Techniques of Linear Prediction for Systems with Periodic Statistics

K. HASSELMANN

Max-Planck-Institut für Meteorologie, Hamburg, West Germany

T. P. BARNETT

Scripps Institution of Oceanography, La Jolla, CA 92093

(Manuscript received 3 November 1980, in final form 15 June 1981)

ABSTRACT

Many parameters that measure climatic variability have nonstationary statistics, that is, they depend strongly on the phase of the annual cycle. In this case normal statistical analysis techniques based on time-invariant models are inappropriate. Generalized methods accounting for seasonal nonstationarity (phase averaged or cyclostationary models) have been developed to treat such data.

The methods are applied to the problem of predicting El Niño off South America. It is shown that El Niños may be predicted up to a year in advance with considerably more confidence and accuracy using phase-averaged models than with time-invariant models.

In a second application surface air temperature anomalies are predicted over North America from Pacific Ocean sea surface temperatures. Again, the phase-averaged models consistently outperform models based on standard statistical procedures.

1. Introduction

Most methods of linear prediction assume the statistics of the data fields to be time invariant. Recent summaries of this work, as it applies to geophysics and particularly climate prediction, may be found in Davis (1978), Hasselmann (1979), and Barnett and Hasselmann (1979, hereafter called BH). In the present paper we modify the latter work to include situations in which the statistics are periodic in time. Similar techniques have been applied in other fields, particularly in econometrics (cf. Parzen and Pagano, 1979; Pierce, 1980; Zellner, 1979; and Swamy and Tinsley, 1976). However, a general application in meteorology does not appear to have been previously published. The generalization of BH to include seasonality yields a method of statistical modeling applicable to climatic phenomena which show strong dependence on the phase of the annual cycle, such as the effect of air/sea heat flux variations on the atmospheric system. Numerous analyses of climatic data sets suggest such methods will be required in many regions of the world.

After the derivation of the general method of modeling systems with periodic statistics in Sections 2–5, the technique is illustrated in Section 6 by application to two examples of general practical interest, the forecasting of El Niño events and the prediction of surface air temperature over North America. The results are compared with those obtained from current time-invariant and simple seasonally stratified (fixed-phase) models.

2. The unfiltered general case

Consider a discrete predictand time series $y(t)$ which we attempt to estimate q time lags into the future using past and present values of a set of n predictor time series $x_i(t)$ by means of a general linear relation

$$\hat{y}(t+q) = \sum_{i=1}^{m'} \sum_{l=0}^{m} D_{itl} x_i(t-l), \tag{1}$$

where $\hat{y}$ denotes the predicted value as opposed to the true value y and the dependence on the lead time index q is implicit in the coefficient D_{itl}. In contrast to BH we do not regard the system as time-invariant (D_{itl} independent of the reference time t), but assume instead that D_{itl} is periodic in t with a basic periodicity p,

$$D_{itl} = D_{i(t+p)l} = D_{i\kappa l}, \tag{2}$$

where $\kappa = t(\mathrm{mod}\,p)$, i.e., $t = p\nu + \kappa$, where ν is an integer and $0 \leqslant \kappa \leqslant p-1$. In our applications the periodicity corresponds to the annual cycle with $p = 12$ or 4 for monthly or seasonal data; κ is then a month or season of the year, and ν is the year index.

Similarly, assume that all statistical quantities are invariant with respect to a time shift p, e.g.,

0022-4928/81/102275-09$06.25

$$\langle x_i(t)x_j(t+l)\rangle$$
$$= \langle x_i(t+p)x_j(t+l+p)\rangle = R_{ij}(\kappa, l), \quad (3)$$

where angle brackets denote ensemble means. Processes whose second moments satisfy the periodicity condition (3) are often referred to as cyclostationary in the wide sense with period p (cf. Gardner, 1978).

The least-squares solution of (1) is obtained by minimizing

$$\langle[\hat{y}(t) - y(t)]^2\rangle = \epsilon_\kappa^2. \quad (4)$$

Since the system is not statistically stationary, the quantity ϵ_κ^2 is not time invariant, but depends on the phase κ of the annual cycle. Thus, in contrast to the time-invariant case, we now have not one but p least-squares solutions, one for each month/season κ. We note that the seasonality can occur independently in the second moments of the predictand, in the predictors, or in the covariances of the predictand and predictors. Only the latter two lead to seasonally dependent coefficients $D_{i\kappa l}$ representing the coupling between the predictors and the predictand.

The problem is separable with respect to κ. For each κ, the least-squares solution is found exactly as in the time-invariant case. This may be seen by rewriting Eq. (1) as

$$\hat{y}^\kappa(\nu) = \sum_\alpha D_\alpha^\kappa x_\alpha^\kappa(\nu), \quad (5)$$

where

$$\hat{y}^\kappa(\nu) \equiv \hat{y}(t_{\kappa+q}^\nu),$$
$$x_\alpha^\kappa(\nu) \equiv x_i(t_{\kappa-l}^\nu),$$

with

$$t_\kappa^\nu \equiv p\nu + \kappa,$$
$$D_\alpha^\kappa \equiv D_{i\kappa l}$$

and $\alpha \equiv (i, l)$.

For fixed phase κ (month, season), Eq. (5) is identical to the time-invariant prediction problem, with the exception that the time ν changes by units of one year. By seeking the minimum of (4) for each κ one obtains a set of optimal prediction coefficients $D_{i\kappa l}$ for each κ. This class of models we call fixed-phase models, for they are solved explicitly for each of the p phase elements of the annual cycle, with no approximation or filtering. The error analysis, model nesting techniques, EOF ordering, etc., all carry over as in the time-invariant case described in BH and Barnett (1981a). However, the estimates of moments are not formed as time averages over all points of the time series, but as averages over consecutive years ν at the same phase κ (month). Thus the estimate of the covariance function (3), for example, is given by

$$\tilde{R}_{ij}(\kappa, l) = N^{-1} \sum_{\nu=1}^{N} x_i(t_\kappa^\nu)x_j(t_{\kappa+l}^\nu), \quad (6)$$

where N is the length of the data sets in years.

The form (6) has to be kept in mind when determining the estimation errors of the covariance matrices given in the Appendix of BH. It yields higher estimation errors than averages taken over all time points and so makes it harder to construct a statistically significant model. This is because the number of statistically independent contributions to the sum in Eq. (6) is generally of order N, whereas for continuous time averaging the appropriate number is N/τ, where τ is the correlation time scale (in years), which in many applications is of order 0.3–0.5.

3. Smoothing the κ-dependence

To construct a statistically significant prediction model, the number of coefficients must be limited. However, if the modulated prediction problem is treated with the full annual-cycle resolution, as described above, a large number of coefficients is unavoidable. It is possible to reduce the number with respect to the predictor and lag indices i and l, for each κ, by the techniques described in BH, but one still requires a separate set of coefficients for each phase value κ. The large number of coefficients is partly offset by the fact that in the modulated prediction problem more information is used, namely, the full κ dependence of the lagged covariance functions (3). However, the net effect of working with the full annual cycle resolution is generally to increase the statistical indeterminacy. Essentially, this is due simply to the larger sampling errors associated with averaging over a time series with annual rather than monthly or seasonal increments.

Improved statistical significance may be obtained by sacrificing some annual-cycle resolution. This may be achieved by representing the κ dependence of the coefficients by a smaller number of basis functions.

An obvious choice of basis functions is the Fourier expansion

$$D_{i\kappa l} = \sum_{\beta=1}^{s} E_{i\beta l} g_\beta(\kappa), \quad (7)$$

with

$$g_1 = 1, \quad g_2 = \cos\left(\frac{2\pi\kappa}{p}\right),$$
$$g_3 = \sin\left(\frac{2\pi\kappa}{p}\right), \cdots \quad (8)$$

and $s \leqslant p$. In many applications a termination of the series at $s = 3$ will be adequate. This procedure yields a phase-averaged model in which the κ dependence of the prediction is smoothed and the number of independent coefficients thereby reduced. [We note that a fixed phase model as defined in Section 2 can be formally included in the definition (7) by defining the basis functions g_β as "box car" functions. However, for small s this generally yields a less accurate description of a continuous system than a Fourier expansion, as is borne out in the examples considered in Section 6.]

If the representation (7) is introduced, all $12 \times N$ months or $4 \times N$ seasons of data must be used simultaneously in the least-squares determination of the coefficients, and the separability of the seasonal dependence is lost. On the other hand, the trade-off between the number of predictor fields, number of lags and resolution of the seasonal cycle can now be combined and varied at will. The main advantage of expanding the seasonal dependence of the model is this flexibility in the choice of trade-off. In keeping with the discussion of BH, however, it is essential that the combination of desired resolution characteristics is defined *a priori* in a one-dimensional model hierarchy.

The phase-averaged model becomes equivalent to the fixed-phase model in the full-resolution limit $s = p$. In many applications, however, the model hierarchy must be broken off before this limit is reached. In this case the fixed-phase model (although of higher apparent skill) must be rejected as statistically insignificant, whereas a less ambitious (although, less skillful) model with $s < p$ may remain statistically significant.

Substituting (7) into (1), one obtains

$$\hat{y}(t^{\nu}_{\kappa+q}) = \sum_{i=1}^{m'} \sum_{l=0}^{m} \sum_{\beta=1}^{s} E_{i\beta l} g_{\beta}(\kappa) x_i(t^{\nu}_{\kappa-l}). \qquad (9)$$

Compressing the subscripts (i, l) again into α, this may be written more concisely as

$$\hat{y}(t^{\nu}_{\kappa+q}) = \sum_{\alpha=1}^{r} \sum_{\beta=1}^{s} E_{\alpha\beta} g_{\beta}(\kappa) x_{\alpha}(t^{\nu}_{\kappa}), \qquad (10)$$

where $r = m' \times m$ and $x_{\alpha}(t^{\nu}_{\kappa}) \equiv x_i(t^{\nu}_{\kappa-l})$.

Introducing the new predictors

$$x_{\alpha}(t^{\nu}_{\kappa}) g_{\beta}(\kappa) \equiv z_{\gamma}(t), \qquad (11)$$

where $\gamma \equiv (\alpha, \beta)$ and t again represents the original discrete time variable running through the entire time series, one finally obtains

$$\hat{y}(t+q) = \sum_{\gamma=1}^{n} a_{\gamma} z_{\gamma}(t), \qquad (12)$$

where $n = m' \times m \times s$ and $a_{\gamma} \equiv E_{\alpha\beta}$.

Eq. (12) is seen to be formally identical to the time-invariant prediction model considered in BH. However, the statistics of the variables $\hat{y}$, z_{γ} are not time invariant. Thus the expectation value of the square error ϵ_{κ}^2, Eq. (4), depends still on the phase κ. The optimal model may be defined then as the solution which minimizes the phase-averaged error $\epsilon^2 = \sum_{\kappa} G_{\kappa} \epsilon_{\kappa}^2$, where G_{κ} is some arbitrarily chosen positive weighting distribution. In the following we set $G_{\kappa} \equiv 1$. The analysis for determining the least-squares solution reduces in this case to the time-invariant case.

Following the standard time-invariant procedure, it is convenient to orthonormalize the predictor time series by defining new predictors

$$z_{\gamma}(t) = \sum_{k} \sigma_k z_k'(t) T_{k\gamma}, \qquad (13)$$

where σ_k^2 represent eigenvalues and $T_{k\gamma}$ are the eigenvectors of $\langle z_k z_j \rangle$. The transformation (13) yields $\langle z_i' z_j' \rangle = \delta_{ij}$. In this coordinate system the prediction coefficients a_{γ}' become simply

$$\alpha_{\gamma}' = \langle y' z_{\gamma}' \rangle, \qquad (14)$$

where y' is the predictand scaled by its standard deviation, so that $\langle (y')^2 \rangle \equiv 1$ (see BH).

The hindcast skill of the model in the coordinate system defined by (13) is given by

$$S_H = \sum_{\gamma=1}^{n} (a_{\gamma}')^2. \qquad (15)$$

Eq. (15) represents the hindcast skill averaged over all seasons. We shall also consider later a fixed-phase hindcast skill S_H^{κ} dependent on the season (month), defined as the average of the skill for given κ over all years.

4. Model order and significance

The model skill S_H [Eq. (15)] increases monotonically with the order n (number of predictors) of a model. However, the statistical significance of a model generally decreases with n. A central problem in constructing models from data, therefore, is arriving at a proper balance between model skill and significance. Methods for doing this are discussed in BH, Hasselmann (1979), and Barnett (1981a). The relevant relations for the present applications are summarized below.

The significance of a model can be expressed in terms of the statistic

$$\rho^2 = \sum_{i,j} \mathbf{M}_{ij}^{-1} \delta a_i' \delta a_j', \qquad (16)$$

where $\delta a_i' = a_i' - a_i^0$ denotes the difference between the least-squares-fit model parameters estimated from a single, finite data realization and the "true" model a_i^0 derived from a (hypothetical) infinite data ensemble. The matrix $\mathbf{M}_{ij} = \langle \delta a_i' \delta a_j' \rangle$ represents the covariance matrix of sampling errors. The statistic ρ^2 can be seen to be a χ^2 variable with n degrees of freedom (orthogonalization of the variables $\delta a_i'$ yields the sum of n squares). A model of order n can accordingly be defined as statistically significant at the confidence level c with respect to the null hypothesis of zero predictability, $a_i' = 0$, if ρ^2 exceeds the appropriate critical value $\chi_{n,c}^2$ of the variable χ_n^2 for this confidence level.

The expression (16) corresponds to the quadratic form with the highest "equivalent number of degrees of freedom" or, expressed geometrically, to the

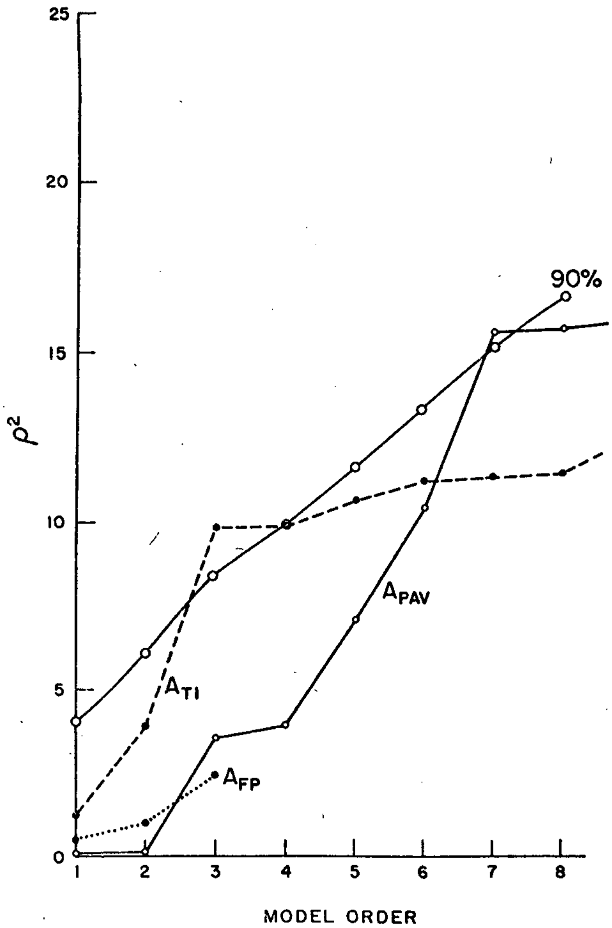

FIG. 1. Model significance tests: El Niño prediction at 12-month lead. The significance parameter (ρ^2) is shown for the time-invariant model A_{TI}, phase-averaged models A_{PAV} and fixed phase models A_{FP}. The 90% confidence limit is also indicated.

quadratic form whose ellipsoidal isosurfaces ρ^2 = const. bound the smallest spatial volume for a given probability volume (cf. Müller *et al.*, 1978; Hasselmann, 1979, BH). In this sense ρ^2 defines an optimal statistic for significance tests. An alternative quadratic form which is also often used to assess model significance is the skill (which can be compared against the artificial skill of the null model $a_i^0 = 0$). However, this statistic generally is characterized by a smaller equivalent number of degrees of freedom and provides less discrimination than the statistic ρ^2.

To arrive at a suitable trade-off between model skill and significance, the significance test is applied not simply to a single model, but to a hierarchy of models. Following BH, the selection of an optimal model from a model hierarchy is carried out in a series of steps:

1) The complete set of potential predictors is first orthogonalized and then ordered with respect to their variance, i.e., with respect to the eigenvalues of $\langle z_i z_j \rangle$.

2) The eigenvectors of this set are then filtered to eliminate higher order predictors that are indistinguishable from spatial/temporal white noise (cf. Preisendorfer and Barnett, 1977). This reduces the original (large) predictor set to a group of N, ordered orthogonal predictors, thereby prescribing an upper limit (N) on the possible order of the model.

3) A nested model hierarchy is then considered, consisting of models using the first, first two, . . . , first n, . . . , all N predictors. For each member (n) of the hierarchy the statistic ρ_n^2 is evaluated.

4) Finally, an optimal model is selected from the hierarchy, and the significance of the selected model is evaluated.

In BH a selection criterion was suggested in which the model with the largest n was chosen for which the model (n) and all lower order models exceeded a prescribed confidence level c: $\rho_m^2 > \chi_{m,c}^2$ for all m in $1 \leq m \leq n$. In practice, the requirement that all models in the range $1 \leq m \leq n$ exceed a given probability level without exception can represent a very stringent condition, requiring particularly good *a priori* insight into the proper ordering of the lowest order predictors. In the examples discussed in BH the condition was actually tacitly relaxed by accepting models for which the higher order members of the hierarchy were all significant up to the cutoff, but the first one or two were insignificant. However, it has been pointed out by R. Davis (private communication) that the number of insignificant models "ignored" at the beginning of the hierarchy can have a non-negligible influence on the true significance of the selected high-order model, and the selection criterion should therefore be clearly defined *a priori* in order to properly evaluate the statistical significance of the final optimal model.

A number of possible selection criteria have been discussed and evaluated statistically (by Monte Carlo simulations) in Barnett *et al.* (1981). In the following examples we adopt the simplest strategy in which the "optimal" model is defined as the model (n) whose ρ_n^2 value yields the highest individual significance level of the entire set $1 \leq n \leq N$. Since the model (n) is selected from a larger set of competitors, the true significance of the optimal model cannot be obtained simply by entering the statistic ρ_n^2 in a standard χ_n^2 table, but must be inferred (as a function of ρ_n^2 and N) from the results of the numerical experiments presented in Barnett *et al.* (1981).

5. Interpreting model results

In BH the analogy between the prediction coefficients a_γ' and Green's functions are discussed.

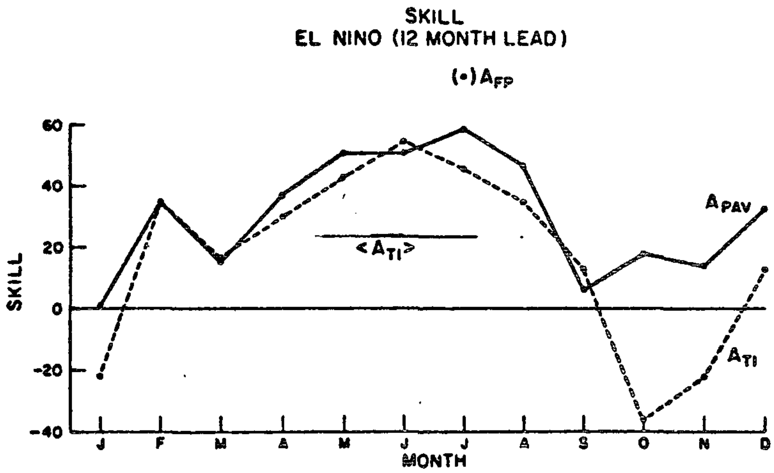

FIG. 2. Skill, by month, for various one-year-in-advance El Niño prediction models. Skill values are in percent.

It was shown that empirical statistical prediction models can generally be related to predictions derived from differential equations with appropriate driving fields. For a similar interpretation of the seasonally modulated case it is convenient to divide the net prediction coefficients (response functions) into their seasonally varying components. Substitution of (7) into (10) yields for a phase-averaged Fourier model, with $s = 3$,

$$\hat{y}(t+q) = \sum_{\alpha} x_a(t)[E_{\alpha 1} + E_{\alpha 2}\cos(2\pi\kappa p^{-1}) + E_{\alpha 3}\sin(2\pi\kappa p^{-1})], \quad (17)$$

where the κ dependence has been written out explicitly in the harmonic terms and α represents, as before, the combined lag subscript l and predictor variable subscript i. The relation between a_γ' and $E_{\alpha\beta}$ is given by (13) and (14).

The coefficients $E_{\alpha 1}$ correspond to the standard time-invariant response functions used in BH, whereas the coefficients $E_{\alpha 2}$ and $E_{\alpha 3}$ describe the seasonality or nonstationarity of the response function. Thus, if $|E_{\alpha 1}| \gg (E_{\alpha 2}^2 + E_{\alpha 3}^2)^{1/2}$ the prediction is essentially stationary to first order and expansions of the form (7) and (8) are unnecessary. If, on the other hand, the seasonal coefficients ($E_{\alpha 2}, E_{\alpha 3}$) are comparable with $E_{\alpha 1}$, the associated time series are nonstationary and the dependence on seasonal phase must be included in the analysis of their properties (or in the construction of differential equations designed to simulate their behavior).

The phase of the seasonal harmonic of the net coefficient $D_\alpha^\kappa = E_{\alpha 1} + E_{\alpha 2}\cos(2\pi\kappa p^{-1}) + E_{\alpha 3}\sin \times (2\pi\kappa p)$ is $\theta = \tan^{-1}(E_{\alpha 3}/E_{\alpha 2})$. We adopt the convention that $\kappa = 0$ corresponds to winter (seasonal data) or January (monthly data), so that $\theta = 0$ or π implies maximal seasonal modulation between winter and summer (or January and July). A strong seasonal modulation of D_α^κ implies that the seasonal variability of the predictand cannot be explained alone by the seasonal variability of the predictor, but must be due to a seasonal change of the physical processes which link the two. In the examples we investigated the modulation of D_α^κ was generally pronounced.

6. Examples

We illustrate the results of the previous sections by two examples of particular practical interest: 1) the prediction of sea surface temperature (SST) anomalies off Peru, i.e., El Niño events and 2) the prediction of air temperature over North America.

a. Predicting El Niño

Barnett (1981a) used the time invariant formalisms of BH to show that El Niño events may be predicted marginally at lead times of one year from knowledge of the trade wind field (TWF) alone. No effort was made to optimize the model. However, it was pointed out in that study and elsewhere that SST in the tropical Pacific has a high level of persistence and is intimately related to the east/west slope of sea level across the ocean basin. It also is known, from a variety of sources, that El Niños tend to occur at certain seasons of the year. We now use these facts to construct a more detailed El Niño prediction model for a lead time of one year.

As predictand, we consider the SST anomaly at Talara, Peru. The predictors consist of several regions of the TWF found useful in predicting ocean/atmosphere variables in the tropical Pacific, a measure of the east/west sea level slope variability, and prior values of the predictand. All data series rep-

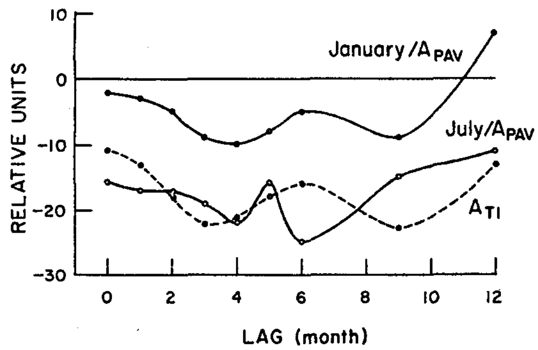

FIG. 3. Prediction coefficients relating SST and trade wind strength from El Niño prediction models at 12-month lead.

resent monthly values extending over 20 years. More details on the predictand/predictors definition and characteristics are given in Barnett (1981a). After orthonormalization [Eq. (13)] and subsequent filtering by the Preisendorfer-Barnett (1977) technique, only eight predictor time series remained for model building.

Three prediction models were constructed. The first (A_{TI}) was time invariant. The second (A_{FP}) was a fixed-phase model designed to forecast Talara SST in July, a time when El Niños normally reach high intensity. The third model (A_{PAV}) was a first-harmonic ($s = 3$) phase-averaged model of the form (9). All predictions were made for a lead time of one year and used the previous 12 months' data in the lagged predictor fields.

The significance measures of the models are shown in Fig. 1. Model A_{TI} (time-invariant) is significant at the 90% level, as is model A_{PAV} (nonstationary). Model A_{FP}, the fixed-phase model for July, is not significant for any order.

The seasonally dependent skill S_H^κ of each model is shown in Fig. 2. The model A_{FP} had a high skill (76%), but was not significant at even the 90% level and must, therefore, be rejected. (This stresses again the fact that model skill and significance are not synonymous.) The mean skill of A_{TI}, indicated by $\langle A_{TI} \rangle$, is 26%. Although the time invariant model assumes a constant skill, the actual evaluation of A_{TI} on a monthly basis shows the skill to vary widely over the course of the year. In October, November and January the skill is negative. The model does well during the summer. It is clear that the average skill normally associated with such a time-invariant model (26%) is misleading in a nonstationary situation. The continuous, nonstationary model (A_{PAV}) shows positive skill for all months. The skill is particularly strong during the summer when El Niños generally reach their highest intensity.

As example of the model transfer functions (Section 5), Fig. 3 shows the coefficients relating the prior years variations of one of the predictors (the zonal component of the southeast trade winds, USET) to the year-in-advance SST fluctuations. For the time-invariant model, only $E_{\alpha 1}$ exists; its lag dependence is indicated by the curve A_{TI}. The seasonally modulated response functions for A_{PAV} are shown for January and July (since the phase $\theta = \arctan(E_{\alpha 3}/E_{\alpha 2})$ is approximately 0 or π for this predictor, the months shown correspond to the maximal and minimal values of the net coefficients). The A_{TI} results show a strong (time-lagged) net correlation between USET and SST averaged over the year. The A_{PAV} results, on the other hand, show that this relation in fact exists for summer fluctuations in SST, but not winter variations (this implies, for $\theta \approx 0$, $E_{\alpha 1} \approx E_{\alpha 2}$). The predictive relationship is clearly determined largely by the covariance of the predictor with the summer SST anomalies. This explains why for summer the skill of the model A_{TI} is nearly as good as the skill from A_{PAV}. It also explains the negative skill of A_{TI} during the winter, for the assumed predictand/predictor relationship does not hold for this season; it is merely an artifact of the strong summertime correlation.

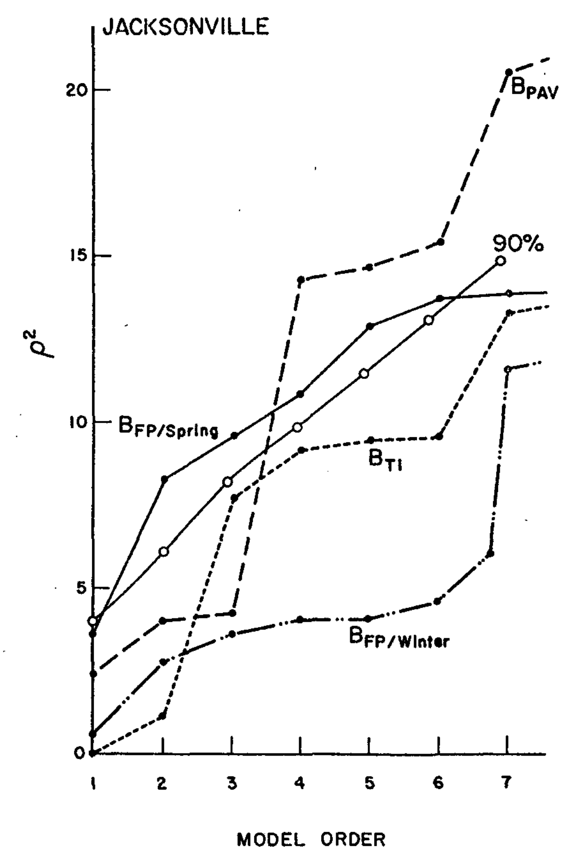

FIG. 4. Model significance tests. North American surface air temperature at one-season lead.

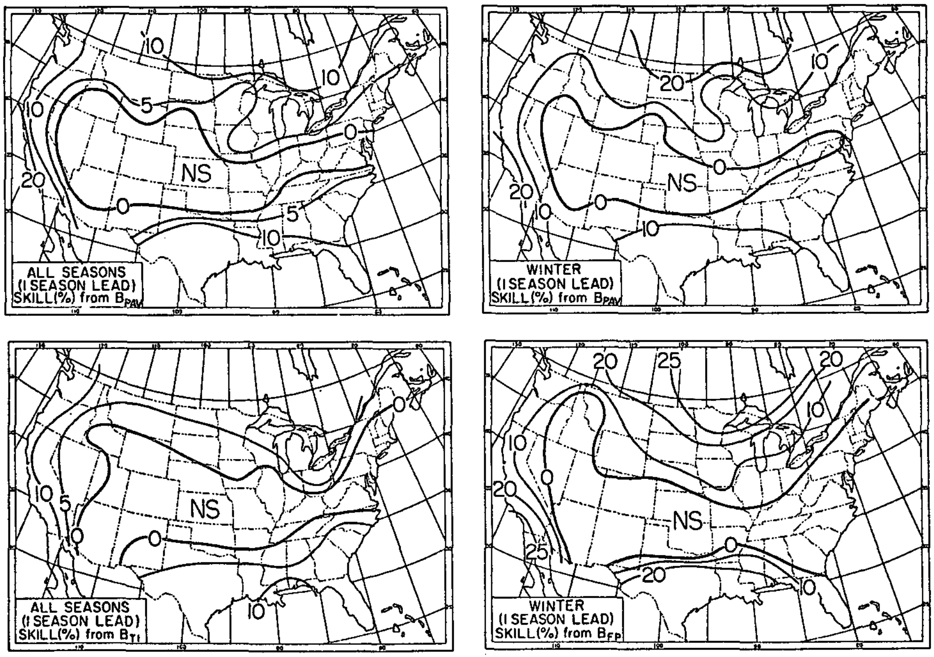

Fig. 5. Skill of hindcasting North American surface air temperature one season in advance. Left: annual average skill for phase-averaged (upper) and time invariant (lower) models; right: winter skill for phase-averaged (upper) and fixed phase (lower) models.

b. Predicting seasonal air temperature over North America

The relation between fluctuations in ocean/atmosphere parameters in the Pacific and subsequent air temperature anomalies over North America further illustrate the ideas of Sections 3 and 4. A detailed analysis of the Pacific/North American interactions is given elsewhere (Barnett, 1981b). Here we excerpt from that study some of the results relating SST to air-temperature prediction.

Three prediction models are again constructed: 1) model B_{TI} assumed the statistics of all data series were stationary; 2) models B_{FP} were fixed phase, designed to predict only one season at a time, and 3) model B_{PAV} was a phase-average model of the form (9). In this case the difference between models B_{FP} and B_{PAV} is rather small because the time series consisted of seasonally averaged data (for the period 1902–72) and the number of free coefficients in B_{PAV} is therefore only one less than in B_{FP}. All models forecast air temperature at a lead time of one season and used the previous four seasons data in the SST predictor field. The SST predictors come from equatorial and midlatitude regions identified by Bjerknes (1966, 1969) and Namias (1975), respectively, as being potentially important to subsequent events over North America.

The significance measure of the models is shown in Fig. 4 for a typical station (Jacksonville). A model of the form B_{TI} could not be constructed at this station with 90% significance. Fixed phase models (B_{FP}) were found to be significant (>90%) for three seasons of the year. The highest and lowest ρ^2 distributions for these models are shown. The other two seasons fall between these extremes. The skill of the significant fixed-phase models ranged from 6–22%. This range, plus the seasonal dependence, explains the poor performance of B_{TI}. The phase averaged model was significant at a higher level than its competitors, as expected.

The skill values for the different classes of model are shown in Fig. 5 for 36 North American stations. The left-hand panels show the annually averaged skill for stationary (B_{TI}) and phase-averaged (B_{PAV}) models. These averages are somewhat misleading, however, since they contain strong seasonal variability. The seasonal character of the ocean/atmosphere interaction that gave the predictive skill (Barnett, 1981b) clearly makes B_{TI} inferior to B_{PAV}. Comparison of the ability of models B_{FP} and B_{PAV} to forecast winter air temperatures is shown in the right-hand panel of Fig. 5. The fixed-phase models have somewhat higher skill at many individual stations, as expected, since the models were tuned for the particular station/season forecast. Note, however, that models of the form B_{FP} could not be constructed at many stations where B_{PAV} shows low, but significant skill.

We note that significant models B_{PAV} occasionally

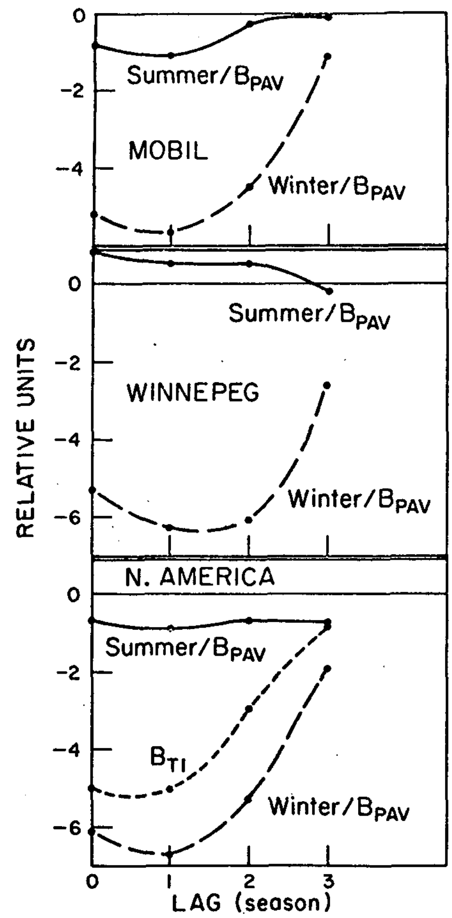

FIG. 6. Seasonally modulated prediction coefficients relating equatorial sea surface temperature and North American surface air temperature at one season lead.

produce negative skill scores at some stations/seasons. In most cases this occurred when the station was rather strongly predictable in only one season, and the representation of the seasonal modulation in terms of the lowest harmonic yielded a rather poor approximation. Models B_{TI} also generally gave negative skills for these stations/seasons. Models of the form B_{FP} could generally not be constructed in these cases at an acceptable level of significance.

Examples of the seasonally modulated response functions with respect to the most important predictor, SST off Peru, are shown in the upper panels of Fig. 6 for two widely separated individual predictand stations, Mobil and Winnepeg. A systematic representation of these response functions for the entire predictand field can be developed in terms of principal predictors (cf. Davis, 1978; Barnett, 1981b). In the present case the technique yields as dominant interaction the response function relating SST off South America (predictor) to the numerically largest pattern of predictable air temperature fluctuations over North America (lower part of Fig. 6). The strong seasonality of the interaction stands out clearly in all three examples. Winter forecast skill is strongly dependent on the equatorial SST during the prior three seasons (fall–spring). The summer forecast, on the other hand, is only weakly dependent on equatorial SST (or, indeed, any other predictor, since summer predictability is much lower than that obtained for winter). The stationary models B_{TI} yield a strong net coupling between the tropics and midlatitudes averaged over the year. This annually averaged correlation is misleading, however, since the neglect of seasonal modulation in this model leads to negative skill scores for many stations/seasons (not shown), as illustrated previously in the example of Section 5a.

7. Conclusions

A general approach for modeling systems with periodic statistics has been developed. The techniques should be particularly useful for short- and medium-range climate prediction, for which the strong annual variability of the basic climatic system cannot be neglected. To achieve maximal statistical significance, *phase-averaged* models were introduced, in which the model coefficients were expanded in a harmonic series and statistical averages were estimated by continuous time averages over the entire time series. The method enables all available data to be used in the estimation of a relatively small number of model coefficients. The technique may be contrasted with *fixed-phase* models constructed for a single time point of the cycle. Although conceptually simpler, these must necessarily be derived from a smaller data base (for a given phase) and therefore are more difficult to construct with adequate statistical significance.

Both model types were tested together with standard time-invariant models in two hindcasting examples: 1) the prediction, one year in advance, of El Niño conditions off South America, and 2) the seasonal forecasting of surface air temperature anomalies over North America. In both cases significant predictability was obtained with phase averaged models, but not always with the time-invariant models. The phase-averaged models were found to be superior to both time-invariant and fixed-phase models in both examples.

Acknowledgments. The authors gratefully acknowledge the support of the National Science Foundation's Office of Climate Dynamics, under Contract ATM-7918206 (TPB), and the Max-Planck-Society visitors program in carrying out this work.

REFERENCES

Barnett, T. P., 1981a: Statistical relations between ocean/atmosphere fluctuations in the tropical Pacific. *J. Phys. Oceanogr.*, **11**, 1043–1058.

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

——, and K. Hasselmann, 1979: Techniques of linear prediction, with application to oceanic and atmospheric fields in the tropical Pacific. *Rev. Geophys. Space Phys.*, **17**, 949–968.

——, L. M. Goldstein, K. Hasselmann and R. W. Preisendorfer, 1981: Significance tests for regression model hierarchies. *J. Phys. Oceanogr.*, **1**, 1150–1154.

Bjerknes, J., 1966: Survey of El Niño 1957–58 in its relation to tropical Pacific meteorology. *Inter-Amer. Trop. Tuna Comm. Bull.*, **12**, 62 pp.

——, 1969: Atmospheric teleconnections from the equatorial Pacific. *Mon. Wea. Rev.*, **77**, 163–172.

Davis, R. E., 1978: Predictability of sea level pressure anomalies over the North Pacific Ocean. *J. Phys. Oceanogr.*, **8**, 233–246.

Gardner, W. A., 1978: Stationarizable random processes. *IEEE Trans. Inform. Theory*, **IT-24**, 8–22.

Hasselmann, K., 1979: Linear statistical models. *Dyn. Atmos. Oceans*, **3**, 501–521.

Müller, P., D. J. Olbers and J. Willebrand, 1978: The IWEX spectrum. *J. Geophys. Res.*, **83**, 479–500.

Namias, J., 1975: *Short Period Climatic Variations. Collected Works of J. Namias 1934 through 1974*, 2 Vols. University of California Press, San Diego, 905 pp.

Parzen, E., and M. Pagano, 1979: An approach to modeling seasonally stationary time series. *J. Econometrics*, **9**, 137–153.

Pierce, D. A., 1980: A survey of recent developments in seasonal adjustment. *Amer. Statist.*, **34**, 125–134.

Preisendorfer, R. W., and T. P. Barnett, 1977: Significance tests for empirical orthogonal functions. *Preprints Fifth Conference on Probability and Statistics*, Las Vegas, Amer. Meteor. Soc., 169–172.

Swamy, P. A. V. B., and P. A. Tinsley, 1976: Linear prediction and estimation methods for regression models with stationary stochastic coefficients. Spec. Stud. Pap. 78, Federal Reserve Board.

Zellner, A., Ed., 1979: Seasonal Analysis of Economic Time Series. Economic Res. Rep. ER-1, U.S. Dept. of Commerce, Bureau of the Census, 485 pp.