Extraction of Mixed Layer Advection Velocities, Diffusion Coefficients, Feedback Factors and Atmospheric Forcing Parameters from the Statistical Analysis of North Pacific SST Anomaly Fields

K. HERTERICH AND K. HASSELMANN

Max-Planck-Institut für Meteorologie, Hamburg, FRG

(Manuscript received 1 September 1985, in final form 15 April 1987)

ABSTRACT

The statistical properties of observed North Pacific sea surface temperature (SST) anomalies are simulated by a simple mixed layer advection and diffusion model with stabilizing feedback and local stochastic forcing by the atmosphere. An optimal fit of the model to the SST auto- and cross-spectra yields the effective temperature advection velocities and diffusion coefficients in the mixed layer, the local feedback factors and the strength and scales of the atmospheric forcing. The results obtained by model fitting are in general agreement with independent direct estimates, where such data are available. The analysis supports previous models in which the origin of midlatitude SST anomalies on time scales of months to a few years is attributed to stochastic forcing by the atmosphere.

1. Introduction

The oceans exert a strong influence on climate through the storage and transport of heat, and since the essential ocean parameter which controls the transfer of heat from the ocean to the atmosphere is the sea surface temperature (SST), the nature of the processes which determine the SST distribution in the ocean has been much discussed in the climate literature. For studies of climate variability, the main focus of interest lies in the structure and origin of SST anomalies. It is generally accepted that in the tropics SST anomalies are generated through large scale ocean–atmosphere feedback processes (cf. review by Anderson, 1983). In midlatitudes, on the other hand, investigations suggest that SST anomalies can be largely explained, at least on time scales smaller than about 2–3 years and outside regions affected by strong boundary currents, simply as the integrated response of the oceanic mixed layer to the external short time scale stochastic forcing associated with the natural internal "weather" variability of the atmosphere. We shall be concerned here with midlatitude SST anomalies of this kind.

From a statistical time series analysis of atmospheric pressure fields and SST anomaly patterns in the North Pacific, Davis (1976) demonstrated that midlatitude SST anomalies generally follow the atmospheric pressure fields, rather than vice versa. Frankignoul and Hasselmann (1977) and Frankignoul (1979) showed further that midlatitude SST anomalies can be described by a Langevin equation (first-order Markov process), in which the atmospheric forcing is represented on the SST time scale as a white noise process. The theoretical red distribution of the SST response spectra agreed well with observations, and where the appropriate input data were available, both the hypothesized white noise input and the predicted form of the input–response cross-spectra were verified. In basinwide studies, Reynolds (1978, 1979) was able to confirm the model for most of the northern midlatitude oceans. However, the model fit tended to degrade in regions of strong currents. Haney (1980, 1985) provided further corroboration of these concepts by successfully simulating the principal features of observed SST anomaly fields in midlatitudes with a numerical ocean model driven by observed ocean–atmosphere heat flux and surface wind stress anomalies.

In the present paper we extend the local stochastic forcing models of Frankignoul–Hasselmann and Reynolds to include the effects of horizontal advection and diffusion. Our motivation is twofold. First, the degradation in the fit of the spatially decoupled local response model in the vicinity of strong currents, together with estimates of the order of magnitudes of the advection and diffusion terms, indicate that these processes cannot generally be neglected. Second, in trying to extract information on the dynamics of the mixed layer from SST anomaly data, the limitation to spatially decoupled local response models restricts the available input data to the SST autovariance spectra. By generalizing the model to include spatial interactions it is possible to use also the more extended information contained in the SST cross-spectra to derive improved parameter estimates and gain additional insight into the structure of the horizontal transport processes.

The second aspect is perhaps the more interesting

of the two. The improvement in the fit of the model to the auto-spectra achieved by the inclusion of horizontal transports is not dramatic: in most grid points, a large percentage of the variance can be explained already by a local model without transports. On the other hand, by applying inverse modeling techniques to the cross-spectra it is possible to derive extended fields of effective horizontal transport quantities in the mixed layer from SST anomaly data. An advantage of this technique is that the transports are estimated with respect to mixed-layer integrated variables, i.e., in the form needed for SST modeling, within the same framework as the other processes (local forcing and feedback) which determine the midlatitude SST anomaly field.

The advection fields derived from the SST anomaly data are found to be generally consistent, within the statistical error limits of the data, with surface current distributions estimated from ship drift and other observations. The other parameters derived from the model are also found to be in order of magnitude agreement with independent estimates, where these are available. The extended model therefore confirms the previous view that midlatitude SST anomalies on time scales of a few months to a few years are generated primarily through the passive response of the ocean to stochastic forcing arising from the natural internal variability of the atmospheric circulation, rather than through large-scale ocean–atmosphere feedback mechanisms. This is supported also by the work of Frankignoul and Reynolds (1982), in which the local stochastic model was extended by including advection estimated from measurements, rather than from model fitting.

It should be stressed, however, that our model is idealized in several respects. We mention only two obvious limitations: First, the effect of large scale, nonlocal feedbacks between the atmosphere and the ocean is ignored. Teleconnection signals will undoubtedly be present on longer time scales, for example, through coupling to El Niño, but pronounced effects have not been found in midlatitudes in the time scales of interest here. Second, our analysis is based on stationary, rather than cyclo-stationary statistics. The annual modulation of all statistical fields and model parameters could have been readily introduced into the analysis, at the cost of some algebraic complexity and degradation of statistical significance. However, the purpose of the present paper is to demonstrate that useful first-order estimates of annually averaged forcing functions, feedback coefficients and transport fields can be derived alone from SST anomaly data using a rather simple stochastic forcing model. The present model, which represents already an extension of the still simpler models of Frankignoul and Hasselmann (1977) and Reynolds (1978, 1979), can clearly be elaborated still further using the model hierarchy and statistical testing methods described below.

A similar stochastic model with horizontal transport processes has been applied in the inverse modeling mode by Lemke et al. (1980) to simulate the statistical properties of Arctic and Antarctic sea ice anomalies.

In section 2 we describe the data used for our analysis. Sections 3–5 describe the model and model-fitting technique, while section 6 is concerned with statistical aspects of the model skill, error and significance. Results are presented in section 7 and are discussed in section 8. Technical background on the derivation of the covariance matrices of the data and model parameters is given in appendices A and B.

2. Dataset

The initial data consisted of time series of monthly mean sea surface temperatures (SST) for the years 1947–75 on a 5° square grid of the North Pacific from the NORPAX dataset (137 grid points). The mean annual cycle was subtracted from each of these time series to yield anomaly fields, which provided the basic data set for our analysis.

Figures 1 and 2 show as example the SST anomaly time series at 35°N, 170°E and the corresponding variance spectrum (on a logarithmic scale), computed with two different frequency resolutions $\Delta\nu$. The high resolution spectrum (a) of Fig. 2 was obtained by averaging the maximum resolution spectrum ($\Delta\nu = (1/320)(\text{months})^{-1} = 0.0375$ cpy) over three neighboring frequency bands, yielding a spectrum with 6 degrees of freedom and $\Delta\nu = 0.1125$ cpy. Spectrum (b) was computed by the Tukey method as the average over the individual variance spectra for ten separate 32-month "chunks" of the 320 month record (20 degrees of freedom, $\Delta\nu = 0.375$ cpy). Figure 3 shows the cross-spectrum (on a linear scale) of the time series of Fig. 1 correlated with the SST anomaly time series of the eastern neighboring grid point, computed again by the Tukey method.

We shall use statistically stable spectra (two sided) of the type Fig. 2b and Fig. 3 throughout this paper. This implies that we limit ourselves to the frequency range between 0.375 cpy (32 month period) and 6 cpy. Thus the longer multiyear time scales evident in Fig.

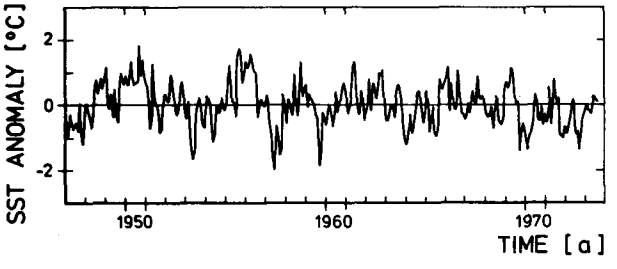

FIG. 1. Time series of SST anomalies at 35°N, 170°E for the years 1947–73.

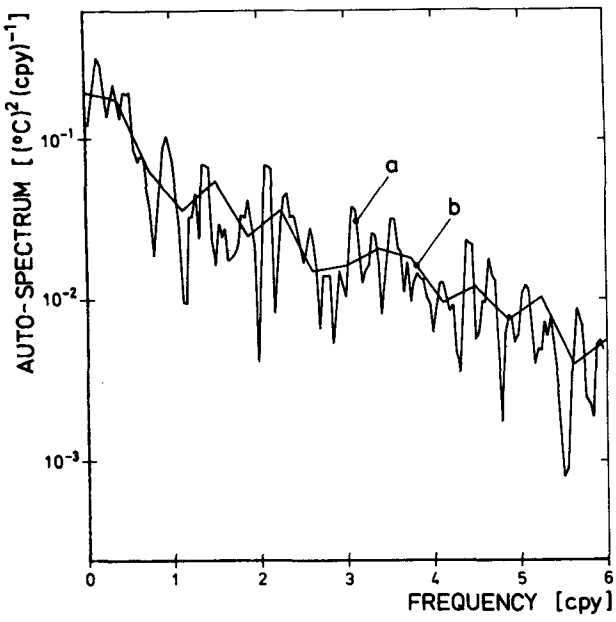

FIG. 2. High (a) and low (b) resolution autospectrum for the time series of Fig. 1.

1 are excluded from our analysis. A comparison with the higher resolution spectrum (a) of Fig. 2 indicates that in averaging over several frequency bands we do not appear to be smoothing out details of the low frequency structures in the spectral bands we have retained. (This procedure would be more questionable for SST spectra in the tropics, which have higher energy levels at periods beyond 1–3 years.) The time scales resolved in our spectra are consistent with the processes we wish to study: the formation of SST anomalies in the mixed layer by atmospheric forcing and local feedback, and the distortion of the anomaly fields by local advection and diffusion processes.

In our model construction we use the autospectra $F_{ii}(\omega)$ and cross-spectra $F_{ij}(\omega)$ for all four nearest neighbor grid points. Since the quadrature spectra and the differences between auto- and co-spectra are not significantly different from zero at high frequencies, we excluded frequencies higher than 3.5 cpy from the analysis. We further excluded the zero frequency band, which lies below our low-frequency limit of (32 months)$^{-1}$. The final dataset used to construct the model consisted then of nine frequency points for each auto- and cross-spectrum, yielding $9 + 4 \times (9 \times 2) = 81$ data values for all interior grid points (84 in total) for which all four nearest neighbor cross-spectra could be constructed.

3. The model

Following Frankignoul (1979) we assume that the statistical behavior of SST anomalies T is governed by a mixed-layer transport equation

$$\frac{\partial T}{\partial t} + \mathbf{v} \cdot \nabla T = -\lambda T + D\nabla^2 T + n, \qquad (3.1)$$

where $\mathbf{v}$ represents the (annual mean) advection velocity, integrated over the depth of the mixed layer, λ is a feedback parameter, D an (isotropic) diffusion coefficient and n the (white noise) atmospheric forcing. The forcing represents primarily fluctuations in the heat flux across the ocean–atmosphere interface, but variable turbulent entrainment effects are also included formally in this term. The model parameters $\mathbf{v}$, λ, D and the statistical properties of n are to be determined from the data.

Since our input data are auto- and cross-spectra, we first derive the relevant theoretical model spectra from (3.1). Discretizing spatially in terms of centered differences, Eq. (3.1) may be written for each grid point j for which all four nearest neighbors (denoted N, S, E and W) exist:

$$\frac{\partial T_j}{\partial t} + \frac{V_x}{2\Delta x}(T_E - T_W) + \frac{V_y}{2\Delta y}(T_N - T_S) = -\lambda T_j + D\left(\frac{T_W + T_E - 2T_j}{\Delta x^2} + \frac{T_N + T_S - 2T_j}{\Delta y^2}\right) + n_j. \qquad (3.2)$$

Equation (3.2) may be written more simply by collecting the advection and diffusion terms into a single transport term characterized by a (sparse) matrix **A**:

$$\frac{\partial T_j}{\partial t} + \lambda_j T_j + \sum_k A_{jk} T_k = n_j. \qquad (3.3)$$

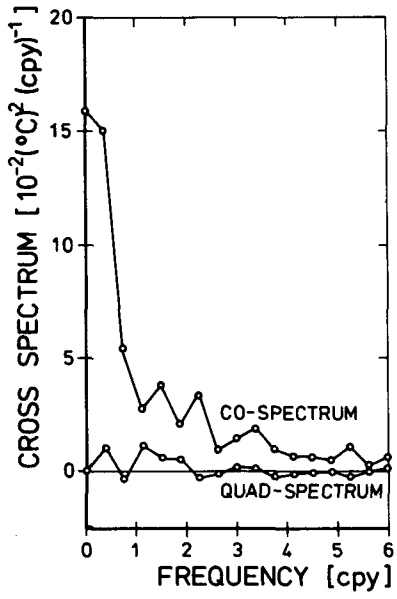

FIG. 3. Low resolution cross-spectrum between the time series of Fig. 1 with its eastern neighboring grid point. (Note that for the imaginary part only three frequency points between 1 and 2 cpy are significantly different from zero at the 1σ level.)

In the frequency domain Eq. (3.3) becomes (retaining the same symbols for the Fourier transforms, with $\omega = 2\pi\nu$):

$$(i\omega + \lambda_j)T_j(\omega) + \sum_k A_{jk}T_k(\omega) = n_j(\omega). \quad (3.4)$$

The formal solution of (3.4) is given by

$$T_j(\omega) = \sum_k (\mathbf{S}^{-1})_{jk} \frac{n_k(\omega)}{(i\omega + \lambda_k)}, \quad (3.5)$$

where the matrix $\mathbf{S}$ is given by

$$S_{jk}(\omega) = \delta_{jk} + \frac{A_{jk}}{i\omega + \lambda_j}, \quad (3.6)$$

and δ_{jk} is the Kronecker symbol. From (3.5) it follows that the model response cross-spectra $F_{jk}(\omega) = \langle (T_j(\omega))^* T_k(\omega) \rangle$ and the atmospheric forcing cross-spectra $N_{jk}(\omega) = \langle (n_j(\omega))^* n_k(\omega) \rangle$ (where the cornered parentheses denote ensemble means and the asterisk indicates the complex conjugate) are related through

$$F_{jk}(\omega) = \sum_{l,m} (\mathbf{S}^{-1})^*_{jl} (\mathbf{S}^{-1})_{km} \frac{N_{lm}}{(\lambda_l - i\omega)(\lambda_m + i\omega)}. \quad (3.7)$$

For a discrete time series, the time derivative $\partial T/\partial t$ is replaced throughout in the above expressions by the (forward first order) finite difference expression

$$[T(t + \Delta t) - T(t)]/\Delta t$$

where t is the time and Δt is the time increment. In the discrete Fourier representation the term $\pm i\omega$ has to be replaced by the discretized equivalent $(e^{\pm i\omega\Delta t} - 1)$.

4. The model hierarchy

We seek the simplest model which is able to reproduce the observed auto- and cross-spectra. For this purpose we construct a hierarchy of models 1 to 4 containing an increasing number of free parameters (cf. Table 1).

The first two models contain no horizontal transport terms and attempt to explain both auto- and cross-spectra alone in terms of the spatial correlation structure of the stochastic forcing, with spatially dependent white noise forcing level and feedback parameter λ. In model 1 the forcing is taken to be isotropic, while in model 2 an elliptic correlation structure is assumed, with principal axes in the north–south, east–west direction. In both cases the dependence on the separation vector $\Delta x_{jk} = (r_1, r_2)$ between point pairs (j, k) is taken as Gaussian:

$$\langle n_j^* n_k \rangle = N_{jj} \exp\left[-\left(\frac{r_1^2}{R_1^2} + \frac{r_2^2}{R_2^2}\right)\right], \quad (4.1)$$

where $R_1 = R_2 = R$ for model 1. A generalization of the form (4.1) to an arbitrarily oriented ellipse yielded no significant improvement over model 2.

TABLE 1. Definition of the model hierarchy (see also section 4).

Model number	Model properties
1	isotropic atmospheric forcing
2	elliptic correlation of atmospheric forcing
3	model 2 + advection
4	model 3 + diffusion

The cross-spectrum of a white noise vector process n_j contains no quadrature component, since the spectrum is identically equal to its value at zero frequency, where the quadrature spectrum, as an odd function, must vanish. Thus models 1 and 2 are in principle unable to simulate the observed quadrature spectra. They are none the less useful as a reference to distinguish between those properties of the SST spectra which can be attributed to the spatial structure of the forcing alone and other features which can be explained only by introducing advective transport processes.

Model 3 represents an extension of model 2 through the inclusion of advection. Since this is an asymmetrical transport operator, it generates quadrature terms in the cross-spectra. Model 4, finally corresponds to the full transport equation (3.1) including both advection and isotropic diffusion. The diffusion term, as a symmetrical operator, affects mainly the co-spectra. It can be distinguished from the influence of the spatial correlation structure of the stochastic forcing through its different frequency dependence. However, the modifications induced by diffusion were found to be relatively minor, and for this reason the model was not generalized further to the case of an anisotropic diffusion tensor.

To compute the SST cross-spectra for models 3 and 4 from Eq. (3.7), the matrix $\mathbf{S}$ containing the transport matrix $\mathbf{A}$ has to be inverted. This was achieved using the expansion:

$$(\mathbf{S}^{-1})_{jk} = \delta_{jk} - \frac{A_{jk}}{i\omega + \lambda_j} + \sum_l \frac{A_{jl}A_{lk}}{(i\omega + \lambda_j)(i\omega + \lambda_l)} - \cdots. \quad (4.2)$$

It was found in practice that the series could be terminated after the second term. Already in this approximation, however, successive application of the matrix $\mathbf{S}^{-1}$ in Eq. (3.7) involves computations of the forcing cross-spectra at point pairs further separated than the closest neighbors. For this reason the white-noise correlation function is specified in Eq. (4.1) for arbitrary pairs (j, k), rather than for only the closest neighbors, which would have sufficed for models 1 and 2.

5. Model fitting

For each grid point j a data vector $\mathbf{f}$ was formed consisting of the autospectrum of a central point and the real and imaginary components of the four cross-

spectra between the central point and its nearest neighbors (N, W, E and S), for all frequencies. Thus $\mathbf{f}^T = (\mathbf{f}^T(\omega_1), \cdots, \mathbf{f}^T(\omega_n))$, where $\mathbf{f}^T(\omega_i) = (\mathrm{Re}F_{j,N}(\omega_i), \cdots, \mathrm{Re}F_{j,S}(\omega_i), \mathrm{Im}F_{j,N}(\omega_i), \cdots, \mathrm{Im}F_{j,S}(\omega_i))$, and T denotes the transpose. The dimension of $\mathbf{f}$ is 81 (cf. section 2). In the same way we define a model vector $\hat{\mathbf{f}}$, which is given as a function of the model parameters.

The model is fitted to the data by minimizing the quadratic form

$$\epsilon = \Delta\mathbf{f}^T\mathbf{M}\Delta\mathbf{f} \tag{5.1}$$

with respect to the model parameters, where $\Delta\mathbf{f} = \mathbf{f} - \hat{\mathbf{f}}$ and $\mathbf{M}$ is a suitably defined weighing matrix. We choose $\mathbf{M}$ equal to the inverse of the covariance matrix $\mathbf{V} = \langle\delta\mathbf{f}\delta\mathbf{f}^T\rangle$ of the data errors $\delta\mathbf{f}$, $\mathbf{M} = \mathbf{V}^{-1}$ (for definition see Appendix A). This matrix has the well-known property of minimizing the statistical errors of the model parameters (maximum likelihood solution, cf. Martin, 1971).

If the data are sufficiently well defined statistically, the covariance matrix $\mathbf{V}$ can be estimated from the observations. However, if a reasonable fit to the data can be achieved by a model, a simpler method is to compute theoretically $\mathbf{V}$ from the model. For model significance tests (cf. next section) it is, in fact, more appropriate to define $\mathbf{V}$ by the model rather than the data. We shall regard $\mathbf{V}$ in the following as given by the model.

The variability covariance matrix $\mathbf{V}$ for spectral estimates can be expressed in terms of the auto- and cross-spectra by standard relations (cf. Jenkins and Watts, 1968, and Appendix A). Since the model spectra are not known initially, the minimum of ϵ, Eq. (5.1), is determined iteratively. Starting from an initial first guess for the parameters, which defines the initial covariance matrix of the data errors, a new set of parameters is obtained by minimizing the quadratic form ϵ and so on. The iteration generally converged rather rapidly.

In computing the model cross-spectra in accordance with (3.7), it was assumed that the fields were locally statistically homogeneous and that the model parameters were constant over the neighboring grid points involved in the fit for a given grid point. Although this is not strictly valid, it simplifies the analysis considerably by decoupling the model fits at different grid points. The approximation is consistent with a local description of the transport processes and incurs only small errors if the grid scale is small compared with typical spatial scales of the model parameters, as is normally the case.

6. Model skill, error and significance

The quality of a model which has been optimally fitted to data may be characterized by three quantities which, though interrelated, describe different properties of the model:

(1) The model *skill*

$$S = 1 - \frac{\sum_i (f_i - \hat{f}_i)^2}{\sum_i f_i^2}$$

provides a normalized measure ($S_{\max} = 1$) of how closely the model vector $\hat{\mathbf{f}}$ and data vector $\mathbf{f}$ lie together.

(2) The model *error* ϵ [Eq. (5.1)] provides a normalized measure of whether the model can be statistically distinguished from a perfect model in which the error is entirely due to data errors. For a perfect model and Gaussian data errors, ϵ is a χ^2-distributed random variable with $n - p$ degrees of freedom, where n is the number of data values and p the number of model parameters. If the error ϵ is larger than ϵ_c (the 95% cutoff of the χ^2-distribution), the model will be rejected.

(3) The *statistical significance* of the model can be inferred from the covariance matrix $\langle\delta\mathbf{p}\delta\mathbf{p}^T\rangle$ of the model parameter errors $\delta\mathbf{p}$ arising from the data errors $\delta\mathbf{f}$. For small errors, $\langle\delta\mathbf{p}\delta\mathbf{p}^T\rangle$ is linearly related to the covariance matrix $\mathbf{V} = \langle\delta\mathbf{f}\delta\mathbf{f}^T\rangle$. (See Appendix B.)

Apart from the differences in the quadratic form the skill is simply one minus the error (the choice of metric for the skill is essentially arbitrary, but it is customary to use the unit matrix).

The model significance is generally related to the model skill in the sense that for a low-order model high skill is normally also associated with small model errors. However, an increase of skill through an increase in complexity of a model is generally accompanied by a decrease in the model significance, so that the simultaneous requirements of model skill and significance generally involve some form of trade-off in model construction (cf. Barnett and Hasselmann, 1979).

To understand the interrelation between model skill (or error) and the statistical significance of the model parameter estimates it is useful to consider not only the overall model skill and error, but also the corresponding quantities defined for subsets of the data. In our case we consider the skill components S_a, S_c and S_q for the auto-, co- and quadrature-spectra. Different model parameters are found to affect the skills for different data subsets quite differently. Generally, a model parameter is well defined if it has a significant impact on at least one skill component.

The dependence of the overall model error and individual error contributions, integrated over all grid points, is shown as a function of the model order in Fig. 4. With respect to the autospectra, the extension of the model hierarchy beyond the original local forcing and feedback model of Reynolds (1978, 1979) clearly yields no net improvement. However, this model, in its simplest extension to an isotropic forcing field, is unable to explain the co- and quadrature-components of the cross-spectra, as indicated by high error levels

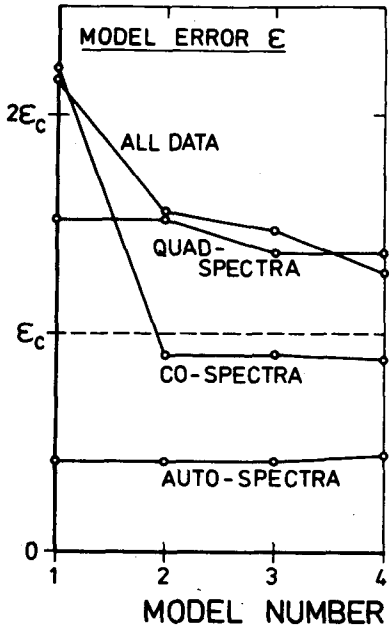

FIG. 4. Mean value of the error ϵ as a function of the model number, for the different subsets of the data (auto-, co- and quad-spectra) averaged over all grid points and normalized by ϵ_c, the critical error level at which the model is rejected.

of model 1 with respect to both of these datasets. The generalization of model 1 to include anisotropic forcing, model 2, significantly improves the simulation of the cospectra, but has no impact on the quadrature spectra, as expected. The introduction of advection, model 3, reduces the error of the quadrature spectra, without significantly affecting the errors in the other spectra. The introduction of the diffusion coefficient, finally has little impact on any of the data errors. The corresponding skill values are shown in Fig. 5.

The regional distribution of the three skill fields is shown in Figs. 6–8. Grid points at which the model is rejected are indicated by open circles. The skill values for the autospectra (model 1) are rather high (above 90% in the eastern North Pacific) and the model is rejected at only 2 grid points. The skill values for the covariance spectra using model 2, Fig. 7, are similarly uniformly high, although the model is now rejected at 20 grid points. It appears at first sight surprising that, in contrast to Fig. 4, the model error is found to be significant for several points in regions of Fig. 7 for which the skill is above 90%. This can be explained simply by the fact that the number of covariance spectra which are available for model testing is four times the number of autospectra, which makes it possible to detect smaller model errors. As anticipated on the basis of the nonsignificant mean error achieved already with model 2 for these data (Fig. 4), the atmospheric forcing and feedback parameters inferred from these models are all statistically significant (Table 2). However, the fields of diffusion coefficients, determined for model 4, were found to be less significant. Figure 8 shows the skill values for the quadrature-spectra, for model 3. As expected from the relatively small (but statistically significant) skill values for these data, and the significant regionally averaged error for this case, the model is now rejected at a large number of grid points (39 out of 84), (i.e., the improvement achieved by including advection is real, but there are also some real processes still missing in the model). Although imperfect, the model yields a statistically significant advection field especially in the eastern North Pacific (see Fig. 12).

7. Parameter values

The regional distribution of the parameters **p** inferred from the best fit model 4 are shown in Figs. 9–12 together with the mean parameter errors $\delta\mathbf{p}$. The mean ratio $\delta p/|p|$ is listed in Table 2. Lower order models yielded almost identical parameters, except that the relaxation times increased slightly when diffusion was included.

The auto spectrum N of the atmospheric forcing (Fig. 9) is found to be largest in the midlatitudes and decreases towards the equator and the pole. There is also a pronounced gradient from maximum values at the western side of the basin to lower values in the east. The spatial correlation scales (Fig. 10) of the atmospheric forcing [Eq. (4.1)] show a marked east–west elongation, and tend to be smaller where the forcing is largest. The principal features of Figs. 9 and 10 are consistent with the usual concept of atmospheric forcing in midlatitudes being due to eastward traveling disturbances, with strong air–sea temperature and humidity contrasts occurring in the western regions of the ocean where the principal air mass transformations take place. The local relaxation time λ^{-1} (Fig. 11) lies in the range 2–3 months and is consistent with previous estimates by Reynolds (1978) for a decoupled model.

The model advection field $\mathbf{v} = (v_x, v_y)$ exhibits some scatter (Fig. 12), but resembles the overall structure of

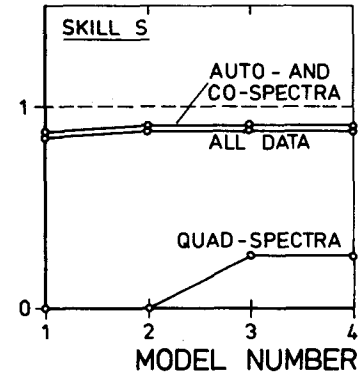

FIG. 5. Mean value of skill as a function of the model number for the different subsets of the data averaged over all grid points.

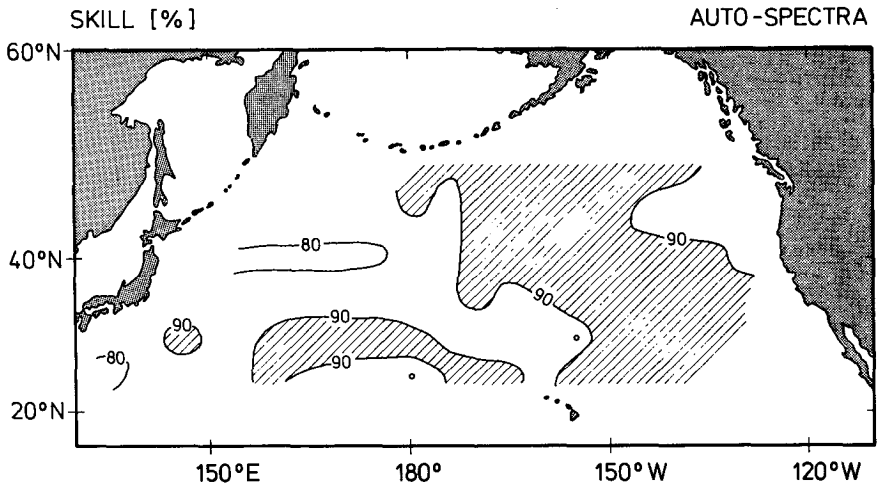

FIG. 6. Distribution of skill for modeling the autospectrum using model 1. At two grid points indicated by open circles, the model is rejected.

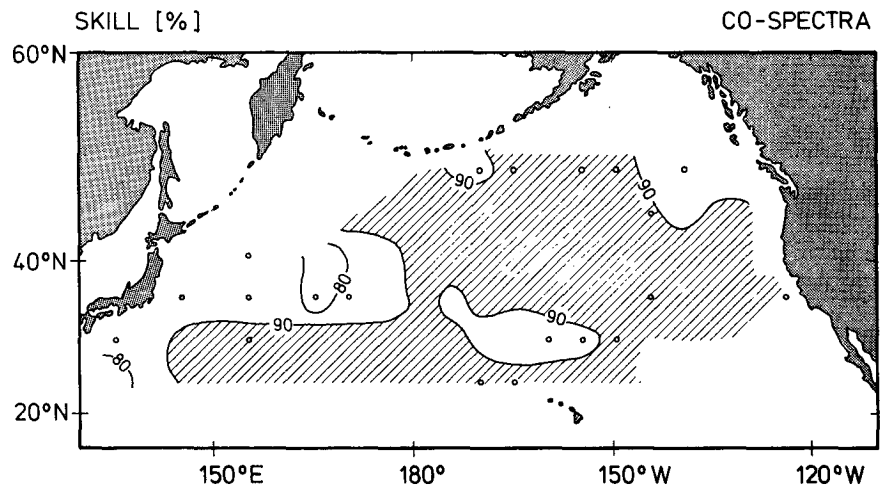

FIG. 7. Distribution of skill for modeling the real parts of the four nearest neighbor cross-spectra using model 2. At 20 grid points ("0") the model is rejected.

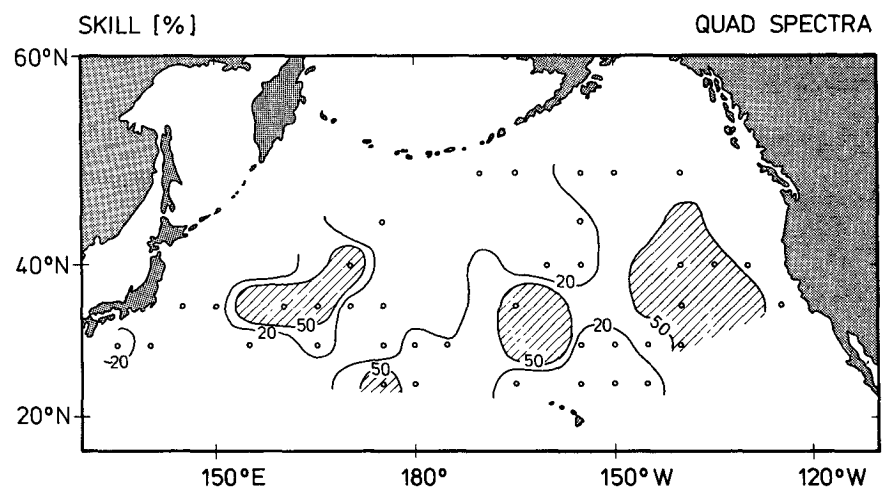

FIG. 8. Distribution of skill for modeling the imaginary part of the four nearest neighbor cross-spectra using model 3. At 39 grid points ("0") the model is rejected.

TABLE 2. Mean relative error of parameters averaged over all grid points for which at least model 4 is accepted. For definition of parameters see text.

	Parameter p						
	N_{ij}	R_1	R_2	λ	V_x	V_y	D
$\left\|\frac{\delta p}{p}\right\|$ [%]	11	38	12	13	71	66	63

the surface current field inferred from ship-drift observations (Fig. 13). The mean value of the advection velocity is about 4 cm s^{-1}, which is also not unrealistic. It should be noted that the model mean advection velocity fields represent 27 year averages over 5° × 5° squares and the depth of the mixed layer, so that features such as the Kuroshio Current are necessarily strongly smoothed. Nevertheless, some discrepancies can be seen, for example, in the northwest corner of the analysed region. These can perhaps be attributed to the fact that the model defines only a net effective advection velocity, including both the advection of anomalous SST patterns by the mean ocean currents and displacements of patterns through interactions with the mean wind.

The diffusion coefficient D (Fig. 14) lies in the range 10^3 to 10^4 m^2 s^{-1}, with lower values occurring towards the northeast. The order of magnitude is consistent with estimates based, for example, on dye-dispersion measurements (Okubo, 1971) and the coefficient is found to be always positive (which is not automatically ensured by the analysis). However, as pointed out in the previous section, the uncertainty of this parameter is of the same order as the estimate itself, so that this result may be regarded only as an order of magnitude estimate.

8. Discussion

We have shown that for a large area of the North Pacific the SST autospectra and cross-spectra between nearest neighbors can be modeled on a 5° square grid by a transport equation including advection and diffusion, a linear feedback term and spatially anisotropic white noise atmospheric forcing. In the case of autospectra, the consideration of the spatial correlation structures of the forcing and the advection and diffusion transports is not necessary to obtain a valid model for most grid points (Reynolds, 1979), and the model skill is not improved significantly for these data when these processes are included (cf. also Frankignoul and Reynolds, 1982). When the dataset is extended to the cross-spectra of neighboring grid point pairs, however, the model fit is successively improved by including the spatial structure of the forcing and the horizontal transport terms. The spatial forcing structure has the strongest impact on the cospectra, while the quadrature spectra are determined primarily by the advection terms. The inclusion of diffusion does not significantly improve the model fit.

The main purpose of considering an extended model, however, was not so much to improve the model fit to the data, but to derive information from SST anomaly data on the effective forcing, feedback and horizontal transport processes in the mixed layer. The parameter fields inferred from the model are found to be reasonably consistent with independent direct estimates, where these exist.

The spatial scales of the atmospheric forcing field agree with the scales of transient midlatitude synoptic disturbances (which must be regarded as the principal short term atmospheric forcing according to the white-noise forcing hypothesis). The level of the white noise forcing (on SST anomaly time scales smaller than 2–

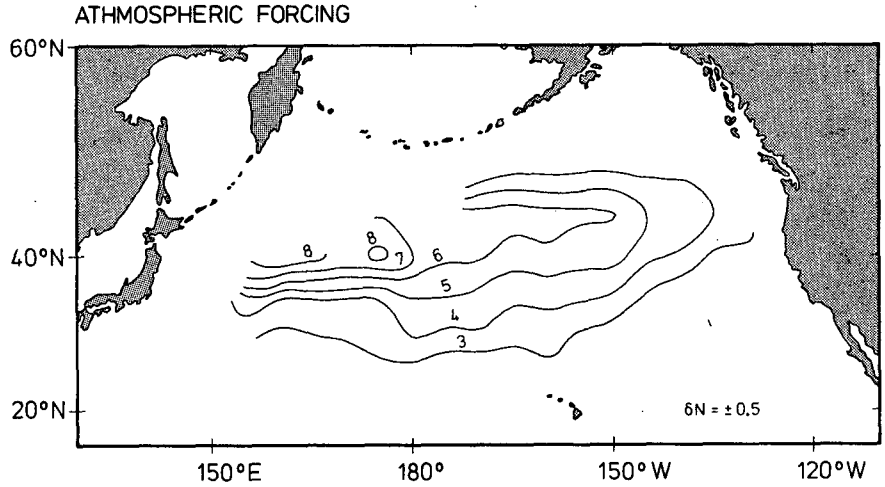

FIG. 9. Distribution of atmospheric forcing white noise autospectrum in $(°C\ yr^{-1})^2(cpy)^{-1}$. The rms error is $\delta N = 0.5$ $(°C\ yr^{-1})^2(cpy)^{-1}$.

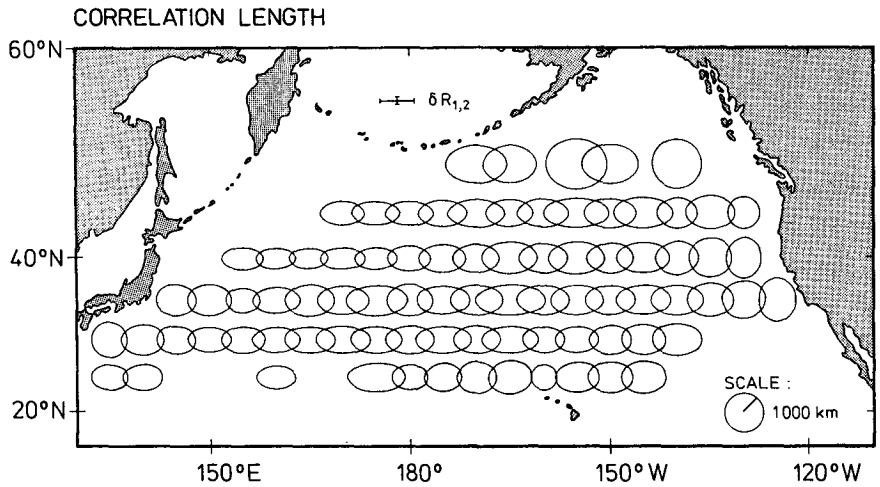

FIG. 10. Distribution of the correlation length scales R_1, R_2 of the model atmospheric forcing (defined as the e-folding distances of the white noise forcing).

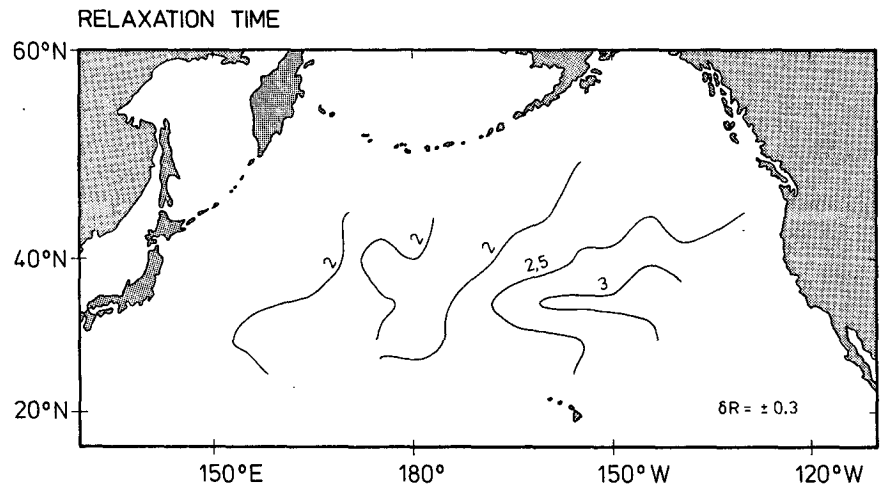

FIG. 11. Distribution of the relaxation time (months) of SST anomalies.

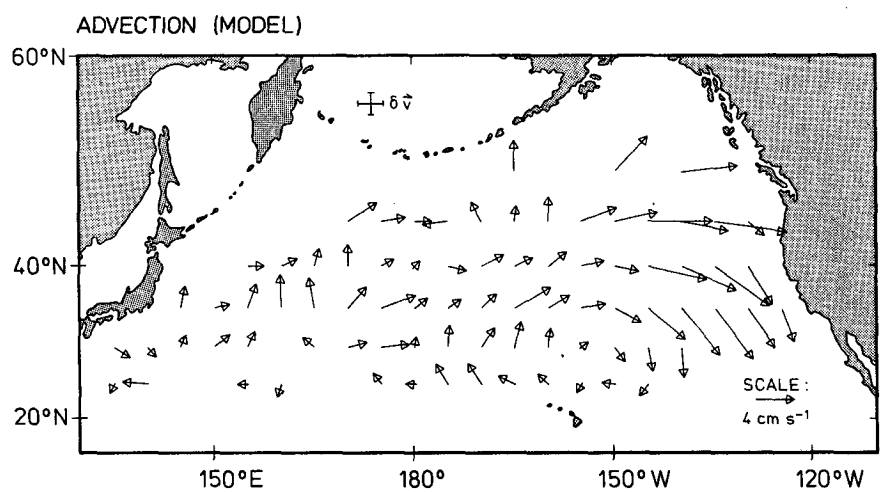

FIG. 12. Distribution of the model advection velocity (cm s^{-1}).

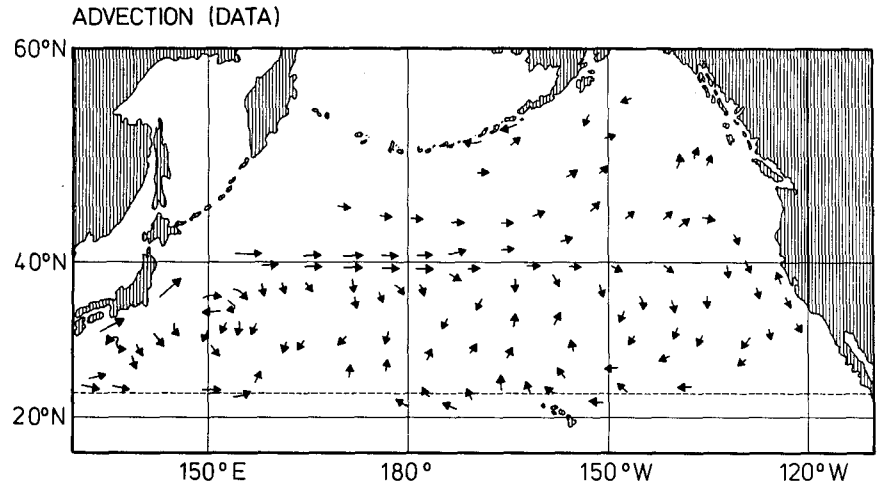

FIG. 13. Distribution of ship drift ocean surface currents in August. Only directional information is shown (after Gorschkov, 1976).

3 years) corresponds to a transient rms heat flux amplitude of order 20–30 W m^{-2}, assuming an atmospheric correlation time scale of 5 days and an average mixed layer depth of 70 m. If the white noise forcing is interpreted alternatively as a fluctuating oceanic advection term, driven by anomalous wind fields, the corresponding rms current velocity amplitude is estimated as 5 cm s^{-1}. Both estimates appear in reasonable agreement with measurements.

The inferred mean advection velocities are also consistent in magnitude and direction with the general surface velocity distributions derived from ship drift data. It should be noted, however, that the net advection velocities contain an unknown contribution from the mean winds, which may be expected to generate an effective advection term through the transport of SST anomaly signals downwind. For the purposes of SST modeling for climate studies, the net mixed layer SST advection fields derived in this manner from the SST anomaly data themselves may nevertheless provide a more appropriate effective advection field than near surface current measurements.

The diffusion coefficient represents the least well determined parameter of the model. It was at least found to be positive everywhere (not automatically guaranteed by the model) and its order of magnitude (10^3–10^4 m^2 s^{-1}) is consistent with previous estimates from diffusion experiments (Okubo, 1971).

The model fitting approach presented in this paper cannot, of course, substitute for more detailed numerical models of the coupled ocean atmosphere system. This follows already from the fact that some of the

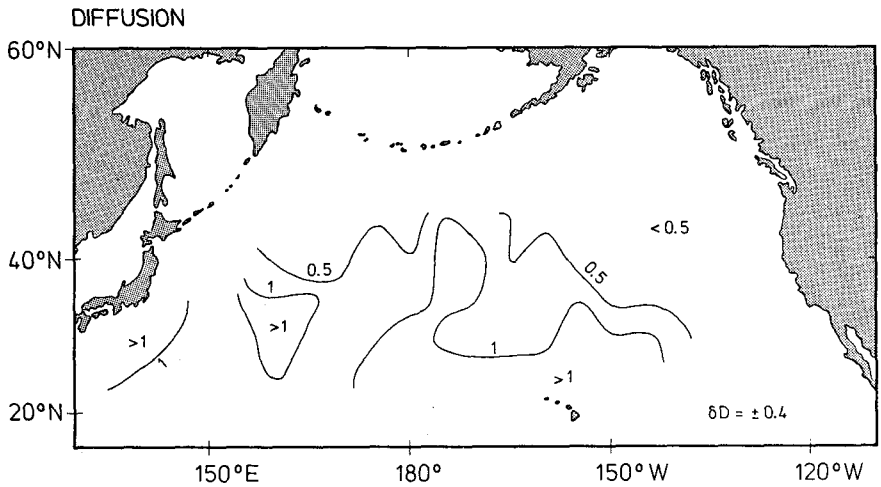

FIG. 14. Distribution of the model diffusion in units 10^4 m^2 s^{-1}.

residual errors, for example in the quadrature spectra, still remained above the critical significance level. However, the method does provide some insight into the orders of magnitude and general interrelationship between some of the stochastic and deterministic processes controlling SST variability in the mixed layer in midlatitudes, which can provide some guidance in the construction of more sophisticated models.

APPENDIX A

Covariance Matrix of the Data

The covariance matrix $\mathbf{V}$ of the data vector $\mathbf{f}$ (the set of auto- and cross-spectra defined in section 5) has the form:

$$\mathbf{V} = \langle (\mathbf{f} - \langle \mathbf{f} \rangle)(\mathbf{f} - \langle \mathbf{f} \rangle)^T \rangle = \begin{pmatrix} \mathbf{V}(\omega_1) & & \\ & & \\ & & \mathbf{V}(\omega_n) \end{pmatrix},$$

where the $\mathbf{V}(\omega_i)$ represent (10×10) submatrices

$$\mathbf{V}(\omega) = \begin{pmatrix} \mathbf{V}_{aa} & \mathbf{V}_{ab} \\ \mathbf{V}_{ab}{}^{\mathrm{T}} & \mathbf{V}_{bb} \end{pmatrix},$$

whose elements consist of the covariance matrices $\mathbf{V}_{aa}$, $\mathbf{V}_{bb}$ of the real and imaginary part of $\mathbf{f}(\omega)$ separately and the mixed products $\mathbf{V}_{ab}$.

We consider first the covariance matrix of the estimated complex cross-spectra:

$$\bar{F}_{ik} = \frac{1}{N} \sum_{j=1}^{N} F^j_{ik} \tag{A1}$$

for the grid points (i, k), where $F^j_{ik} = \langle (z_i{}^j)^* z_k{}^j \rangle$ is the cross spectral estimate for the "chunk" j and $\bar{F}_{ik}$ is the mean over all chunks.

The components of the complex covariance matrix for the estimates $\bar{F}_{ik}$ are defined by

$$\begin{aligned} \langle \delta\bar{F}_{ik} \delta\bar{F}_{jl} \rangle &= \langle (\bar{F}_{ik} - \langle \bar{F}_{ik} \rangle)(\bar{F}_{jl} - \langle \bar{F}_{jl} \rangle) \rangle \\ &= \langle \bar{F}_{ik} \bar{F}_{jl} \rangle - \langle \bar{F}_{ik} \rangle \langle \bar{F}_{jl} \rangle. \end{aligned} \tag{A2}$$

From Eq. (A1) we have

$$\langle \bar{F}_{ik} \bar{F}_{jl} \rangle = \frac{1}{N^2} \sum_{n,m=1}^{N} \langle (z_i{}^n)^* z_k{}^n (z_j{}^m)^* z_l{}^m \rangle.$$

Assuming that the $z_i{}^n$ are Gaussian random variables, the fourth moment can be decomposed into a sum of products of second moments:

$$\begin{aligned} \langle \bar{F}_{ik} \bar{F}_{jl} \rangle = \frac{1}{N^2} \sum_{n,m=1}^{N} & [\langle (z_i{}^n)^* z_k{}^n \rangle \langle (z_j{}^m)^* z_l{}^m \rangle \\ & + \langle (z_i{}^n)^* (z_j{}^m)^* \rangle \langle z_k{}^n z_l{}^m \rangle + \langle (z_i{}^n)^* z_l{}^m \rangle \langle z_k{}^n (z_j{}^m)^* \rangle]. \end{aligned} \tag{A3}$$

The first term in (A3) is equal to $\langle \bar{F}_{ik} \rangle \langle \bar{F}_{jl} \rangle$, cancelling the last term in (A2), and the second term is identically zero, leaving only the last term as a contribution to the complex covariance matrix:

$$\langle \delta\bar{F}_{ik} \delta\bar{F}_{jl} \rangle = \frac{1}{N^2} \sum_{n,m=1}^{N} \langle (z_i{}^n)^* z_l{}^m \rangle \langle z_k{}^n (z_j{}^m)^* \rangle.$$

If the chunks (n, m) are statistically independent, the right-hand side reduces further to

$$\begin{aligned} \langle \delta\bar{F}_{ik} \delta\bar{F}_{jl} \rangle &= \frac{1}{N} \langle (z_i{}^n)^* z_l{}^n \rangle \langle z_k{}^n (z_j{}^n)^* \rangle \\ &= \frac{1}{N} \langle F^n_{il} \rangle \langle F^n_{jk} \rangle = \frac{1}{N} \hat{F}_{il} \hat{F}_{jk}, \end{aligned}$$

for any chunk n and where we have assumed that the expectation values of the data can be approximated by the model.

Decomposing the spectra into its real and imaginary part:

$$\bar{F}_{ik} = a_{ik} + i b_{ik}, \quad \hat{F}_{ik} = \hat{a}_{ik} + i \hat{b}_{ik},$$

we have

$$a_{ik} = \frac{1}{2} [\bar{F}_{ik} + (\bar{F}_{ik})^*]$$

$$b_{ik} = \frac{1}{2i} [\bar{F}_{ik} - (\bar{F}_{ik})^*],$$

and obtain then by straightforward algebra:

$$\langle \delta a_{ik} \delta a_{jl} \rangle = \frac{1}{2N} (\hat{a}_{il} \hat{a}_{jk} - \hat{b}_{il} \hat{b}_{jk} + \hat{a}_{ij} \hat{a}_{lk} - \hat{b}_{ij} \hat{b}_{lk})$$

$$\langle \delta b_{ik} \delta b_{jl} \rangle = \frac{1}{2N} (-\hat{a}_{il} \hat{a}_{jk} + \hat{b}_{il} \hat{b}_{jk} + \hat{a}_{ij} \hat{a}_{lk} - \hat{b}_{ij} \hat{b}_{lk})$$

$$\langle \delta a_{ik} \delta b_{jl} \rangle = \frac{1}{2N} (\hat{a}_{il} \hat{b}_{jk} + \hat{a}_{jk} \hat{b}_{il} + \hat{a}_{kl} \hat{b}_{ji} + \hat{a}_{ji} \hat{b}_{kl}).$$

APPENDIX B

Covariance Matrix of the Parameters

The optimal model is defined as the minimum of the quadratic form

$$\epsilon = (\mathbf{f} - \hat{\mathbf{f}})^{\mathrm{T}} \mathbf{M} (\mathbf{f} - \hat{\mathbf{f}}),$$

and therefore satisfies the system of equations:

$$\frac{\partial \epsilon}{\partial \mathbf{p}} = -2 \mathbf{D}^{\mathrm{T}} \mathbf{M} (\mathbf{f} - \hat{\mathbf{f}}) = 0, \tag{B1}$$

where $\mathbf{p}$ denotes the set of model parameters and $\mathbf{D}^{\mathrm{T}}$ is the transpose of the rectangular matrix:

$$\mathbf{D} = \frac{\partial \hat{\mathbf{f}}}{\partial \mathbf{p}}. \tag{B2}$$

We split the model and the data vector into a sum of the expectation value and a small deviation:

$$\mathbf{f} = \langle \mathbf{f} \rangle + \delta\mathbf{f}, \quad \hat{\mathbf{f}} = \langle \hat{\mathbf{f}} \rangle + \delta\hat{\mathbf{f}}, \tag{B3}$$

with $\delta\hat{\mathbf{f}} = \mathbf{D}\delta\mathbf{p}$ from (B2).

Inserting (B3) into (B1) and assuming that $\langle \mathbf{f} \rangle = \langle \hat{\mathbf{f}} \rangle$, we obtain:

$$\mathbf{D}^{\mathrm{T}}\mathbf{M}(\delta\mathbf{f} - \mathbf{D}\delta\mathbf{p}) = 0.$$

Solving for $\delta\mathbf{p}$ we have

$$\delta\mathbf{p} = (\mathbf{D}^{\mathrm{T}}\mathbf{M}\mathbf{D})^{-1}\mathbf{D}^{\mathrm{T}}\mathbf{M}\delta\mathbf{f},$$

and the covariance matrix of the parameters is then given by

$$\langle \delta\mathbf{p}\delta\mathbf{p}^{\mathrm{T}} \rangle = (\mathbf{D}^{\mathrm{T}}\mathbf{M}\mathbf{D})^{-1}.$$

REFERENCES

Anderson, D. L. T., 1983: The oceanic general circulation and its interaction with the atmosphere. *Large Scale Dynamical Processes in the Atmosphere.* B. Hoskins and R. Pearce, Eds., Academic Press, London.

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

Davis, R. E., 1976: Predictability of sea surface temperature and sea level pressure anomalies over the North Pacific Ocean. *J. Phys. Oceanogr.*, **6,** 249–266.

Frankignoul, C., 1979: Large scale air–sea interactions and climate predictability. *Marine Forecasting,* J. C. J. Nihoul, Ed., Elsevier, 35–55.

——, and K. Hasselmann, 1977: Stochastic climate models, Part 2, Application to sea surface temperature anomalies and thermocline variability. *Tellus,* **29,** 284–305.

——, and R. W. Reynolds, 1982: Testing a dynamical model for mid-latitude sea surface temperature anomalies. *J. Phys. Oceanogr.*, **13,** 1131–1145.

Gorschkov, C. G., 1976: *World Ocean Atlas,* Vol. 1, *Pacific Ocean* (in Russian with English introduction). Pergamon.

Haney, R. L., 1980: A numerical case study of the development of large scale thermal anomalies in the Central North Pacific Ocean. *J. Phys. Oceanogr.*, **10,** 541–556.

——, 1985: Mid-latitude sea surface temperature anomalies: A numerical hindcast. *J. Phys. Oceanogr.*, **15,** 787–799.

Jenkins, G. M., and D. G. Watts, 1968: *Spectral Analysis and its Application.* Holden-Day, San Francisco, 525 pp.

Lemke, P., E. W. Trinkl and K. Hasselmann, 1980: Stochastic dynamic analysis of polar sea ice variability. *J. Phys. Oceanogr.*, **10,** 2100–2120.

Martin, B. R., 1971: *Statistics for Physicists.* Academic Press, 209 pp.

Okubo, A., 1971: Oceanic diffusion diagrams. *Deep-Sea Res.*, **18,** 789–802.

Reynolds, R. W., 1978: Sea surface temperature in the North Pacific Ocean. *Tellus,* **30,** 97–103.

——, 1979: A stochastic forcing model of sea surface temperature anomalies in the North Pacific and North Atlantic. Climatic Research Institute, Rep. No. 8, Oregon State University.