Tellus (1995), 47A, 118–131
Printed in Belgium – all rights reserved

TELLUS
ISSN 0280–6495

An analysis of North Pacific SST anomalies by means of a linear thermodynamic stochastic two-dimensional model

By J. M. GARCÍAORTIZ and A. RUIZDEELVIRA*, *Departamento de Física, Universidad de Alcalá de Henares, Apdo. 20, 28880 Alcalá de Henares, Madrid, Spain*

(Manuscript received 1 February 1993; in final form 3 February 1994)

ABSTRACT

The evolution of sea surface temperature anomalies (SSTA) in the North Pacific is studied with a 2-D linear heat transport equation with seasonally varying feedback and constant diffusion. The extension of the model to include an average cyclo-stationary SSTA advection field, estimated from seasonal ship-drift data of surface currents, does not improve the fit to the data. The model parameters are determined by an optimal fit of the observed SSTA lag-time auto-correlation function and the optimal model turns out to be statistically satisfactory (at the 95% confidence level) in most (83%) of the analyzed region. The model is also able to explain the lag-time SSTA auto-covariances. Additionally, by applying the model operator to the observed monthly SSTA time series we obtain the actual stochastic forcing. This forcing is consistent with the usual concept of midlatitude atmospheric behaviour. In particular, the basic hypothesis of the stochastic approach, that the forcing is uncorrelated in the SSTA time scale, is verified. Both the success of the model hierarchy fitting procedure and the comparison of the best-fit value of the feedback parameter with mixed-layer depth data suggest that the feedback can be seen as an essentially local process closely related to the thermal inertia of the ocean mixed-layer.

1. Introduction

As the number of experiments performed with general circulation models (GCM) increases, the importance of simple models as a complementary tool becomes more evident. In simple models the relation between the physics and the parameters of the model is easy to understand, and the relevance of the few individual processes included is easy to evaluate (by establishing a hierarchy that successively includes more processes); therefore the results of simple model simulations can help to analyze the complex output of a GCM.

In analyzing complex non linear systems that follow known equations a possible way is to try integrating these equations, using numerical and analytical techniques. But the results are not guaranteed to be better than those of other simpler approaches. The best known example can be found in the theory of gases. The particles that constitute an ideal gas, like Helium, for instance, follow perfectly known non linear laws of motion. These equations can be integrated, at least in principle. But after a lenghty and costly integration, the results are rather similar, if not equal, to those proposed by thermodynamics. The laws of thermodynamics do not include the essential physics, but they predict the evolution of systems as well as, and for a fraction of the cost of, molecular dynamics. Therefore, it seems worthwhile to use simple, statistical models, for trying to understand complex systems, in parallel with other more sophisticate dynamical models.

Simple models of parts of the climate system are based on the assumption that the number of degrees of freedom required is lower than that used in GCMs. In some simple models, the introduction of a series of plausible hypotheses about the physics of the system allows us to derive a system of equations simpler than the equations of a GCM.

In the stochastic climatic models proposed by

* Corresponding author.

Hasselmann (1976) the complexity is reduced on the basis of the different time scales of the subsystems or processes included. For these simple models the slow variables representing the "climate" (long term evolution of the ocean or the ice cover) are viewed as the response of the climate system to the stochastic forcing provided by the rapidly varying "weather" variables (usually identified with the atmosphere).

Gill and Niiler (1973) indicate the importance of the feedback between ocean and atmosphere for the evolution of SSTA. They indicate the possibility that a variation in the seasonal amplitude of the forcing factors produce seasonal changes in the SST anomalies.

Frankignoul and Hasselmann (1977), proposed a stochastic model of SSTA evolution. Their analysis, based on single station and spatially averaged data, shows that the spectra and the structure function of SSTA as well as the cross-correlation function of the first EOF of SSTA and SLPA could be explained by a local Langevin equation, that is, a linear first order differential equation forced by noise, in which the damping coefficient represents a negative feedback. Reynolds (1978) tested the same model on a $5° \times 5°$ grid in the midlatitude North Pacific, finding that the auto-spectra of observed SSTA were consistent with the model in about half the region. By including seasonality in the feedback parameter, OrtizBeviá and RuizdeElvira (1985) were able to explain the observed seasonally-dependent time-lag auto-covariances of SSTA for almost all (98%) of the region. Frankignoul and Reynolds (1983) introduced the first step to non-locality by including an advection term estimated from observed mean currents, but this did not improve the fit of the model SSTA auto-spectra to that of observations. Finally, Herterich and Hasselmann (1987) used a multivariate inverse modeling procedure for the SSTA auto- and cross-spectra based upon a stochastic equation that took into account advection and diffusion in addition to the usual feedback mechanism, thereby estimating the annual mean advection from the SSTA data.

Work performed with dynamic models of the atmosphere or the ocean has supported some of the hypotheses and/or conclusions of these stochastic models. For instance, Haney (1985) forced a multilevel primitive equation ocean with observed winds and climatological values of all the other atmospheric parameters, over several years. His hindcast SSTA were significantly correlated with the observed SSTA in a broad area of the North Pacific. He found that in the region where hindcast and observed SSTA were well correlated, the hindcast anomalies were clearly associated with thermal structures mainly confined to the mixed layer, supporting the use of a simple mixed layer equation to predict the SSTA evolution.

But the Haney (1985) hindcast SSTA had amplitudes too small (about one third in terms of standard deviation). Luksch et al. (1990) and Luksch and Storch (1991) used an OGCM forced with observed winds and coupled to a simple advective model for the air temperature to parameterize anomalous heat fluxes. The exclusion of the anomalous heat flux produced a similar underestimate of the SSTA variability. Additionally, their results agreed with those of Frankignoul and Reynolds (1983) in suggesting that the forcing due to anomalous heat transport is more important for the evolution of the SSTA than the contribution by the wind-induced Ekman transport.

Basic to the stochastic models is the assumption of the spectral characteristics (white noise) of the forcing. The analysis of the auto-correlation of the time series of the first EOF of the observed SLPA field (Davis, 1976) and of the air temperature and wind anomalies (Luksch and Storch, 1991), as well as a recent experiment by Large et al. (1991) using wind stresses derived from twice daily analysis from the ECMWF, support the assumption that, if the forcing is provided by the high frequency motions of the atmosphere, this forcing can be considered white noise.

In the present work we continue the analysis of stochastic models of SSTA in the Pacific region, by including into a linear, non local thermodynamic model the possibility of seasonal varying feedback and constant diffusion.

We include also an analysis of the forcing driving the data, estimated not as the residual of the model fit, but as that part of the data variability not explained by the linear model. This estimated forcing is uncorrelated for lag times of $\mathcal{O}$(month), and the length scale of its spatial correlation at zero time lags is reasonable for midlatitude atmospheric phenomena. Additionally, the structure of the cross-correlations between forcing and SSTA is similar to the one found by Davis (1976) between SSTA and observed SLPA, in that high correla-

tion exists between SSTA and forcing when SSTA lags SLPA by 1 month, and low correlations occur when the SSTA leads the forcing also by one month. Finally, the fitted feedback parameter values compare well with observations of the ocean surface mixed-layer depth, and the diffusion coefficient is comparable to previous estimates.

The plan of the paper is as follows. The model hierarchy is considered in Section 2, while the data and the model fitting procedure based on the observed SSTA correlations are described in Section 3. The results are presented and discussed in Section 4 and the conclusions are summarized in Section 5.

2. The model

Because of the mixing caused by wind and turbulent convection, the surface layer of the ocean is nearly homogeneous. This layer undergoes big seasonal changes in temperature and depth, but below its maximum depth, the temperature of the ocean is practically constant. Therefore, to understand the SSTA variability on the time scale of months to years, it is sufficient to study just the surface layer.

Following the arguments of GarcíaOrtiz (1989), we assume that the SSTA are governed by a mixed layer bidimensional heat transfer equation

$$\partial_t T + \boldsymbol{V}\cdot\nabla T - D\nabla^2 T = -\lambda T + w, \tag{2.1}$$

where $\boldsymbol{V}$ is the (verticallly averaged) horizontal advection velocity, D is the diffusion coefficient, and λ is a linear feedback parameter.

w represents the forcing term that includes the fluctuations in the heat exchange through the air-sea interface and the turbulent variations of the surface layer depth and the drag caused by the wind stress.

We are interested in a broad region, 10°–45° N, of the North Pacific, and therefore it is important to take into account the sphericity of the Earth. We will write eq. (2.1) in spherical coordinates:

$$(\partial_t + u(x, y, t)\,\partial_x + v(x, y, t)\,\partial_y - \alpha(y)\,\partial^2_{xx} - \beta(y)\,\partial^2_{yy})\,T = -\lambda T + w, \tag{2.2}$$

with

$$u(x, y, t) = \frac{U(x, y, t)}{R\cos y},$$

$$v(x, y, t) = \frac{V(x, y, t)}{R} + \frac{D\sin y}{R^2\cos^2 y},$$

$$\alpha(y) = \frac{D}{R^2\cos^2 y}, \qquad \beta(y) = \frac{D}{R^2\cos y},$$

where U and V are the components of $\boldsymbol{V}$, x and y are the geographic longitude (positive towards the east) and latitude (positive towards the north), and R is the Earth's mean radius.

We have used an advection field estimated from ship drift data of seasonal mean currents (Stidd, 1974). At each point, the zonal and meridional components of the velocity have been expanded into Fourier series up to the seasonal component.

In each experiment we used a space-time constant D, but we changed this value from experiment to experiment.

The depth (and therefore the thermal capacity) of the mixed layer and the strength of the atmospheric forcing, change with point and season. OrtizBeviá and RuizdeElvira (1985), have shown that the seasonal change in the forcing can be well-modelled by choosing the correct seasonal dependence for the feedback parameter λ. We regard the spatially dependent feedback parameter λ as a cyclo-stationary function of time, and we will use the two first terms of its Fourier expansion

$$\lambda(x, y; t) = \lambda_0(x, y) + \lambda_1(x, y)\cos(\omega t + \phi(x, y)),$$
$$\omega = 2\pi/T, \qquad T = 1 \text{ year}, \tag{2.3}$$

where the amplitudes λ_0 and λ_1 and the phase ϕ are spatial functions of the point (x, y).

Model (2.2) is bidimensional. To integrate it we need boundary conditions along meridians and parallels. These conditions are specified as the solutions of a local version of (2.2) with the feedback given by (2.3), that is, we assume that the boundary values T_B obey a cyclo-stationary Langevin equation:

$$\frac{dT_B}{dt} = -\lambda T_B + w, \tag{2.4}$$

where here the derivative is total, as we assume

that the boundary condition T_B does not have a spatial dependence. To solve (2.2) we have used an alternating direction implicit (ADI) method, equivalent to the Crank-Nicholson scheme. The procedure is stable and its accuracy is $\mathcal{O}(\Delta x^2, \Delta t^2)$, where Δx is the spatial step in the numerical model and Δt the corresponding temporal one.

With (2.2) and (2.3) we establish a model that takes into account various phenomena. This allows us to construct a hierarchy of models, as recommended by Herterich and Hasselmann (1987), by using several of the terms of the model equation and looking at the different results associated with these specifications. Except for the inclusion of advection, which is estimated from data, the inclusion of more terms into the model means that we have more parameters free to be fitted.

The first two models considered are local ones without any spatial coupling. They are given by an ordinary differential equation of the type (2.4) and their parameters are fitted to the data at each grid point.

Model I is a simple local model with constant feedback, i.e., a Markov process of first order for which $\lambda = \lambda_0(\boldsymbol{r})$. Model II takes into account the seasonality of the relevant processes mentioned above, leading to a cyclo-stationary Langevin equation with the feedback given by (2.3). It has three free parameters $\lambda_0(\boldsymbol{r})$, $\lambda_1(\boldsymbol{r})$ and $\phi(\boldsymbol{r})$ to be determined.

Once the parameters of these local models have been determined, we make two model extensions that introduce spatial coupling by transport terms. It seems reasonable to assume that the spatial structure of the most relevant local property that contributes to the feedback, the thermal capacity of the ocean mixed layer (closely connected to its depth), is well captured by the local models. This can also be true with respect to the local aspects of the atmosphere-ocean coupling. Therefore, we maintain the spatial structure of the feedback parameters λ_0 and λ_1.

Model III includes diffusion and is introduced in order to account for the turbulent horizontal heat transport in the mixed layer, as well as for a possible contribution of the spatial structure of the atmospheric forcing not captured by the locally fitted feedback term. A single constant isotropic diffusion coefficient D is added in this model to the previously fitted set of parameters.

Inclusion of the diffusion produces an increase of the thermal relaxation time. If we maintain the spatial structure of λ_0 and λ_1, we can rescale their numerical values by introducing a new constant parameter c:

$$\lambda^c(\boldsymbol{r}, t) = c\lambda(\boldsymbol{r}, t), \tag{2.5}$$

where λ is the feedback found by fitting model II.

Finally, model IV adds the possible contribution of the estimated advection field yielding the complete model eq. (2.2).

3. Data and model fitting

The first studies in this series (Frankignoul and Hasselmann (1977), Reynolds (1978), Frankignoul and Reynolds (1983), OrtizBeviá and RuizdeElvira (1985), and Herterich and Hasselmann (1987)), used the data set available at that time, the NORPAX data set (SST in a grid of $5° \times 5°$ in the North Pacific for the interval 1947–1974). This set has a considerable quantity of missing data and extends for 28 years.

We decided to use COADS (Comprehensive Ocean-Atmosphere Data Set) (Fletcher et al. (1983)) for roughly the same region as the NORPAX data but for an interval of 37 years (1950–1987), which should improve the quality of the statistical fits.

We analyzed the low-to-mid latitude region (15°N–40°N, 135°E–130°W), comprising 120 grid points. The mean annual cycle was subtracted from each series to produce SST monthly anomalies (SSTA).

Since our model is linear and stochastic we cannot aim to characterize more than the possible statistical linear relationships present in the data and these are represented by the lag time correlations. Using correlations is equivalent to first normalizing the time series by its variance (which depends on the month of the year for periodic data) and then calculating the expected linear relationship.

To the extent that SSTA vary seasonally, their correlations $r(t, \tau)$ are functions of two times, a seasonal time t and a lag time τ. We have chosen four seasonal times t_i corresponding to the months of February, May, August and November as representatives of the annual cycle, and we have

used lag times up to 15 months forwards and backwards, $\tau = -15, ..., 0, ..., 15$ months.

From the 124 correlations (31 correlations per month $\times$ 4 months) obtained, we have disregarded the four zero-lag correlations $r(t_i, 0)$, that are unity by definition, and those corresponding to positive lags that are multiple of three because these are redundant, i.e., $r(t_1, 3j) = r(t_{j+1}, -3j)$. Then, for each grid point, a set of 100 data correlations is used.

The forcing term w in (2.2) is modeled as a white noise process. Its variance should present spatial and seasonal dependence, but this dependence is unknown a priori. We could guess, but this would introduce some severe bias on the results. We decided to use the less stringent assumption on this variance and take it as a constant in both space and time. An analysis of the actual data forcing, as opposed to the model forcing term w, performed after the model was integrated (see below), allowed us to determine the spatial structure of this estimated data forcing.

The numerical solution has been obtained using the same grid spacing as that of the data, $\Delta x = 5^\circ$, and a time step $\Delta t = 2$ days. At each point and each time step a sample of the model forcing w is used. This means that it is uncorrelated with the model forcing at other steps.

The model was integrated for the same time interval as that of the data (37 years). We built a statistical ensemble by repeating this integration 80 times. For each grid point we thus had a set of 80 statistically independent time series or realizations of the model. Then the same steps described above for the data were used to evaluate the correlations of any realization of the model, and we obtained 80×4 sample correlation functions per grid point, each one involving 30 time-lags for the 4 months of February, May, August, and November

$$\begin{aligned} &r_{kl}(t, \tau), \quad k = 1, 120 \text{ (grid points)}, \\ &\quad l = 1, 80 \text{ (realizations)}, \\ &\quad t = 1, 4 \text{ (mid season months)}, \\ &\quad \tau = -15, 15 \text{ (lags)}. \end{aligned} \tag{3.1}$$

Once the model and data correlation sets have been built, we use the Fisher z-transformation to transform correlations $r(t, \tau)$ into new functions $z(t, \tau)$

$$z(t, \tau) = \frac{1}{2} \ln \left(\frac{1 + r(t, \tau)}{1 - r(t, \tau)} \right). \tag{3.2}$$

For each set of values of λ_0, λ_1, D and c, we determine the average, across the 80 model realizations, of the 4×25 Fisher transforms of the correlations for each grid point k:

$$m_k(t_j, \tau_i) = \sum_{l=1}^{80} r_{kl}(t_j, \tau_i).$$

If we assume that our model is correct, then the data should be a realization of the model, and each data correlation Fisher transform $d_i(t, \tau)$ should be a sample of an approximately Gaussian-distributed variable whose expected value is the Fisher transform of the corresponding model correlation, $m_i(t, \tau)$, with variance σ given by (3.3).

Then, for each grid point i, we define an error measure ε_i^2 as

$$\varepsilon_k^2 = \sum_{j=1}^{4} \sum_{i=1}^{25} \left(\frac{d_k(t_j, \tau_i) - m_k(t_j, \tau_i)}{\sigma} \right)^2 \tag{3.4}$$

(note that in the sum to the time lags the values $\tau = 3n$, n being an integer, are missing).

The model is fitted to the data at each grid point by minimizing the error measure ε_k^2 with respect to the parameters λ_0, λ_1 and D, to obtain their best-fit values, which then define the optimal model, and produce an optimun error $\tilde{\varepsilon}^2$.

As a reference, we define a trivial model that predicts a white noise response and thus null correlations, $m_i = 0$. If we substitute in the sum (3.4) the results of any sample of this trivial model for the data correlations d, then the terms in the sum of (3.4) will be squares of independent gaussian variables with unit variance, and the error ε^2 for any realization of the model would be a random variable distributed as a χ^2 variable with 100 degrees of freedom. For 100 degrees-of-freedom, the 95% χ^2 critical value is $\chi_c^2 = 124$.

To test if any non trivial model (characterized by a set of values of the parameters λ_0, λ_1 and D) could be considered valid, given the data, we propose the null hypothesis H_0 that the data can be considered a realization of the model. An optimum value of $\tilde{\varepsilon}_k^2$ higher than 124 at any

grid point k would then indicate that the null hypothesis should be rejected, and that the model does not represents the data at this grid point.

This test is conservative, as the null hypothesis is not independent from the data used to perform the test. We carried a series of Montecarlo experiments, by generating fake data time series and fitting model (2.1) to them. The results of using the above test were not much different of the Monte Carlo results, and as this test was used in all the other studies in the series, we decided to employ it.

With the data correlations d and the trivial model $m = 0$ we determined that the trivial model can be rejected everywhere with a very high statistical significance level.

4. Results

We present the results according to the following plan. In Subsection (4.1) we are concerned with the lag-time auto-correlations of model and data. We show in Subsection (4.2) that model III can be made to fit the data auto-covariances with the required statistical significance. Next (Subsection 4.3), we study the statistical properties of the forcing needed for the model to reproduce the observed SSTA time series, and finally, we discuss the optimal values of the model parameters and compare them with observations in Subsection (4.4).

4.1. Test on correlations

We have performed the test for models I to III. Model IV, including seasonal advection, is nearly undistinguishable from model III. The extension of model II to include only advection also failed to improve the model fitting. Frankignoul and Reynolds (1983) found that adding advection to a constant feedback model produces a slight modification of the response auto-spectra at low frequencies (say, less than two cycles per year). Thus the effect of cyclo-stationary advection seems to be captured by the seasonally varying feedback. Therefore we restrict ourselves to models I to III.

The simple local model with constant feedback, model I, can be rejected in 35% of the grid points with a statistical confidence of at least of 95%. The regions where the model fails are the northwest and southeast regions as well as around the Date Line north of the 23°N parallel. The values of the errors $\tilde{\varepsilon}^2$ (not shown) indicate that this very simple model I represents a significant improvement over the trivial model.

Model II can also be rejected in the northwest and southeast areas but it is accepted in the central area and the results at low latitudes represent an improvement over model I. The inclusion of seasonal variation in the feedback improves the model fit everywhere and makes it acceptable in points where model I fails.

The introduction of diffusive transport reduces the model error in the regions where model II still fails. Model III can be rejected in 20 grid points at the 95% confidence level.

The model extensions produce successive improvements in terms of data and model correlations, and model III fits the data reasonably well in a statistical sense and in 83% of the region in study.

But the data have sources of variability other than linear effects, such as measurement errors and the non-linear terms disregarded in the model equations. Note that the variance of the data correlations is estimated as that of the model realizations and that we do not take into account any additional data variance due to measurement errors.

Consider the lag-time correlations of the data, model and forcing (see Subsection (c) below) for model II at any arbitrary point. We present in Fig. 1 those at the grid point (175°W, 30°N), in one of the regions where the optimal error is big. We can observe that both model and data correlations depend on the seasonal time and that, with the exception of a few peaks present in the data for large lags, the decay in the data correlation is followed rather correctly by the model. The asymmetry, with respect to the sign of the lag, for the data in spring and fall is also present in the model. The symmetry in the two other seasons and the long-lasting character of winter correlations, which can be associated with the higher thermal inertia of the mixed-layer in the winter, are also reproduced by the model.

This is not the case for the significant correlations for negative lags in February and November. We are fitting a complete field of data, with a simple model. We should be satisfied if we get an acceptable fit for most of the points in the field. In any case it is interesting to note that also in this case when there are some important errors, there

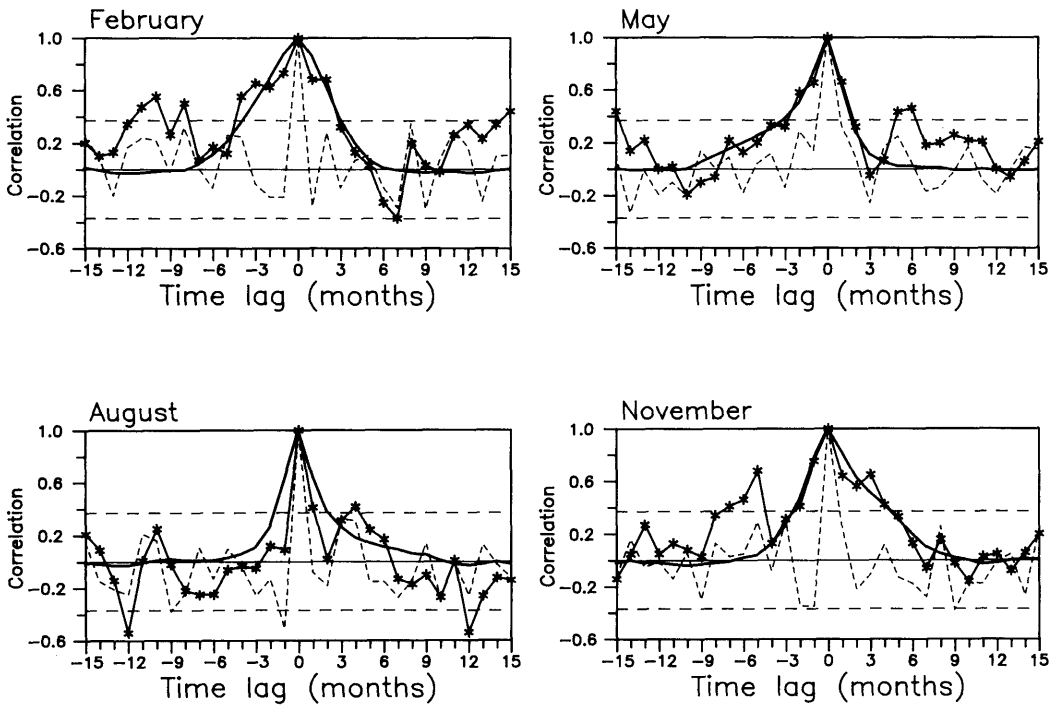

Fig. 1. Correlations of the data (full line with stars), Model II (full line) and forcing (dashed line), as functions of lag time at the grid point (175° W, 30°) for four seasons. Long dashed lines enclose the band where correlations are not significative at the 95% level.

are some features reasonably well captured by the model.

4.2. *Test on covariances*

A further test for the model is to consider the fitting of the covariances with the parameters determined by the best fit to the correlations. We would like to see if a model that fits lag-time data autocorrelations does simultaneously fit the lag-time data autocovariances, in particular, the significant secondary peaks in the covariances for lag times around 12 months that are not present, in general, in the lag-time autocorrelations.

We have built the temporal auto-covariances of model III results and of the data for each grid point and for time lags between 0 and 42 months. The model is always run with noise whose variance is arbitrary but constant. The lag-time autocovariances of the model results were scaled, at each grid point, with the adequate factor in order to make their zero-lag value equal to the data variance. Then, for each grid point k, we define an error measure based on the relative errors as:

$$\varepsilon'_k = \sum_{j=1}^{4} \sum_{i=1}^{30} \left(\frac{v_{ji} - M_{ji}}{v_{ji}} \right)^2 g_i, \tag{4.1}$$

where v_{ji} are the i-lag covariances of data, M_{ji} that of model and g_i is a weight function designed to avoid the possible singularities when $v_{ji} \simeq 0$:

$$g_i \equiv (1 - \exp(-0.5(v_i/v)^2))^2, \tag{4.2}$$

where v is a convenient empirical normalization factor whose effect is to make $g_i \simeq 0$ for non-significant covariance values and $g_i \to 1$ outside the error band around zero.

The variance of ε', $\sigma^2_{\varepsilon'}$, was obtained by applying (4.1) to 20 realizations of the model, v_i being in this case a realization and M_i the average across the 20 realizations. Note that estimating $\sigma_{\varepsilon'}$ from the model realizations with the weight function

defined in (4.2) avoids the effect of the inclusion in the test of many near-zero covariance data values that are well predicted by any model with big enough damping.

An example of model and data covariances is given in Fig. 2. A model sample is also included to show that not only are the data covariances coherent with the expected model ones, but also that the model realizations are similar to the data. Note that oscillations with large time-lags, present in the data covariances, are also present in the model realizations.

4.3. *Observed-forcing analysis*

A linear model cannot explain the non-linear sources of variability. Therefore we have assumed that non-linear processes, that cannot be modelled by the deterministic part of (2.1), act as an uncorrelated forcing in the time scale of the response.

To get some information about these non-linear processes we could take another point of view, and study the auto-correlations of the residuals. That is, if our model does contain the linear processes that govern the SSTA evolution, then the forcing needed to produce the observed time series of the SSTA should be linearly uncorrelated for time-lags of $\mathcal{O}$(month). The forcing will be consequence of short time scale, $\mathcal{O}$(day), non-linear processes that are out of the scope of linear analysis.

If we rewrite the equation (2.1) defining the model as

$$\mathscr{L}T = w' \tag{4.2}$$

with

$$\mathscr{L} \equiv (\partial_t + u(x, y, t)\,\partial_x + v(x, y, t)\,\partial_y - \alpha(y)\,\partial^2_{xx} - \beta(y)\,\partial^2_{yy} - \lambda)$$

and apply $\mathscr{L}$ to the *observed data values of* T at every grid point k, and each time i, we can obtain the actual forcing w'_{ik} that produces the observed data variability.

In Fig. 3 we show a time series of data and of one randomly chosen realization of model III, normalized by their standard deviations, together with the estimated forcing at the point chosen above. This estimated forcing is not the white noise forcing used to run the model.

Both data and model realization time series have been band pass filtered between 6 and 56 months in order to make the comparison easier. We would like to point out that the fact that both model and data correlations have statistically significant non-zero correlations can be related to

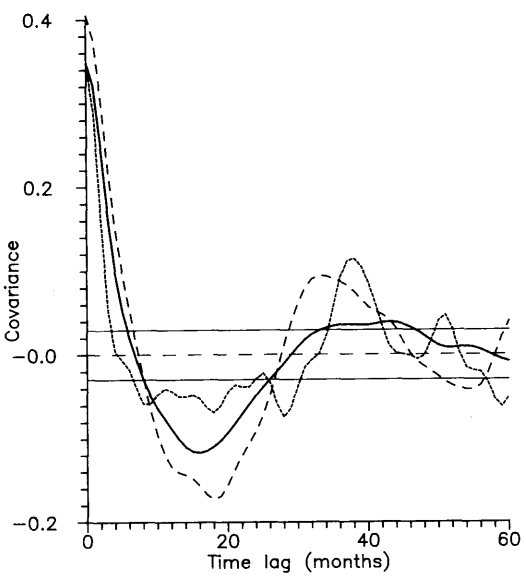

Fig. 2. Covariances of data (short dashed line), model III (full line) and an arbitrary realization of this model III as functions of lag time at the grid point at (175°W, 30°). Horizontal lines mark the region where the covariances cannot be considered significant.

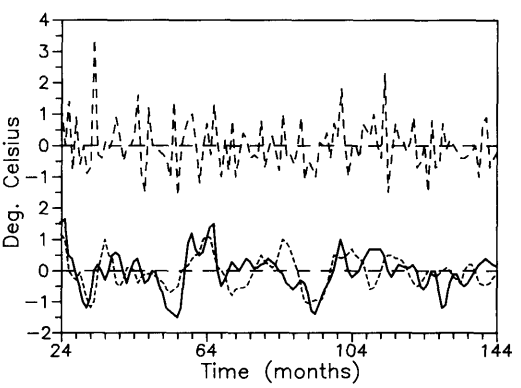

Fig. 3. An arbitrary portion of the time series (normalized by its standard deviation) of SSTA data (solid line) and of a realization of model III (dashed line), at the grid point at (175°W, 30°). The computed time series of the forcing at this point is shown vertically displaced (long-dashes).

the ocurrence of extended periods when data or model keep the same sign. On the other hand, the forcing w' estimated from the data shows much higher frequency oscillations producing near zero correlations for time lags equal or greater than one month.

For model III it is possible to determine the forcing only in the interior of the domain in study, that is for 70 grid points excluding the boundaries. For this interior region the forcing can be considered white noise at the 95% confidence level in 69 of the 70 grid points.

An analysis similar to the previous one has been carried ou for the cross-correlations between the forcing at each point and the forcings at the four nearest neighbour grid points. Excluding the zero-lag value, the reslts are nearly the same as those for the auto-correlations.

For zero time-lags the cross-correlations of the forcing are non-zero, as expected if part of the forcing is to be due to a real atmosphere, because atmospheric variables at a point are correlated with those at other somewhat distant places for lag times of $\mathcal{O}$(day). The e-folding correlation length of the forcing obtained with model III is about 1000 to 1500 km in the zonal direction and about 500 km in the meridional one. These distances are reasonable for the latitudes under study and the longer W–E correlation lengths are consistent with the presence of eastward travelling disturbances and trade winds.

We can confirm that, as expected, the model captures the linear signal present in the data and that the idea of an atmosphere acting like a white noise forcing for times of the order of the time scale characteristic of the SSTA evolution, is correct. From the fact that we find no correlations distinct from zero for time lags bigger that one month in the forcing, we can conclude that if non-linear processes in the ocean act as forcing for eq. (2.1), these processes should have time scales either shorter than the two days used as time step in the model or bigger than the 15 months used as time-lag limit for the correlations.

An EOF analysis of the estimated observed-forcing (Fig. 4) shows two markedly different characteristics with the first EOF indicating a set of zonal winds changing direction along parallel 25, and the second EOF indicates a nucleus of winds centered around the date line.

4.4. *Model parameters*

Assuming then that our model is valid, we show in Fig. 5 the spatial structure of the relaxation time λ_0^{-1} of model I. Except for the eastern region, there is a northward gradient that is coherent with the deepening of the mixed layer and the consequent increase of its heat capacity. There is a broad area, roughly the western half of the latitudinal belt 15–25°N, where the gradient is very small and north of it there is a zonal succession of troughs and ridges. These features are similar to those present in the structure of the mixed-layer depth based on a density criterion for March and September (Figs. 6b, 7b).

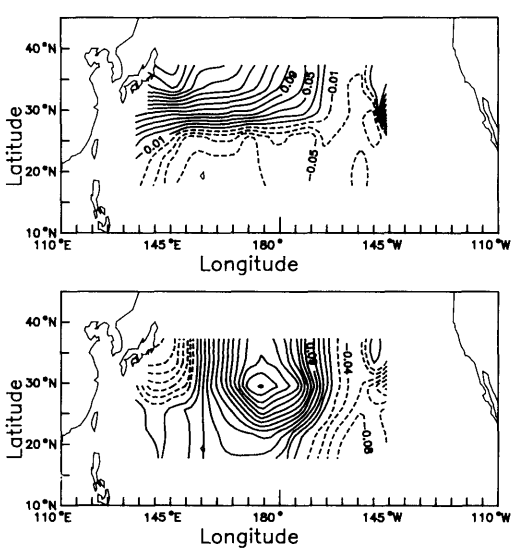

Fig. 4. EOFs of the forcing derived from the real data by using eq. (4.2). Upper panel: First EOF, lower: second EOF.

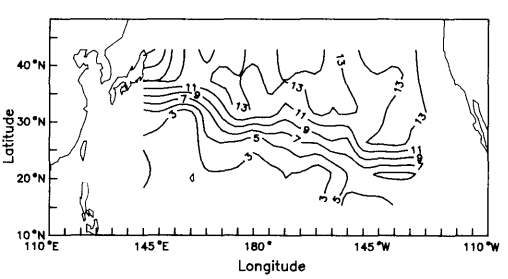

Fig. 5. Spatial distribution of the best fit values of the feedback parameter λ_0^{-1}, in months, corresponding to model III.

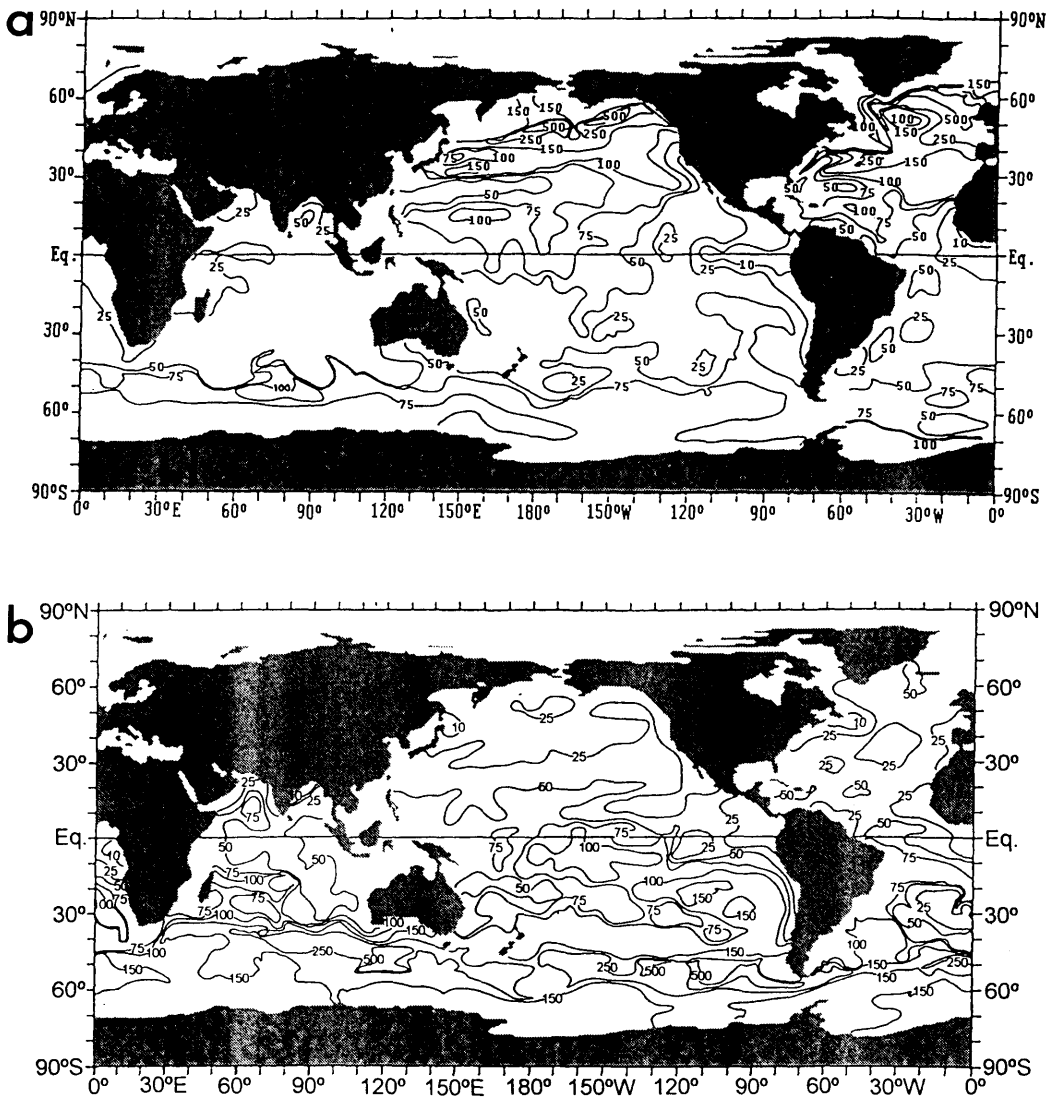

Fig. 6. Mixed layer depths from Levitus, 1982. Month: March. (a) Based on a temperature criterium of 0.5°C. (b) Based on a σ_t criterium of 0.125.

In particular, the maximum at about 160° E north of 30° N appears in both months when a temperature criterion for the estimation of the depth is used (Figs. 6a, 7a). The range of space variation of Levitus depths (about a factor of three) also coincides with the range of space variation of our relaxation times. Because in these areas the minimum relaxation time is generally reached in summer or fall (Fig. 8), those months approximately correspond to extreme depth values and features of λ_0 we have indicated above.

As could be expected, the inclusion of a cyclostationary modulation in the feedback λ does not appreciably change its average value λ_0, which is nearly the same for models I and II. While model I correlations are the same for every month and for positive and negative lags, in model II correlations the seasonal varying λ introduces a seasonal

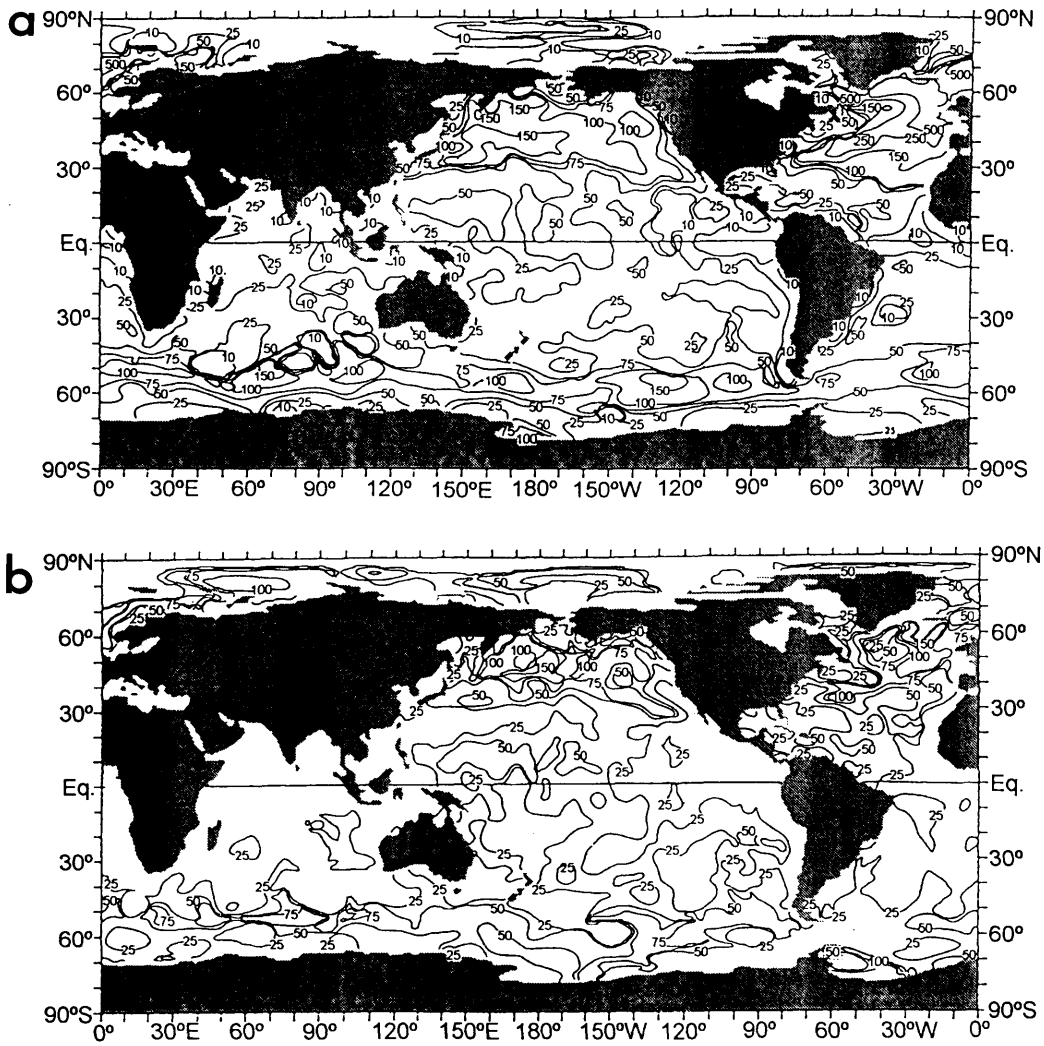

Fig. 7. Mixed layer depths from Levitus, 1982. Month: September. (a) Based on a temperature criterium of 0.5°C. (b) Based on a σ_t criterium of 0.125.

dependence of the model correlations and an asymmetry with respect to the time lag, but does not modify the appropriate mean decay of correlations that is a direct consequence of the feedback term (Fig. 1).

In Fig. 9 we show the relative importance of the seasonal modulation of the feedback λ_1/λ_0, that goes up to 80% at some points, and that can be considered to be around 50% in general. The region of bigger seasonal modulation roughly coincides with the region where model II shows the biggest improvement over the trivial one.

Our estimated months of minimul relaxation time for the points (144°E, 31°N), southeast of the Kuroshio, and (179°E, 25°N), located in a subtropical region, coincide with the months of minimum mixed-layer depth (minimum heat capacity) in Levitus data (Levitus, 1982, Fig. 97), and our relative amplitudes of seasonal modulation are similar to those of the mixed-layer depth.

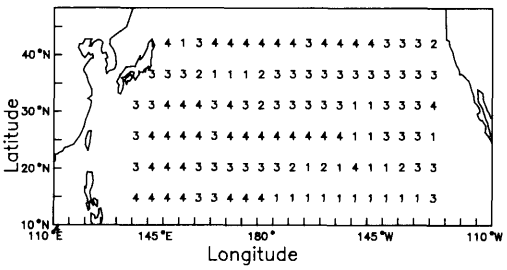

Fig. 8. Percent seasonal modulation of the feedback λ_1/λ_0. Contour interval 20%.

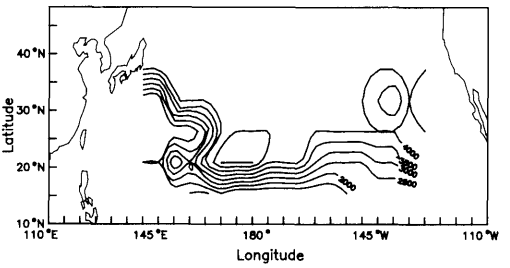

Fig. 10. Spatial distribution of diffusion coefficient D. Contour interval: 0.6×10^4 m^2 s^{-1}.

This discussion supports the idea that the feedback is dominated by the thermal inertia of the mixed-layer of the ocean and indicates that the space and time structures of our estimated parameters are realistic.

The best fit value overall of the diffusion coefficient was found to be $D = 2.3 \times 10^4$ m^2 s^{-1}. But, as can be expected, using the values of λ_0 and λ_1 found with this value of D, the best fit, not overall but at each point, was obtained with experiments differing on the value of D. The distribution in space of the value of D for each of these best fits is presented in Fig. 10. In view of these results, a model with spatially varying D should perhaps be considered necessary. However, a model including spatially varying diffusion would be considerably more complex than the present one. As our aim is to use simple models, we decided not to include varying diffusion. Nevertheless, the information given by these partial fits is of some interest, in that it is consistent with the idea that the diffusive term in the present model takes into account the turbulent transport which should be much more intense in the north part of the basin.

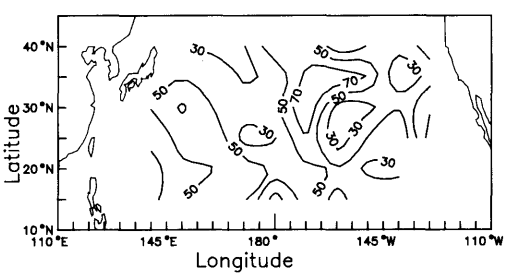

Fig. 9. Seasons of maximum feedback. 1: Winter, 2: Spring, 3: Summer, 4: Autumn.

The numerical values of the diffusion coefficient D should not be compared with these of the horizontal turbulent thermal diffusivity used in ocean models that treat explicitly the dynamics, which is about two orders of magnitude smaller (Luksch et al., 1990). Our model is merely a thermodynamic balance equation and so D must also account for the heat transport associated with the mass flux. Our space resolution is about 500 km whereas an important part of the oceanic transport (i.e., that related to quasi-geostrophic gyres) seems to have a much smaller spatial scale, say 50 km, whith a time scale of a few months (Frankignoul, 1979). Actually, experiments with a simplified version of our model (constant feedback and no space variation of the coefficients) show an order of magnitude decrease of the best fit value of D when the grid spacing is reduced by a factor of three (GarcíaOrtiz, 1989).

5. Conclusions

We have analyzed a 2-dimensional stochastic model of the SST anomalies of the low-to-mid latitudes of the North Pacific ocean. One of the basic assumptions of the model, that the driving force (the source term in the equation) is not linearly auto-correlated in time at each grid point, is verified. We have also found that there are no linear cross-correlations between the forcing at different grid points except for zero lag time (we should remember that our time step is one month). The e-folding (simultaneous) correlation length is around 1500 km for the W–E direction, and around 500 km for the N–S direction. This seems

Tellus A/47/1-9

reasonable if the forcing is related to the atmospheric motion in midlatitudes.

The model proposed here fits reasonably well the statistics (auto- correlations and covariances) of the observed SSTA. Introduction of cyclostationarity into the feedback improves the fitting, but a bigger improvement is obtained when diffusive transport is included. On the other hand, advection by mean currents does not bring any improvement.

The fitted feedback coefficients are coherent with the observed evolution of the depth of the mixing layer, thus supporting the idea that this feedback is a purely local effect representing the thermal inertia of the layer. However, no effort to identify the correct physical mechanisms for the feedback or interactions is made. The aim of these stochastic models is to explain the statistics of some climate variables in terms of simple stochastic equations, and not to investigate their detailed dynamics.

When the feedback λ_0, obtained by fitting the local models is used, we obtain values of the same order of magnitude than the previous estimates based on auto- and cross-spectra of SSTA (Frankignoul and Reynolds, 1983) and Herterich and Hasselmann (1987)). However, when diffusion is included, the best fit value of λ_0 is 3 times smaller and becomes comparable to the one found by OrtizBeviá and RuizdeElvira (1985) when they fitted the first three Fourier coefficients of the two-time SSTA auto-covariance function. The results for λ_0 and its cyclostationary corrections agree reasonably well with the mixed layer data of Levitus for this region.

We obtained the forcing by applying eq. (4.3) to the data at each grid point and for each time interval. This forcing shows significant non-zero positive cross-correlation with posterior SSTA, for some 7–9 months. For zero lag time there is a small but significant negative cross-correlation, that decays steeply for negative lag times τ (forcing lags), being undistinguishable from zero for $\tau = -2$ months. This is coherent with the results of Davis (1976).

6. Acknowledgments

This work was started with support from project 2428/83 of the Spanish CICYT, and has been finished in the frame of projects E4VC and EPOC003 of the European Community. We want to thank Dr. M. J. OrtizBeviá for very helpful discussions and comments on this paper.

REFERENCES

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.

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

Fletcher, J. O., Slutz, R. J. and Woodruff, S. D. 1983. Towards a comprehensive ocean-atmosphere data set. *Trop. Ocean-Atmos. Newslett.* **20**, 13–14.

Frankignoul, C. 1979. Large scale air-sea interactions and climate predictability. *Marine forecasting*, J. C. J. Nihoul (ed.), Elsevier, pp. 35–55.

Frankignoul, C. and Hasselmann, K. 1977. Stochastic climate models. Part II. Application to sea-surface temperature anomalies and thermocline variability. *Tellus* **29**, 289–305.

Frankignoul, C. and Reynolds, R. W. 1983. Testing a dynamical model for mid-latitude sea surface temperature anomalies. *J. Phys. Oceanogr.* **13**, 1131–1145.

GarcíaOrtiz, J. M. 1989. *Un modelo estocástico con advección y difusión de las anomalías de temperatura superficial del Océano Pacífico Norte.* PhD thesis. Universidad de Alcalá de Henares, 214 pp., Alcalá de Henares, Spain.

Gill, A. E. and Niller, P. P. 1973. The theory of seasonal variability in the ocean. *Deep-Sea Res.* **20**, 141–177.

Haney, R. L. 1985. Midlatitude sea surface temperature anomalies: a numerical hindcast. *J. Phys. Oceanogr.* **15**, 787–799.

Hasselmann, K. 1976. Stochastic climate models. Part I. Theory. *Tellus* **28**, 473–485.

Herterich, K. and Hasselmann, K. 1987. Extraction of mixed layer advection velocities, diffusion coefficients, feedback factors and atmospheric forcing parameters from the statistical analysis of the North Pacific SST anomaly fields. *J. Phys. Oceanogr.* **17**, 2146–2155.

Large, W. G., Holland, W. R. and Evans, J. C. 1991.

Quasi-geostrophic ocean response to real wind forcing. The effects of temporal smoothing. *J. Phys. Ocean.* **21**, 998–1017.

Levitus, S. 1982. *Climatological atlas of the world ocean.* NOAA Professional Paper, 3. US Government Printing Office, Washington, DC, 173 pp.

Luksch, U., Von Storch, H. and Maier-Raimer, E. 1990. Modeling North Pacific SST anomalies as a response to anomalous atmospheric forcing. *J. Mar. Syst.* **1**, 155–168.

Luksch, U. and Von Storch, H. 1991. Modelling the low-frequency sea surface temperature variability in the North Pacific. *J. Climate* **5**, 893–906.

OrtizBeviá, M. J. and RuizdeElvira, A. 1985. A cyclostationary model of sea surface temperatures in the Pacific Ocean. *Tellus* **37A**, 14–23.

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

Stidd, C. K. 1974. *Ship-drift components: means and standard deviations.* SIO Ref. Ser. 74. Scripps Institution of Oceanography, 57 pp., La Jolla, San Diego, USA.