

The “Shukla–Gutzler” Method for Estimating Potential Seasonal Predictability

TIMOTHY DELSOLE

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

XIA FENG

George Mason University, Fairfax, Virginia

(Manuscript received 4 January 2012, in final form 24 July 2012)

ABSTRACT

This paper reexamines a procedure proposed by Shukla and Gutzler for estimating potential seasonal predictability. Certain subtle and unverified assumptions required for the method to work are clarified, and Monte Carlo experiments are used to demonstrate that these assumptions are adequate even for autocorrelated processes in typical applications, provided the effective time scale T_0 of the stochastic process is known. This paper highlights the fact that the time scale T_0 is difficult to estimate reliably (as noted in other papers) and can be biased by an order of magnitude. This bias can seriously compromise the reliability of the Shukla–Gutzler method.

1. Introduction

Variability arising from the chaotic, turbulent nature of the atmospheric system irrespective of changes in radiative forcing or boundary conditions is often called *weather noise*. It is generally believed that weather noise is predictable on time scales less than a month (Lorenz 1963; Shukla 1981). Beyond a month, slowly varying components of the climate system, such as sea surface temperature, soil moisture, or sea ice thickness, are predictable owing to their longer memory. If these slower components influence the atmosphere, then they can give rise to atmospheric predictability beyond a month. Such predictability arising from “boundary forcings” is called potential predictability. The degree to which variability of seasonal means exceeds variability due to weather noise determines the degree of potential predictability.

There are two fundamentally different ways of estimating potential predictability. The first is to perform idealized experiments with a general circulation model, as pioneered by Chervin (1986) and continuing to the present day (Zwiers 1987; Kumar and Hoerling 1995; Phelps et al. 2004, to name a few). Of course, this approach is limited by the ability of the dynamical model

to capture all dynamical processes that are relevant to the potential predictability. The second is to estimate potential predictability directly from observational data, as pioneered by Madden (1976) and continuing to the present day (Zheng et al. 2000; Feng et al. 2011, 2012). Of course, the empirical approach is limited by the statistical assumptions invoked to separate weather noise from boundary forced variability. A comparison between various empirical estimates of potential predictability has been conducted recently by Feng et al. (2012, manuscript submitted to *J. Geophys. Res.*). In the course of this comparison, an error in the method proposed by Shukla and Gutzler (1983) was discovered. The purpose of this paper is to correct this error, and to clarify certain aspects of the method proposed by Shukla and Gutzler (1983).

2. Background

In this section we summarize the mathematical basis for potential predictability and discuss a significance test for potential predictability that serves as the foundation of the Shukla–Gutzler method. Let X be a stationary stochastic process with mean μ_X and variance σ_X^2 . We call this process “weather noise.” The observed variability is assumed to contain an additional component of variability that is independent of, and varies on much longer time scales than, the stochastic process X . This

Corresponding author address: Timothy DelSole, 4041 Powder Mill Rd., Calverton, MD 20705.
E-mail: delsole@cola.iges.org

additional component of variability is associated with boundary forcing. Let different boundary forcings be identified with the indices $b = 1, 2, \dots, B$. In the context of seasonal predictability, boundary forcing is often identified with the state of the sea surface temperature, in which case the index b is effectively the “year.” The above assumptions imply that the observed variable O on the d th day can be modeled by

$$O_{d,b} = \mu_b + X_{d,b}, \tag{1}$$

where $X_{d,b}$ denotes the stochastic process X , but with an extra index b to indicate an independent realization of X for each value of b . More precisely, we assume that $X_{d,b}$ is independent of $X_{d,b'}$ when $b \neq b'$. We further assume that $X_{d,b}$ is independent of μ_b , and has a mean, variance, and autocorrelation that are independent of b . The parameter μ_b is the change in population mean due to boundary forcing and is assumed to be uncorrelated in b . Because the two terms on the right-hand side of (1) are independent, it follows that the variance of *daily* observations is

$$\text{var}(O_{d,b}) = \text{var}(\mu_b) + \sigma_X^2. \tag{2}$$

In other words, the variance of the observed daily variable can be decomposed into two terms: a boundary forced term $\text{var}(\mu_b)$ and a weather noise term σ_X^2 .

Model (1) serves as the foundation in many studies for estimating potential predictability based on observations (Madden 1976; Zwiers 1987). Also, the use of additive stochastic forcing to understand and model dynamical systems can be formally justified in certain systems for which there exists a strong separation in time scales (Majda et al. 2001). In the context of seasonal predictability, the time scale of weather variability (e.g., 5 days) is well separated from ENSO variability (e.g., 6 months), suggesting that the model is plausible. Nevertheless, the assumption that weather noise is independent of boundary forcing is not likely to be strictly true in realistic systems; for instance, the variability of weather may vary with sea surface temperature. Studies based on large ensemble sizes find that the probability of extremes can depend on sea surface temperature (Sardeshmukh et al. 2000; Peng and Kumar 2005). Recently, Feng et al. (2012) tested the independence assumption within an analysis of covariance framework and found that the assumption holds well for seasonal predictability. For the purpose of this paper, we proceed under the assumption that (1) is an adequate model of observations and consider a particular estimate of potential predictability when the model is true.

In the context of (1), the observed variable $O_{d,b}$ is said to have no potential predictability when $\text{var}(\mu_b) = 0$. It follows that the hypothesis of no potential predictability is equivalent to the following hypothesis:

$$H_0 : \mu_1 = \mu_2 = \dots = \mu_b. \tag{3}$$

The standard procedure for testing hypothesis (3) is analysis of variance (ANOVA), provided that $X_{d,b}$ in model (1) is uncorrelated in time (i.e., is a white noise process). It is well known that temperature is auto-correlated on daily time scales and consequently does not satisfy the assumptions needed to apply ANOVA. Nevertheless, ANOVA provides the basis for much of potential predictability analysis and, hence, is worth reviewing briefly.

An estimate of the boundary forced term μ_b is the time mean of $O_{d,b}$. Given daily values $O_{1,b}, O_{2,b}, \dots, O_{D,b}$, where D is the number of days, the time mean is defined as

$$\bar{O}_b = \frac{1}{D} \sum_{d=1}^D O_{d,b}. \tag{4}$$

Assuming $O_{d,b}$ and $O_{d,b'}$ are independent for $b \neq b'$, an unbiased estimate of the variance of time means is

$$\hat{\sigma}_M^2 = \frac{1}{B-1} \sum_{b=1}^B (\bar{O}_b - \bar{\bar{O}})^2, \tag{5}$$

where the “grand mean” is defined as

$$\bar{\bar{O}} = \frac{1}{B} \sum_{b=1}^B \bar{O}_b. \tag{6}$$

To test the null hypothesis in (3), we compare the estimated variance of time means in (5) to the variance expected under the assumption that the null hypothesis is true. This alternative estimate is based on the intraseasonal variance defined by

$$\hat{\sigma}_I^2 = \frac{1}{\text{dof}_I} \sum_b \sum_d (O_{d,b} - \bar{O}_b)^2, \tag{7}$$

where dof_I is the “degrees of freedom” chosen to yield an *unbiased* estimate of intraseasonal variance. In ANOVA, which assumes $X_{d,b}$ is white noise, the correct degrees of freedom is

$$\text{dof}_I = B(D - 1). \tag{8}$$

If the null hypothesis in (3) is true and $X_{d,b}$ is Gaussian white noise, then the statistic

$$F = D \frac{\hat{\sigma}_M^2}{\hat{\sigma}_I^2} \tag{9}$$

has an F distribution with $B - 1$ and $B(D - 1)$ degrees of freedom. A test of the hypothesis in (3) is performed by evaluating the statistic in (9) derived from a sample time series, and then comparing this value to the critical values of the appropriate F distribution.

3. Accounting for autocorrelations

To account for autocorrelations in $X_{d,b}$, we start with some well-established facts. The time mean of the weather noise is defined as

$$\bar{X}_b = \frac{1}{D} \sum_{d=1}^D X_{d,b}. \tag{10}$$

Let the autocorrelation function of $X_{d,b}$ be ρ_τ , where τ is time lag between days. It is well known (Brockwell and Davis 1991, chapter 7) that the time mean is an unbiased estimate of the mean μ_X with variance:

$$\text{var}(\bar{X}_b) = \frac{T_0}{D} \sigma_X^2, \tag{11}$$

where

$$T_0 = 1 + 2 \sum_{\tau=1}^{D-1} \left(1 - \frac{\tau}{D}\right) \rho_\tau. \tag{12}$$

The T_0 often is called the *effective time scale* of the stochastic process. No b index appears on the right-hand side of (11) because the statistics of X are independent of b . The governing equation for the time mean can be derived from (1) as

$$\bar{O}_b = \mu_b + \bar{X}_b. \tag{13}$$

Since the two terms on the right-hand side of (13) are independent, it follows that

$$\text{var}(\bar{O}_b) = \text{var}(\mu_b) + \frac{T_0}{D} \sigma_X^2. \tag{14}$$

Thus, given the above assumptions, the variance of *time means* can be decomposed into two terms: a boundary forced term $\text{var}(\mu_b)$ and a weather noise term $T_0 \sigma_X^2 / D$.

If we were to consider a single realization of the weather noise process, denoted X_d , it is well known (see, e.g., Trenberth 1984) that an unbiased estimate of σ_X^2 is

$$\hat{\sigma}_X^2 = \frac{1}{D - T_0} \sum_{d=1}^D (X_d - \bar{X})^2, \tag{15}$$

where we use the caret symbol $\hat{}$ to denote sample quantities. For the more general case of $X_{d,b}$, in which different realizations for different b are independent, we have the estimate

$$\hat{\sigma}_X^2 = \frac{1}{B(D - T_0)} \sum_{b=1}^B \sum_{d=1}^D (X_{d,b} - \bar{X}_b)^2, \tag{16}$$

Note that T_0 occurs in these equations in two fundamentally different ways: it *divides* the number of days in (11), while it *subtracts* from the number of days in (16).

In general, we do not observe $X_{d,b}$ directly. However, it is readily verified from (1) that

$$O_{d,b} - \bar{O}_b = X_{d,b} - \bar{X}_b. \tag{17}$$

Substituting this expression into (16) gives

$$\hat{\sigma}_X^2 = \frac{1}{B(D - T_0)} \sum_{b=1}^B \sum_{d=1}^D (O_{d,b} - \bar{O}_b)^2, \tag{18}$$

Comparison with (7) implies that the degrees of freedom required to account for autocorrelations in weather noise is

$$\text{dof}_I = B(D - T_0). \tag{19}$$

To derive a test statistic, we compute the ratio of two independent, chi-squared distributed quantities. If the null hypothesis in (3) is true, then according to (13), O and X differ by a constant, in which case $\text{var}(\bar{O}_b) = \text{var}(\bar{X}_b)$. Moreover, different realizations of \bar{O}_b can be considered to be independent and identically distributed. It follows that the quantity

$$\theta_M = \frac{\sum_{b=1}^B (\bar{O}_b - \bar{\bar{O}})^2}{\text{var}(\bar{O}_b)} = \frac{\sum_{b=1}^B (\bar{O}_b - \bar{\bar{O}})^2}{\text{var}(\bar{X}_b)} = \frac{(B - 1) \hat{\sigma}_M^2}{\frac{T_0}{D} \sigma_X^2}, \tag{20}$$

has a chi-squared distribution with $B - 1$ degrees of freedom. Unfortunately, the quantity

$$\begin{aligned} \theta_I &= \frac{\sum_{b=1}^B \sum_{d=1}^D (X_{d,b} - \bar{X}_b)^2}{\sigma_X^2} = \frac{\sum_{b=1}^B \sum_{d=1}^D (O_{d,b} - \bar{O}_b)^2}{\sigma_X^2} \\ &= \frac{B(D - T_0) \hat{\sigma}_X^2}{\sigma_X^2}, \end{aligned} \tag{21}$$

does not have an exact chi-squared distribution because $X_{d,b}$ is not white noise (i.e., not independently and identically distributed), nor is it independent of \bar{X}_b , when $X_{d,b}$ is autocorrelated (Thiébaux and Zwiers 1984, see their appendix). Nevertheless, for large sample sizes, we assume that this quantity has an approximate chi-squared distribution (this will be checked shortly). A rational approach for choosing the degrees of freedom is to ensure that the mean of θ_I is consistent with the mean of the chi-square distribution. Recall that the mean of a chi-squared distribution equals its degrees of freedom. Since $\hat{\sigma}_X^2$ defined in (18) is unbiased, it follows that θ_I has a mean of $B(D - T_0)$. We therefore assume that θ_I has an approximate chi-squared distribution with $B(D - T_0)$ degrees of freedom. Assuming that θ_M and θ_I are approximately independent, and assuming that the null hypothesis (3) is true, then the statistic

$$F = \frac{\frac{\theta_M}{B-1}}{\frac{\theta_I}{B(D-T_0)}} = \frac{D \hat{\sigma}_M^2}{T_0 \hat{\sigma}_X^2} \tag{22}$$

has an approximate F distribution with $B - 1$ and $B(D - T_0)$ degrees of freedom.

4. The Shukla–Gutzler method

In contrast to the above analysis, Shukla and Gutzler (1983) effectively propose, in their Eq. (2), estimating intraseasonal variance in (7) using the degrees of freedom:

$$\text{dof}'_I = B[(D/T_0) - 1]. \tag{23}$$

This choice differs from the proper dof_I defined in (19) by a factor of T_0 , implying that this choice does not produce an unbiased estimate. Shukla and Gutzler (1983) also propose estimating intraseasonal variance using standard formulas [e.g., see their equation for σ^2 after their (3)], but the above considerations imply that this estimate will also be biased.

It turns out that the F statistic in (22) corresponds *exactly* to Eq. (1) of Shukla and Gutzler (1983). Therefore, the error in the degrees of freedom does not affect the *value* of the F statistic proposed by Shukla and Gutzler (1983), but it does affect the significance level with which the F value is compared. Moreover, since $\text{dof}'_I < \text{dof}_I$, the incorrect degrees of freedom leads to conservative significance levels. Hence, after correcting the degrees of freedom, we expect to identify more potential predictability.

To assess the various approximations, we perform the following Monte Carlo experiments. First, we generate time series from an autoregressive (AR) model:

$$X_{d,b} = \phi X_{d-1,b} + w_{d,b}, \tag{24}$$

where $w_{d,b}$ is Gaussian white noise with zero mean and unit variance. As is well known (Box et al. 2008), the resulting random variable has autocorrelation function $\rho_\tau = \phi^\tau$ and variance:

$$\sigma_X^2 = \frac{1}{1 - \phi^2}. \tag{25}$$

In practice, the initial condition $X_{0,b}$ was chosen to be 0 and the first 100 time steps were discarded to avoid the influence of the initial condition. We choose parameters consistent with Shukla and Gutzler (1983), namely, $D = 31$ days and $B = 15$ yr.

Perhaps the most questionable assumption in accounting for autocorrelations discussed in section 3 is that the estimated intraseasonal variability, as measured by θ_I in (21), has a chi-squared distribution. To assess this assumption, we generated a synthetic time series from the autoregressive model in (24) and used this time series to evaluate θ_I . This procedure was repeated 1000 times for three separate values of the autoregressive parameter ϕ . The results of these experiments are shown as histograms in Fig. 1. Superimposed on the histograms are two chi-squared distributions: one with degrees of freedom given by (19), which is $\text{dof}_I = 15 \times (31 - T_0)$ (solid curve); and another with degrees of freedom given by (23), which is $\text{dof}'_I = 15 \times [(31/T_0) - 1]$ (dotted curve). In each case, the population value of T_0 , derived from the population autocorrelation function $\rho_\tau = \phi^\tau$, is used. The figure clearly shows that the degrees of freedom dof_I gives a much better match to θ_I than dof'_I . There is some evidence that the chi-squared assumption breaks down for large ϕ , but nevertheless the chi-squared distribution with corrected degrees of freedom gives a better description of the Monte Carlo results than with the chi-squared distribution with Shukla–Gutzler degrees of freedom.

Although the difference between the degrees of freedom in (19) and (23) is large, this fact may not lead to substantially different tests of potential predictability. To understand this, let us take a specific case of $\phi = 0.7$, which corresponds to a time scale of approximately 5 days and degrees of freedom for dof_I and dof'_I of 390 and 78, respectively. Although these values differ by a factor of 5, the F distribution is relatively insensitive to this change in the degree of freedom, because both numbers

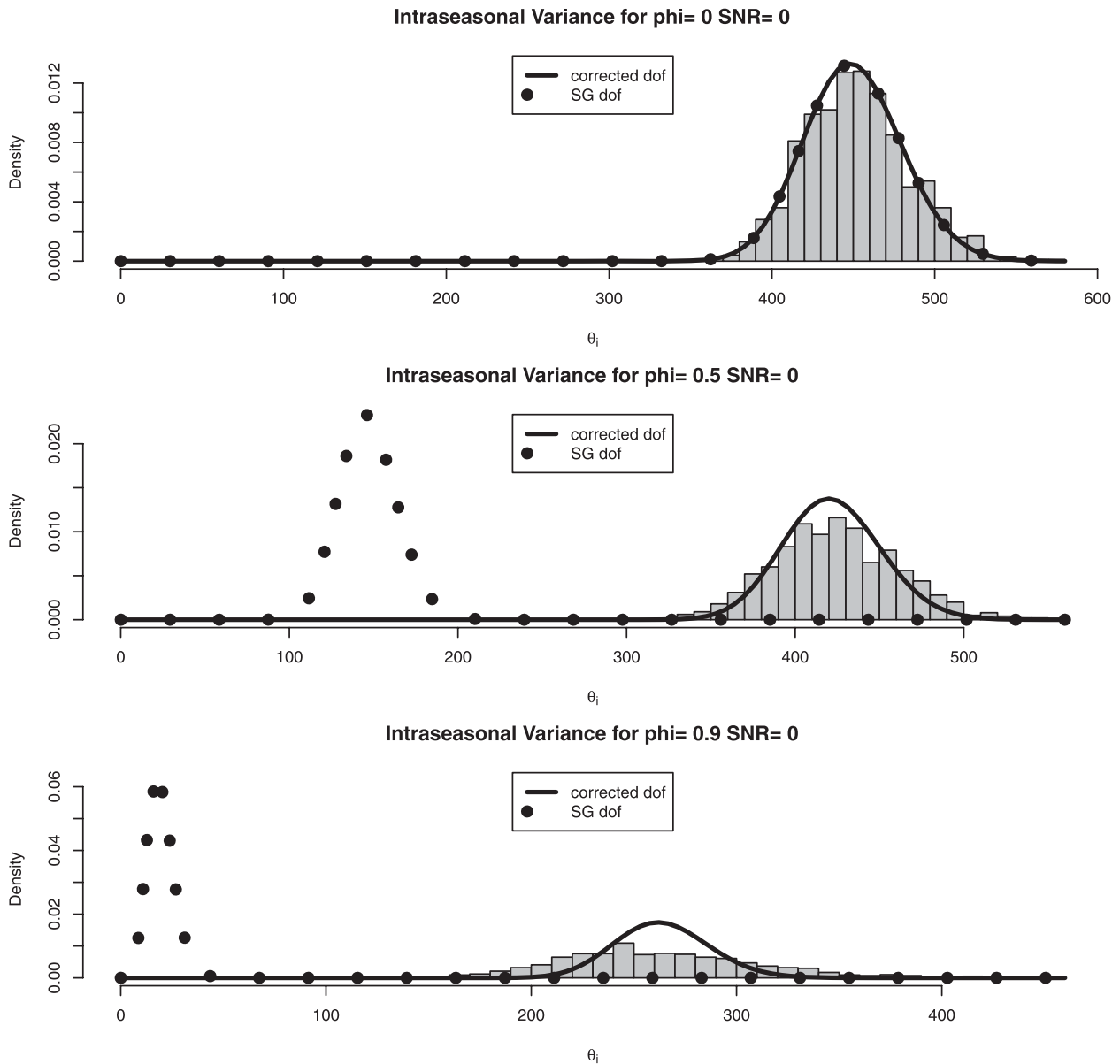


FIG. 1. Histogram of the parameter θ_I defined in (21) derived from 1000 independent realizations of time series generated by a first-order autoregressive model for three different values of the AR parameter: (top) $\phi = 0$, (middle) $\phi = 0.5$, and (bottom) $\phi = 0.9$. Also shown are the chi-squared distributions with degrees of freedom dof_I defined in (19) (solid; “corrected dof”), and the value suggested by Shukla–Gutzler, dof'_I , defined in (23) (dotted; “SG dof”). The experiments use the parameters $D = 31$ days and $Y = 15$ yr, meaning that 15 independent time series of length 31 were generated from the AR model.

are relatively large (e.g., above 50). This point is illustrated in Fig. 2, which shows the histogram of the F ratio in (22) superposed with two analytic F distributions, where the first degree of freedom is $B - 1$ for both, but the second degree of freedom is either dof_I (solid curve) or dof'_I (dotted curve). The figure clearly shows that the F distribution is nearly the same even though the degrees of freedom are very different. Moreover, the standard F distribution matches the Monte Carlo results even for

large ϕ . Thus, despite the large change in degrees of freedom, there is only a slight change in the significance threshold of the F distribution, because of the relatively large sample size. However, the impact of the change of degrees of freedom will be more significant for smaller sample sizes. As a possible example, the error may be important for estimating potential predictability of a short run from a very high-resolution atmospheric general circulation model.

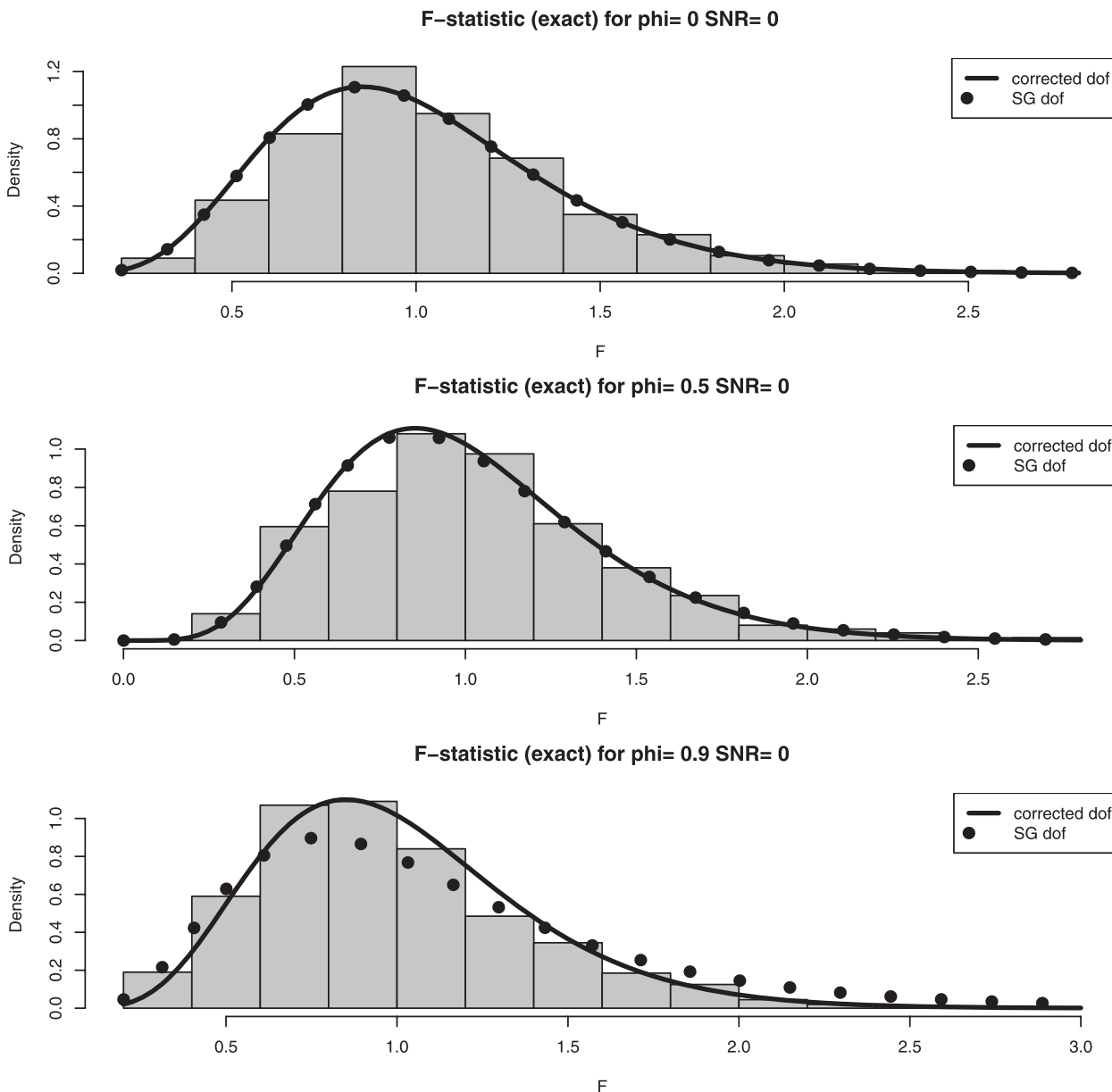


FIG. 2. As in Fig. 1, but for the F statistic for testing potential predictability. Also shown are the standard F distributions with the first degree of freedom being $Y - 1$, and the second degree of freedom being the corrected value dof_t (solid), or being the SG value dof'_t (dotted).

5. Practical issues

The comparisons illustrated in Figs. 1 and 2 show that the theoretical distributions derived in section 3 provide a reasonable description of the variability of a first-order autoregressive process, provided the population parameter T_0 is known. Unfortunately, this parameter is unknown in practical applications. One natural idea is to substitute an estimate of the time scale T_0 derived from the sample autocorrelation function. When this substitution is done, the histogram of F values derived

from the Monte Carlo experiments no longer match the predicted distribution (not shown). This result implies that estimation of the time scale T_0 is a critical step in the Shukla-Gutzler method. Thiébaux and Zwiers (1984) examine the problem of estimating T_0 and conclude that it “is quite difficult to estimate reliably.” Trenberth (1984) has also shown that, in the absence of potential predictability, improved estimates of the autocorrelation function can be obtained by subtracting the mean of the entire dataset, rather than subtracting a different mean for each subsample (i.e., season). Unfortunately, this approach is

problematic if potential predictability exists, for then the potentially predictable signal contaminates the autocorrelation function and leads to overestimates of T_0 .

We have investigated a variety of approaches to estimating T_0 , but none of them proved satisfactory. Nevertheless, some robust results were obtained that are worth reporting. These results also can be found in other papers (Thiébaux and Zwiers 1984; Trenberth 1984), but do not appear to be widely appreciated. The above considerations suggest at least four methods for estimating T_0 :

ACF-SEAS: Estimate T_0 by substituting in (12) the sample autocorrelation function $\rho_\tau = c_\tau^s/c_0^s$, where the covariances are computed from residuals from the seasonal mean:

$$c_\tau^s = \frac{1}{DB} \sum_{b=1}^B \sum_{d=1}^{D-\tau} (O_{d+\tau,b} - \bar{O}_b)(O_{d,b} - \bar{O}_b). \quad (26)$$

AR1-SEAS: Estimate T_0 by fitting seasonal anomalies to a first-order autoregressive model, and then substituting the autocorrelation function derived from the model into (12). In this case, T_0 is estimated by substituting into (12):

$$\rho_\tau = \left(\frac{c_1^s}{c_0^s} \right)^\tau. \quad (27)$$

ACF-GRAN: Estimate T_0 by substituting in (12) the sample autocorrelation function $\rho_\tau = c_\tau^g/c_0^g$, where covariances are computed from residuals relative to the grand mean:

$$c_\tau^g = \frac{1}{DB} \sum_{b=1}^B \sum_{d=1}^{D-\tau} (O_{d+\tau,b} - \bar{O})(O_{d,b} - \bar{O}). \quad (28)$$

AR1-GRAN: Estimate T_0 by fitting “grand anomalies” to a first-order autoregressive model, and then substituting the autocorrelation function derived from the model into (12). In this case, T_0 is estimated by substituting into (12):

$$\rho_\tau = \left(\frac{c_1^g}{c_0^g} \right)^\tau. \quad (29)$$

The resulting estimates computed from 1000 realizations of the autoregressive process for three different values of the AR parameter are shown in Fig. 3. The figure reveals that estimating the autocorrelation function (ACF) from residuals with respect to the season mean systematically underestimates the true time scale, often by a factor of 3 or more. This bias can be reduced (but not eliminated) by

fitting seasonal anomalies to a first-order autoregressive (AR1) model and then determining the time scale of the resulting AR1 model. Similar procedures applied to residuals with respect to the grand mean give even more accurate results, consistent with Trenberth (1984).

To explore the case of potential predictability, we add to $X_{d,b}$ a signal term of the following form:

$$O_{d,b} = X_{d,b} + \mu_b, \quad (30)$$

where μ_b is a random variable drawn independently for each b from a normal distribution with zero mean and variance $\text{SNR}/(1 - \phi^2)$, where SNR is a specified signal-to-noise ratio. The corresponding estimated values of T_0 are shown in Fig. 4. As anticipated, estimations based on residuals with respect to the grand mean consistently overestimate the true time scale, by as much as a factor of 2–10, depending on how the autocorrelation is estimated. In addition, estimates based on residuals with respect to the seasonal mean systematically underestimate the true time scale, as found in the case with no potential predictability.

The above biases in estimating T_0 have important implications on the Shukla–Gutzler method. Specifically, when T_0 is underestimated, then the F statistic in (22) is *inflated*, and the degrees of freedom are overestimated, both of which cause the Shukla–Gutzler method to overstate the degree of predictability. However, overestimation of the degrees of freedom has marginal impact on the significance test, as shown in the previous section. Much more important is the fact that the F statistic in (22) is inversely proportional to T_0 , so a factor of 2–10 error in T_0 leads to a similar error in F , which could easily compromise the significance test. For a perspective on the seriousness of this bias in practical applications, see X. Feng et al. (2011, unpublished manuscript).

6. A spectral interpretation of the Shukla–Gutzler method

To understand why the time scale T_0 tends to be underestimated, it is useful to consider a spectral interpretation. Furthermore, it is useful to consider the more general estimate

$$\hat{T}_0 = 1 + 2 \sum_{\tau=1}^M \hat{\rho}_\tau \left(1 - \frac{\tau}{D} \right), \quad (31)$$

where $M < D$. Shukla and Gutzler (1983) proposed this estimate and suggested using $D = 30$ days and an upper limit of summation between $M = 15$ and $M = 30$ days. Given a single time series of length N , the sample autocorrelation function $\hat{\rho}_\tau$ is related to the normalized

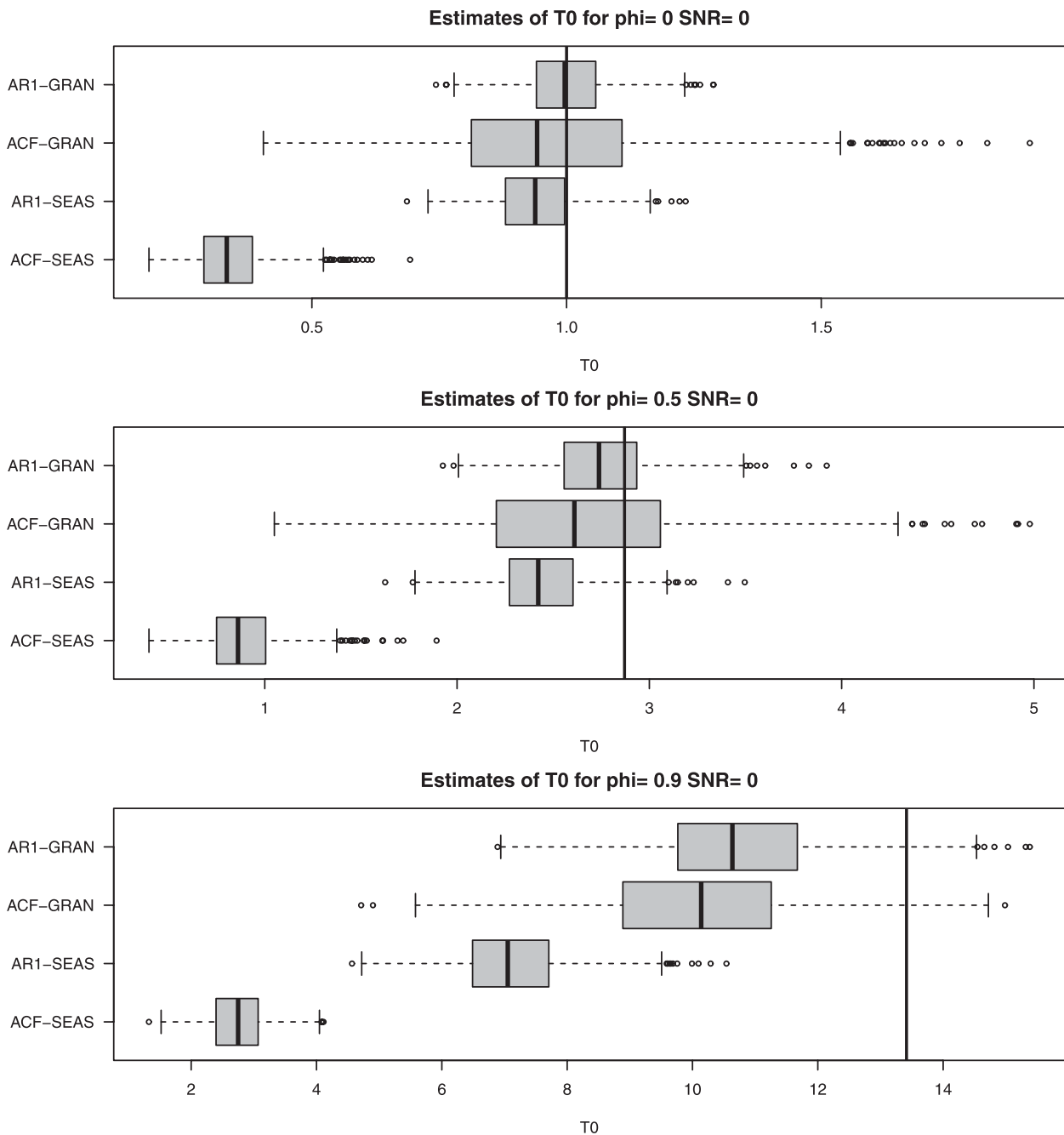


FIG. 3. Box-and-whisker plots of estimates of the effective time scale T_0 computed by four different methods from 1000 realizations of time series generated by an autoregressive model. The four methods are defined in section 5, and results are shown for three different values of the AR parameter: (top) $\phi = 0$, (middle) $\phi = 0.5$, and (bottom) $\phi = 0.9$. The box-and-whisker plots show the median as the centerline in the rectangle, the first and last quartile as the ends of the rectangle, and 1.58 times the interquartile range as the ends of the error bars.

periodogram $\hat{p}(\omega)$ at frequency ω by a discrete Fourier transform pair (Cryer and Chan 2010):

$$\hat{p}(\omega) = \sum_{\tau=-(N-1)}^{N-1} \hat{\rho}_\tau \cos(\omega\tau) \quad (32)$$

$$\hat{\rho}_\tau = \frac{1}{\pi} \int_{-\pi}^{\pi} \hat{p}(\omega) \cos(\omega\tau) d\omega. \quad (33)$$

Assuming a single time series, substituting (33) into (31) and manipulating it gives

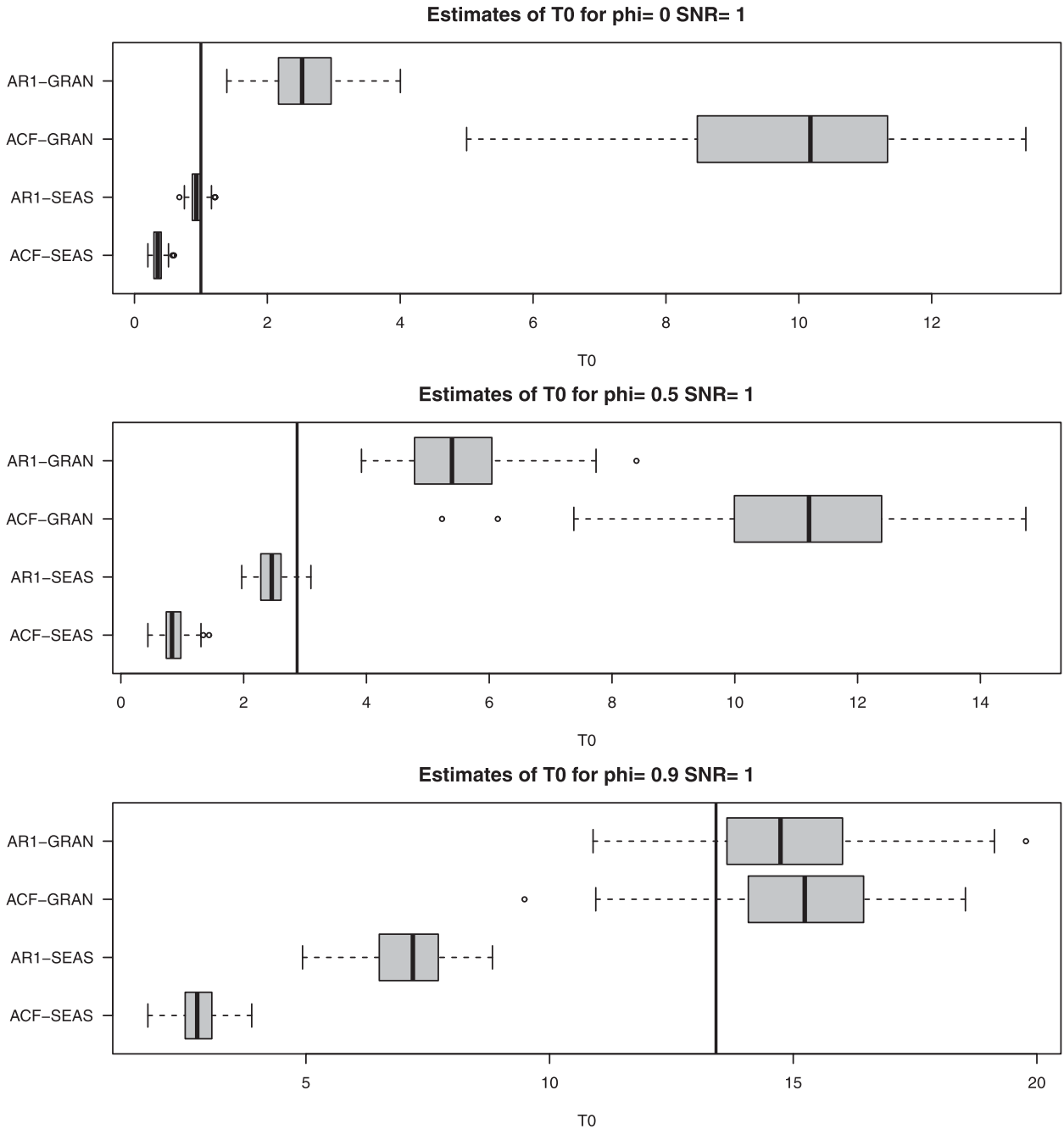


FIG. 4. Box-and-whisker plots of estimates of the effective time scale T_0 as in Fig. 3 but for time series generated from (30), with a signal-to-noise ratio of 1.

$$\hat{T}_0 = \int_{-\pi}^{\pi} \hat{p}(\omega) \lambda_M(\omega) d\omega, \quad (34)$$

where $\lambda_M(\omega)$ is a set of spectral weights, called the *spectral window*:

$$\lambda_M(\omega) = \frac{1}{\pi} \sum_{\tau=-M}^M \cos(\omega\tau) \left(1 - \frac{|\tau|}{D}\right). \quad (35)$$

The result in (34) shows that time scale T_0 estimated from (31) can be interpreted as a *smoothed* periodogram estimate of the power spectrum. The spectral window for selected values of M is shown in Fig. 5. We see that while the window applies the largest weights to the periodogram at low frequencies, it also applies *negative* weights at moderately low frequencies. Negative weights are problematic because they lead to *underestimates* of the true time scale

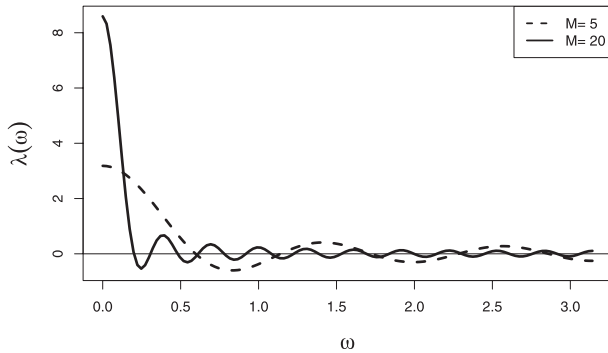


FIG. 5. The scaled spectral window associated with the Shukla–Gutzler method.

T_0 , and in some cases to negative values of T_0 . The negative weights are an artifact of the sharp cutoff in the summation in (31) and can be avoided by applying weights that gradually decrease to zero (Priestley 1981, p. 434). This approach, in turn, can be interpreted as modifying the spectral window $\lambda_M(\omega)$, which in turn can be viewed as a generalization of the method of Madden (1976), which puts all weight on the lowest possible resolved frequency.

7. Summary and conclusions

This paper reexamined a procedure proposed by Shukla and Gutzler (1983) for testing potential predictability of seasonal means using techniques based on analysis of variance. Certain subtle and unverified assumptions that are required for the Shukla–Gutzler method to work, such as the fact that certain quantities follow the chi-square distribution and that the time mean is independent of its residuals, even for autocorrelated data, were clarified. Moreover, Monte Carlo experiments demonstrate that these assumptions are adequate even for autocorrelated data, provided the effective time scale T_0 is known. We derive the same F statistic in (22) for testing potential predictability as Shukla and Gutzler (1983), but find that the degrees of freedom differ from that proposed by Shukla and Gutzler (1983). On the other hand, in many practical applications, the degrees of freedom are already so large that the correction has relatively little effect on the overall significance test. Unfortunately, the effective time scale is generally not known and hence must be estimated from data. This estimation problem has received considerable attention in previous papers with no clear solution. Our Monte Carlo experiments highlight the fact that estimates of T_0 can be biased by a factor of 3–10, depending on the method. The question of how the potential predictability estimated from the corrected Shukla–Gutzler method compares with other estimates will be considered in Feng et al. (2012, manuscript submitted to *J. Geophys. Res.*).

Acknowledgments. Support is gratefully acknowledged from grants from the NSF (Grant 0830068), the National Oceanic and Atmospheric Administration (Grant NA09OAR4310058), and the National Aeronautics and Space Administration (Grant NNX09AN50G).

REFERENCES

- Box, G. E. P., G. M. Jenkins, and G. C. Reinsel, 2008: *Time Series Analysis: Forecasting and Control*. Wiley-Interscience, 746 pp.
- Brockwell, P. J., and R. A. Davis, 1991: *Time Series: Theory and Methods*. 2nd ed. Springer-Verlag, 577 pp.
- Chervin, R. M., 1986: Interannual variability and seasonal climate predictability. *J. Atmos. Sci.*, **43**, 233–251.
- Cryer, J. D., and K.-S. Chan, 2010: *Time Series Analysis with Applications in R*. Springer, 491 pp.
- Feng, X., T. DelSole, and P. Houser, 2011: Bootstrap estimated seasonal potential predictability of global temperature and precipitation. *Geophys. Res. Lett.*, **38**, L07702, doi:10.1029/2010GL046511.
- , —, and —, 2012: A method for estimating potential seasonal predictability: Analysis of covariance. *J. Climate*, **25**, 5292–5308.
- Kumar, A., and M. P. Hoerling, 1995: Prospects and limitations of seasonal atmospheric GCM predictions. *Bull. Amer. Meteor. Soc.*, **76**, 335–345.
- Lorenz, E. N., 1963: Deterministic nonperiodic flow. *J. Atmos. Sci.*, **20**, 130–141.
- Madden, R. A., 1976: Estimates of the natural variability of time-averaged sea-level pressure. *Mon. Wea. Rev.*, **104**, 942–952.
- Majda, A. J., I. Timofeyev, and E. Vanden-Eijnden, 2001: A mathematical framework for stochastic climate models. *Commun. Pure Appl. Math.*, **54**, 891–974.
- Peng, P., and A. Kumar, 2005: A large ensemble analysis of the influence of tropical SSTs on seasonal atmospheric variability. *J. Climate*, **18**, 1068–1085.
- Phelps, M. W., A. Kumar, and J. J. O’Brian, 2004: Potential predictability in the NCEP CPC dynamical seasonal forecast system. *J. Climate*, **17**, 3775–3785.
- Priestley, M. B., 1981: *Spectral Analysis and Time Series*. Academic Press, 890 pp.
- Sardeshmukh, P. D., G. P. Compo, and C. Penland, 2000: Changes of probability associated with El Niño. *J. Climate*, **13**, 4268–4286.
- Shukla, J., 1981: Dynamical predictability of monthly means. *J. Atmos. Sci.*, **38**, 2547–2572.
- , and D. S. Gutzler, 1983: Interannual variability and predictability of 500 mb geopotential height over the Northern Hemisphere. *Mon. Wea. Rev.*, **111**, 1273–1279.
- Thiébaux, H. J., and F. W. Zwiers, 1984: The interpretation and estimation of effective sample size. *J. Climate Appl. Meteor.*, **23**, 800–811.
- Trenberth, K. E., 1984: Some effects of finite sample size and persistence on meteorological statistics. Part I: Autocorrelations. *Mon. Wea. Rev.*, **112**, 2359–2368.
- Zheng, X., H. Nakamura, and J. A. Renwick, 2000: Potential predictability of seasonal means based on monthly time series of meteorological variables. *J. Climate*, **13**, 2591–2604.
- Zwiers, F. W., 1987: A potential predictability study conducted with an atmospheric general circulation model. *Mon. Wea. Rev.*, **115**, 2957–2974.