Prediction of Seasonal Climate in a Low-Dimensional Phase Space Derived from the Observed SST Forcing

RISHENG WANG

Recherche en Prévision Numérique, Environment Canada, Dorval, Quebec, Canada

(Manuscript received 16 February 1998, in final form 10 October 1999)

ABSTRACT

A methodology is presented to make an optimal use of the global SST for the prediction of seasonal climates. First, the space–time extended principal component analysis was applied to the key SST forcing regions, such as the tropical Pacific and the Atlantic, to establish a low-dimensional phase space model. This allows a nonlinear prediction, in terms of analogs found in the nearest neighborhood of the state associated with the initial time of prediction. Second, the predicted results derived independently from those different SST forcing regions are then linearly combined using the best linear unbiased estimates based on all available verification periods under a cross-validation scheme. This enables optimal use of the predictive skills inherent to each of the key SST forcing regions for each climate zone. The proposed methodology is justified by the analysis of the origins of predictive skills for seasonal predictions based on SST predictors (the geographical distribution of the skill scores and their time changes). Application was made to the prediction of winter (December–January–February) surface air temperatures over North America, based on the observed monthly mean data from January 1949 to December 1996. Significant skill scores were found over most parts of North America. The superiority of nonlinear prediction was demonstrated. It is concluded that the low-dimensional phase space approach may be used as an effective tool for seasonal forecasting.

1. Introduction

Seasonal climate may be considered as the statistical state of the daily weather for a given season over a specific geographical region. In practice, 3-month means are commonly used. Thus, the timescale of seasonal climate goes far beyond the predictability that is defined in terms of sensitive dependence on the initial atmospheric conditions. Instead, the predictability of seasonal climate is often connected with a forcing field such as the SST associated with the ENSO. The key to a truly successful application of an empirical model lies in the understanding of the underlying physical mechanism for the relation between the predictor and the predictand fields. Unlike the dynamic models that try to answer how a certain anomaly occurs by simulating the detailed processes that are necessary to produce the observed seasonal anomaly, the statistical models directly try to determine the probability that a certain anomaly will occur over a specific geographical region under a known condition, in particular, the space–time structure of a given forcing, such as the El Niño signals in the SST field. The effectiveness of empirical models, therefore, depends crucially on whether the relevant components (with respect to space and time scales) to be used as predictors are suitably incorporated into the prediction models, and whether the relationships between the predictors and predictands (which may not necessarily be linear, particularly when mid- and high latitudes are concerned) are properly established.

The space–time extended principal component (ST-PC) models, proposed by Vautard et al. (1996) and based on multiple singular spectra analysis (MSSA, Plaut and Vautard 1994), take the advantages of both space–time extended EOF analysis (Weare and Nasstrom 1982; Fraedrich et al. 1993) and the analog approaches (Barnett and Preisendorfer 1978; Bergen and Harnack 1982; Livezey and Barnston 1988). First, the major source of information contained in the predictor field is condensed in a relatively low-dimensional phase space, which is spanned by the leading ST-PCs. Second, analogs are sought in the phase space spanned by these ST-PCs to keep the most relevant information for the subsequent prediction. The most significant feature of the ST-PC models is the combination of ST-PC analysis and analog approach. However, the analog method is effective only when the dimension is sufficiently low because in a high-dimensional phase space, an extremely large number of samples is required to define a good analog (Van den Dool 1994a). Natural analogs are therefore difficult

Corresponding author address: Dr. Risheng Wang, % Prof. Bin Wang, Department of Meteorology and IPRC, SOEST, University of Hawaii, 2525 Correa Rd., Honolulu, HI 96822.
E-mail: rwang@philmultic.com

to find. In fact, even if a good natural analogue could be found, it would not necessarily yield good forecasting on a seasonal scale, due to the intrinsic limit of predictability and to the fact that the natural analogs found in physical space involve many different space–time scales that are not essential to seasonal climate. This, however, does *not* mean that the idea of analog is invalid. For instance, Barnett and Preisendorfer (1978) proposed a theoretical frame, on the basis of analogs for short-term climate prediction, which is conceptually clear and physically based.

The present paper attempts to apply the idea of analogs in a low-dimensional phase space. In terms of low-dimensional phase space approach, the analog takes the form of nearest neighborhood of the initial time of prediction, with a clear awareness of the geometrical structure of the trajectories and the associated probability density distribution. As specific meaning is assigned to the variability represented by the structure of the trajectories in the phase space, the statistical forecasting, in this sense, goes beyond an application of some ''black box'' statistical tools such as the EOF-based linear models with arbitrary truncations or the neural network. For convenience, the term ''nearest neighborhood'' will be used as a synonym for the analogs that fall into a prescribed neighborhood of the state associated with the initial time of prediction (see section 3 for details).

The motivation for such an approach came from the following facts. First, when the global SST is used directly in an EOF-based prediction scheme, the ENSO signal is so dominant that the contribution (in terms of loading) from other regions is often underestimated. Moreover, one has to deal with a more or less arbitrary truncation of the EOFs and a relatively large number of principal components, in order to include signals from other key regions of SST forcing. This leads to a rather high-dimensional phase space, which prevents an effective application of an analog method, due to the limitation of sample size. Second, we have observed that the time evolution of a specific key SST forcing region, relevant to the seasonal scales, can often be effectively described in a low-dimensional phase space. In particular, the observed ENSO signals can be described in a three-dimensional phase space derived from tropical Pacific SST fields, implying that the time evolution of the ENSO may be basically a low-dimensional attractor, as indicated by the results derived from the dynamic modeling and theoretical studies (see Wang and Fang 1996 and references therein). The possibility for constructing a low-dimensional phase space from the predictor fields allows a closer examination of their impact on the seasonal climate over a specific climate zone, in terms of phase relations, leading to the possibility of nonlinear and more dynamically based statistical prediction of seasonal climates (in contrast to pure statistical approaches). Finally, the geographical distribution of the predictive skills, as well as their time behavior, varies from one key region of SST forcing to another, such that the final prediction can be optimized through a linear or nonlinear combination of the predicted results derived from different key forcing regions.

The methodology introduced in the present paper consists of the following two steps. First, the ST-PC models (Vautard et al. 1996, 1999) are applied separately to individual SST forcing regions, such as the tropical Pacific and the Atlantic, so as to take advantage of nonlinear predictions in low- (typically three-) dimensional phase space. Second, an optimal linear combination of the predictions that resulted from each key SST forcing region is carried out, to minimize the prediction error, using the best linear unbiased estimates (BLUE; see Sarda et al. 1996 and the reference therein). As in Vautard et al. (1996), the predictive skills are evaluated under a cross-validation scheme. Our preliminary application of the present method to the prediction of the winter [December–January–February (DJF)] surface air temperature (SAT) over North America showed significant improvement in the skill scores, both in geographical distribution and in time behavior. The causes for the present methodology to yield the observed improvements were explored.

The paper is structured as follows. The methodology will be presented in section 3. Analyzed in sections 4 and 5 are the predictions of the SATs over North America based on predictors derived from the tropical Pacific, Atlantic, and extratropical North Pacific SSTs, respectively. In section 6, issues regarding the predictors derived from global SST are discussed, and the improved skill scores are presented as a result of optimal linear combination of the predictions derived from the above three regions. Concluding remarks are given in section 7.

2. Data

The global monthly mean sea surface temperature and the surface air temperature used in the present study are derived from the same source as in Vautard et al. (1999). The predictand is the seasonal SAT observed at 164 stations in North America (105 in the United States and 59 in Canada). The original daily data are provided by the Climate Prediction Center. Monthly mean SATs were first derived by averaging the daily measurements. A monthly record is regarded as missing when more than one-third of the daily records are missing. The seasonal SAT is the consecutive 3-month mean. In a similar way, a seasonal SAT record is considered as missing if any of the three months' records are missing. The missing data are not interpolated. Because the prediction is carried out station by station, this does not affect the model building. The seasonal dataset is quite complete, covering a time period of 48 years (January 1949–December 1996). The amount of missing seasonal records for all 164 station is 0.43%.

The predictor field is the monthly mean global SST, derived originally from the Comprehensive Ocean–Atmosphere Data Set archive, which were interpolated on

$10^\circ \times 10^\circ$ grids, as in Vautard et al. (1996). The dataset covers a period from January 1949 to July 1997. The annual cycle and the interannual linear trends are removed from both the predictor (SST) and predictand (SAT) fields, as in Vautard et al. (1999). The anomalies (departure from the climate means) are used for all prediction models.

The data used are divided into learning and verification sets, denoted by Ω_L and Ω_V, respectively. All calculations for the subsequent model building are derived from Ω_L only. The data in Ω_V are used only for evaluating the predictive skill.

It should be pointed out that the linear trend, calculated on each grid after the annual cycle is removed, is related either to artificial factors (associated with measurements) or to variabilities whose timescales are equivalent to or larger than a century, which is not relevant to seasonal prediction. The amplitude of the linear trend is very small, however, and may give rise to a shifting of the center of the trajectories in phase space and thus affect the search for closest neighbors for a nonlinear prediction. Therefore, it is a necessary step, in the present context, to remove the linear trend. This does not influence the dynamics of the seasonal to interannual scales that are relevant to seasonal prediction.

3. Methodology

The basic ideas of the present methodology are derived from the ST-PC models (Vautard et al. 1996, 1999) and the hybridization of different models using the BLUE (see Sarda et al. 1996 and references therein), except that the nonlinear prediction plays a central role, in the present context, due to the use of a low-dimensional phase space constructed from the key SST forcing regions. Nonlinear prediction in low-dimensional space has been an active research area in recent decades [see, e.g., Elsner and Tsonis (1992) and Abarbanel et al. (1993) for a review]. In the present context, however, the phase space is a coordinate system describing the time evolution of the predictors, which in turn are used to forecast the predictand, according to their phase relation and probability measure. In the following, we shall present the methodology in three steps: namely, the construction of a proper low-dimensional phase space from a given predictor field, the establishment of the prediction models, and the optimization of the predictions from different sources.

a. Construction of a low-dimensional phase space

Many studies indicate that the major source of the predictive skills for seasonal climate is derived from the SST field (Barnett 1981; Barnett and Preisendorfer 1987; Barnston 1994; Barnston and Smith 1996; Vautard et al. 1996). However, it is difficult to embed the global SST variability in a low-dimensional phase space, because each key SST forcing region has its own relative independent part of variability on the seasonal scales. Thus, to work in a low-dimensional phase space, it is necessary to divide the global SST field into a few key forcing regions, according to their dominant space–time patterns that have a certain impact on seasonal climate.

The concept of phase space, used in the present context, is derived from the dynamical system theory and the associated time series analysis [see Eckmann and Ruelle (1985), Sauer et al. (1991), and Abarbanel et al. (1993) for reviews, and Wallace et al. (1993), Fraedrich et al. (1993), and Wang et al. (1995) for atmospheric applications]. The starting point, assuming that the observed low-frequency variabilities are derived from a low-dimensional attractor, is what we can learn about the underlying system based on observational data and how much we can benefit for the prediction of seasonal climate.

As the original phase space is unknown, the first step is the reconstruction of a phase space based on observed variables. The clue for such a possibility is provided by the embedding theorems (Whitney 1936; Takens 1981; Sauer et al. 1991). Because the variabilities observed in the atmosphere–ocean system are wave motions of various scales, it is reasonable to consider a phase space as some manifolds. According to the embedding theorem of Whitney (1936), for a low-dimensional manifold, any linearly independent variables derived from the system can serve as coordinates that span the phase space. In another words, a phase space can be understood as arbitrary coordinates that describe the underlying system, with the following two constraints [see Sauer et al. (1991), for a precise definition].

1) The trajectories in the reconstructed phase space should be approximately on a manifold. There should be no intersection of trajectories, that is, two trajectories can be very close but never meet (uniqueness). This also means that distinctively different states in physical space must be distinctively separated in phase space. To avoid crossing of trajectories of a chaotic system, the minimum dimension should be three.
2) The direction of time must be preserved such that the trajectories can be described on a manifold with definite orbit structures. In this way, the direction in which the trajectory evolves in phase space is a one-to-one correspondence to the direction in which an observed wave propagates in physical space (e.g., Wang 1994; Fraedrich et al. 1993; Wang et al. 1995).

In reality, such a low-dimensional attractor may not even exist to the precision of its very definition, or it may exist, but in a higher-dimensional phase space. However, this does not need to deter us from using a low-dimensional manifold to approximate a specific part or a particular aspect of the observed dynamics or physical reality (Sauer et al. 1991; R. Wang 1994). The subsequent application in section 4 indeed suggests that the

variability in the SST field associated with El Niño–La Niña can be effectively embedded into a three-dimensional phase space [see Wang and Wang (2000) for details]. The resulting prediction may serve here as a measure of how good such an assumption is.

Apparently, there are many different ways to construct a phase space (e.g., Sauer et al. 1991; Fraedrich and Wang 1993; Fraedrich et al. 1993; R. Wang 1994). The phase space used in the present context is considered as a hyperplane (or a manifold), spanned by the relevant space–time extended principal components derived from the predictor field. The ST-PCs are the projection of the SST anomaly field, over a specific region, onto the space–time extended EOFs (ST-EOFs), which are derived from diagonalizing the cross-time-lagged covariance matrix (Weare and Nasstrom 1982; Wang 1991; Fraedrich et al. 1993), which is equivalent to the MSSA (Plaut and Vautard 1994). The ST-EOFs serve here only as a prescribed basis (coordinates) for the phase space, $\mathbf{\Omega}_L \subset \mathbf{R}^m$, containing all the observed states of the SST. In the subsequent discussion, $\mathbf{y}_\zeta \in \mathbf{\Omega}_L \subset \mathbf{R}^m$ represents the projection of the SST anomaly field onto m ST-EOFs, that is, a point in the m-dimensional phase space spanned by these ST-EOFs at the time ζ.

The maximum time lag of covariance matrix is also referred to as the time window w, which is chosen to be 12 months for seasonal forecasting purposes. As the ST-EOF patterns contain both spatial and temporal dependencies, the corresponding ST-PCs reflect the spatial–temporal correlated variability of the monthly SST field. Thus the values of the ST-PCs at a specific month, $\mathbf{y}_\zeta \in \mathbf{\Omega}_L \subset \mathbf{R}^m$, define a state point in the m-dimensional phase space, equivalent to a space–time coherent structure in physical space, reflecting the statistical state of the SST for the past 12 months corresponding to the ST-EOFs. This may be what we need for seasonal prediction; the seasonal climate is a statistical state of the daily weather for the whole season, which may not necessarily be influenced so much by an SST anomaly at a particular time, but by how the anomaly evolves during the past year. Note that, in such a long time span, the observed time evolution of the SST field also carries information about the influence from the atmospheric motion.

Now the question is, Which ST-PCs should be considered and how many variables should be retained? It must be stressed that, unlike the classical (ST-)EOF-based statistical analysis, with a more or less arbitrary truncation concerning the number of EOFs retained, in the present context of phase space reconstruction, only those ST-PCs are chosen which are found to be relevant to the seasonal scales by proper diagnostic tools, such as spectra or cross-time-lagged correlation analyses. Thus the order of the (ST-)EOFs or (ST-)PCs just serves as a convenient mark for notation and does not play any role in phase space reconstruction. The only decisive factor is the dynamics that the ST-PCs represent. Considering that the coordinates of the same manifold are coupled variables in the same sense as a simple linear oscillator, $u = \cos(t)$ and $v = \sin(t)$ in the phase space spanned by (u, v), the ST-PCs that describe the same manifold should be significantly correlated at the lag time equivalent to a quarter of its averaged cycle (Wang et al. 1995 and the references therein). This is the basic reason for using the cross-time-lagged correlation of the ST-PCs to determine which ST-PCs should be retained for phase space reconstruction (R. Wang 1994).

The nearest neighborhood of a given state $\mathbf{y}_t \in \mathbf{\Omega}$ is defined as the subset $\Delta(\mathbf{y}_t) \subset \mathbf{\Omega}_L$, whose elements $\{\mathbf{y}_\zeta \in \Delta(\mathbf{y}_t), \zeta = 1, 2, \ldots, N(\delta)\}$ satisfy the condition $|\mathbf{y}_t - \mathbf{y}_\zeta| < \delta$, where $\zeta \neq t$ is a generic time, $\delta > 0$ is a prescribed value, and $N(\delta)$ is the maximum number of neighbors. To ensure that $\Delta(\mathbf{y}_t) \subset \Omega_L$ is as much along the manifold as possible, the phase space is renormalized such that the geometrical shape of $\Delta(\mathbf{y}_t) \subset \Omega_L$ becomes a three-dimensional elliptic ball, with the flat side perpendicular to the radius.

b. Prediction models and predictability

The prediction models can be generalized as a relation that projects the initial state $\mathbf{y}_t$ onto the future state of the predictand, $\Theta_{t+\tau}$, under a certain function Γ;

$$\Theta_{t+\tau} = \Gamma(\mathbf{y}_t) + \epsilon, \qquad \mathbf{y} \in \mathbf{\Omega} \subset \mathbf{R}^m, \tag{1}$$

where $\mathbf{\Omega}$ is the phase space containing all possible states of $\mathbf{y}$. The initial time of prediction is denoted by t, with the lead time $\tau \geq 0$, defined as the time span between the end of the latest observation and the beginning of the predictand (Barnston 1994). The parameter ϵ is associated with the prediction errors. Note that, in the present context, Θ is the anomaly of a seasonal SAT at a given station. The value of Θ can be continuous or categorical numbers, depending on whether the prediction is on a real or categorical basis. The establishment of Γ changes according to whether the prediction is on a linear or nonlinear basis.

For linear prediction, the projection function Γ is commonly specified by a linear regression:

$$\hat{\Theta}_{t+\tau} = A + \mathbf{B} \cdot \mathbf{y}_t^{\mathrm{T}}, \tag{2}$$

where $\hat{\Theta}$ is the predicted value of Θ (in contrast to the observed), $\mathbf{y}^{\mathrm{T}}$ is the transpose of $\mathbf{y} = (y_1, y_2, \ldots, y_m) \in \mathbf{\Omega}$, $\mathbf{B} = (b_1, b_2, \ldots, b_m)$ is the coefficient matrix, and A is a constant. Both A and $\mathbf{B}$ are usually derived from minimizing the forecast error ϵ in (1), based on the *whole* learning set Ω_L. The source of predictability here is derived directly from the linear correlation between $\mathbf{y}$ and $\Theta_{t+\tau}$. There are other forms of linear models, but most of them can be described in the form of (2), by an appropriate preparation (e.g., transformation) of $\mathbf{y}$ and $\Theta_{t+\tau}$. Thus, the linear prediction for a given predictand (Θ, at a station) relies on a linear relation (2), optimized along a single line in the phase space Ω,

which distinguishes only two different phases of opposite poles.

In contrast, the nonlinear prediction realizes that the relation (1) may depend on the current phase, which may not happen to be on the line derived from the linear regression (2). In particular, there may be regime behavior that linear relation cannot handle properly. To tackle this problem, the relation (1) has to be established such that Γ changes according to the geometrical position of $\mathbf{y}_t \in \mathbf{R}^m$. This is possible when the establishment of the relation (1) is based on the nearest neighborhood $[\Delta(\mathbf{y}_t)]$ of the initial state $\mathbf{y}_t \in \Omega_L$, instead of the entire learning set Ω_L, as in (2).

There are apparently different ways to formulate a nonlinear prediction scheme [see Eckmann and Ruelle (1985); Elsner and Tsonis (1992); Abarbanel et al. (1993) for a review]. The nearest neighborhood approach is, however, one of the most straightforward methods, with clear physics in terms of regimes and phase relations, which enables us to analyze the origin of predictive skills, as well as how and why the skills change with the phase. Neural network, for instance, is a powerful nonlinear scheme based on ''black box'' statistics, where one can tune the parameters to arrive at a good prediction (e.g., Elsner and Tsonis 1992; Tangang et al. 1998), but can see neither the phase relation between the predictand and predictors, nor the origin of skills.

Assume that there is a one-to-one correspondence between a given state $\mathbf{y}_\zeta \in \Omega_L$ and the value of the predictand $\Theta_{\zeta+\tau}$ at a given station. The predictability at $\mathbf{y}_t$ may be defined as the probability of finding $\Theta_{\zeta+\tau}$ in the same categorical state as $\Theta_{t+\tau}$ for all $\mathbf{y}_\zeta \in \Delta(\mathbf{y}_t) \subset \Omega_L$ and $\zeta \neq t$. It is thus clear that the predictability may change with the geometrical position of $\mathbf{y}_t$ or with the time t (as the state $\mathbf{y}_t$ evolves with t), which can be easily demonstrated in phase space [see Ziehmann-Schlumbohm et al. (1995) for an example with the Lorenz model]. The predictability for Θ on the basis of $\mathbf{y}$ may be measured by the ensemble average of the local predictability for all available $\mathbf{y}_t \in \Omega$.

The central point for nonlinear prediction is the establishment of a linear relation based on all $\mathbf{y}_\zeta \in \Delta(\mathbf{y}_t) \subset \Omega_L$ $(\zeta \neq t)$ and the corresponding predictand $(\Theta_{t+\tau})$ to give a realistic description of the phase relation between the forcing field and the predictand. This relation typically changes with the initial state y_t. In another words, a nonlinear model (1) in the present context is locally linear but globally nonlinear. As a preliminary application, the most simple kind of linear relation, the composite anomaly, is assumed in the present study, which means that the output of the nonlinear prediction is the composite anomaly of the seasonal SAT $(\Theta_{\zeta+\tau})$, corresponding to all neighbors in $\Delta(\mathbf{y}_t)$:

$$\hat{\Theta}_{t+\tau} = \frac{1}{N(\delta)} \sum \Theta(\mathbf{y}_\zeta), \quad \text{for all } \mathbf{y}_\zeta \in \Delta(\mathbf{y}_t) \subset \Omega_L,\ \zeta \neq t, \tag{3}$$

where $N(\delta)$ is the number of nearest neighbors in $\Delta(\mathbf{y}_t) \subset \Omega_L$. As $N(\delta)$ changes with δ, one often prefers a fixed number of neighbors to fixed size δ. For this simple model, we found that the prediction is optimal when the size of $\Delta(\mathbf{y}_t)$ is taken as 15% of the total samples in Ω_L [section 4b(2)].

For categorical forecasting, the probability of each of the three categories, A, N, and B (representing above, near, and below normal, respectively), is estimated using all $\mathbf{y}_\zeta \in \Delta(\mathbf{y}_t)$ $(\zeta \neq t)$. This is used in the present paper as a diagnostic tool to reveal the origin of the predictive skill. For real application, we shall focus only on the composite anomaly (3).

Now, one may ask, is there any dynamic or physical basis for the proposed nonlinear scheme? The answer is apparent; each state, $\mathbf{y}_t$, in the phase space corresponds to a space–time coherent structure in the physical space. As the ST-PCs are projections of the predictor field onto its ST-EOF basis, with a time window of w months, $\mathbf{y}_t \in \mathbf{R}^m$ defines a space–time coherent structure of the predictor field for the time period from $t - w + 1$ to t. Given a specific time evolution of the spatial distribution of the predictor field for the last $w = 12$ months, the prediction equations (1) or (4) yield a spatial distribution of the predictand Θ (station by station) at lead time τ. In this sense, the proposed methodology is a kind of ''down-scaling'' forecasting scheme. The dynamic processes that relate SST anomaly and the seasonal mean SAT are not explicitly incorporated in the scheme. However, they are accounted for, in a sense, by the use of the nearest neighbors, under the assumption that similar (space–time) SST time evolution induces similar climate anomaly. For a definite set of predictors, we cannot expect a good prediction for every station. What we can hope for is that at least the stations over some specific geographical regions will be predicted with useful skills (see sections 4–6).

Predictions based on (1) are referred to as one-step models in Vautard et al. (1996). Alternatively, this can be decomposed into two steps by first extrapolating $\mathbf{y}_t \in \boldsymbol{\Omega}$ up to $\mathbf{y}_{t+\tau} \in \boldsymbol{\Omega}$ and then establishing the relation between $\mathbf{y}_{t+\tau} \in \Omega$ and $\Theta_{t+\tau}$:

$$\mathbf{y}_{t+\tau} = \Gamma_1(\mathbf{y}_t) + \boldsymbol{\epsilon}_1, \quad \text{and} \quad \Theta_{t+\tau} = \Gamma_2(\mathbf{y}_{t+\tau}) + \boldsymbol{\epsilon}_2. \tag{4}$$

Note that Γ_1 is a projection of $\Omega \subset \mathbf{R}^m$ onto itself, called the extrapolating stage, which can be carried out by nonlinear (see, e.g., Farmer and Sidorovich 1987; Casdagli 1991; Abarbanel et al. 1993) or linear methods (Vautard et al. 1996, 1999). Because there is no lead time involved, Γ_2 is called the specification stage. Accordingly, the methods following these two steps are referred to as two-step ST-PC models (Vautard et al. 1996). It should be pointed out that, though the one-step (1) and two-step (4) ST-PC models draw information from the same source, the two-step model may profit from the predictability inherent to the predictors whose timescales are typically larger (and thus more predictable) than those of the predictand. This is par-

ticularly true for longer lead time. In practice, however, the difference in the predictive skills between the two types of models may not be statistically significant for limited datasets. In certain instances, with linear (or close to linear) ST-PC models, the two-step models appear to be better than the one-step models (Vautard et al. 1996, 1999). Without losing generality, we shall stick to the one-step model (1) in the present study.

c. Optimal linear combination

At the final stage, the predictions are optimized according to

$$\hat{\Theta} = \alpha\hat{\Theta}_1 + (1 - \alpha)\hat{\Theta}_2 \quad \text{or}$$
$$\hat{\Theta} = \hat{\Theta}_2 + \alpha(\hat{\Theta}_1 - \hat{\Theta}_2), \tag{5}$$

where $\hat{\Theta}_1$ and $\hat{\Theta}_2$ are the predicted values based on predictors derived from two separate SST key forcing regions. Here, α is a weight derived from minimizing the prediction error $\epsilon = (\Theta - \hat{\Theta})^2$, based on K verification periods,

$$\alpha = \frac{\sum_{k=1}^{K} (\Theta^k - \hat{\Theta}_2^k)(\hat{\Theta}_1^k - \hat{\Theta}_2^k)}{\sum_{k=1}^{K} (\hat{\Theta}_1^k - \hat{\Theta}_2^k)^2}, \tag{6}$$

where K is the total number of verification periods under the cross-validation scheme and Θ is the observed value. So, α, calculated for each station, reflects the contribution of the corresponding key forcing regions to the skill score at that station [see section 6b(3) for further discussion].

d. Measures of predictive skills

Evaluation of the prediction skill is carried out under a cross-validation scheme (Michaelsen 1987; Barnston and Van den Dool 1993; Vautard et al. 1996) by successively removing a part of the data for verification (Ω_V), while using the rest (Ω_L) to build the models. In each case, one of the successive 2-yr periods starting from the initial time of prediction is removed from the available data, to guarantee independence of the experiments. For prediction of winter (DJF) SAT with 0 lead, the verification period is from October 1952 to September 1954; the next is from October 1953 and to September 1955, and so on, leading to a total number of $K = 44$ verification periods. Based on this total, the skill scores are evaluated and parameters (α) for the BLUE (5) are estimated for each of the $L = 164$ stations. Note that in evaluating the BLUE skill, the year to be predicted is not used for estimating α in (6), in order to avoid optimal ''fitting.''

The correlation skill score is used for examining the spatial behavior of a prediction model. Since the prediction is carried out station by station for all $L = 164$ stations over North America, and the resulting prediction is expressed in terms of anomaly, the skill score for station i is defined as the correlation between the observed (Θ) and the predicted $\hat{\Theta}$ SAT over all the K available verification periods:

$$\rho_i = \frac{\sum_{k=1}^{K} \Theta_{ki}\hat{\Theta}_{ki}}{\sqrt{\sum_{k=1}^{K} \Theta_{ki}^2 \sum_{k=1}^{K} \hat{\Theta}_{ki}^2}} \quad \text{for } i = 1, 2, \ldots, L. \tag{7}$$

For a significant test, Vautard et al. (1999) have estimated the correlation of two uncorrelated random processes using 1000 randomly generated realizations with 40 cases in each realization. It was found that the 95% significance level is at 0.26. Figures 1a and 1b give the results for persistence forecasting of winter (DJF) SATs

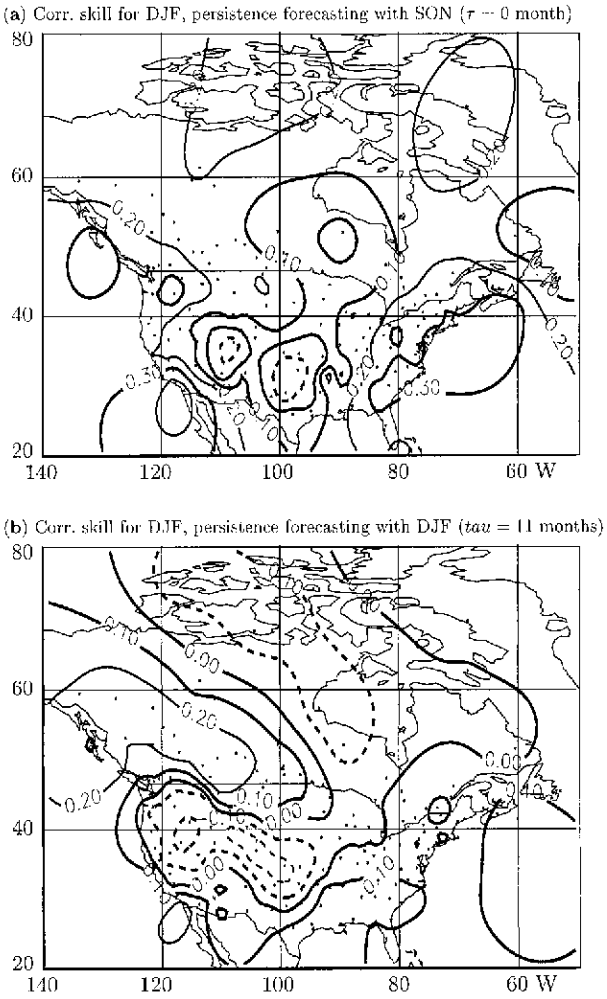

FIG. 1. Geographical distribution of the correlation skill scores derived from the persistence forecasting of the winter (DJF) SAT using (a) the preceding autumn (SON) SAT anomaly ($\tau = 0$) and (b) the preceding winter SAT anomaly ($\tau = 11$). The 95% significant level is 0.26. The stations used are marked with dots.

using the anomalies of the preceding autumn [September–October–November (SON), $\tau = 0$] and winter (DJF, $\tau = 11$), respectively. When judged by the above criterion, the persistence method displays a significant skill for regions along the Pacific and the Atlantic coasts for $\tau = 0$, whereas for $\tau = 11$ months, hardly any skill regions are observed. In the subsequent application, persistence forecasting with $\tau = 0$ (Fig. 1a) will be used as an objective model to compare with the predictive skills of the proposed models.

In order to show the time changes of the predictive skill, the pattern correlation between the predicted and observed SAT anomalies are calculated for the kth verification period:

$$s_k = \frac{\sum_{i=1}^{L} \Theta_{ki}\hat{\Theta}_{ki}}{\sqrt{\sum_{i=1}^{L} \Theta_{ki}^2 \sum_{i=1}^{L} \hat{\Theta}_{ki}^2}}, \tag{8}$$

where L is the total number of stations (predictands). It should be pointed out that the pattern correlation (8) is used only as a relative measure to reveal the time changes of predictive skills. Therefore, we did not consider the degree of freedom. In fact, pattern correlation is a rather misleading measure, particularly for seasonal prediction based on statistical methods. For instance, one set of predictors may give very good prediction for one climate zone, while the other zones are poorly predicted. In this case, the pattern correlation may be very bad, due to the parts which are poorly predicted. However, this does not need to invalidate the model, for no one would expect that such a simple statistical model could predict every zone correctly. This can be easily understood, as the key SST forcing region for one zone may not be the same for the other. This cannot be properly distinguished by pattern correlation. For such a purpose, the geographical distribution of the predictive skills is more relevant.

4. ENSO signals and the predictability of seasonal SAT

a. Extracting the ENSO signals

The ENSO signal is typically exhibited in the tropical Pacific SST field, with a large-scale SST anomaly pattern expanding and retreating from the east coast during its time evolution (e.g., Rasmusson and Carpenter 1982). The importance of the ENSO in seasonal prediction has been well established both from statistical studies (see, e.g., Walker and Bliss 1932; Bjerknes 1969; Barnett 1981; Barnett and Preisendorfer 1987; Barnston 1994) and from dynamic modeling [see Palmer and Anderson (1994), for a review]. In order to examine the predictability of seasonal climate associated with the ENSO, it is necessary to extract the ENSO signals from the observational data. For this purpose, the ST-EOF analysis is applied to the tropical Pacific SST (30°S–30°N, on 101 grids) with the time window of 12 months, because this timescale gives sufficient time resolution for seasonal changes and at the same time provides a suitable ''background condition'' of the current state, with respect to interannual variabilities. The space–time patterns of the resulting ST-EOFs show the physical aspects of the ENSO evolution. The projection of the SST field onto their ST-EOF basis leads to the ST-PCs, a new set of variables that describe the ENSO evolution. The major issue concerning the determination of the minimum number of variables that describe the ENSO signals is presented in Wang and Wang (2000) with detailed diagnostic analysis and dynamic justifications.

The resulting analysis shows that the first three leading ST-EOFs (Fig. 2) and the corresponding ST-PCs (Fig. 3) can describe the essential part of the ENSO evolution. ST-EOF 1 (Fig. 2, left-hand panels) at every time lag within the window, is characterized by a positive anomaly along the eastern and central Pacific that is almost symmetric about the equator, with its center moving slightly westward with the time; in ST-EOF$_2$, the mature phase of El Niño is changed gradually to the mature phase of La Niña. The transition takes about three to four months, which is short compared to the mature phase. Correspondingly, the first two ST-PCs represent, respectively, the mature phase of the El Niño (or La Niña) and the transition from La Niña to El Niño (or vice versa). The cross-time-lagged correlation between the ST-PCs shows that the first two are significantly correlated at the time lag (or lead) of a quarter of the average ENSO cycles (Fig. 4), implying that the first two ST-PCs jointly depict a wavelike motion of the ENSO in a way not unlike the sine and cosine jointly describing a linear oscillator. The ENSO signals, however, appear to be more than a simple linear oscillation, as is evident from the phase portrait (Fig. 5) and the broad-banded nature of the power spectra of the corresponding ST-PCs (Fig. 3, right-hand panels). The most significant variabilities are found around the cycles of two to six years. The phase portrait also shows strong evidence of phase-locking of the ENSO onset to the annual cycle (see Fig. 5).

ST-PC$_3$ is characterized by interdecadal changes in the onset phase of the ENSO, which, however, cannot be revealed by power spectrum analysis (see Fig. 3, left-hand panels). The time series of the ST-PC$_3$ displays a significant regime behavior in time. For instance, there is a significant difference before and after 1972, as is noticed by Wang (1995). The corresponding ST-EOF$_3$ (Fig. 2, right-hand panels) displays a space–time structure on a different stage of the ENSO onset that is revealed by the composite study in Wang (1995). The time-lagged correlation analysis (Fig. 4) shows that ST-PC$_3$ leads ST-PC$_1$ by a quarter of the average ENSO life cycle, implying that they, too, describe the same manifold.

Thus the first three leading ST-PCs derived from the

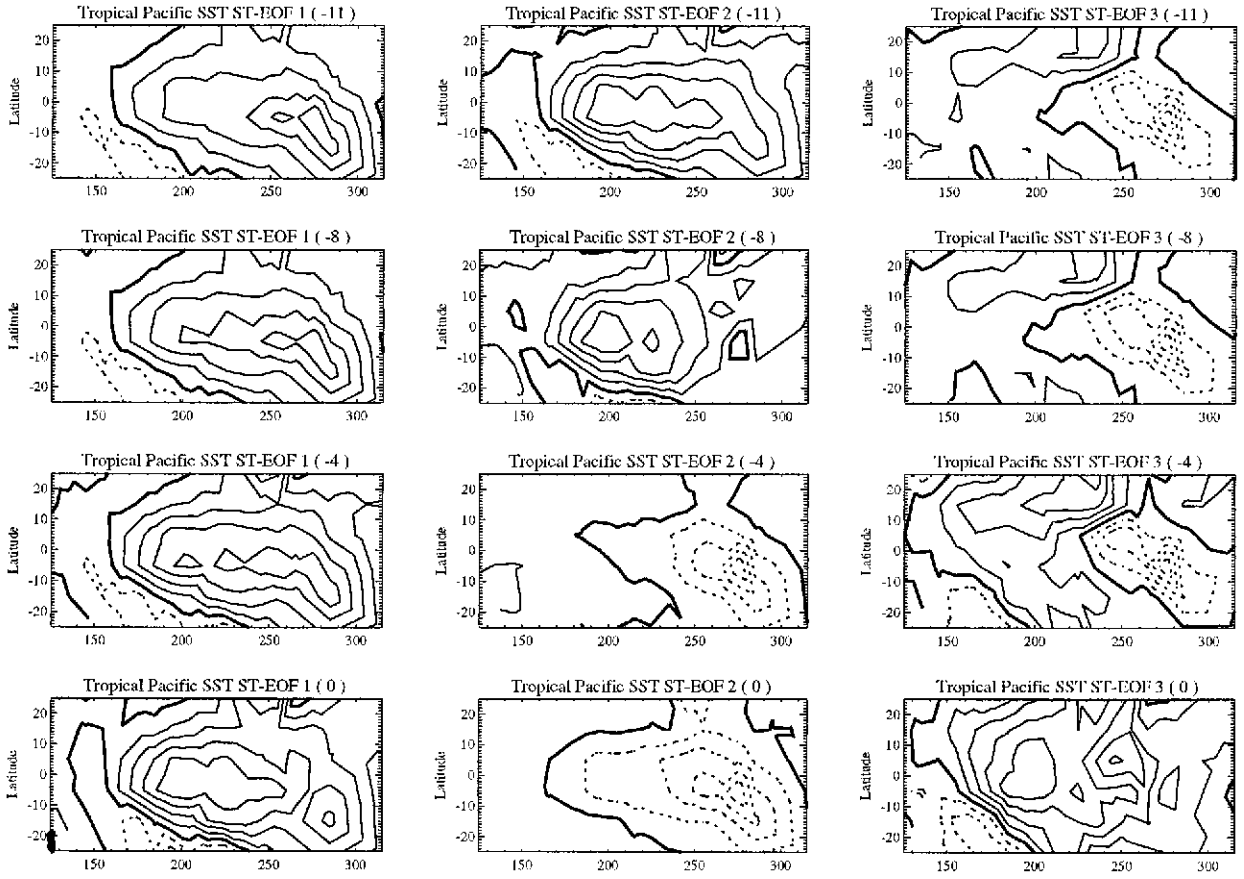

FIG. 2. The spatial patterns of ST-EOFs 1–3 at lag 11, 8, 4, and 0, respectively. The maximum time lag of 11 months is equivalent to a time window of 12 months. Note that only 4 of the 12 time lags are displayed. Note also that ''0 lag time'' here and ''0 lead time'' in the subsequent application means the same initial time of prediction.

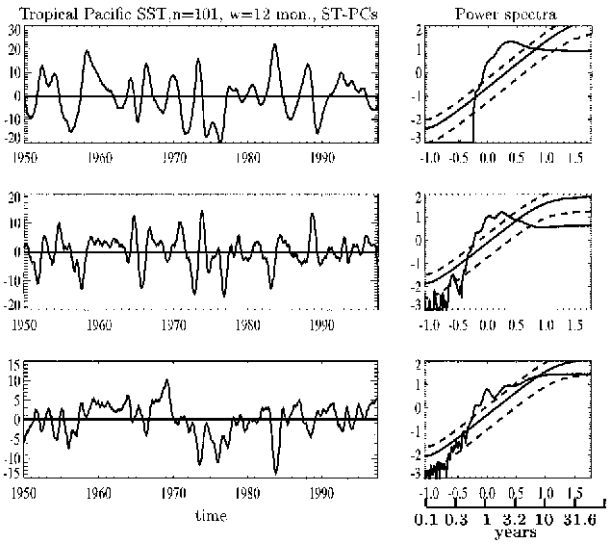

FIG. 3. ST-PCs 1–3 derived from the projection of the monthly tropical Pacific SST (after seasonal cycle and interannual trend are removed) onto the corresponding ST-EOFs in Fig. 2. The right panel shows their power spectra with both coordinates being in log10 scales.

tropical SST give a full picture of the ENSO evolution, and there seems to be no need to include more in this context. Indeed, only these three ST-PCs are significantly correlated at the lead and lag time of about a quarter of the ENSO life cycle (Fig. 4), implying that they describe the same manifold. The phase space spanned by these three variables, $\mathbf{y} = (y_1, y_2, y_3)$, contains all the observed states of the ENSO system and sufficiently describes its time evolution (Wang and Wang 2000). Figure 5 shows that the basic dynamics of the ENSO display near-cyclic trajectories in a two-dimensional plane. The third dimension (ST-PC_3) pushes the trajectory up and down the 2D plane, in particular during the 1982–83 extraordinary El Niño year and the abnormal El Niño–La Niña events from 1974 to 1977 (Wang 1995).

It should be pointed out that the use of the first three ST-PCs has nothing to do with an arbitrary truncation of the ST-EOFs based on a certain mathematical criterion about the eigenvalues, nor with the purpose of data condensation. We started by assuming that the ENSO is a low-dimensional attractor. Then we set out

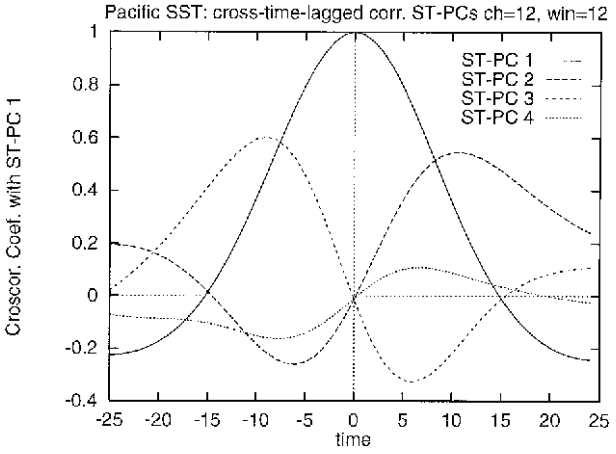

FIG. 4. Autocorrelation of ST-PC$_1$ (solid line) and its cross-time-lagged correlations with ST-PCs 2–4, respectively.

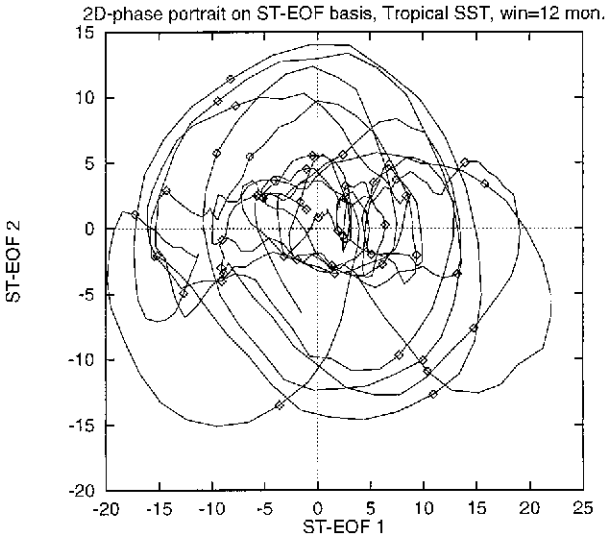

FIG. 5. Phase portrait of the ENSO signals on a two-dimensional plane spanned by ST-EOF 1 and ST-EOF 2. Marked are the positions of the initial state of prediction, $\mathbf{y}_t$, with t = Nov for each verification period.

to look for suitable variables that best describe the ENSO evolution. The assumption that the ENSO is a low-dimensional attractor is rooted in the studies based on dynamic modeling. For instance, using a low-order dynamic system model derived from the first principles, Wang and Fang (1996) are able to show the essential characteristics of the ENSO evolution, such as the irregularity and phase-locking to the annual cycle. The well-known Zebiak–Cane ENSO model (Zebiak and Cane 1987) produces basically a low-dimensional oscillation (Chang et al. 1995). However, there is no general agreement concerning the exact dimensionality, which is typically the case when dealing with a complex natural system such as the ENSO and global or regional climate systems. Indeed, it is impossible to prove a unique dimensionality for such a system, because the dimension may change due to many factors, including, for instance, the time- and space scale range and resolution, sophistication of the model, noise level, and the method used for dimension estimation. In the present study, the monthly observational data are used for constructing the phase space. Thus, our major concern is timescales larger than a month. In the subsequent analysis, a three-dimensional phase will be used to describe the ENSO signals, assuming that the SST associated with the basic dynamics of the ENSO represented by the given dataset can be effectively described in this 3D phase space. The observed dynamics are well explained in simple models such as Wang and Fang (1996). This should be sufficient to justify the above assumption. However, this does not need to confuse the conclusions derived from other models showing higher dimensionality, for a model is, after all, an approximation to the reality, and its dimension typically depends on factors such as those mentioned above, particularly when there are various kinds of noise involved. One may still argue about the dimensionality of the ENSO, but a good model should be able to produce valuable predictions. For instance, a recent paper by Tangang et al. (1998) indicates that optimal skill can be achieved for the prediction of ENSO using a neural network method in a low-dimensional ST-PC space.

Now we have a three-dimensional phase space spanned by the first three leading ST-PCs, $\mathbf{y} = (y_1, y_2, y_3)$, derived from the tropical Pacific monthly mean SST field. In the following, we shall deal with this question: From this three-dimensional phase space where the ENSO signals sit, what can we learn about the seasonal SAT over North America based on (1)? In the present study, we focus only on the prediction of the winter (DJF) SAT with lead time $\tau = 0$, that is, using November as the initial time, t, to predict the following winter (DJF) mean SAT. The initial positions for all the available $\mathbf{y}_t$ are marked in Fig. 5.

b. Geographical distribution of skill scores

1) LINEAR REGRESSION MODEL

For linear prediction of the SAT at a given station denoted by Θ, the Γ in (1) is specified by a linear regression equation (2) using the ENSO signals, $\mathbf{y} = (y_1, y_2, y_3)$ (Fig. 3), as predictors. The source of predictive skill for (2) is derived directly from the linear correlation between Θ and $\mathbf{y} = (y_1, y_2, y_3)$. Figure 6 shows the spatial distribution of correlation coefficients between the winter (DJF) Θ and $\mathbf{y}_t = (y_1, y_2, y_3)$ for t = November. The significance of the correlations in Fig. 6 gives, in part, a measure for the predictability based on the linear regression model. Note that ST-PC$_1$ and ST-PC$_2$ represent different phases of the ENSO. Therefore, the spatial distribution of correlation coefficients reflect, in part, how the SATs in different geographical regions respond linearly to different phases of the

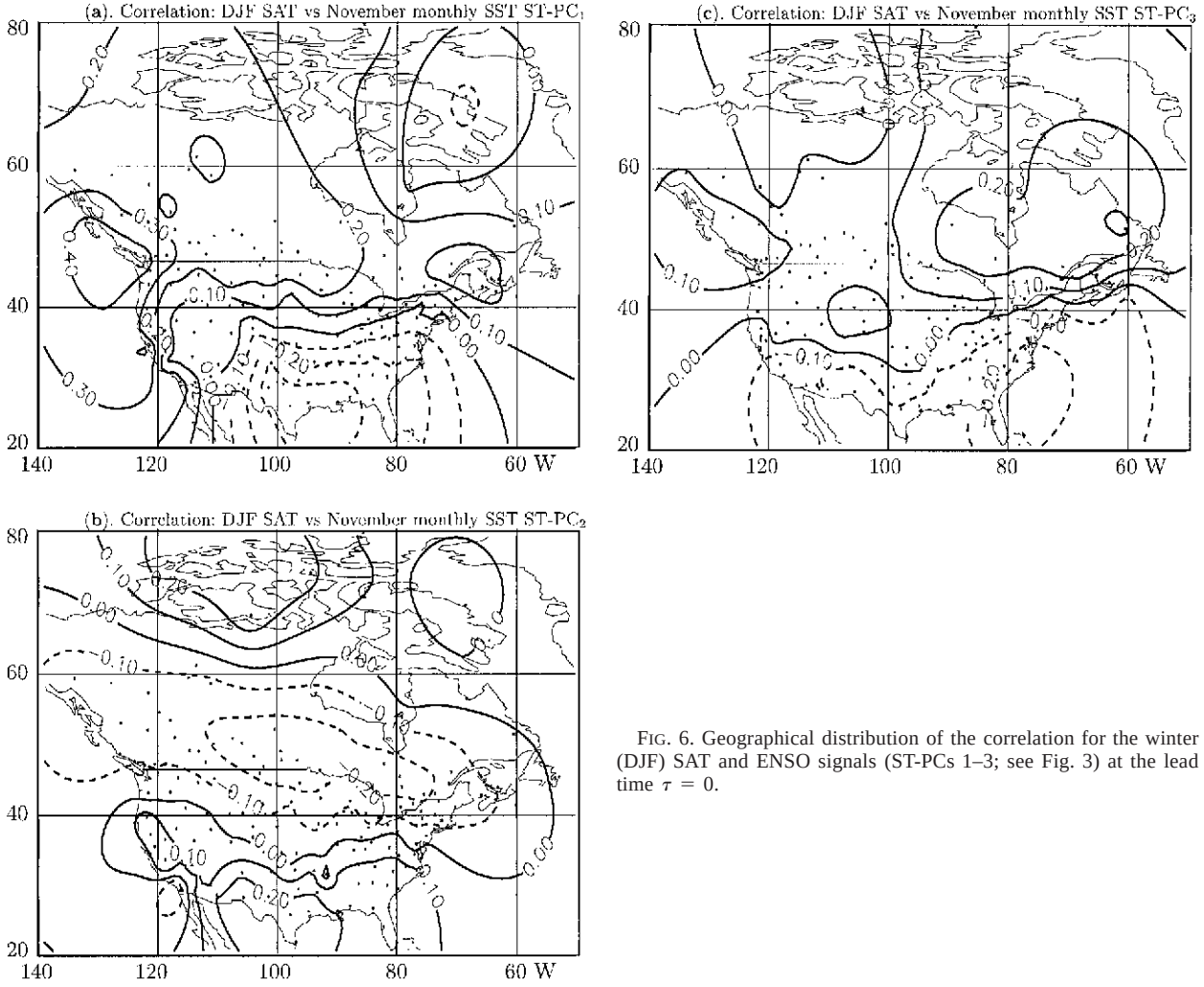

FIG. 6. Geographical distribution of the correlation for the winter (DJF) SAT and ENSO signals (ST-PCs 1–3; see Fig. 3) at the lead time $\tau = 0$.

ENSO. There is a low correlation area along 35°–40°N for both ST-PC_1 and ST-PC_2, implying bad predictive skill in this region. The correlation pattern with ST-PC 3 (Fig. 6c) is similar to that of ST-PC 1 (Fig. 6a) but less significant. However, the correlation seems to be more significant on the eastern part of the continent, which is in contrast to the correlation pattern with ST-PC 1.

Figure 7a shows the geographical distribution of the correlation skill scores derived from (7) for the prediction of the winter SATs based on (2). When compared with Fig. 6, it becomes immediately clear that only those highly correlated regions are significantly (linearly) predictable. When using the 0.3 isoline as a significance criterion, the most significantly predicted areas are along the Pacific coast and the Gulf of Mexico, quite similar to those obtained by persistence forecasting (see Fig. 1a). Note that there is a large region with negative skill scores in the linear model (Fig. 7a) that coincides with the area where the linear correlations in Fig. 6 are close to 0. From a statistical view point, the negative skill score in the linear model is the result of covariance degeneracy (Barnston and Van den Dool 1993). From a dynamic viewpoint, however, this degeneracy may come from the fact that the linear model cannot handle the phase relation of the underlying system. Note that, in the viewpoint of dynamic system theory, near-zero linear correlation comes from the fact that two variables are sometimes positively and sometimes negatively correlated, depending on their phase relation. To better deal with this situation, a nonlinear relationship needs to be established. It is thus expected that the phase space approaches provide new possibilities for producing better predictive skills.

2) NONLINEAR MODELS IN PHASE SPACE

As discussed in section 3b, the predictability for a given lead time τ is defined as the probability of finding the same state of $\Theta_{t+\tau}$ for all $\Theta_{\zeta+\tau}$ corresponding to $\mathbf{y}_\zeta \in \Delta(\mathbf{y}_t) \subset \Omega_L$ and $t \neq \zeta$. In general, the predictability defined in this way changes with the size of the nearest

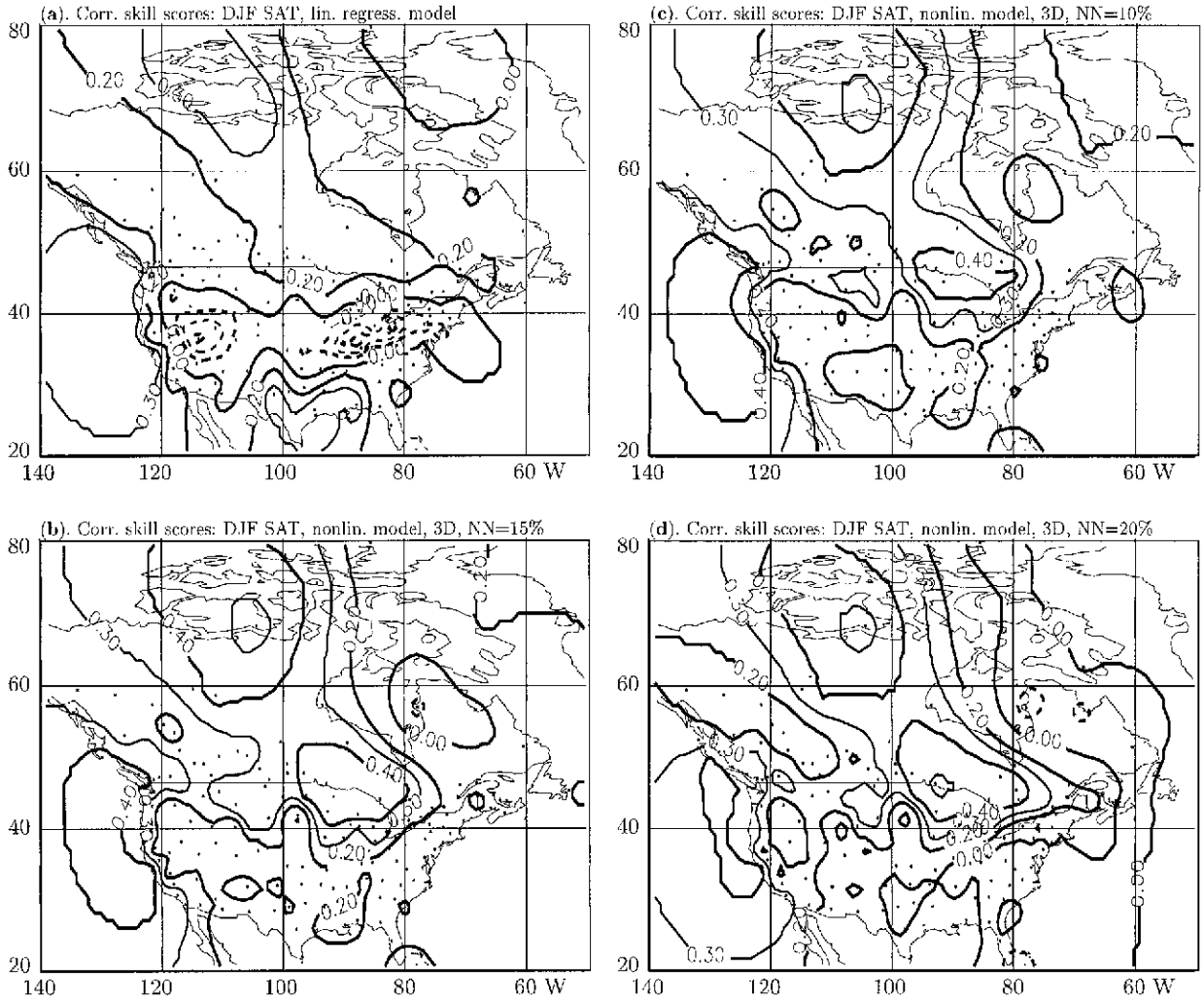

FIG. 7. The geographical distributions of the correlation skill score (ρ) for the winter (DJF) SAT for lead time $\tau = 0$ using (a) linear regression model and nonlinear model (3) with the nearest neighborhood [$\Delta(\mathbf{y}_t)$], respectively composed of nearest (b) 15%, (c) 10%, and (d) 20% of the total samples. The same ENSO signals (ST-PCs 1–3; see Fig. 3) are used as predictors for both linear and nonlinear models. Areas with $\rho > 0.26$ (roughly in accordance with the 0.3 isolines) are considered as displaying significant skill.

neighborhood $\Delta(\mathbf{y}_t)$ for nonlinear systems, but not for linear systems. Figure 8a gives an example of diagnostic results about predictability as a function of the size of $\Delta(\mathbf{y}_t)$, for a regional mean winter (DJF) SAT over the southwestern United States, with 47×3 samples (three samples per year). The predictand for this particular example is classified into one of three categorical states: below (B), near (N), or above (A) normal. For an arbitrary initial state of the ENSO, $\mathbf{y}_t$, there is an observed state of SAT, say Θ_t = A. The predictability can then be understood as the probability of finding Θ_ζ = A, corresponding to each $\mathbf{y}_\zeta \in \Delta(\mathbf{y}_t) \subset \Omega_L$ ($\zeta \neq t$). The same is true for Θ_t = B and N. The results in Fig. 8 show clearly that the predictability is derived from a small neighborhood and that, when the neighborhood is increased to 30% of the total sample space, the resulting prediction is no different from random noise.

The same result can be obtained by examining the chance of making incorrect predictions for a given initial state, $\mathbf{y}_t$, corresponding to the true value of Θ_t = B. In this case, the chance of finding Θ_ζ = A or N, corresponding to each $\mathbf{y}_\zeta \in \Delta(\mathbf{y}_t) \subset \Omega_L$ ($\zeta \neq t$), increases with the size of $\Delta(\mathbf{y}t)$, as is evident from Fig. 8b. The same is true for the initial state being N or A (Figs. 8c,d). The probability of making correct and incorrect predictions for each of the three categories converges to 0.33, a probability expected from random noise, as the size of $\Delta(\mathbf{y}t)$ increases. These reflect a certain nonlinearity of the underlying phase relation. Note that for Θ = A (Fig. 8d), the probability of making correct forecasting is higher than expected from random noise, even with very large $\Delta(\mathbf{y}t)$, due to the linear correlation between the El Niño and the SAT over that region. Thus, above-normal cases (A) in that region are relatively easy

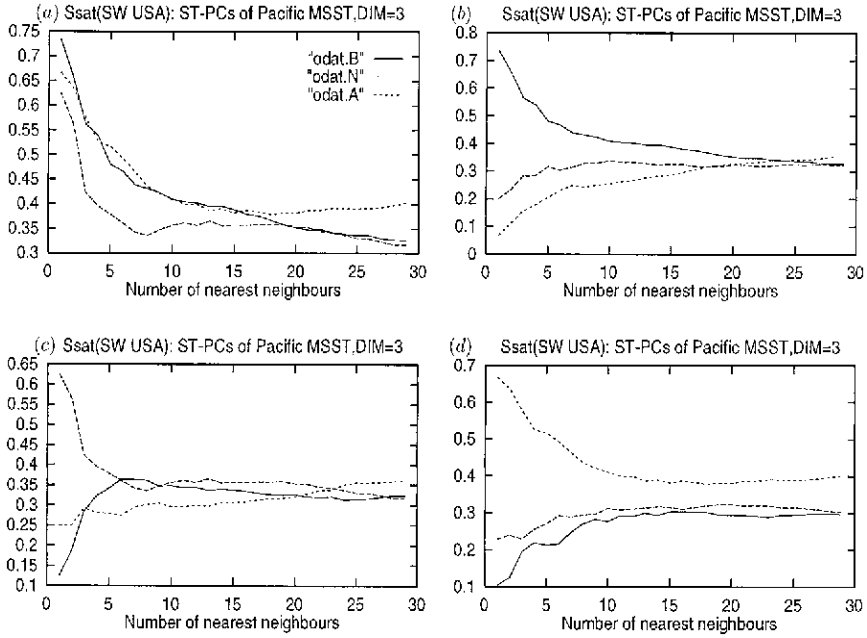

FIG. 8. (a) The probability of finding an arbitrary winter (DJF) SAT Θ_ξ in the same categorical state as Θ_t corresponding to each $\mathbf{y}_\xi \in \Omega_L \subset \mathbf{R}^3$ in the nearest neighborhood of $\mathbf{y}_t \in \Omega \subset \mathbf{R}^3$, where $\mathbf{y}_t$ represents the ENSO signals for t = Nov (see Fig. 5) and Θ = A (above normal), N (near normal), or B (below normal), as in Vautard et al. (1996). The total available number of samples is 131 (three per year). The figure shows how the probability of correctly predicted number of Θs changes as the number of neighbors increased from 1 to 30. (b)–(d) The same as (a) but also plotted are the probabilities of the incorrectly predicted categories (see text).

to predict (Figs. 8a,d), even with linear models. The predictability, however, is smaller than what is expected from a nonlinear model with a smaller neighborhood. Note also that the near-normal state (N) is the most difficult to predict correctly (Figs. 8a,c), for both linear and nonlinear models, which seems to be consistent with the conclusion in Van den Dool and Toth (1991). It is apparent that all of these characteristics may change with the climate zone to be predicted and with the key SST forcing regions from which the predictors are derived.

From the above results, it seems reasonable to conclude that the closest neighborhood yields the best prediction. It is true, on average, based on categorical prediction schemes. However, as the system is not perfect, there are still chances to produce very bad forecasting once in a while, as Figs. 8b–d show. To derive a statistically stable forecast, a locally linear relation should be established between the predictors ($\mathbf{y} \in \Omega_L$) and the predictand (Θ), instead of the one-to-one correspondence used by the categorical prediction schemes (or analogs), as discussed in section 3b. For this, we need to use all members in $\Delta(\mathbf{y}_t) \subset \Omega_L$, as in Vautard et al. (1996, 1999), instead of the single closest one. Herein lies the essential difference between a linear and a nonlinear model: for the former, relation (2) is established based on the whole learning set Ω_L, whereas the latter is derived only from $\Delta(\mathbf{y}_t) \subset \Omega_L$, thus locally linear, but globally nonlinear (section 3b). For the present application, we used a simple composite anomaly (3) of all Θ's corresponding to $\Delta(\mathbf{y}_t) \subset \Omega_L$ as the output of the prediction model (1). We found that 15% of nearest neighbors extracted from all available data in Ω_L often yields optimal results. Figure 7b gives the geographical distribution of the skill scores for the prediction of the DJF SATs at the lead time $\tau = 0$, based on the three-dimensional phase space $\mathbf{y}$, spanned by $\mathbf{y} = (y_1, y_3, y_3)$, ST-PCs (Fig. 3). The initial states, $\mathbf{y}_t$ (t = November), are marked with points in the two-dimensional phase portrait in Fig. 5, which shows apparent uneven clustering of the state points, representing regime behavior and phase-locking (Wang and Wang 2000).

Two striking facts are observed from Fig. 7b, as compared with the linear model (Fig. 7a), as well as persistence forecasting (Fig. 1a): first, the average skill score is much improved with the highest predictable regions covering a large area of the continent, except for the southeastern part of the subtropical region and the regions along the northeastern coast. This is in sharp contrast to the linear regression and persistence models. Second, the regions with negative skill scores in the linear regression model (Fig. 7a) almost disappeared in the nonlinear model (Fig. 7b). Moreover, it seems that the linear prediction produces slightly better skill scores for the subtropical region (along the border between Mexico and the United States and the Gulf of Mexico),

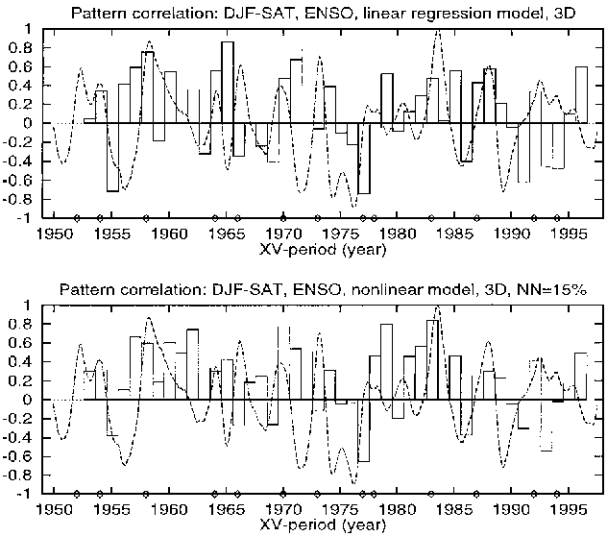

FIG. 9. Pattern correlation (8) between the predicted and observed winter (DJF) SATs over North America with 164 stations for each cross-validation period based on the linear regression method (upper panel) and the nonlinear composite model (lower panel), with the size of nearest neighborhood [$\Delta(\mathbf{y}_t)$] composed of 15% of the available data in Ω_L. The models used are the same as in Fig. 7. For comparison, ST-PC$_1$ from Fig. 3 is also plotted (dash lines), and the El Niño years are marked with dots at the bottom.

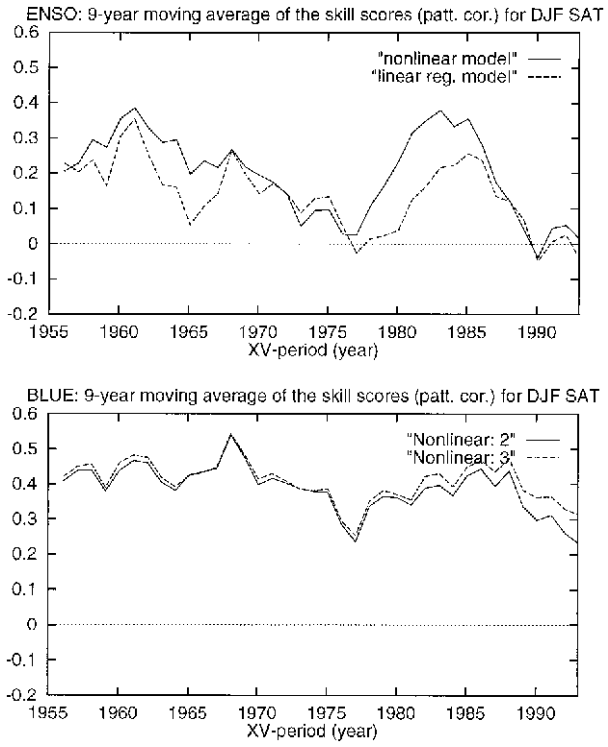

FIG. 10. (Top) 9-yr running mean of the pattern correlation derived from Fig. 9 for linear (dashed line) and nonlinear (solid line) predictions. (Bottom) The same but derived from Fig. 17, the results after the application of the BLUE to two (ENSO and Atlantic SST, solid line) and three predictions (ENSO, Atlantic, and North Pacific SSTs, dashed line), respectively (see Fig. 15).

while the nonlinear method works better in mid- and high latitudes, implying a stronger nonlinear response (phase relation) of the extratropical SATs to the tropical ENSO.

To demonstrate that the nonlinear prediction is robust with respect to the choice of the $\Delta(\mathbf{y}_t)$ size, Figs. 7c, d give the same prediction as Fig. 7b, but with $\Delta(\mathbf{y}_t)$ containing respectively 10% (Fig. 7c) and 20% (Fig. 7d) of the available data in Ω_L. Note that they are almost the same as Fig. 7b. The gain in the predictive skill compared with the linear and the persistence models is derived from the improved identification of the phase relation between the predictor and predictand fields, which is not difficult to understand when considering that the relation between the predictors and the predictand is a function of the state $\mathbf{y} \in \Omega$. In general, the predictability may also change with $\mathbf{y} \in \Omega$.

3) THE TIME BEHAVIOR OF THE PREDICTIVE SKILL

Figure 9 gives the pattern correlation (8) between the predicted and observed winter SATs for each of the $K = 44$ verification periods from 1953 to 1996. For comparison with the ENSO signal, ST-PC$_1$ from Fig. 3 is also plotted as dashed lines and the El Niño years are marked at the bottom of the figures. The following can be observed.

1) The nonlinear method is better than the linear regression method in that the correlation skill scores of the former are higher, on average, and more stable in time (Fig. 10, upper panel) and that there are fewer negative scores (Fig. 9).
2) The change of the skill scores with time does not appear to be particularly related to the ENSO phase, indicating that some other factors (such as the Atlantic SST) may have stronger influence on the winter SATs in the years when the skills are small or negative. This may particularly be the case when the SST distribution in the tropical Pacific region is close to climate mean.
3) There are strong indications of an interdecadal variability of the correlation scores in both linear and nonlinear models, which is more clearly demonstrated in Fig. 10 (upper panel) by their 9-yr moving average. This implies that the interdecadal variability of the atmospheric–oceanic system may possibly modulate the influence of ENSO on the atmosphere. For instance, in Trenberth (1990), interdecadal climate change is linked to the frequency of El Niño–La Niña events; Wang (1995) shows there is an interdecadal change of the ENSO onset and that there is a significant difference in the ENSO onset before and after 1972. Note that both the linear and the nonlinear models do indeed display different time behaviors before and after 1972. This result does not

seem to be influenced by the detrending in the SST field.

The change of the predictive skill with time may generally be attributed to the fact that there is an intrinsic change of predictability with the location of the initial points (see Fig. 5 for the above example). Similar results are also observed in other dynamic models (Ziehmann-Schlumbohm et al. 1995). It is thus not surprising that the lowest predictive skills are found when the initial state ($\mathbf{y}_t$) is close to the climate mean ($\mathbf{y}_t = 0$). The predictive skill may also be modulated by other low-frequency variabilities that are relatively independent of that of the ENSO, such as the interdecadal changes in the Atlantic SST. It is expected that the predictability based on the ENSO signal is lower when the amplitude of the Atlantic SST anomaly is high, and that the ENSO signal plays a full role when all other parts of SST are close to climate mean. This is partially demonstrated by the apparent coincidence between the low skill from 1966 to 1979 (Fig. 10) and the significant anomaly of the Atlantic SST during this period [see section 6b(3) for further discussion].

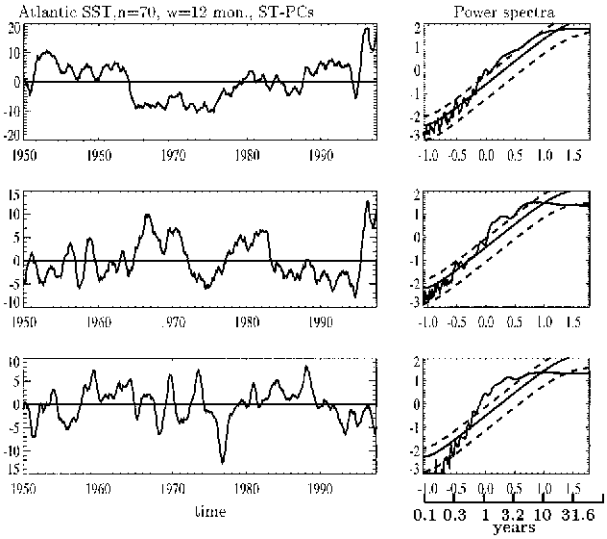

FIG. 11. The first three leading ST-PCs derived from Atlantic monthly SST after the seasonal cycle and interannual linear trend are removed. The time window used is 12 months. The right-hand panels show their power spectra in log–log plots.

5. Other key SST forcing regions and the predictability of seasonal SAT

a. Atlantic SST

In order to investigate other key SST forcing regions that influence the seasonal SAT and thus contribute to the predictive skill scores, we carried out a similar analysis with the Atlantic monthly SST using the same data source from 30°S to 65°N. The first three leading ST-EOFs (not shown) represent the SST anomaly on the western part of mid- and high-latitudes (ST-EOF_1), the contrast between the western and eastern parts of the mid- and high-latitudes (ST-EOF_2), and the coupled anomaly between western subtropical SST (Gulf Stream) and the eastern part of the high-latitude SST. The fourth ST-EOF reflects mainly the variability of the Gulf Stream region and the high-latitude Atlantic. The corresponding ST-PCs (Fig. 11, left-hand panel) capture the main feature of the variability of the Atlantic SST. Unlike the ENSO signals, the Atlantic SST is dominated by interdecadal variabilities, as is evident from the spectra of the ST-PC_1 (Fig. 11, right-hand panel). Thus the timescale of this ST-PC goes far beyond the seasonal scale. Tests with this component show that no significant contribution to the seasonal predictive skill can be observed, particularly for the nonlinear model. On the other hand, ST-PC_2, ST-PC_3, and ST-PC_4 have timescales relevant to the seasonal climate, and they are cross-correlated with each other at different time lags and all related more or less to SST anomaly associated with the Gulf Stream. Thus we took a subspace spanned by the second, third, and fourth ST-PCs, denoted by $\mathbf{y} = (y_2, y_3, y_4)$, for the subsequent analysis.

The linear correlation patterns (Fig. 12) between the SST ST-PCs and SAT appear to be very different from those obtained with the ENSO signals. There is no significantly correlated areas with ST-PC_1 (figure not shown), due to the fact that the timescale of this ST-PC goes far beyond seasonal scales. However, the correlation with the other three ST-PCs (see Fig. 12) are significant in the midlatitudes. Note that in these regions, the correlation with the ENSO signal is generally poor (cf. Figs. 6 and 12). It seems that the ENSO signals are *linearly* correlated with the SATs of the subtropical and the more or less higher latitude regions, while Atlantic SSTs are more related to the middle latitude (around 40°–60°N). Thus, in this region, the linear model is expected to produce better predictions with the Atlantic SST than with the ENSO signal based on (2). Figure 13a shows, indeed, that around 50°N the correlation skill (ρ) based on (7) exceeds 0.40, far above the persistence skill (Fig. 1).

Similarly, we performed nonlinear predictions of the winter SAT with the nearest neighborhood approach, using $\mathbf{y} = (y_2, y_3, y_4)$ as the coordinates. The resulting geographical distribution of the correlation skill scores (ρ) and the time changes of the pattern correlation (ρ) are given in the lower panels of Figs. 13 and 14, respectively. Tests with different $\Delta(\mathbf{y}_t)$ size indicate that the results are robust. Two striking facts can be observed in comparison with the results based on the ENSO signals (Figs. 7 and 9).

First, there are fewer differences between linear and nonlinear predictions with Atlantic SST than with the ENSO signals. The nonlinear model is better than the linear model, on average, but the most skillful regions, found between 40° and 60°N to the west of 90°W, are significantly improved in the nonlinear model, with the

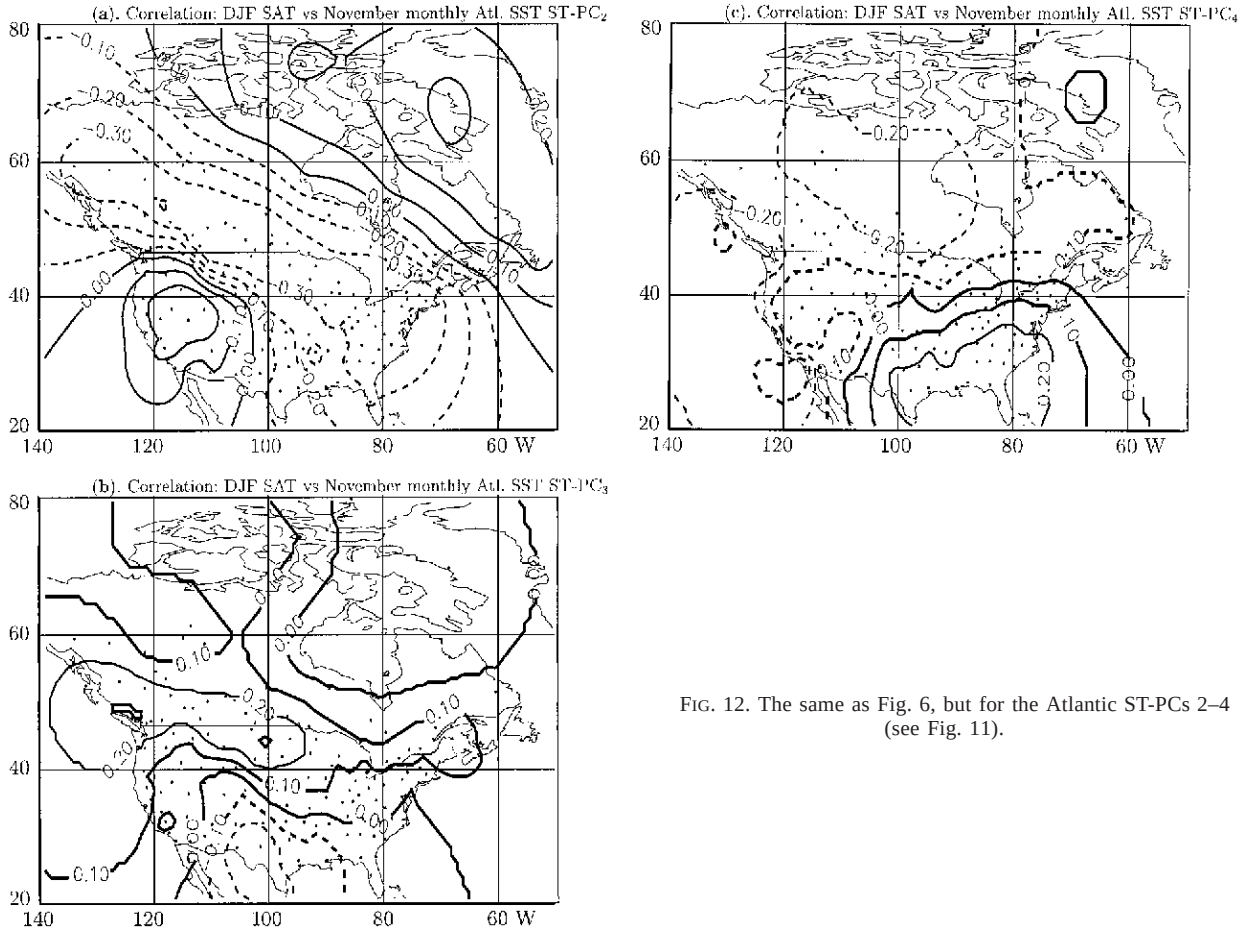

FIG. 12. The same as Fig. 6, but for the Atlantic ST-PCs 2–4 (see Fig. 11).

highest skill score exceeding 0.50. However, this improvement is less significant compared to that observed in the linear and nonlinear predictions with the ENSO signal (see Fig. 7). This may be due to the fact that the dominant timescales of the Atlantic SST variability are larger than the typical ENSO scales. In general, projecting much larger timescales onto seasonal scales based on (1) leads to relations that are closer to linear.

Second, the predictive skill displays larger variability both in geographical location and in verification period than that obtained with the ENSO signals. The significant differences in geographical distribution and time behavior of the skill scores indicate that the origin of the predictive skill, with respect to the Atlantic SST, is very different from that of ENSO predictors. This implies a possibility for further improving the final skill by an optimal linear combination (see section 6b for details).

b. Extratropical North Pacific SST

We have also carried out linear and nonlinear predictions of seasonal SAT based on the ST-PCs derived from extratropical North Pacific SST (30°–75°N). Significant skill scores were found for the linear model in the subtropical regions. The source of predictive skill is the same as for the ENSO predictors, since an important part of the variability in the midlatitude Pacific is related to the ENSO signal, as is evident from the first ST-EOF derived from the global SST (figure omitted). For the nonlinear model, the most significant regions lie in the northwest of the continent and along the northeast coast, indicating that the mid- and high-latitude Pacific SST influences principally the mid- and high-latitude winter SATs over North America. However, the pattern correlation skills (figure omitted) change quite abruptly from year to year, indicating instability of the predictive skill, both in space and in time. This may be partially due to the fact that the stations located in the midlatitude and subtropical regions have low average skill scores, resulting in large year-to-year changes. It may also reflect the complications of the influence of the midlatitude SST on the seasonal SAT. Thus the information derived from this SST forcing region can hardly be used directly for practical forecasting without further constraints from the atmospheric components. However, it may be used to improve the average skill score by a suitable linear (or nonlinear)

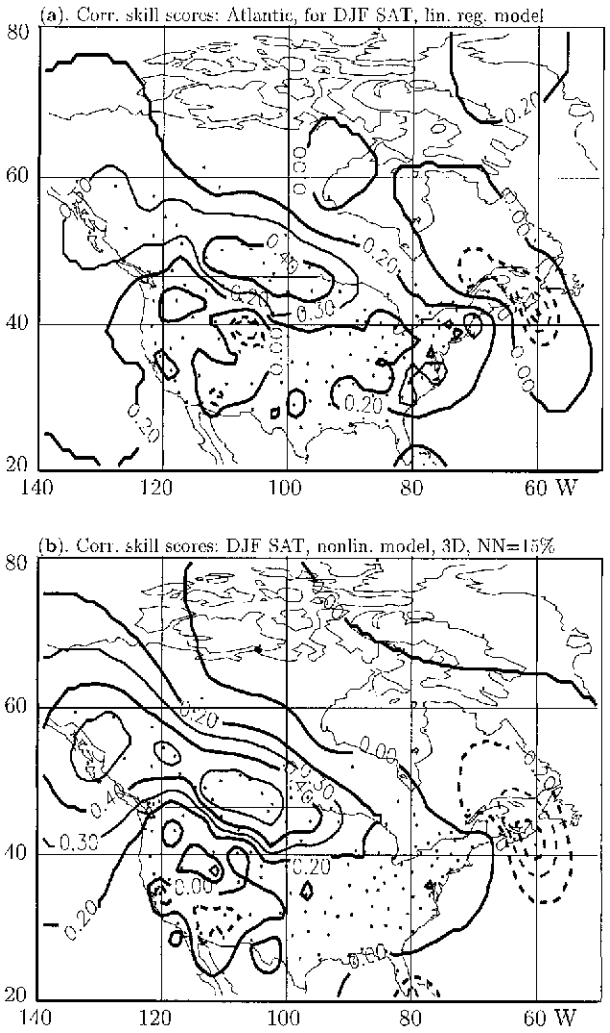

FIG. 13. The same as Fig. 7, but derived from ST-PCs 2–4 of the Atlantic SST.

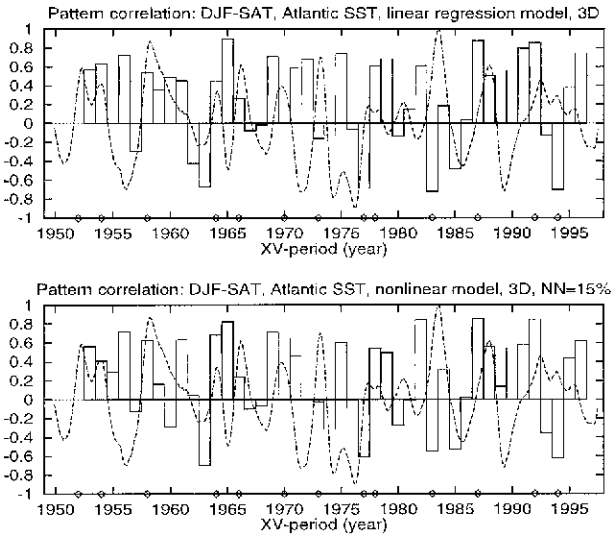

FIG. 14. The same as Fig. 9, but the predictions are derived from Fig. 13.

combination with other predictors, though not very often in the present application (see section 6.2).

6. Global SST and the predictability of seasonal SAT

a. Characteristics of the global SST field

The global SST is dominated by ENSO signals, as is reflected by its ST-EOF$_1$ (figure omitted), which is characterized by two anomaly centers with opposite signs, indicating a significant negative correlation between the tropical and midlatitude Pacific SSTs. The spatial structure of global ST-EOF 1 in the subtropical Pacific is almost identical to that obtained from the tropical Pacific SST alone (see Fig. 2a). Accordingly, the first ST-PC represents the ENSO signal on its mature phase, like the ST-PC 1 derived from the tropical Pacific (Fig. 3). The rest of the ST-PCs represent the different phases of the ENSO and other key forcing regions. It is not difficult to have a one-to-one correspondence between the ST-PCs derived from regional SST and those from the global SST, with the help of a suitable rotation. In the present context, rotation of the ST-PCs is unnecessary since the phase space they span is the same. It is sufficient to know how many ST-PCs derived from the global monthly SST contain the most relevant information for seasonal forecasting. Our analysis showed that about 10 ST-PCs derived from the global SST with a time window of 12 months may reasonably well represent all the key SST forcing regions discussed earlier, as suggested in Vautard et al. (1996, 1999). This number is still too high for nearest neighborhood approaches. In Vautard et al. (1999), the size of $\Delta(\mathrm{y}_t)$ is taken as 50% of the available data in Ω_L, which is a reasonable choice for a dimension larger than five, with the data available. This is considered too large for nonlinear methods used in the present context. When using regression models based on $\Delta(\mathbf{y}_t) \subset \Omega_L$, a larger $\Delta(\mathbf{y}_t)$ size leads to predictions closer to the linear model, though not the same.

In practice, one has to make compromises concerning the choice for the number of ST-PCs for seasonal prediction. In general, including more ST-PCs in the prediction model means higher-dimensional phase space, which may give better representation of the variability for the SST field, but not necessarily yield good results for seasonal forecasting, as high-dimension space requires a larger number of samples to arrive at a robust conclusion (Van den Dool 1994a). Moreover, for seasonal prediction on an empirical basis, to excluded unwanted information is almost as important as to include relevant information. Therefore, the choice for the predictors is very important for statistical predictions. The understanding of physical relations is the key to making

good choices. For instance, the significant correlation between the Atlantic SST and the midlatitudes' SAT (Fig. 12) reflects influence of the former on the winter jet, and the jet's position in turn determines the SAT anomaly in the midlatitudes. On the other hand, the ENSO signals are less correlated with the midlatitudes' SAT (Fig. 12), indicating that the position of the midlatitude jet is less sensitive to the ENSO than to the Atlantic SST. The nonlinear prediction model used in the present study does not explicitly contain the physical laws that connect the SST anomaly with the seasonal SAT. However, the use of the analogs is based on the assumption that such laws exist. These are partially manifested by the success of the resulting predictions in the previous sections and in what follows.

b. Optimal linear combination of the predictions derived from different key SST forcing regions

From the analysis in the previous sections it can be concluded that each key forcing region has its own typical contribution to the skill scores for both linear and nonlinear models. When using the global SST field, the dimension may not need to be very high, but it has to be high enough to include signals from these key SST forcing regions. This is almost fatal for nonlinear predictions based on analog methods, due to the limitation of the available data (Van den Dool 1994a). Moreover, one has to be aware of the fact that the ENSO signal dominates the global SST so much that the amplitude of the loading patterns in other regions becomes less significant for almost all of the EOF-based prediction schemes. To make good use of the predictive skills from each SST key forcing region, and to profit from a low-dimensional phase space approach, we use the BLUE (see Sarda et al. 1996) to extract the optimal predictions, which are based on predictors respectively derived from tropical Pacific (ENSO signals), Atlantic, and extratropical Pacific monthly mean SSTs.

1) PREDICTIONS BASED ON ENSO SIGNALS AND ATLANTIC SST

Figure 15a shows the correlation score for the winter (DJF) SAT after optimal combination of the predictions derived independently from ENSO signal (Θ_1) and Atlantic SST (Θ_2), using the BLUE according to (5). As expected, significant improvements in the predictive skills are observed compared to the prediction based on either of the two predictors in Figs. 7 and 13. The correlation skill scores exceeding 0.3 cover most of the continent, and there is a large area around 50°N where the skill score is close to 0.6. [The time behavior of the skill scores, shown in Fig. 17 (upper panel), is more stabilized, although there is still an interdecadal change (see also Fig. 10, lower panel)]. Note that the evaluation of the skill is based on 44 independent cases, and in each case, the year to be predicted is not used for estimating α [see(6)], in order to avoid optimal fitting of the two predictions.

To a certain extent, the weight α for each station, derived from (6), on the basis of the observed and predicted values from the 44 verification periods, gives a measure of the contribution to the predictive skill from both SST forcing regions. From (5) it becomes immediately clear that $\alpha > 0.5$ means ENSO contributes more to the predictive skill than Atlantic SST and vice versa. Thus the geographical distribution of α provides an important clue to the origin of predictive skill or the impact of the SST forcing on the seasonal climate. Figure 16a shows that, aside from the western part (to the west of 90°W) of the midlatitudes (around 50°N), most of the predictive skills are derived from the ENSO signal.

2) PREDICTIONS BASED ON ENSO SIGNAL, THE ATLANTIC SST, AND THE EXTRATROPICAL NORTH PACIFIC SST

The BLUE method (5) can be easily extended to include more predictions. In order to test whether we can profit from the information derived from the extratropical North Pacific SST, we simply used the above BLUE results as Θ_1 and the prediction derived from the extratropical North Pacific SST (see section 5b) as Θ_2, for a further application of the BLUE based on (5). The resulting predictive skills are shown in Figs. 15b and 17 (lower panel). There is a slight improvement of the skill in that the area when the correlation skill score $\rho > 0.5$ is increased (Fig. 15b). However, there is no significant sign of further stabilizing the time behavior of ρ (Figs. 17 and 10, lower panel). Figure 16b shows that the weight α is greater than 0.5 everywhere, indicating that the ENSO and the Atlantic SST contribute the most to the predictive skill.

3) DISCUSSION

The above results show that a conceptually clear and dynamically based statistical method can be developed for seasonal prediction based on the low-dimensional phase space. Due to the fact that different climate zones respond differently to the same SST forcing, the proposed method can often yield good prediction for certain areas, based on a given key SST forcing region. Thus it is not surprising that the BLUE applied to predictions based on different key SST forcing regions can lead to significant improvement of the final skill scores, compared to the individual predictions based on either of the three key SST forcing regions. The correlation skill score averaged over all 164 stations is about 0.4 (see Fig. 15), with a large area of the continent exceeding 0.5, implying that there is a great potential for practical applications to seasonal forecasting along this line. Note that the 95% significance level is at 0.26 (see section 3d and Fig. 1).

It should be pointed out that the application of the

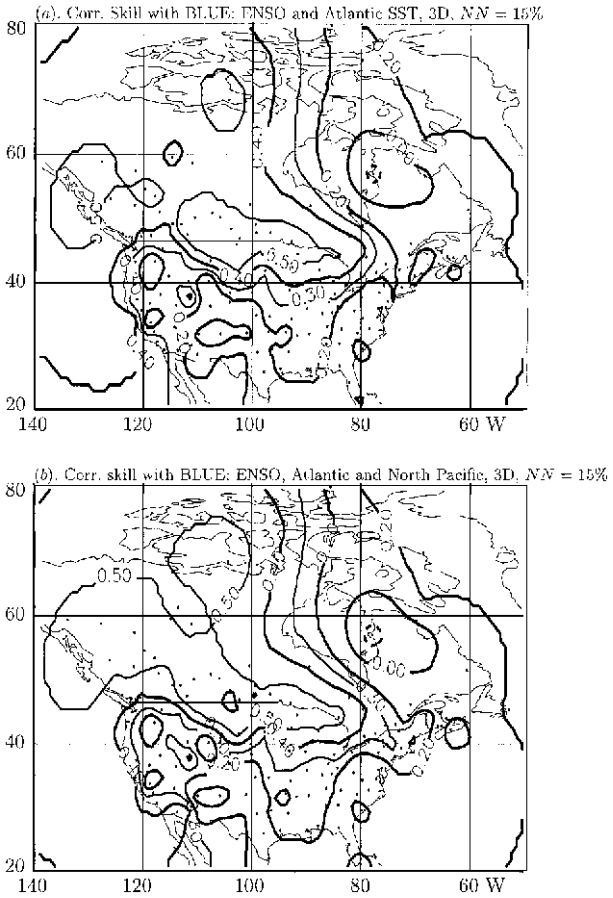

FIG. 15. The same as Fig. 7, but for the results after the application of the BLUE (5) to (a) two and (b) three independent analog predictions carried out in three-dimensional phase space with $\Delta(y_t)$ composed of 15% of Ω_L. The predictors of these models are derived from tropical Pacific (Fig. 7), Atlantic (Fig. 13), and North Pacific SSTs, respectively. The weight coefficient α for each station is derived from 44 verification periods under a cross-validation scheme according to (6).

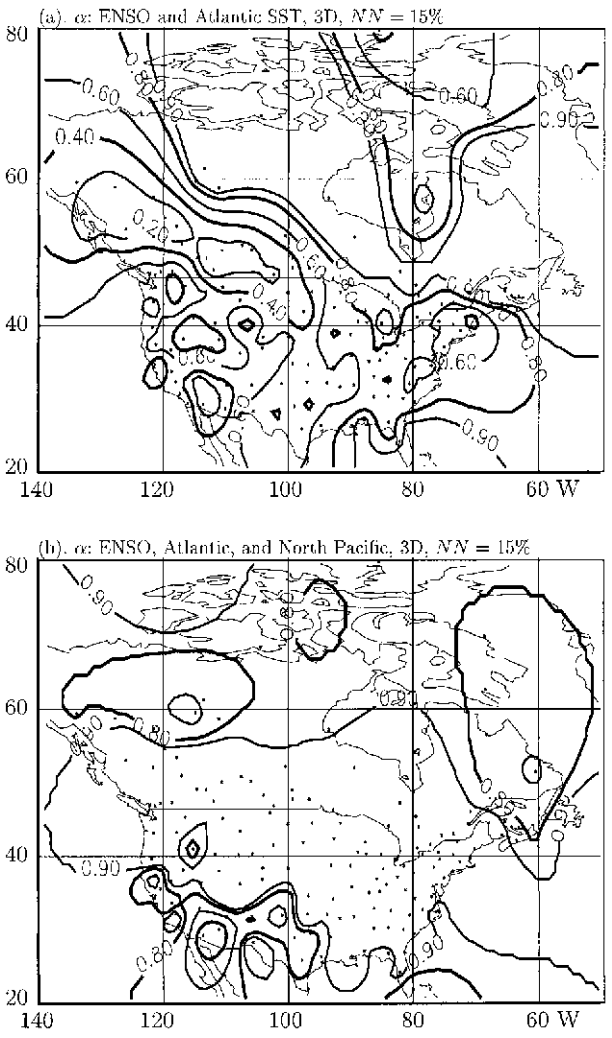

FIG. 16. The weights (α) used for the BLUE in Fig. 15.

BLUE can sometimes result in false skills if there are stations with the weight $\alpha < 0$, because negative weights in (5) can convert negative scores into positive ones (Livezey and Neumeister 1990). This often happens when the individual prediction models produce systematic negative skill scores. It is particularly the case for some linear models, due to statistical degeneracy when the correlation between the predictors and predictand is close to 0 (cf. Figs. 6 and 7a; also see Barnston and Van den Dool 1993). Thus the negative weight (α) should be avoided to diminish false skills. This can be done by rendering $\alpha = 0$ in (5) whenever $\alpha < 0$. In this case, stations with $\alpha < 0$ will not be used, which is a reasonable decision. Another, more direct, way to avoid false skill induced by the BLUE may be to use only one prediction from one model for a given station or climate zone, by choosing either $\alpha = 0$ or $\alpha = 1$ in (5) instead of using $0 \leq \alpha \leq 1$ to combine predictions from different models. This is the same as using different models to predict different climate zones. The same principle can also be applied to different seasons.

In our present study, however, the nonlinear model does not produce systematic negative skill scores. So the above-mentioned issue is not a problem. Indeed, although there are some negative scores in individual models (see Figs. 7b and 13b), there are no negative values of weight in the BLUE (see Fig. 16), implying that the proposed nonlinear models are not likely to produce systematic negative skill scores. The success of the BLUE, in this context, lies in the fact that different climate zones have different responses to each key SST forcing region, leading to uneven geographical distribution of predictive skills. The BLUE just picks up the best predicted stations derived from each key SST forcing to form an optimal ''global'' distribution (see Fig. 15).

Understanding the influence of interdecadal variability on the seasonal predictive skill can undoubtedly improve the skill for seasonal prediction. For instance,

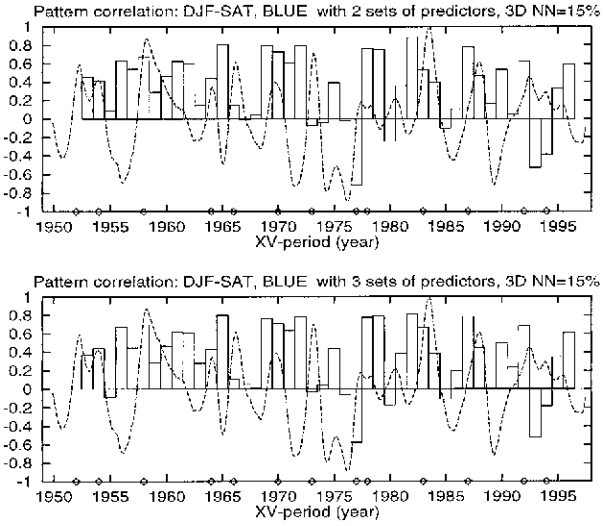

FIG. 17. The same as Fig. 9, but the pattern correlations are calculated after application of the BLUE corresponding to Figs. 15 and 16. See Fig. 10 (lower panel) for their 9-yr tuning mean.

closer examination of long-term changes of the skill score derived from ENSO (Fig. 10, upper panel) and the ST-PC 1 of the Atlantic SST (Fig. 11) reveals that the predictions based on the ENSO signals are the best when the ST-PC 1 of the Atlantic SST is close to 0. The worst skills were found in the 1970s when the negative values of the ST-PC 1 reached maximum. Considering that the ST-PC 1 represents the SST anomaly in the North Atlantic, this implies that the ENSO predictors work most effectively when the North Atlantic SST is close to climate mean. This indicates that the North Atlantic SST, closely associated with the Northern Atlantic Oscillation NAO observed in the atmosphere, may change the impact of ENSO on the SAT over some particular regions, through influencing the climate conditions of the atmospheric circulation, such as the mean positions of the trough and the midlatitude winter jet. Thus it is not surprising that the incorporation of the atmospheric components into the prediction scheme based on SST forcing may lead to improvements of the seasonal forecasting skill (Ting et al. 1996; Wang and Van den Dool 2001, unpublished manuscript). Meanwhile, one may benefit from the low-frequency variability for seasonal predictions based on dynamically based statistical models (e.g., Brunet and Vautard 1996). All of these may also imply that the SST-based prediction can be further improved by selecting prediction models according to climate zone and season.

Being an effective method for optimal predictions, the BLUE method is, after all, a linear combination of the two (or more) predictions based on all verification periods. Therefore, the joint impact (or phase relation of the two forcing regions) in some specific years cannot be fully captured based on the BLUE. There is a possibility of designing a nonlinear version of the BLUE, given a sufficient number of independent verification periods. Here, the results derived from dynamical models may help.

7. Concluding remarks

The central point of the present paper is that statistical prediction of the seasonal climate can be made on a dynamical basis in terms of a low-dimensional phase space approach. The basis for such an approach is the assumption that variabilities of the key SST forcing regions relevant to seasonal prediction can be effectively represented by a suitably constructed low-dimensional phase space. In a low-dimensional phase space, the time evolution of the predictors, such as the ENSO signal, can be closely studied in terms of orbit structure and probability measure, which allows a closer examination of the phase relation with the predictand. The nearest neighborhood approach profits from such a dynamical basis, which allows an analysis of the origin of predictive skill. Because the sources of the predictive skills are different, the final prediction can be optimized based on the BLUE applied to predictions derived from different SST forcing regions. Both the spatial distribution (Fig. 15) and the time behavior (Fig. 17 and Fig. 10, lower panel) of the skill scores indicate that meaningful forecasting for practical applications can be expected from the present scheme.

The seasonal predictability associated with a forcing field can be conveniently examined on a low-dimensional basis, leading to a clearer conceptual understanding of the underlying system. The SST anomaly associated with the ENSO signal can be effectively represented by a phase space spanned by three ST-PCs derived from the tropical Pacific SST. The low-dimensionality thus established is consistent with the results derived from the dynamic system model (Wang and Fang 1995, 1996) and the coupled numerical model (Chang et al. 1995). When this three-dimensional phase space is used for the prediction of winter SAT over North America, based on a nonlinear method, significant skills with correlation scores over 0.5 can be observed for most of the continent, which is very significant when compared with the most significant area of the persistence forecasting (about 0.3). In the nonlinear model, regions with negative correlation skill scores, which appear in most linear models, are almost reduced to 0 (see Fig. 7).

Significant skill scores are also observed from the Atlantic SST associated mainly with the changes of the Gulf Stream using three-dimensional phase space spanned by ST-PCs 2–4. The geographical distribution, as well as the time behavior of the skill scores, are different from those obtained with ENSO signals, indicating the possibility of increasing the predictive skill scores and stabilizing their time behavior by an optimal linear combination of the predictions derived from these two key SST forcing regions (Fig. 15). Using the best linear unbiased estimates [see Sarda et al. (1996) and

references therein], the mean correlation skill score is increased by 46%. The area with a skill score higher than 0.5 is increased significantly (cf. Figs. 7 and 15). The increase of the skill scores due to extratropical North Pacific SST is not so significant as the Atlantic predictors, though still considerably significant (Fig. 15b). The gain of the predictive skill via BLUE is genuine, for there are no systematic negative skill scores in the individual phase space model and no negative value is observed in the weight (α) for the BLUE (5), as shown in Fig. 16.

It was also suggested that there is a great potential to improve further the seasonal prediction by selecting models and their predictors according to climate zone and season. In the present application for winter SATs, linear regression models can produce significant skill in tropical and subtropical areas when using ENSO signals as predictors. In midlatitudes, the nonlinear methods produce the best prediction, while the linear regression is the worst with negative skill scores (Fig. 7). In higher latitude, though significant skill can also be observed with linear models, the skill scores are more significant with nonlinear models that are based on the nearest neighborhood method. Moreover, the latter yield more robust skill, in the sense that their significant skills change less with geographical location or with the verification period, than those of linear models. These changes are derived from the phase relation of the predictors and predictand which is subjected to modulations by low-frequency fluctuation, such as interdecadal changes [cf. Figs. 10 and 11, and see, e.g., Trenberth (1990) and Wang (1995)].

The prediction of seasonal climate in phase space relies very much on the possibility of constructing a low-dimensional phase space based on observational data. The present application indeed suggests such a possibility with the help of physical insights derived from diagnostic analysis and dynamic models (See section 4a). It is expected that the seasonal prediction can be further improved with the development of our knowledge about the coupled dynamics of the climate system, in particular, the atmosphere–ocean system (e.g., Wang and Fang 1996; Chang et al. 1995) and their long time behavior [e.g., Ghil et al. (1991); Pandolfo (1993) and Haines (1994), for reviews], for which the role of observational study and GCM experiments is not to be underestimated (Van den Dool 1994b; Palmer and Anderson 1994). Meanwhile, looking for better (and supplementary) predictors for a particular season and region, such as sea ice, snow cover, soil moisture, and so on, is a never ending effort in the realm of statistical forecasting.

Acknowledgments. The author gratefully acknowledges Drs. A. Barnston (CPC, Washington, DC), G. Brunet (RPN, Dorval), J. Derome (McGill University, Montreal), G. Plaut (Institut Non-Lineaire de Nice), H. Ritchie (RPN, Dorval), H. Van den Dool (CPC, Washington, DC), R. Vautard (LMD, Paris), B. Wang (University of Hawaii), and F. Zwier (Victoria University) for helpful discussions and constructive comments. In particular, Drs. R. Vautard and G. Plaut kindly provided the SST data and the ANAXV programs, without which the present work would have been impossible. Drs. G. Brunet, H. Van den Dool, and B. Wang carefully read the manuscript and gave invaluable comments and suggestions which have led to substantial improvements of the present paper. Dr. G. Plaut has kindly checked the calculations independently. The daily SAT data were kindly provided by the Climate Prediction Center (A. Barnston and J. Hoopingarner). The author also wishes to thank the editor and two anonymous reviewers for constructive comments. Special thanks are also due to Drs. B. Dugas (RPN) and A. Plante (RPN) for technical and graphic support, and to the RPN for the hospitality. The research has been supported by an NSERC fellowship from Environment Canada, in association with the Historical Forecast project (HFP) and CLIVAR project.

REFERENCES

Abarbanel, H. D. I., R. Brown, J. J. Sidorowich, and S. Tsimring, 1993: The analysis of observed chaotic data in physical systems. *Rev. Mod. Phys.,* **65,** 1331–1392.

Barnett, T. P., 1981: Statistical prediction of North American air temperatures from Pacific predictors. *Mon. Wea. Rev.,* **109,** 1021–1041.

——, and R. W. Preisendorfer, 1978: Multifield analog prediction of short-term climate fluctuations using a climate vector. *J. Atmos. Sci.,* **35,** 1771–1787.

——, and ——, 1987: Origin and levels of monthly and seasonal forecast skill for United States surface air temperatures determined by canonical correlation analysis. *Mon. Wea. Rev.,* **115,** 1825–1850.

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

——, and H. M. van den Dool, 1993: A degeneracy in estimated skill in forecasts using regression-based cross-validation designs. *J. Climate,* **6,** 963–977.

——, and T. M. Smith, 1996: Specification and prediction of global surface temperature and precipitation from global SST using CCA. *J. Climate,* **9,** 2660–2697.

Bergen, R. E., and R. P. Harnack, 1982: Long-range temperature prediction using a simple analog approach. *Mon. Wea. Rev.,* **110,** 1083–1099.

Bjerknes, J., 1969: Atmospheric teleconnections from the equatorial Pacific. *Mon. Wea. Rev.,* **97,** 163–172.

Brunet, G., and R. Vautard, 1996: Empirical normal modes versus empirical orthogonal functions for statistical prediction. *J. Atmos. Sci.,* **53,** 3468–3489.

Casdagli, M., 1991: Nonlinear prediction of chaotic time series. *Physica D,* **35,** 335–356.

Chang, P., L. Ji, B. Wang, and T. Li, 1995: Interaction between the seasonal cycle and El Niño–Southern Oscillation in an intermediate coupled ocean–atmosphere model. *J. Atmos. Sci.,* **52,** 2353–2372.

Eckmann, J. P., and D. Ruelle, 1985: Ergodic theory of chaos and strange attractors. *Rev. Mod. Phys.,* **57,** 617–656.

Elsner, J. B., and A. A. Tsonis, 1992: Nonlinear prediction, chaos, and noise. *Bull. Amer. Meteor. Soc.,* **73,** 49–60.

Farmer, J. D., and J. J. Sidorovich, 1987: Predicting chaotic time series. *Phys. Rev. Lett.,* **59,** 845–851.

Fraedrich, K., and R. Wang, 1993: Estimating the correlation di-

mension from noisy and small data set based on re-embedding. *Physica D,* **65,** 373–398.
——, ——, and S. Pawson, 1993: An EOF analysis of the vertical–time delay structure of the quasi-biennial oscillation (QBO). *J. Atmos. Sci.,* **50,** 3357–3365.
Ghil, M., M. Kimoto, and J. D. Neelin, 1991: Nonlinear dynamics and predictability in the atmosphere sciences. *Rev. Geophys.* (Suppl.), 46–55.
Haines, K., 1994: Low-frequency variability in atmospheric middle latitudes. *Surv. Geophys.,* **15,** 1–61.
Livezey, R. E., and A. G. Barnston, 1988: An operational multifield analog/antianalog prediction system for United States seasonal temperatures. 1. System design and winter experiments. *J. Geophys. Res.,* **93,** 10 953–10 974.
——, and B. K. Neumeister, 1990: Mixed analog/persistence prediction of United States seasonal mean temperatures. *J. Climatol.,* **10,** 329–340.
Michaelsen, J., 1987: Cross-validation in statistical climate forecast models. *J. Climate Appl. Meteor.,* **26,** 1589–1600.
Palmer, T. N., and D. L. T. Anderson, 1994: The prospects for seasonal forecasting—A review paper. *Quart. J. Roy. Meteor. Soc.*, **120,** 755–793.
Pandolfo, L., 1993: Observational aspects of the low-frequency intraseasonal variability of the atmosphere in middle latitudes. *Advances in Geophysics,* Vol. 34, Academic Press, 93–174.
Plaut, G., and R. Vautard, 1994: Spells of low-frequency oscillations and weather regimes in the Northern Hemisphere. *J. Atmos. Sci.,* **51,** 210–236.
Rasmusson, E. M., and T. H. Carpenter, 1982: Variation in tropical sea surface temperature and surface wind fields associated with the Southern Oscillation/El Niño. *Mon. Wea. Rev.,* **110,** 354–384.
Sarda, J., G. Plaut, C. Pires, and R. Robert, 1996: Statistical and dynamical long-range atmospheric forecasts: Experimental comparison and hybridization. *Tellus,* **48A,** 518–537.
Sauer, T., J. A. Yorke, and M. Casdagli, 1991: Embedology. *J. Stat. Phys.,* **65,** 579–616.
Takens, F., 1981: Detecting strange attractors in turbulence. *Proceedings of the Warwick Symposium 1980, Lecture Notes in Mathematics,* D. A. Rand and L.-S. Young, Eds., Springer-Verlag, 366–381.
Tangang, F. T., B. Tang, A. H. Monahan, and W. W. Hsieh, 1998: Forecasting ENSO events: A neural network–extended EOF approach. *J. Climate,* **11,** 29–41.
Ting, M.-F., M. P. Hoerling, T.-Y. Xu, and A. Kumar, 1996: Northern Hemisphere teleconnection patterns during extreme phases of the zonal-mean circulation. *J. Climate,* **9,** 2614–2633.
Trenberth, K. E., 1990: Recent observed interdecadal climate changes in the Northern Hemisphere. *Bull. Amer. Meteor. Soc.,* **71,** 988–993.
van den Dool, H. M., 1994a: Searching for analogues, how long must we wait? *Tellus,* **46A,** 314–424.
——, 1994b: Long-range weather forecasts through numerical and empirical methods. *Dyn. Atmos. Oceans,* **20,** 247–270.
——, and Z. Toth, 1991: Why do forecasts for ''near normal'' often fail? *Wea. Forecasting,* **6,** 76–85.
Vautard, R., C. Pires, and G. Plaut, 1996: Long-range atmospheric predictability using space–time principal components. *Mon. Wea. Rev.,* **124,** 288–307.
——, G. Plaut, R. Wang, and G. Brunet, 1999: Seasonal prediction of North American surface air temperatures using space–time principal components. *J. Climate,* **12,** 380–394.
Walker, G. T., and W. Bliss, 1932: World Weather V. *Mem. Roy. Meteor. Soc.,* **4,** 53–84.
Wallace, J. M., R. L. Panetta, and J. Estberg, 1993: Representation of the equatorial stratospheric quasi-biennial oscillation in EOF phase space. *J. Atmos. Sci.,* **50,** 1751–1762.
Wang, B., 1994: On the annual cycle in the tropical eastern central Pacific. *J. Climate,* **7,** 1926–1942.
——, 1995: Interdecadal changes in El Niño onset in the last four decades *J. Climate,* **8,** 267–285.
——, and Z. Fang, 1996: Chaotic oscillation of the tropical climate: A dynamic system theory for ENSO. *J. Atmos. Sci.,* **53,** 2786–2802.
Wang, R., 1991: Propagating waves in the EOF-decomposition of space–time-delay diagrams. *Ann. Geophys.,* **9** (Suppl.), C588–C599.
——, 1994: A phase-space approach to atmospheric dynamics based on observational data—Theory and applications. *Meteorologishe Abhandlungen: Neue Folge, Serie A: Monographien,* Vol. 8, No. 2, Verlag von Dietrich Reimer, 227 pp.
——, and B. Wang, 2000: Phase space representation and characteristics of El Nino–La Niña. *J. Atmos. Sci.,* **57,** 3315–3333.
——, K. Fraedrich, and S. Pawson, 1995: Phase–space characteristics of the tropical stratospheric quasi-biennial oscillation. *J. Atmos. Sci.,* **52,** 4482–4500.
Weare, B. C., and J. S. Nasstrom, 1982: Examples of extended empirical orthogonal function analysis. *Mon. Wea. Rev.,* **110,** 481–485.
Whitney, H., 1936: Differentiable manifolds. *Ann. Math.,* **37,** 645–680.
Zebiak, S. E., and M. A. Cane, 1987: A model El Niño–Southern Oscillation. *Mon. Wea. Rev.,* **115,** 2262–2278.
Ziehmann-Schlumbohm, C., L. Smith, and K. Fraedrich, 1995: Ein internes Vorhersagbarkeitsexperiment im Lorenz Modell. *Meteor. Z.,* **4,** 16–21.