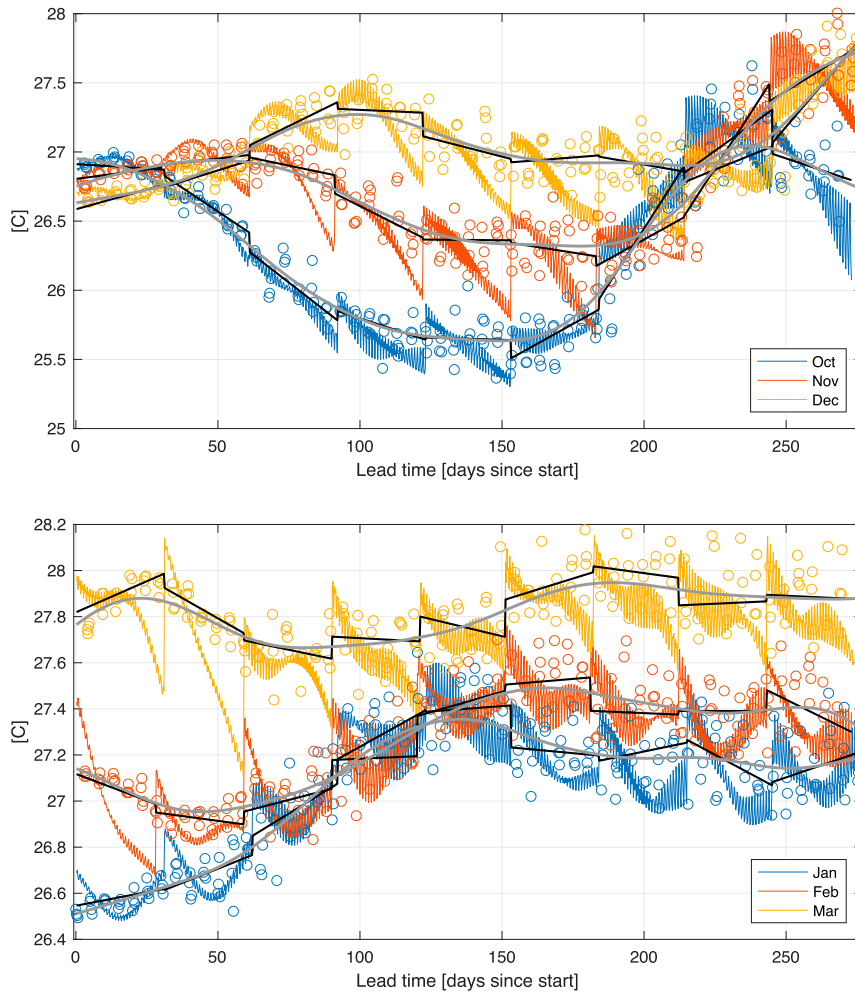


FIG. 5. The CFSv2 Niño-3.4 index forecast climatology for (top) October, November, and December targets and (bottom) January, February, and March targets as a function of lead time as provided by EMC (jagged colored lines), fit to be periodic in target month T /linear in lead time L (black line segments) and estimated by local linear regression (smooth gray curve). Circles are hindcast averages.

not shown). The periodic-in- T /linear-in- L model from Eq. (7) is a piecewise linear function of lead time and has small jumps at calendar-month boundaries that are due to fitting hindcasts separately for each value of k . A solution to this problem is to fit all of the hindcast data together rather than fitting separately for each value of k . For a fixed target month, the dependence of the forecast climatology on lead time is not linear, although the piecewise-linear approximation is fairly good. Fitting the hindcast data to a polynomial in lead time for each target gives reasonable results, but with some excessive variations at the lead-time endpoints (not shown). Instead we propose fitting the hindcasts for each target by local linear regression in lead time as described in section 3a (Hastie et al. 2009; Tippett and DelSole 2013). The resulting forecast climatology matches

the jumps in hindcast averages as the target month changes (Fig. 3) and is a smooth function of lead time for fixed target months (Fig. 5).

Qualitatively similar improvements are seen when the 2-m temperature forecast climatology is computed using the periodic-in- T /linear-in- L and local linear regression methods. Using the periodic-in- T /linear-in- L model [Eq. (7)], the March value of the monthly forecast climatology changes little as the start time moves from 1800 31 January to 0000 1 February (Figs. 6a,b), in contrast to the jump seen in the EMC forecast climatology (Figs. 2a,b). Nonetheless, the change (Fig. 6c) is larger than would be expected from increasing the lead time by 6 h and is a consequence of fitting $k = 1$ and $k = 2$ hindcasts separately. March values of the monthly forecast climatology

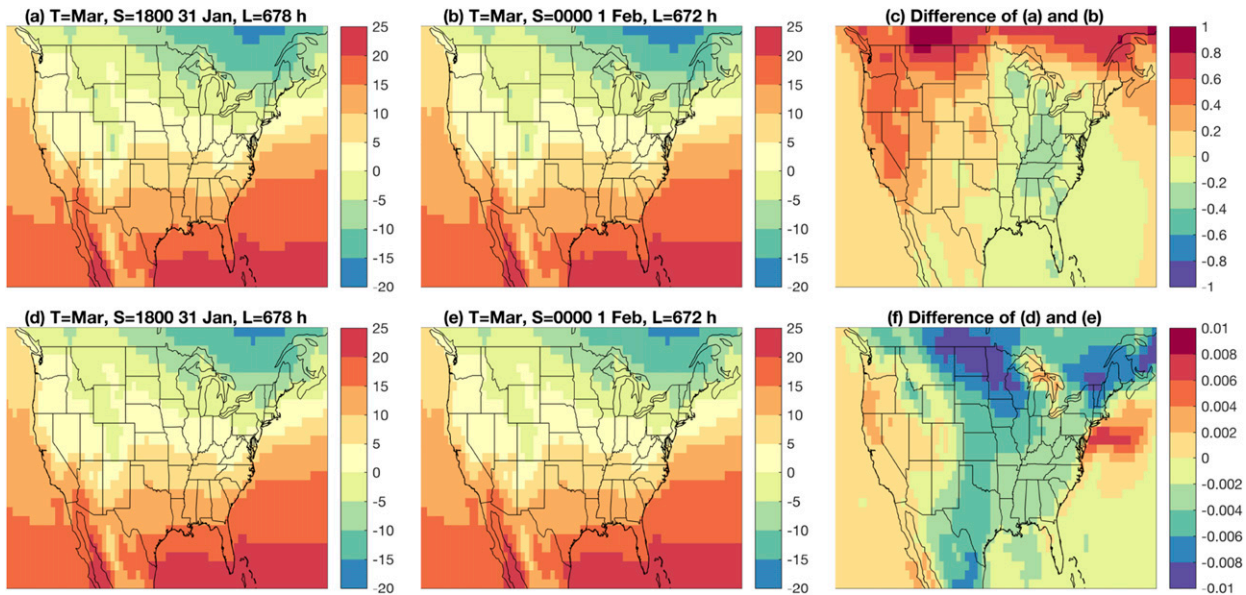


FIG. 6. March values of 2-m temperature from the periodic-in- T /linear-in- L monthly forecast climatology for starts at (a) 1800 31 Jan and (b) 0000 1 Feb, and (c) their difference. Also shown are March values of 2-m temperature from the local linear regression monthly forecast climatology for (d) starts at 1800 31 Jan and (e) starts at 0000 1 Feb, and (f) their difference. The T, S, and L labels denote target month, start time, and lead time, respectively.

that are based on local linear regression (Figs. 6d,e) include $k = 1$ and $k = 2$ hindcast data in their calculation and show substantially smaller changes as the start time moves from 1800 31 January to 0000 1 February (Fig. 6f).

The choice of forecast climatology can have a large impact on the resulting forecast anomalies and, therefore, forecast skill. For example, consider CFSv2 forecasts of the March 2017 2-m temperature anomaly with start times at the end of January 2017 and the beginning of February 2017. Using the EMC forecast climatology, the forecast anomaly from the 1800 31 January 2017 CFSv2 integration (Fig. 7a) is substantially cooler than that from the CFSv2 integration starting 6 h later (Fig. 7b). This difference in forecast anomaly is primarily due to the discontinuity in the EMC forecast climatology, however, rather than to the change in the forecast. The forecast anomalies computed with respect to the local linear forecast climatology show much less variation from one start time to the next (Figs. 7c,d). Observed March 2017 2-m temperature anomalies (not shown) were warm in the west and cool in the east, meaning that, although none of the forecasts captured the observed anomaly pattern, using the local linear regression forecast climatology resulted in smaller errors.

5. Summary and conclusions

Forecast climatologies are an essential part of a forecast system, especially for extended-range predictions that are expressed as anomalies or that have systematic errors in

their climatology. When the arrangement of hindcasts differs from that of real-time forecasts, some interpolation or fitting method is needed to produce a forecast climatology that matches the real-time forecasts. Fitting methods also help to reduce errors due to sampling variability. Improved estimates of forecast climatologies can lead to more accurate forecasts and assessments of forecast skill.

Here we have examined factors that lead to bias in the monthly CFSv2 forecast climatology provided by the NCEP Environmental Modeling Center. Real-time seasonal CFSv2 initializations are more frequent (every 6 h) than hindcast ones (every fifth day), and some fitting procedure is necessary to estimate a forecast climatology for start times that are not in the hindcast. The EMC forecast climatology fits the hindcast data to annual and diurnal harmonics in start time for each of the nine calendar-month target periods (indexed by k , $k = 1, \dots, 9$) that follow the start time. As a result, the EMC forecast climatology is a smooth function of start time for each value of k . The forecast target period for each k changes discontinuously as the start time crosses calendar-month boundaries, however. For instance, February averages are the targets of $k = 1$ forecasts starting in January, whereas March averages are the targets of $k = 1$ forecasts starting in February. These jumps in target month are not captured by the harmonic expansion in start time, and the excessive smoothness results in considerable bias between the EMC forecast climatology and hindcast averages near calendar-month transitions, which we have

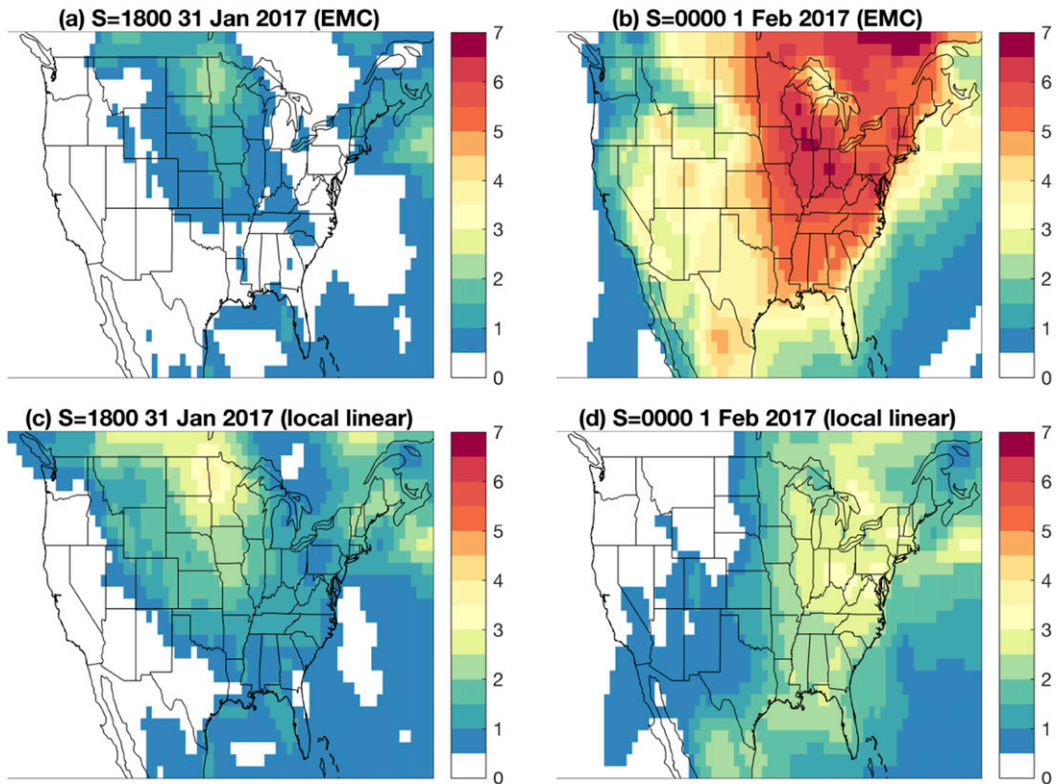


FIG. 7. March 2017 2-m temperature anomalies with respect to the (a),(b) EMC climatology and the (c),(d) local linear forecast climatology for forecasts starting at (left) 1800 31 Jan 2017 and (right) 0000 1 Feb 2017. The label S denotes start time.

demonstrated for 2-m temperature and the Niño-3.4 index. A further problem with the EMC forecast climatology is that fitting the forecast climatology separately for each value of k does not constrain the forecast climatology to be a smooth function of lead time for fixed target month because lead times corresponding to different values of k are fit independently. As a result, the EMC forecast climatology has discontinuities with respect to lead time for fixed target months.

Here we have proposed two alternative methods for computing monthly forecast climatologies that better match the CFSv2 hindcast data. In the first, the forecast climatology is assumed to be piecewise linear in lead time and annually periodic in target month. This model better fits the data but has small jumps as lead times cross calendar-month boundaries because it is fit separately for each value of k . The second proposed method uses local linear regression to fit the lead-time dependence for each target month using all of the hindcast data with that target. The resulting forecast climatology is a smooth function of lead time for each target month, similar to that obtained in Trenary et al. (2018) by smoothing the EMC-provided forecast climatology for fixed target months. The proposed methods provide a clearer and more accurate representation of lead-time dependence

(model climate drift) in the Niño-3.4 forecast climatology. CFSv2 forecasts starting in March and April tend toward lower Niño-3.4 values as lead time increases, whereas forecasts starting in June, July, and August tend toward higher Niño-3.4 values.

Acknowledgments. Useful comments and suggestions from Tony Barnston, Huug van den Dool, and anonymous reviewers are gratefully acknowledged. Author Tippet was partially supported by NOAA's Climate Program Office's Modeling, Analysis, Predictions, and Projections program award NA14OAR4310184. Authors DelSole and Trenary were supported by National Oceanic and Atmospheric Administration, under the Climate Test Bed program (NA10OAR4310264).

REFERENCES

- Barnston, A. G., and M. K. Tippet, 2013: Predictions of Niño-3.4 SST in CFSv1 and CFSv2: A diagnostic comparison. *Climate Dyn.*, **41**, 1615–1633, <https://doi.org/10.1007/s00382-013-1845-2>.
- , —, M. Ranganathan, and M. L. L'Heureux, 2018: Deterministic skill of ENSO predictions from the North American Multimodel Ensemble. *Climate Dyn.*, <https://doi.org/10.1007/s00382-017-3603-3>, in press; Corrigendum, <https://doi.org/10.1007/s00382-017-3814-7>.

- Epstein, E. S., 1988: A spectral climatology. *J. Climate*, **1**, 88–107, [https://doi.org/10.1175/1520-0442\(1988\)001<0088:ASC>2.0.CO;2](https://doi.org/10.1175/1520-0442(1988)001<0088:ASC>2.0.CO;2).
- , 1991: Determining the optimum number of harmonics to represent normals based on multiyear data. *J. Climate*, **4**, 1047–1051, [https://doi.org/10.1175/1520-0442\(1991\)004<1047:DTONOH>2.0.CO;2](https://doi.org/10.1175/1520-0442(1991)004<1047:DTONOH>2.0.CO;2).
- Hastie, T., R. Tibshirani, and J. Friedman, 2009: *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer, 745 pp.
- IRI, 2011: Climate Forecast System version 2. IRI Data Library, accessed 12 October 2017, <http://iridl.ldeo.columbia.edu/SOURCES/.NOAA/.NCEP/.EMC/.CFSv2/>.
- Kumar, A., M. Chen, L. Zhang, W. Wang, Y. Xue, C. Wen, L. Marx, and B. Huang, 2012: An analysis of the nonstationarity in the bias of sea surface temperature forecasts for the NCEP Climate Forecast System (CFS) version 2. *Mon. Wea. Rev.*, **140**, 3003–3016, <https://doi.org/10.1175/MWR-D-11-00335.1>.
- Narapusetty, B., T. DelSole, and M. K. Tippett, 2009: Optimal estimation of the climatological mean. *J. Climate*, **22**, 4845–4859, <https://doi.org/10.1175/2009JCLI2944.1>.
- Saha, S., and Coauthors, 2014: The NCEP Climate Forecast System version 2. *J. Climate*, **27**, 2185–2208, <https://doi.org/10.1175/JCLI-D-12-00823.1>.
- Tippett, M. K., and T. DelSole, 2013: Constructed analogs and linear regression. *Mon. Wea. Rev.*, **141**, 2519–2525, <https://doi.org/10.1175/MWR-D-12-00223.1>.
- Trenary, L., T. DelSole, M. K. Tippett, and K. Pegion, 2018: Monthly ENSO forecast skill and lagged ensemble size. *J. Adv. Model. Earth Syst.*, <https://doi.org/10.1002/2017MS001204>, in press.
- van den Dool, H. M., S. Saha, J. Schemm, and J. Huang, 1997: A temporal interpolation method to obtain hourly atmospheric surface pressure tides in reanalysis 1979–1995. *J. Geophys. Res.*, **102**, 22 013–22 024, <https://doi.org/10.1029/97JD01571>.