

# Stochastic Analysis of Southern and Pacific Ocean Sea Surface Winds

Philip Sura <sup>1</sup>

*Physical Oceanography Research Division,  
Scripps Institution of Oceanography, La Jolla, California*

August 26, 2002

Manuscript submitted to *JAS*

---

<sup>1</sup>Current affiliation: NOAA-CIRES, Climate Diagnostics Center  
*Corresponding author address:* Philip Sura, NOAA-CIRES, Climate Diagnostics Center,  
R/CDC1, 325 Broadway, Boulder, CO 80305-3328, E-Mail: psura@cdc.noaa.gov

## Abstract

This paper shows that the synoptic variability of zonal and meridional mid-latitude Pacific and Southern Ocean sea surface winds can be well described by a univariate stochastic dynamical system directly derived from data. The method used to analyze blended QuikSCAT/NCEP winds is a general method to estimate drift and diffusion coefficients of a continuous stationary Markovian system. Almost trivially, the deterministic part consists of a simple, nearly linear damping term. More importantly, the stochastic part appears to be a state-dependent white noise term, that is, multiplicative noise. The need for a multiplicative noise term to describe the variability of midlatitude winds can be interpreted by the fact that the variability of midlatitude winds increases with increasing wind speed. The results indicate that a complete stochastic description of midlatitude winds requires a state-dependent white noise term, that is, multiplicative noise. A simple Ornstein-Uhlenbeck process is not sufficient to describe the wind data within a stochastic framework. The method used fails for tropical regions, suggesting that tropical variability might be non-Markovian.

# 1 Introduction

The description of atmospheric multiscale nonlinear dynamics by stochastic differential equations (SDEs) is becoming increasingly popular. The general idea of *stochastic climate models* was introduced by Hasselmann (1976) and is based on the Brownian motion analog: the observed red spectrum of oceanic fluctuations is a consequence of the amplification of low-frequency weather fluctuations. Stochastic climate models have been surprisingly successful in describing, for example, a broad frequency band of oceanic variability. The success of this concept has inspired researchers to consider stochastic atmospheric forcing as a possible source of more complex ocean dynamics, for instance for midlatitude climate variability (e.g., Barnett et al. 1999; Junge et al. 2000; Czaja and Frankignoul 1999; Weng and Neelin 1998; Sura et al. 2000, 2001; Sura and Penland 2002; Münnich et al. 1998; Saravanan and McWilliams 1997, 1998; Mikolajewicz and Maier-Reimer 1990, and others) and ENSO variability (e.g., Blanke et al. 1997; Kleeman and Moore 1997; Eckert and Latif 1997, and others).

*Empirical stochastic models* are derived by fitting stochastic models to time series of observed data. This is a widespread method in climate re-

search (see e.g., von Storch and Zwiers 1999). For example, using Principal Oscillation Patterns (POPs), Hasselmann (1988) fitted the time evolution of a system to a simple linear low-order SDE. Other well-known examples are the linear inverse models (e.g., Penland and Sardeshmukh 1995; DelSole and Hou 1999; Winkler et al. 2001). Empirical stochastic models are often used as forecast tools. Most empirical stochastic models introduce the stochastic perturbations as additive noise. That is, the strength of the noise is held constant and does not depend on the state of the system. However, the strength of the noise may also depend on the state of the system. In that case, the stochastic term appears as multiplicative noise. The use of multiplicative noise might be useful or even necessary to describe the turbulent atmospheric motion in a stochastic framework.

The importance of multiplicative noise to describe small-scale turbulence has been recently shown by Friedrich and Peinke (1997a,b) and Renner et al. (2001). By analyzing experimental data the authors showed that the statistics of velocity increments  $\Delta u_r = u(x+r) - u(x)$  ( $u$  is the velocity in the direction of  $x$ , and  $r$  is distance measured along  $x$ ;  $\Delta u_r$  is a measure of the turbulent cascade) in a turbulent jet can be described by a univariate Markov process determined by a drift and a diffusion coefficient. The deterministic

drift acts as a damping term. Yet, the most prominent feature is that the noise term is not a constant, but appears to be state-dependent: higher velocity increments correspond to higher variability, that is noise, and vice versa. Thus, the stochastic term appears as multiplicative noise. See Sreenivasan and Antonia (1997) for a recent review of small-scale turbulence.

These findings motivate the present paper. Using the same method, the deterministic drift and the noise term are determined for observed wind velocities, to explore the stochastic properties of the atmospheric motion. In particular, this paper shows that a proper stochastic description of univariate midlatitude sea surface winds requires a multiplicative noise component, a feature which has hardly been emphasized in related studies. In section 2 the method is introduced. The data are described in section 3, and results are presented in section 4. Finally, section 5 provides a summary and a discussion.

## **2 Method**

In this study a general method is used to estimate the drift and diffusion coefficients of the Fokker-Planck equation for a continuous stationary Markovian

stochastic process (Siegert et al. 1998; Friedrich et al. 2000b,a; Gradišek et al. 2000). A wide class of physical processes can be described by a Markovian system. Thus, we consider the dynamics of a  $n$ -dimensional system governed by the following Itô-SDE:

$$\frac{d\vec{x}}{dt} = \vec{A}(\vec{x}) + \tilde{B}(\vec{x})\vec{\eta} \quad (1)$$

with the  $n \times n$  matrix  $\tilde{B}$ . In the subsequent discussion all stochastic components  $\eta_i$  are assumed to be independent Gaussian white noise processes:

$$\langle \eta_i(t) \rangle = 0, \quad \langle \eta_i(t) \eta_j(t') \rangle = \delta_{ij} \delta(t - t'), \quad (2)$$

where  $\langle \dots \rangle$  denotes the averaging operator. For a detailed discussion of stochastic integration and the differences between Itô and Stratonovich SDEs see, for example, Horsthemke and Lefever (1984) or Gardiner (1985). To briefly summarize, the Stratonovich calculus best represents situations where rapidly fluctuating quantities with small but finite correlation times are parameterized as white noise. The Itô stochastic calculus is used when discrete uncorrelated fluctuations are approximated as continuous white noise. That means, that continuous physical systems, as the atmosphere, are normally described by the Stratonovich calculus. Nevertheless, in the Itô interpretation the deterministic term  $\vec{A}(x)$  can simply be interpreted as the so called

effective drift (the sum of the deterministic and the noise-induced drift in Stratonovich systems; see Horsthemke and Lefever (1984) or Gardiner (1985) for details).

The probability density function  $p(\vec{x}, t)$  (PDF) of the Itô-SDE (1) is governed by the corresponding Fokker-Planck equation (e.g., Gardiner 1985; Horsthemke and Lefever 1984; Paul and Baschnagel 1999), which reads

$$\frac{\partial p(\vec{x}, t)}{\partial t} = - \sum_i \frac{\partial}{\partial x_i} A_i p(\vec{x}, t) + \frac{1}{2} \sum_{i,j} \frac{\partial^2}{\partial x_i \partial x_j} (\tilde{B} \tilde{B}^T)_{ij} p(\vec{x}, t). \quad (3)$$

The Fokker-Planck equation describes the conservation of the probability density  $p(\vec{x}, t)$  of the system described by the SDE. The first term on the right describes the dynamics of the deterministic system and is called the deterministic drift. The remaining term causes the diffusion of the system.

Now, it is possible to determine the deterministic and stochastic part directly from data by simply using their statistical definition:

$$\vec{A}(\vec{x}) = \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \langle \vec{X}(t + \Delta t) - \vec{x} \rangle |_{\vec{X}(t) = \vec{x}} \quad (4)$$

$$\tilde{B}(\vec{x}) \tilde{B}^T(\vec{x}) = \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \langle (\vec{X}(t + \Delta t) - \vec{x})(\vec{X}(t + \Delta t) - \vec{x})^T \rangle |_{\vec{X}(t) = \vec{x}} \quad (5)$$

where  $\vec{X}(t + \Delta t)$  is a solution, that is, a single stochastic realization of the SDE (1), that starts at  $\vec{X}(t) = \vec{x}$  at time  $t$ . At every point  $\vec{x}$  in the state

space spanned by the data that is visited often enough by the trajectory, deterministic and stochastic parts of the underlying dynamics can be estimated. Note that the theoretical limit  $\Delta t \rightarrow 0$  must be replaced by a finite-difference approximation (see the appendix for the univariate error estimation in the case of a finite time increment  $\Delta t$ ). If required, analytical functions can be fitted to the numerically estimated functions  $\vec{A}(\vec{x})$  and  $\tilde{B}(\vec{x})\tilde{B}^T(\vec{x})$  to formulate analytical model equations to describe the system under consideration. In order to verify the results, the estimated functions  $\vec{A}(\vec{x})$  and  $\tilde{B}(\vec{x})\tilde{B}^T(\vec{x})$  can be inserted into the Fokker-Planck equation (3), and the resulting PDF predicted by (3) can be compared with the PDF obtained directly from the data. Note that  $\tilde{B}(\vec{x})\tilde{B}^T(\vec{x})$  is estimated from data. In general it is impossible to find a unique expression for  $\tilde{B}(\vec{x})$  in the multivariate case, because it is not guaranteed that  $\tilde{B}(\vec{x})$  is invertable. However, in the univariate case  $B(x) = \sqrt{B(x)^2}$ . The sign of the square root is arbitrary because  $B(x)$  is multiplied by Gaussian white noise with zero mean. Thus, in the univariate case even the SDE (1) can be used to test the estimates of  $A(x)$  and  $B(x)$  by simply comparing the properties (e.g., moments, spectra etc.) of the original time series with the properties of the time series obtained by integrating (1).

The technique described here has been successfully applied to a wide class



of problems. For example, Friedrich and Peinke (1997a), Friedrich and Peinke (1997b), and Renner et al. (2001) describe statistical properties of a turbulent cascade. Friedrich et al. (2000a) quantify deterministic and stochastic influences on the foreign exchange market. Geophysical examples are provided by Ditlevsen (1999), who fitted a Fokker-Planck equation to ice core data, Egger (2001), who stochastically described the equatorial components of global angular momentum and related torques using a Master equation, and Egger and Jonsson (2002), who stochastically describe Icelandic meteorological data sets. Note that Egger and Jonsson (2002) analyzed two-dimensional wind data in the same way as it is done in this study. Nevertheless, as discussed above, it is impossible to find a unique expression for  $\tilde{B}(\vec{x})$  in the two-dimensional case. The advantage in Egger and Jonsson (2002) is that the dynamics of the winds come out more clearly. Nevertheless, the noise terms can not be analyzed in a straightforward way, as it is possible in the univariate case. Therefore, Egger and Jonsson (2002) do not focus on multiplicative noise.

### 3 Data

In this study observed global 6-hourly ocean surface winds are used. The data are derived from a space and time blending of QuikSCAT scatterometer observations and NCEP reanalyses (Chin et al. 1998; Milliff et al. 1999). The blending method creates global fields by retaining QuikSCAT wind retrievals in swath regions, and in the unsampled regions augmenting the low-wavenumber NCEP fields with a high-wavenumber component that is derived from monthly regional QuikSCAT statistics. 6-hourly maps of 10 m zonal (u) and meridional (v) wind components are available at a resolution of  $0.5^\circ \times 0.5^\circ$  from  $88^\circ\text{S}$  to  $88^\circ\text{N}$ . The data used cover the period July 20, 1999 to December 31, 2000. The data files are available from the National Center for Atmospheric Research (NCAR) Data Support Section (DSS) at the URL <http://dss.ucar.edu/datasets/ds744.4/>. A detailed description of the blended wind product similar to the one used in the present paper can be found in Milliff et al. (1999), and the details of the blending method are described in Chin et al. (1998). The spectral properties of the QuikSCAT surface wind were recently analyzed by Patoux and Brown (2001).

In reanalyzed datasets (NCEP or ECMWF) observations are assimilated

within an atmospheric model to produce the analyzed fields that are consistent with the model dynamics. Nevertheless, the spatial and temporal coverage of observations over the oceans (in particular over the Southern Ocean) is very limited. That is, the used QuikSCAT based dataset contains much more wind observations over the oceans than the reanalyzed products. Therefore, in contrast to NCEP or ECMWF data, the QuikSCAT based wind data over the oceans are regarded as observations, not as model output. Indeed, scatterometer winds are shown to be tremendously accurate (Bourassa et al. 1997) and are expected to be more reliable than reanalyzed datasets (Patoux and Brown 2001). Nonetheless, the disadvantage of the QuikSCAT dataset is that it covers a rather short period.

Because an estimation of the drift and diffusion functions  $\vec{A}(\vec{x})$  and  $\tilde{B}(\vec{x})\tilde{B}^T(\vec{x})$  requires long time series, it is not possible to obtain reliable estimates at each grid-point. 17 months of 6-hourly wind fields yield 2072 data points, which is insufficient to obtain stable estimates of  $\vec{A}(\vec{x})$  and  $\tilde{B}(\vec{x})\tilde{B}^T(\vec{x})$ . To circumvent this problem, artificial “time series”, which retain the stochastic properties of the wind fields as a function of latitude, are constructed in the following way. Estimates of spatial autocorrelations for Southern Ocean QuikSCAT winds indicate decorrelation scales of about

1000 km in the zonal direction and 1500 km in the meridional direction (not shown). A comparable zonal spatial decorrelation scale is valid for the mid-latitude North Pacific storm track. For tropical and subtropical regions the zonal spatial decorrelation scale is smaller, because there the atmospheric variability is dominated by relatively small scale convective events.

Therefore, grid-points that are separated by 1000 km or more in the zonal direction can be treated as stochastically independent locations representing independent realizations of the same stochastic process. The zonal dependence, as well as the annual cycle, of the wind field are neglected by normalizing the time series to have zero mean and unit standard deviation. This is done by normalizing the data for each month and every chosen zonal location. Finally, several 17 months long time series can be concatenated to obtain a longer “time series” which retains the overall stochastic dynamical properties of the actual wind field as a function of latitude. The resulting time series can be interpreted as normalized zonally averaged wind speeds. This procedure is carried out for the Southern Ocean (20 locations in longitude) and the Pacific Ocean (between 6 and 14 points in longitude, depending on the latitude). The zonal extension of the Atlantic and the Indian Ocean is too small to obtain enough stochastically independent locations. Neverthe-

less, it is expected that the results discussed in the next section are general properties of the atmospheric wind field, and are, therefore, valid for regions other than those analyzed in this paper.

## 4 Results

In this section the method described in section 2 is applied to univariate wind data sets for the Southern and Pacific Oceans. The governing SDE is

$$\frac{dx}{dt} = A(x) + B(x)\eta, \quad (6)$$

where  $x$  is the zonal wind velocity ( $u$ ), and the meridional wind component ( $v$ ) respectively. To evaluate  $A(x)$  and  $B(x)$  using the one-dimensional finite-difference versions of Eqs. (4) and (5), the interval spanned by the data is divided into 50 equal bins (sensitivity experiments with different numbers of bins were performed as well, but the general results discussed below did not change). The error up to the order  $\Delta t$  made by using a finite time increment  $\Delta t$  is calculated in the appendix and given by (A.5) and (A.6). In the following, the smallest possible discrete time step of 6 hours is used for  $\Delta t$  in the finite-difference approximation, except for an error estimation. Actually, the most practical way to detect the error made by using a finite

time step is to change  $\Delta t$  and to compare the results. Therefore, the use of different time steps  $\Delta t$  ensures that the error made by using a finite-difference approximation of (4) and (5) is small and neglectable (see the appendix for a more detailed discussion). Local estimates of  $A(x)$  and  $B(x)$  are significant if the trajectory of the data visits a bin more than 100 times. That is, border points where the significance is automatically low are neglected. For 100 visits Gaussian noise with zero mean and unit variance is detected within a confidence interval  $\pm 0.2$  on a 95% confidence level. This is a very stringent assumption, because the data are assumed to be purely stochastic, so that the deterministic part  $A(x)$  is assumed to be zero. The confidence interval decreases, and the confidence level increases with more visits in a particular bin. However, one has to keep in mind that for this method the data are expected to be Markovian; the method fails if the data are non-Markovian. Thus, one has to check if the method yields stable and physically reasonable results for different numbers of bins and, more importantly, for different discrete time steps  $\Delta t$ . A lack of convergence can be regarded as an indication of the non-Markovian properties of the process.

## 4.1 Southern Ocean

For the Southern Ocean the method yields stable results. That is, the finite-difference estimates of  $A(x)$  and  $B(x)$  do not depend sensitively on the time step  $\Delta t$ . In particular, the results are the same for time steps of  $\Delta t = 6, 12,$  and 18 h. That is, the error terms proportional to  $\Delta t$  in (A.5) and (A.6) are small and neglectable for those time steps. The estimates diverge for time steps equal to or larger than 24 h. Furthermore, it turns out that the agreement of the solutions of the Fokker-Planck equation (3) with the PDFs obtained directly from the data, and the comparison of the properties of the original time series with those of the data obtained by integrating the SDE (1) is very good (see below). Thus, the following results are significant.

As a representative result for the Southern Ocean the estimated functions  $A(x)$  and  $B(x)$  for the zonal and meridional winds at 50°S are shown in Figs. 1 and 2. Other latitudes within the Southern Ocean yield similar results and are, therefore, not shown. The dimensional zonally averaged zonal and meridional wind speeds are  $\bar{u} = 6.6 \text{ m s}^{-1}$  and  $\bar{v} = -1.1 \text{ m s}^{-1}$ . The corresponding zonally averaged standard deviations are  $\bar{\sigma}_u = 5.7 \text{ m s}^{-1}$  and  $\bar{\sigma}_v = 6.5 \text{ m s}^{-1}$ .  $A(x)$  and  $B(x)$  are approximated by fourth-order

polynomial fits:  $A(x) = \sum_{i=0}^4 a_i x^i$  and  $B(x) = \sum_{i=0}^4 b_i x^i$ . Note that the normalized wind speeds, for example the zonal wind  $u$ , can be transformed into dimensional winds through the relation  $x^* = x \bar{\sigma}_u + \bar{u}$ , whereby  $x$  denotes normalized winds and  $x^*$  dimensional winds.

For the zonal wind (Fig. 1) it can be seen that the deterministic term  $A(x)$  acts to damp the wind stress. The damping time scale (the inverse slope of  $A(x)$ ) is not constant but depends somewhat on the zonal wind speed. Near the origin the damping time scale is about 1.5 days. The slope becomes steeper for high absolute values of  $x$ . There, the damping time scale is about 0.5 days. As a first approximation the damping term can be regarded as a linear function with a damping time scale of about 1 day. Note that  $A(x)$  is zero at the origin. That means, the deterministic drift actually acts as a relaxation term towards the mean zonal flow. In the linear case, and if the the function  $B(x)$  is treated as constant, the wind would obey an Ornstein-Uhlenbeck process, also called red noise process or Brownian motion. Nevertheless, the function  $B(x)$  (Fig. 1b) is not constant, but depends on the state of the system. It can be seen that the function  $B(x)$  almost obeys a parabola. Thus, the stochastic forcing appears as multiplicative noise. The need for a parabolic multiplicative noise term to describe



the variability of the Southern Ocean winds can be qualitatively interpreted by the well known fact that, in general, the variability of midlatitude winds increases with increasing wind speed; the multiplicative noise accounts for the gustiness of the synoptic winds. Thus, an additional multiplicative noise term in a stochastic description of the Southern Ocean winds appears to be reasonable. The minimum of  $B(x)$  is not at  $x = 0$ , but is found for a positive value of  $x \approx 0.6$ . Moreover, the absolute gradient of  $B(x)$  is larger to the right of the minimum than it is to the left. Note that zero zonal flow is attained for  $x = -\bar{u}/\bar{\sigma}_u = -1.2$ . Furthermore, in dimensional units the minimum of  $B(x)$  is at  $x^* = 0.6\bar{\sigma}_u + \bar{u} = 10 \text{ m s}^{-1}$ . That is, the structure of the function  $B(x)$  reveals a remarkable characteristic of the underlying time series: the variability of eastward winds *decreases* for *increasing* wind speeds, until the eastward flow exceeds speeds of  $10 \text{ m s}^{-1}$ . Only after this threshold the variability of the zonal wind increases with further increasing wind speeds.

This behavior might be understood in terms of instability in the presence of friction (e.g., Pedlosky 1987). Without friction an unstable wave can use the total amount of energy extracted out of the mean flow for its growth. The presence of friction alters this behavior, because for the wave to grow it must

drain energy from the basic flow in excess of its dissipative loss. For small but increasing wind speeds, the dissipative loss of energy increases more than the gain of energy from the main flow. That means, the variability of the flow decreases with increasing wind speeds. Only beyond a certain threshold wind speed does the energy gain from the basic flow exceed the dissipative loss. After this point, the variability increases with further increasing wind speeds. Nevertheless, this mechanism has to be regarded as a hypothesis.

Likewise for the meridional wind (Fig. 2) it can be seen that the deterministic term  $A(x)$  acts to dampen the wind. Again the damping is not constant but actually depends somewhat on the zonal wind speed. The damping time scales are nearly the same as for the zonal wind. Once more the function  $B(x)$  almost obeys a parabola, although the minimum of  $B(x)$  is now at  $x \approx 0.2$ . Thus, in dimensional units the minimum of  $B(x)$  is at  $x^* = 0.2\bar{\sigma}_v + \bar{v} = 0.2 \text{ m s}^{-1}$ . Note that for the meridional wind zero flow is attained for  $x = -\bar{v}/\bar{\sigma}_v = 0.17$ . That means, that the previously discussed effect for the variability of the zonal wind, is not as strong for the variability of the meridional wind.

In order to verify the results, the estimated functions  $A(x)$  and  $B(x)$

are, on the one hand, inserted into the Fokker-Planck equation (3), and the solution is compared with the PDF obtained directly from the data. Analytic solutions of the Fokker-Planck equation (3) can only be found for limited cases; for more general cases numerical methods must be used. In this paper the semi-implicit Chang-Cooper method is implemented to solve (3) (Chang and Cooper 1970; Park and Petrosian 1996). The PDFs for the zonal and meridional winds are shown in Fig. 3. It can be seen that the solutions of the Fokker-Planck equation (3) are in very good agreement with the PDFs obtained from the data for both the zonal (Fig. 3a) and the meridional wind (Fig. 3b). Even the skewness of the zonal winds is well captured by the solution of the Fokker-Planck equation. This skewness of the zonal winds is a result of the nonsymmetric behavior of  $B(x)$  relative to its minimum at  $x \approx 0.6$ . As seen in Fig. 1b, the absolute gradient of  $B(x)$  is larger right of the minimum than left of it. For example,  $\pm 1$  units away from the minimum,  $|\partial B/\partial x|_{1.6} \approx 0.8$  and  $|\partial B/\partial x|_{-0.4} \approx 0.3$ . Dynamically speaking, right of the minimum the steeper absolute gradient of  $B(x)$ , as compared to the region left of the minimum, gives rise to stronger stochastic kicks than on the left hand side. Thus, the probability that the trajectory  $x$  is found right of the minimum is lower than the probability it is found left of the

minimum. This mechanism is well known to be at work in stochastic system with multiplicative noise (Sura 2002). Because for the meridional wind the absolute gradient of  $B(x)$  is more symmetric relative to its minimum at  $x \approx 0.2$  as compared to the zonal wind, the corresponding PDF is more symmetric as well ( $\pm 1$  units away from the minimum,  $|\partial B/\partial x|_{1.2} \approx 0.4$  and  $|\partial B/\partial x|_{-0.8} \approx 0.5$ ).

On the other hand, the SDE (1) is used to test the estimates of  $A(x)$  and  $B(x)$  by comparing the original time series with the time series obtained by integrating (1). This is done by calculating autocorrelations and spectra of the original and the corresponding “artificial” data (see Fig. 4). The SDE (1) is numerically solved by the stochastic Euler scheme (see e.g., Kloeden and Platen 1992). For the zonal wind the original autocorrelations and the spectra are in good agreement with the artificial ones. Nevertheless, a striking difference can be seen in the spectra at very high frequencies. The stochastic model does not reproduce the steep spectral dip of the original data at frequencies above about 0.3 cycles per 0.25 days (the highest resolved frequency is 0.5 cycles per 0.25 days). It is known that Markov models are not capable of reproducing steep spectral dips at very high frequencies (Del-Sole 2000). For the meridional wind the autocorrelations and spectra are in

slightly poorer agreement. Again, the stochastic model does not reproduce the steep spectral dip of the original data at very high frequencies. Nevertheless, even for the meridional wind the stochastic model reproduces the general characteristics of the original data.

To summarize, a complete stochastic description of the Southern Ocean winds requires a state-dependent white noise term, that is, multiplicative noise. A simple Ornstein-Uhlenbeck process is not sufficient to describe the wind data within a stochastic framework.

## 4.2 Pacific Ocean

For the Pacific Ocean the method yields stable results only in the extratropics (that is, the regions poleward of about  $30^\circ$  N/S). There, the estimates of  $A(x)$  and  $B(x)$  do not depend sensitively on the time step  $\Delta t$ . The successful verification of the results by comparing the original data with the SDE (1) and the Fokker-Planck equation (3) are not presented for the extratropical Pacific Ocean. In the tropics the finite-difference estimates of  $A(x)$  and  $B(x)$  depend on the time step  $\Delta t$ . Thus, in the following only the results for the extratropics are presented. Possible reasons for the lack of convergence in

the tropics are discussed in the final section.

The estimated functions  $A(x)$  and  $B(x)$  for the zonally-averaged zonal and meridional surface winds over the Pacific Ocean as a function of latitude are shown in Figs. 6, 7, 8, and 9. Again,  $A(x)$  and  $B(x)$  are approximated by fourth-order polynomials. For convenience the actual estimates of  $A(x)$  and  $B(x)$  are not shown. In this section only the fitted polynomials are presented, because in the previous section it has been shown that  $A(x)$  and  $B(x)$  can be very well represented by fourth-order polynomials. The corresponding dimensional zonally-averaged zonal and meridional wind speeds and standard deviations are shown in Fig. 5 (for convenience the tropics are included in Fig. 5).

$A(x)$  and  $B(x)$  for the zonally-averaged zonal wind as a function of latitude between  $30^\circ$  N/S and  $55^\circ$  N/S for the North and South Pacific are shown in Figs. 6 and 7. The damping of the zonal winds is nearly constant for the latitudes considered here. On average the damping time scale is about 1 day and, as for the Southern Ocean, the slope of  $A(x)$  becomes steeper for higher absolute values of  $x$ . Again  $A(x)$  becomes zero in the origin of the coordinate system. Thus, the deterministic drift actually acts as a relaxation

term towards the mean zonal flow. At  $30^\circ$  N/S the noise term  $B(x)$  has its minimum for negative values of  $x$ . Note that for  $30^\circ$  N/S dimensional zero zonal flow is now attained for positive values of  $x$ , because the zonal flow is westward at  $30^\circ$  N/S (see Fig. 5). Further poleward the winds change their sign to form the midlatitude Westerlies. There, zero zonal flow is attained for negative values of  $x$ , as previously discussed for the eastward winds over the Southern Ocean. That means, that the variability of westward respectively eastward winds *decreases* for *increasing* wind speeds, until the westward respectively eastward wind exceeds a certain threshold value. Moreover, on both hemispheres  $B(x)$  is nearly constant within the Westerlies.

$A(x)$  and  $B(x)$  for the zonally averaged meridional wind as a function of latitude for the North- and South-Pacific are shown in Figs. 8 and 9. The deterministic term  $A(x)$  acts to damp the wind, whereas again the damping is not constant but depends somewhat on the meridional wind speed. The damping time scales are nearly the same as for the zonal wind. However, for the North Pacific the damping does not increase for large negative values of  $x$ , but the slope is instead nearly constant. Again, the function  $B(x)$  reveals the tendency of the winds to become more variable for increasing wind speeds.

## 5 Summary and Discussion

This paper shows that the synoptic variability of midlatitude Pacific and Southern Ocean sea surface winds can be well described by a stochastic dynamical system directly derived from data. The method used to analyze blended QuikSCAT/NCEP winds is a general method to estimate drift and diffusion coefficients of a continuous stationary Markovian system. The deterministic part consists of a simple damping term. More importantly, a proper description of the analyzed winds requires a state-dependent white noise term, that is, multiplicative noise. An Ornstein-Uhlenbeck process is not sufficient to describe the wind data within a stochastic framework. The need for a parabolic multiplicative noise term to describe the variability of the midlatitude winds can be qualitatively interpreted by the fact that the variability of midlatitude winds increases with increasing wind speed. Therefore, an multiplicative noise term in a stochastic description of midlatitude winds appears to be reasonable. Moreover, the method used reveals another remarkable characteristic of the underlying time series: the variability of westward and eastward winds decreases for increasing wind speeds, until the winds exceed a certain threshold value. This behavior may be understood in



terms of an instability mechanism in the presence of friction. Nevertheless, this notion has to be regarded as a hypothesis, which has to be tested in a modeling study.

The method fails for tropical regions. In the Pacific Ocean the method yields stable results only in the extratropics. There, the finite-difference estimates of the drift and diffusion coefficients do not depend on the time step, whereas in the tropics the estimates do not converge. There is one likely reason for the lack of convergence in the tropics. The tropical variability might be non-Markovian. Yano et al. (2001) present evidence that tropical convective variability behaves as  $1/f$  noise. However,  $1/f$  noise can not be modeled by a simple Markovian process (Hooge 1976). This notion is supported by the spectral analyzes of the tropical wind fields: near the equator the wind fluctuations behave as  $1/f$  noise (see Fig. 10). The  $1/f$  dependence, that is, the non-Markovian behavior in the tropics and its impact on ocean dynamics deserves further research.

The implications of multiplicative noise in sea surface wind data are clear. Sura and Penland (2002) show in the context of a reduced gravity, double-gyre ocean model driven by wind stress how “minor” details of stochastic

forcing can have a significant effect on the temporal and spatial response by the forced system: the regime behavior of the double-gyre model depends on the short-scale temporal aspects of the stochastic forcing. Thus, the accuracy of a stochastic model crucially depends on how appropriate the stochastic representation is to the physical process it is meant to represent. Therefore, the impact of a multiplicative wind forcing on the midlatitude ocean circulation has to be studied in the future.

In summary, the results indicate that it might be necessary to include multiplicative noise terms in stochastic models to fit them to reality or to more complex models. In particular, the structure of the multiplicative noise term  $B(x)$  reveals physical processes not captured by additive noise. Thus, multiplicative noise is essential to capture the complex behavior of given data within an empirical stochastic model.

*Acknowledgments.* I would like to thank Sarah Gille for helpful discussions and support. The author is very grateful to Joe Barsugli and Cécile Penland for lively and instructive discussions about stochastic dynamical systems. The data were kindly made available by Jan Morzel and Ralph Milliff. Furthermore, I thank Joseph Egger, Judith Berner and two anonymous reviewers for their valuable comments that led to significant improvements of the original manuscript. This work was funded through the NASA Ocean Vector Wind Science Team, JPL Contract number 1222984.

# Appendix

## Error Estimation: Stochastic Taylor Expansion

Estimating the deterministic drift (4) stochastic diffusion (5) from data is fraught with the potential for error. In this appendix a simple way of estimating the error in these parameters for an univariate system is derived. A more detailed discussion can be found in Sura and Barsugli (2002). The governing Itô-SDE reads

$$\frac{dx}{dt} = A(x) + B(x)\eta(t) \quad (\text{A.1})$$

with  $\langle \eta(t) \rangle = 0$  and  $\langle \eta(t)\eta(t') \rangle = \delta(t-t')$ . The correct definitions of  $A(x)$  and  $B^2(x)$  are given by (4) and (5). Nevertheless, the definitions are only correct in the limit  $\Delta t \rightarrow 0$ . For a given time increment  $\Delta t$  the finite-difference approximations  $\tilde{A}(x)$  and  $\tilde{B}^2(x)$  of (4) and (5) become

$$\tilde{A}(x) = \frac{1}{\Delta t} \langle X(t + \Delta t) - x \rangle_{|X(t)=x} \quad (\text{A.2})$$

$$\tilde{B}^2(x) = \frac{1}{\Delta t} \langle (X(t + \Delta t) - x)^2 \rangle_{|X(t)=x}. \quad (\text{A.3})$$

To estimate the error made by using a finite time increment  $\Delta t$ ,  $X(t + \Delta t)$  has to be expanded in a stochastic Itô-Taylor series (see e.g., Kloeden and

Platen 1992). The weak stochastic Itô-Taylor approximation up to the order  $\Delta t^2$  reads:

$$\begin{aligned}
X(t + \Delta t) &= X(t) + A I_{(0)} + B I_{(1)} + \left( AA' + \frac{1}{2} B^2 A'' \right) I_{(0,0)} \\
&\quad + \left( AB' + \frac{1}{2} B^2 B'' \right) I_{(0,1)} + BA' I_{(1,0)} + BB' I_{(1,1)} \\
&\quad + R.
\end{aligned} \tag{A.4}$$

The Itô integrals  $I_{(i,j)}$  are defined as

$$\begin{aligned}
I_{(0)} &= \int_t^{t+\Delta t} dt' \\
I_{(1)} &= \int_t^{t+\Delta t} dW(t') \\
I_{(0,0)} &= \int_t^{t+\Delta t} \int_t^s dt' ds \\
I_{(0,1)} &= \int_t^{t+\Delta t} \int_t^s dt' dW(s) \\
I_{(1,0)} &= \int_t^{t+\Delta t} \int_t^s dW(t') ds \\
I_{(1,1)} &= \int_t^{t+\Delta t} \int_t^s dW(t') dW(s)
\end{aligned}$$

where  $W$  denotes a Wiener process. Inserting the expansion of  $X(t + \Delta t)$  in (A.2) and (A.3), and keeping the terms up the order  $\Delta t$  yields the finite-difference estimates  $\tilde{A}$  and  $\tilde{B}^2$ :

$$\begin{aligned}
\tilde{A} &= \frac{1}{\Delta t} \langle X(t + \Delta t) - x \rangle |_{X(t)=x} \\
&= A + \left( \frac{AA'}{2} + \frac{B^2 A''}{4} \right) \Delta t + O(\Delta t^2)
\end{aligned} \tag{A.5}$$

$$\begin{aligned}
\tilde{B}^2 &= \frac{1}{\Delta t} \langle (X(t + \Delta t) - x)^2 \rangle |_{X(t)=x} \\
&= B^2 + \left( A^2 + B^2 A' + BAB' + \frac{1}{2} (B^2 B'^2 + B^3 B'') \right) \Delta t \\
&\quad + O(\Delta t^2).
\end{aligned} \tag{A.6}$$

From (A.5) and (A.6) one can calculate the expected errors for a given finite time increment if  $A(x)$  and  $B(x)$  are known. Note that, of course, for  $\Delta t \rightarrow 0$  the estimates  $\tilde{A}(x)$  and  $\tilde{B}(x)^2$  converge to  $A(x)$  and  $B(x)^2$ . Because it is impossible to know the structure of the terms  $A(x)$  and  $B(x)$  in advance, the most practical way to detect the error made by using a finite time step is to change  $\Delta t$  and to compare the results. Nevertheless, it can be seen immediately that is problematic to detect pure additive noise with finite time steps  $\Delta t$  because a parabolic error emerges. Furthermore, the errors in  $\tilde{A}(x)$  and  $\tilde{B}(x)^2$  depend on nonlinear combinations of  $A(x)$ ,  $B(x)$  and the corresponding derivatives.

To visualize the problem detecting additive noise with finite time steps  $\Delta t$ , a linear deterministic damping term  $A = -x$  is used in combination with a constant stochastic term  $B = 1$ . The known functions  $A(x)$  and  $B(x)$  are inserted into (A.5) and (A.6) to calculate the theoretically expected errors  $\tilde{A}(x) - A(x)$  and  $\tilde{B}(x) - B(x)$ . Then, the theoretical results are compared with the analysis of numerically generated time series. The results are shown in Fig. 11. It can be seen that it is indeed problematic to detect pure additive noise with finite time steps  $\Delta t$  because a parabolic error emerges.

One might ask if the structure of the multiplicative noise discussed in this study is only due to the error terms in (A.5) and (A.6). As already mentioned, it is impossible to know the structure of  $A(x)$  and  $B(x)$  in advance. The practical way to detect the error made by using a finite time step is to change  $\Delta t$  and to compare the results. This has been done, and it appears that the error is neglectable for  $\Delta t = 6, 12,$  and  $18$  h for midlatitude winds. The estimates of  $A(x)$  and  $B(x)$  begin to diverge for time steps equal to or larger than  $24$  h. To test the numerically estimated functions  $\tilde{A}$  and  $\tilde{B}$  for consistency, we assume that the estimated functions are actually correct. Then, the “correct” estimates are inserted into (A.5) and (A.6). If the estimates are consistent with the analytical error estimation, the error terms

in (A.5) and (A.6) should be small. This has been done with the numerical estimates, and a representative example (Southern Ocean zonal winds at 50°S, as in Fig. 1) is shown in Fig. 12. The error is indeed relatively small, and, therefore, the estimates of  $\tilde{A}$  and  $\tilde{B}$  are consistent with the error formulae. This, in combination with the finding that for midlatitude winds the results do not depend on the time step  $\Delta t$ , proves that the multiplicative noise found in midlatitude wind data is not a spurious signal.



## References

- Barnett, T. B., D. W. Pierce, R. Saravanan, N. Schneider, and D. Dommenget, 1999: Origins of the midlatitude Pacific decadal variability. *Geophys. Res. Lett.*, **26**, 1453–1456.
- Blanke, B., J. D. Neelin, and D. Gutzler, 1997: Estimating the effect of stochastic wind stress forcing on ENSO irregularity. *J. Climate*, **10**, 1473–1486.
- Bourassa, M. A., M. H. Freilich, D. M. Legler, W. T. Liu, and J. J. O’Brien, 1997: Wind observations from new satellite and research vessels agree. *EOS Trans. of Amer. Geophys. Union*, **78**, 597, 602.
- Chang, J. S., and G. Cooper, 1970: A practical difference scheme for Fokker-Planck equations. *J. Comput. Phys.*, **6**, 1–16.
- Chin, T. M., R. F. Milliff, and W. G. Large, 1998: Basin-scale, high-wavenumber sea surface wind fields from a multiresolution analyses of scatterometer data. *J. Atmos. Oceanic Technol.*, **15**, 741–763.
- Czaja, A., and C. Frankignoul, 1999: Decadal buoyancy forcing in a simple model of the subtropical gyre. *J. Phys. Oceanogr.*, **29**, 3145–3159.

- DelSole, T., 2000: A fundamental limitation of Markov models. *J. Atmos. Sci.*, **57**, 2158–2168.
- DelSole, T., and A. Y. Hou, 1999: Empirical stochastic models for the dominant climate statistics of a general circulation model. *J. Atmos. Sci.*, **56**, 3436–3456.
- Ditlevsen, P. D., 1999: Observation of  $\alpha$ -stable noise induced millennial climate changes from an ice-core record. *Geophys. Res. Lett.*, **26**, 1441–1444.
- Eckert, C., and M. Latif, 1997: Predictibility of a stochastically forced hybrid coupled model of El Nino. *J. Climate*, **10**, 1488–1504.
- Egger, J., 2001: Master equations for climatic parameter sets. *Climate Dynamics*, **18**, 169–177.
- Egger, J., and T. Jonsson, 2002: Dynamic models for icelandic meteorological data sets. *Tellus A*, **54**, 1–13.
- Friedrich, R., and J. Peinke, 1997a: Description of a turbulent cascade by a Fokker-Planck equation. *Phys. Rev. Lett.*, **78**, 863–866.
- Friedrich, R., and J. Peinke, 1997b: Statistical properties of a turbulent cascade. *Physica D*, **102**, 147–155.

- Friedrich, R., J. Peinke, and C. Renner, 2000a: How to quantify deterministic and random influences on the statistics of the foreign exchange market. *Phys. Rev. Lett.*, **84**, 5224–5227.
- Friedrich, R., S. Siebert, J. Peinke, S. Lück, M. Siefert, M. Lindemann, J. Raethjen, G. Deusch, and G. Pfister, 2000b: Extracting model equations from experimental data. *Phys. Lett. A*, **271**, 217–222.
- Gardiner, C. W., 1985: *Handbook of Stochastic Methods for Physics, Chemistry and the Natural Science, Second Edition*. Springer-Verlag, 442 pp.
- Gradišek, J., S. Siebert, R. Friedrich, and I. Grabec, 2000: Analysis of time series from stochastic processes. *Phys. Rev. E*, **62**, 3146–3155.
- Hasselmann, K., 1976: Stochastic climate models. Part I. Theory. *Tellus*, **28**, 473–484.
- Hasselmann, K., 1988: PIPs and POPs: the reduction of complex dynamical systems using principal interaction and oscillation patterns. *J. Geophys. Res.*, **93**, 11015–11021.
- Hooge, F. N., 1976: 1/f noise. *Physica B,C*, **83**, 14–23.
- Horsthemke, W., and R. Lefever, 1984: *Noise-Induced Transitions: Theory*

- and Applications in Physics, Chemistry, and Biology*. Springer-Verlag, 318 pp.
- Junge, M., J.-S. von Storch, and J. Oberhuber, 2000: Large-scale variability of the main thermocline excited by stochastic wind stress forcing. *J. Climate*, **12**, 2833–2840.
- Kleeman, R., and A. M. Moore, 1997: A theory for the limitation of ENSO predictability due to stochastic atmospheric transients. *J. Atmos. Sci.*, **54**, 753–767.
- Kloeden, P., and E. Platen, 1992: *Numerical Solution of Stochastic Differential Equations*. Springer-Verlag, 632 pp.
- Mikolajewicz, U., and E. Maier-Reimer, 1990: Internal secular variability in an ocean general circulation model. *Climate Dyn.*, **4**, 145–156.
- Milliff, R. F., W. G. Large, J. Morzel, and G. Danabasoglu, 1999: Ocean general circulation model sensitivity to forcing from scatterometer winds. *J. Geophys. Res., Oceans*, **104**, 11337–11358.
- Münnich, M., M. Latif, S. Venske, and E. Maier-Reimer, 1998: Decadal oscillations in a simple coupled model. *J. Climate*, **11**, 3309–3319.

- Park, B. T., and V. Petrosian, 1996: Fokker-Planck equations of stochastic acceleration: A study of numerical methods. *The Astrophysical Journal Supplement Series*, **103**, 255–267.
- Patoux, J., and R. A. Brown, 2001: Spectral analysis of QuikSCAT surface winds and two-dimensional turbulence. *J. Geophys. Res., Atmospheres*, **106**, 23995–24005.
- Paul, W., and J. Baschnagel, 1999: *Stochastic Processes: From Physics to Finance*. Springer-Verlag, 231 pp.
- Pedlosky, J., 1987: *Geophysical Fluid Dynamics, Second Edition*. Springer Verlag, New York, 710 pp.
- Penland, C., and P. D. Sardeshmukh, 1995: The optimal growth of tropical sea surface temperature anomalies. *J. Climate*, **8**, 1999–2024.
- Renner, C., J. Peinke, and R. Friedrich, 2001: Experimental indications of Markov properties of small-scale turbulence. *J. Fluid Mech.*, **433**, 383–409.
- Saravanan, R., and J. C. McWilliams, 1997: Stochasticity and spatial resonance in interdecadal climate fluctuations. *J. Climate*, **10**, 2299–2320.
- Saravanan, R., and J. C. McWilliams, 1998: Advective ocean-atmosphere

- interaction: An analytical stochastic model with implications for decadal variability. *J. Climate*, **11**, 165–188.
- Siegert, S., R. Friedrich, and J. Peinke, 1998: Analysis of data sets of stochastic systems. *Phys. Lett. A*, **243**, 275–280.
- Sreenivasan, K. R., and R. A. Antonia, 1997: The phenomenology of small-scale turbulence. *Annu. Rev. Fluid Mech.*, **29**, 435–472.
- Sura, P., 2002: Noise induced transitions in a barotropic  $\beta$ -plane channel. *J. Atmos. Sci.*, **59**, 97–110.
- Sura, P., and J. J. Barsugli, 2002: A note on estimating drift and diffusion parameters from timeseries. *Phys. Lett. A*, submitted.
- Sura, P., K. Fraedrich, and F. Lunkeit, 2001: Regime transitions in a stochastically forced double-gyre model. *J. Phys. Oceanogr.*, **31**, 411–426.
- Sura, P., F. Lunkeit, and K. Fraedrich, 2000: Decadal variability in a simplified wind-driven ocean model. *J. Phys. Oceanogr.*, **30**, 1917–1930.
- Sura, P., and C. Penland, 2002: Sensitivity of a double-gyre ocean model to details of stochastic forcing. *Ocean Modelling*, **4**, 327–345.

- von Storch, H., and F. W. Zwiers, 1999: *Statistical Analysis in Climate Research*. Cambridge University Press, 484 pp.
- Weng, W., and J. D. Neelin, 1998: On the role of ocean-atmosphere interaction in midlatitude interdecadal variability. *Geophys. Res. Lett.*, **25**, 167–170.
- Winkler, C. R., M. Newman, and P. D. Sardeshmukh, 2001: A linear model of wintertime low-frequency variability. Part I: Formulation and forecast skill. *J. Climate*, **14**, 4474–4493.
- Yano, J. I., K. Fraedrich, and R. Blender, 2001: Tropical convective variability as  $1/f$  noise. *J. Climate*, **14**, 3608–3616.

## Figure captions

Figure 1: a) The estimated deterministic drift  $A(x)$  and b) the estimated noise  $B(x)$  for the zonal wind at 50°S (Southern Ocean). The dashed line with circles shows the actual estimated function, the solid line is a fourth-order polynomial fit.

Figure 2: a) The estimated deterministic drift  $A(x)$  and b) the estimated noise  $B(x)$  for the meridional wind at 50°S (Southern Ocean). The dashed line with circles shows the actual estimated function, the solid line is a fourth-order polynomial fit.

Figure 3: PDFs predicted by the Fokker-Plank equation using the estimated functions  $A(x)$  and  $B(x)$  (solid line), and the PDFs obtained directly from the data (dashed line with circles): a) for the zonal wind, and b) for the meridional wind at 50°S (Southern Ocean).

Figure 4: Autocorrelations and spectra of original data (solid lines) and data obtained by the SDE using the estimated functions  $A(x)$  and  $B(x)$  (dashed lines) for a-b) the zonal wind, and c-d) the meridional wind at 50°S (Southern



Ocean).

Figure 5: Zonal mean and standard deviation of the surface winds over the Pacific Ocean for a) the zonal wind, and b) the meridional wind.

Figure 6: The estimated deterministic drift  $A(x)$  and the estimated noise  $B(x)$  for the zonal wind over the Pacific Ocean. a) and b): between  $30^\circ\text{N}$  and  $40^\circ\text{N}$ . c) and d): between  $45^\circ\text{N}$  and  $55^\circ\text{N}$ .

Figure 7: The estimated deterministic drift  $A(x)$  and the estimated noise  $B(x)$  for the zonal wind over the Pacific Ocean. a) and b): between  $30^\circ\text{S}$  and  $40^\circ\text{S}$ . c) and d): between  $45^\circ\text{S}$  and  $55^\circ\text{S}$ .

Figure 8: The estimated deterministic drift  $A(x)$  and the estimated noise  $B(x)$  for the meridional wind over the Pacific Ocean. a) and b): between  $30^\circ\text{N}$  and  $40^\circ\text{N}$ . c) and d): between  $45^\circ\text{N}$  and  $55^\circ\text{N}$ .

Figure 9: The estimated deterministic drift  $A(x)$  and the estimated noise  $B(x)$  for the meridional wind over the Pacific Ocean. a) and b): between  $30^\circ\text{S}$  and  $40^\circ\text{S}$ . c) and d): between  $45^\circ\text{S}$  and  $55^\circ\text{S}$ .

Figure 10: Spectrum (solid line) of the zonal wind at the equatorial Pacific. The meridional wind has the same spectral dependence. The  $1/f$  slope is indicated by the dashed line.

Figure 11: Error estimates of the finite-difference ( $\Delta t = 0.25$ ) approximations a)  $\tilde{A}(x)$  and b)  $\tilde{B}(x)$  in the case of  $A = -x$  and  $B = 1$ .  $A(x)$ ,  $B(x)$ : solid lines;  $\tilde{A}(x)$ ,  $\tilde{B}(x)$ : dashed lines;  $\tilde{A}(x) - A(x)$ ,  $\tilde{B}(x) - B(x)$ : dotted lines. The corresponding numerical estimates are indicated by the +-signs.

Figure 12: Consistency check of the finite-difference ( $\Delta t = 0.25$ ) approximations in the case of observed data (Southern Ocean zonal winds at  $50^\circ\text{S}$ , as in Fig. 1): The “correct” functions a)  $A(x)$  and b)  $B(x)$  are indicated by solid lines. The theoretically predicted functions  $\tilde{A}(x)$  and  $\tilde{B}(x)$  are indicated by the dashed lines. The errors  $\tilde{A}(x) - A(x)$  and  $\tilde{B}(x) - B(x)$  are indicated by the dotted lines.

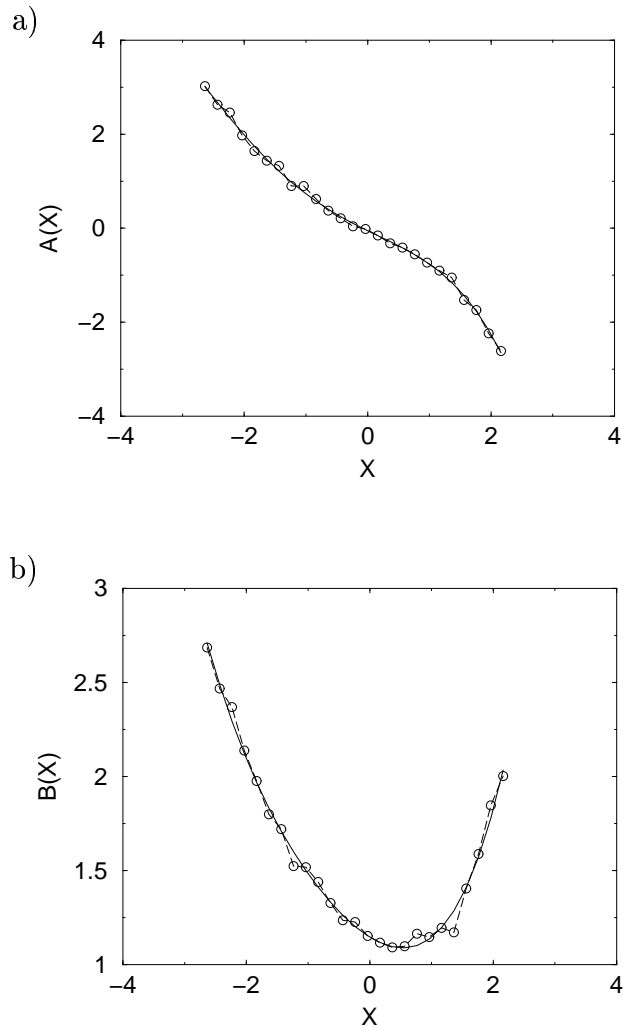


Figure 1: a) The estimated deterministic drift  $A(x)$  and b) the estimated noise  $B(x)$  for the zonal wind at 50°S (Southern Ocean). The dashed line with circles shows the actual estimated function, the solid line is a fourth-order polynomial fit.

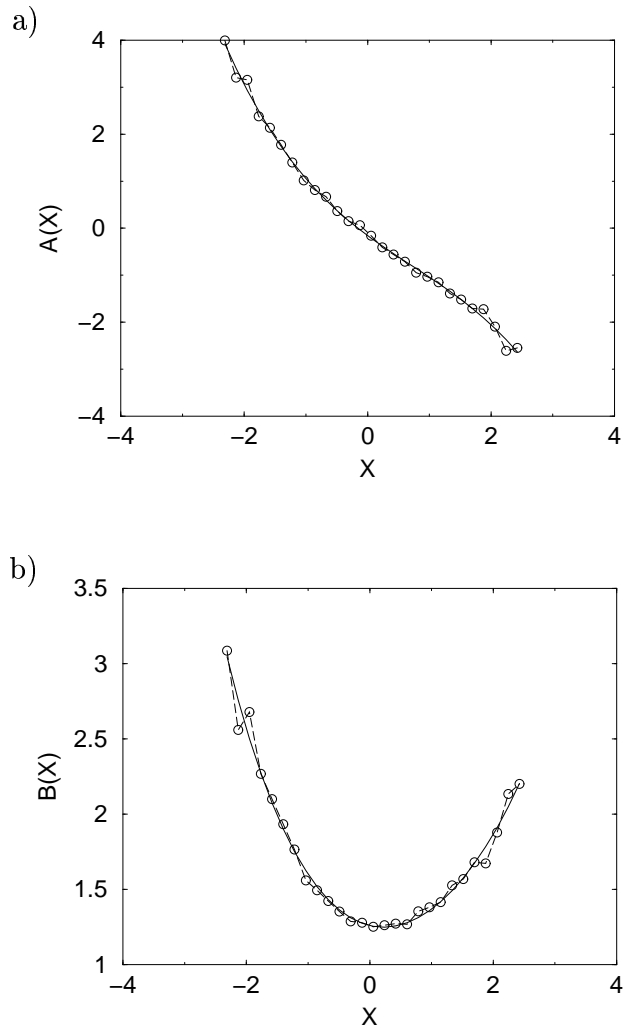


Figure 2: a) The estimated deterministic drift  $A(x)$  and b) the estimated noise  $B(x)$  for the meridional wind at 50°S (Southern Ocean). The dashed line with circles shows the actual estimated function, the solid line is a fourth-order polynomial fit.

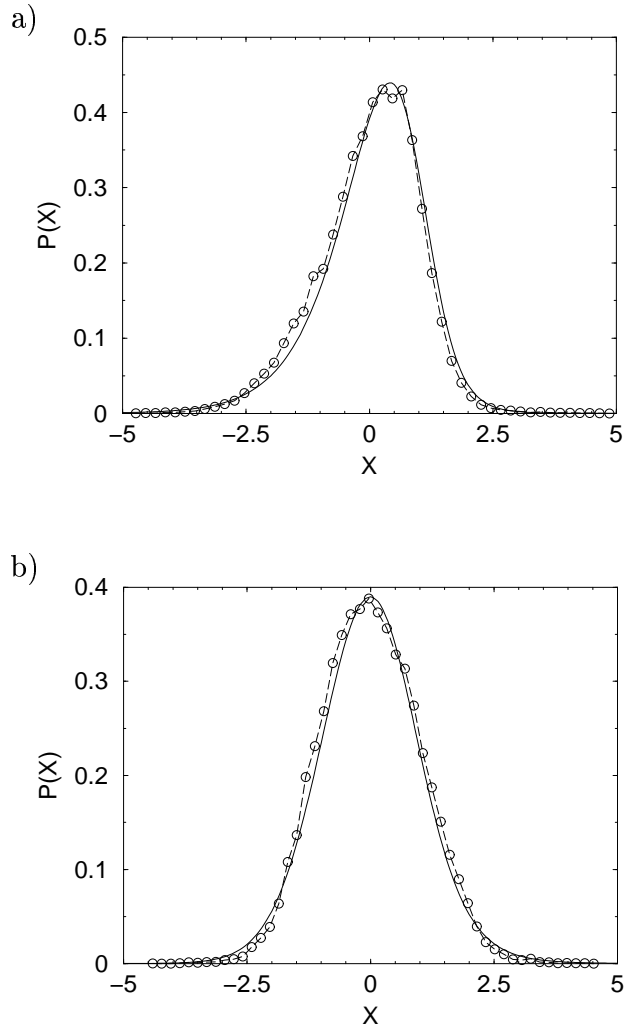


Figure 3: PDFs predicted by the Fokker-Plank equation using the estimated functions  $A(x)$  and  $B(x)$  (solid line), and the PDFs obtained directly from the data (dashed line with circles): a) for the zonal wind, and b) for the meridional wind at  $50^\circ\text{S}$  (Southern Ocean).

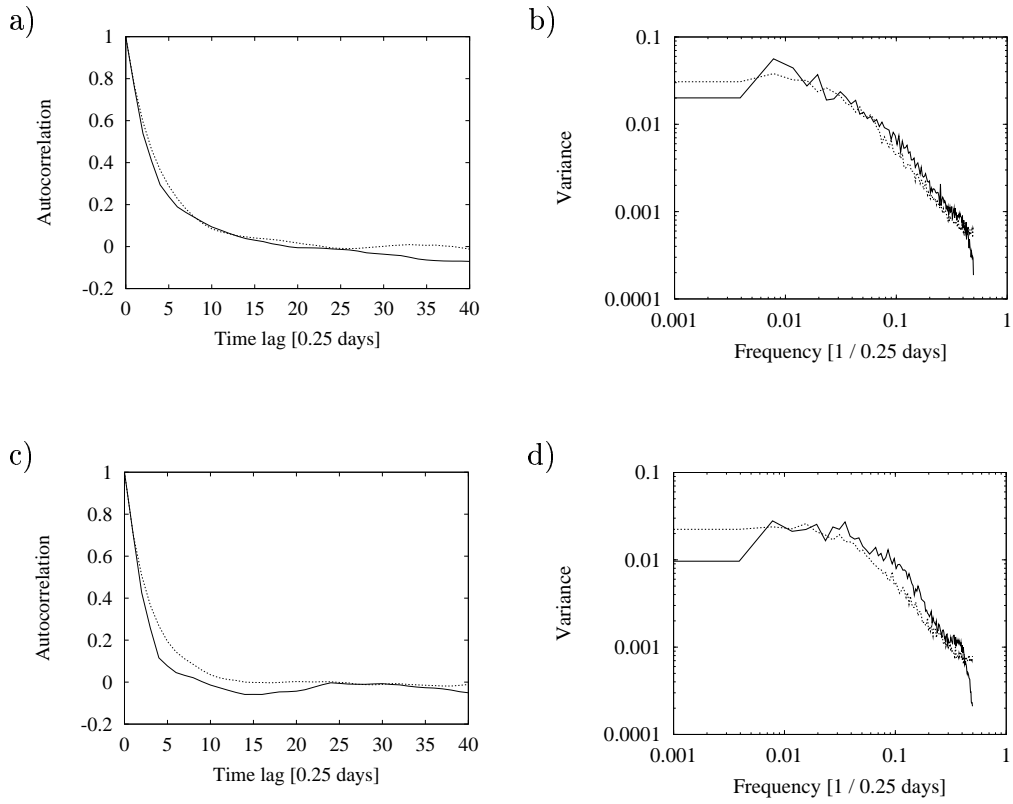


Figure 4: Autocorrelations and spectra of original data (solid lines) and data obtained by the SDE using the estimated functions  $A(x)$  and  $B(x)$  (dashed lines) for a-b) the zonal wind, and c-d) the meridional wind at  $50^\circ\text{S}$  (Southern Ocean).

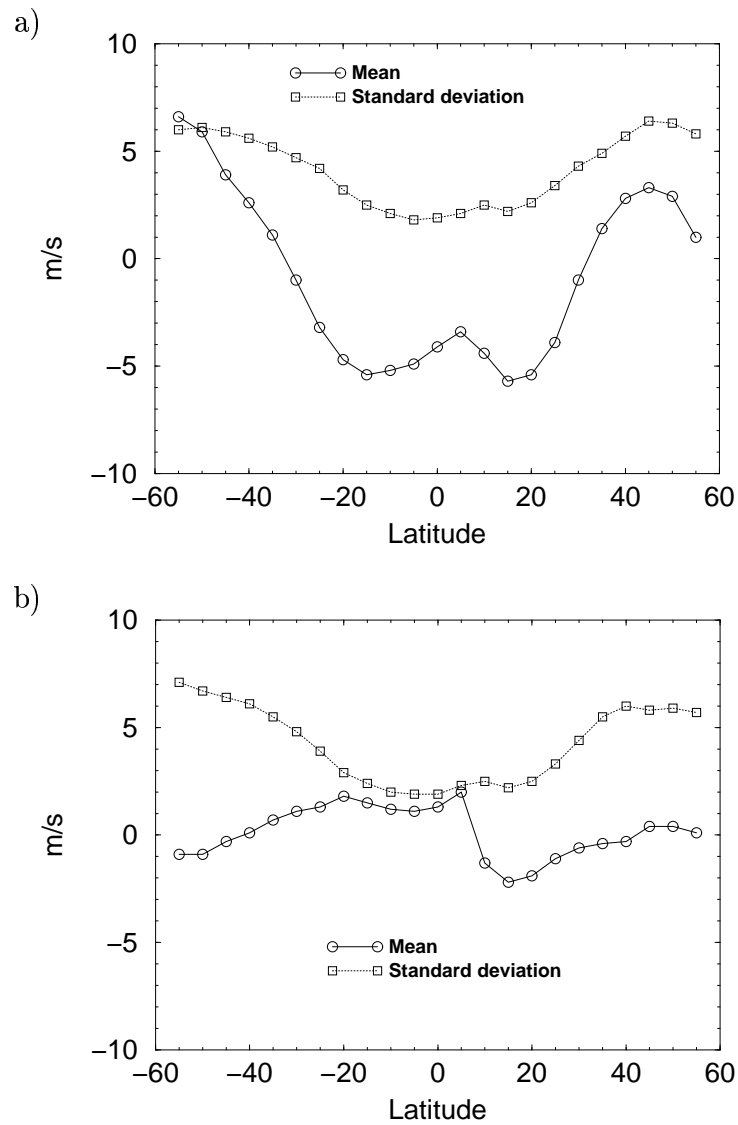


Figure 5: Zonal mean and standard deviation of the surface winds over the Pacific Ocean for a) the zonal wind, and b) the meridional wind.

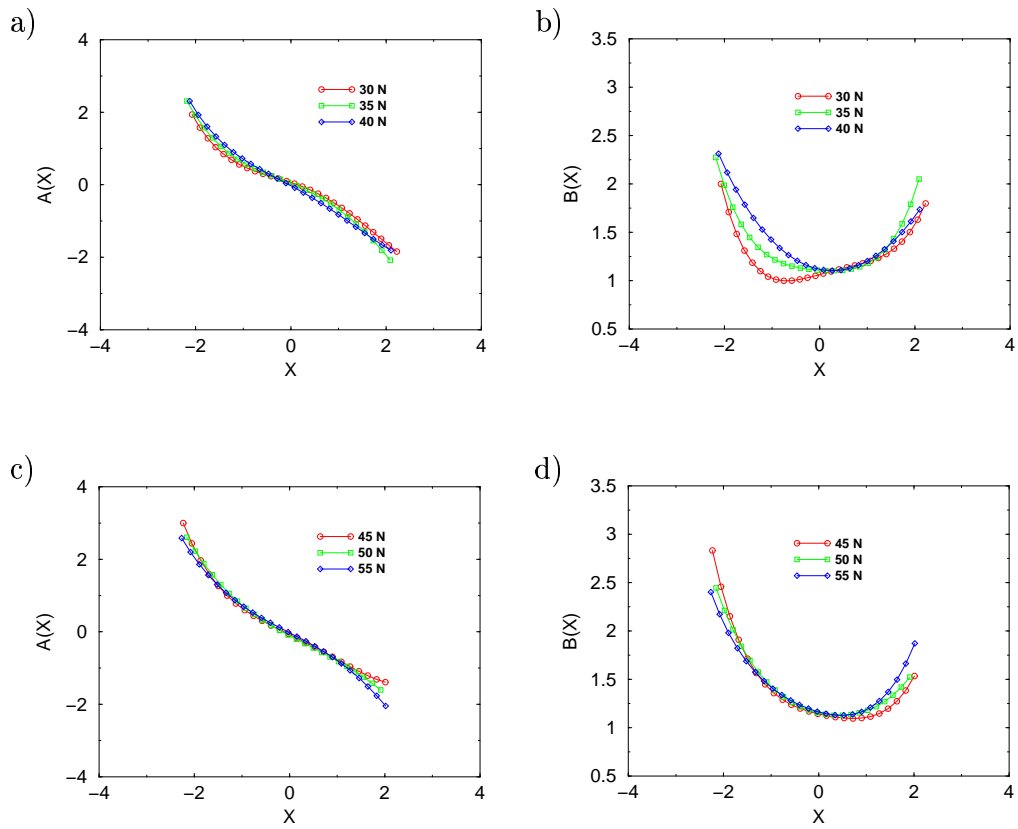


Figure 6: The estimated deterministic drift  $A(x)$  and the estimated noise  $B(x)$  for the zonal wind over the Pacific Ocean. a) and b): between 30°N and 40°N. c) and d): between 45°N and 55°N.



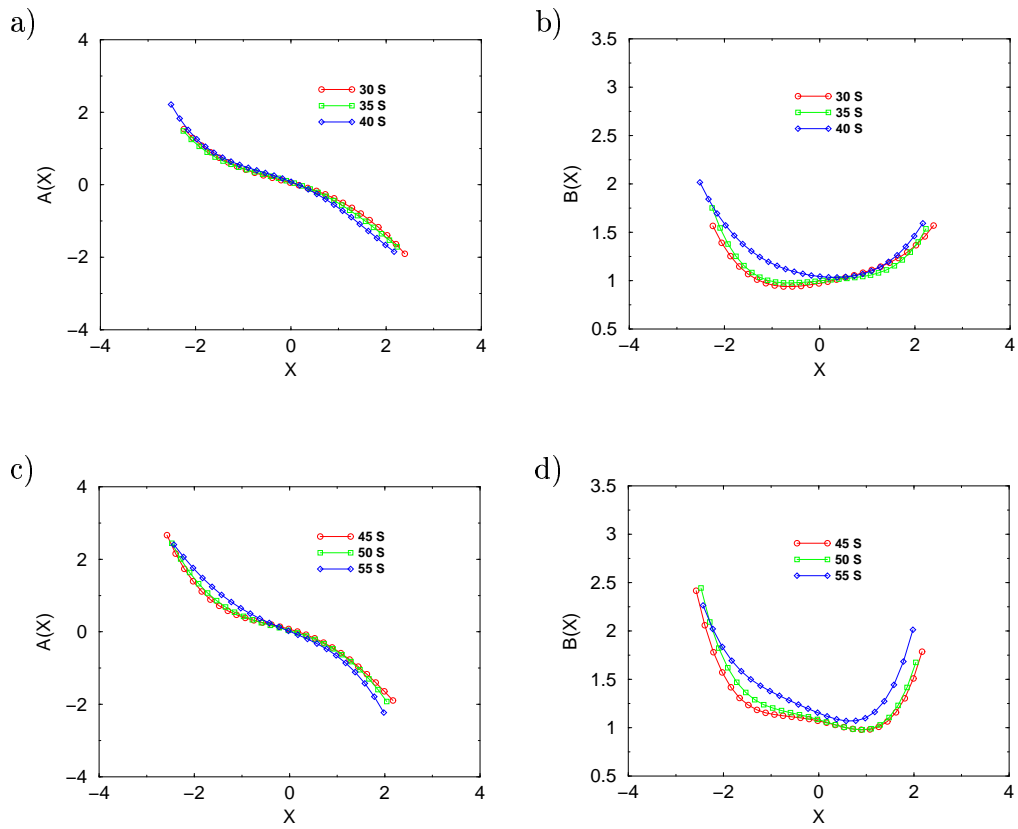


Figure 7: The estimated deterministic drift  $A(x)$  and the estimated noise  $B(x)$  for the zonal wind over the Pacific Ocean. a) and b): between 30°S and 40°S. c) and d): between 45°S and 55°S.

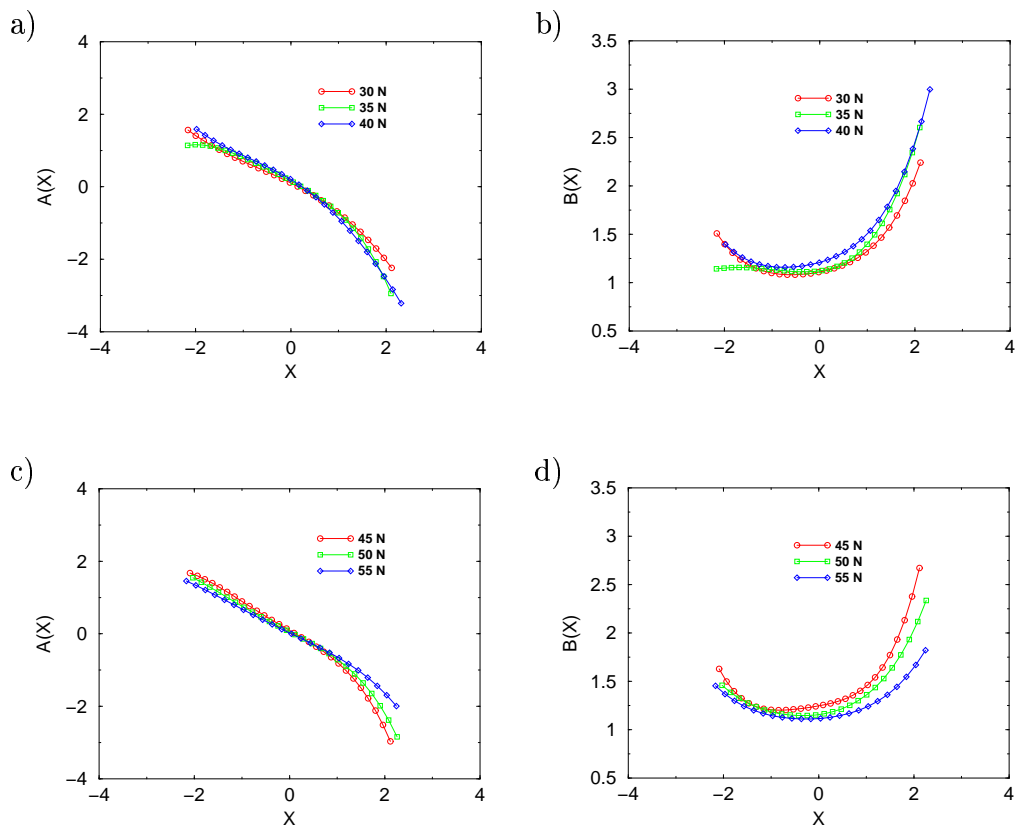


Figure 8: The estimated deterministic drift  $A(x)$  and the estimated noise  $B(x)$  for the meridional wind over the Pacific Ocean. a) and b): between 30°N and 40°N. c) and d): between 45°N and 55°N.

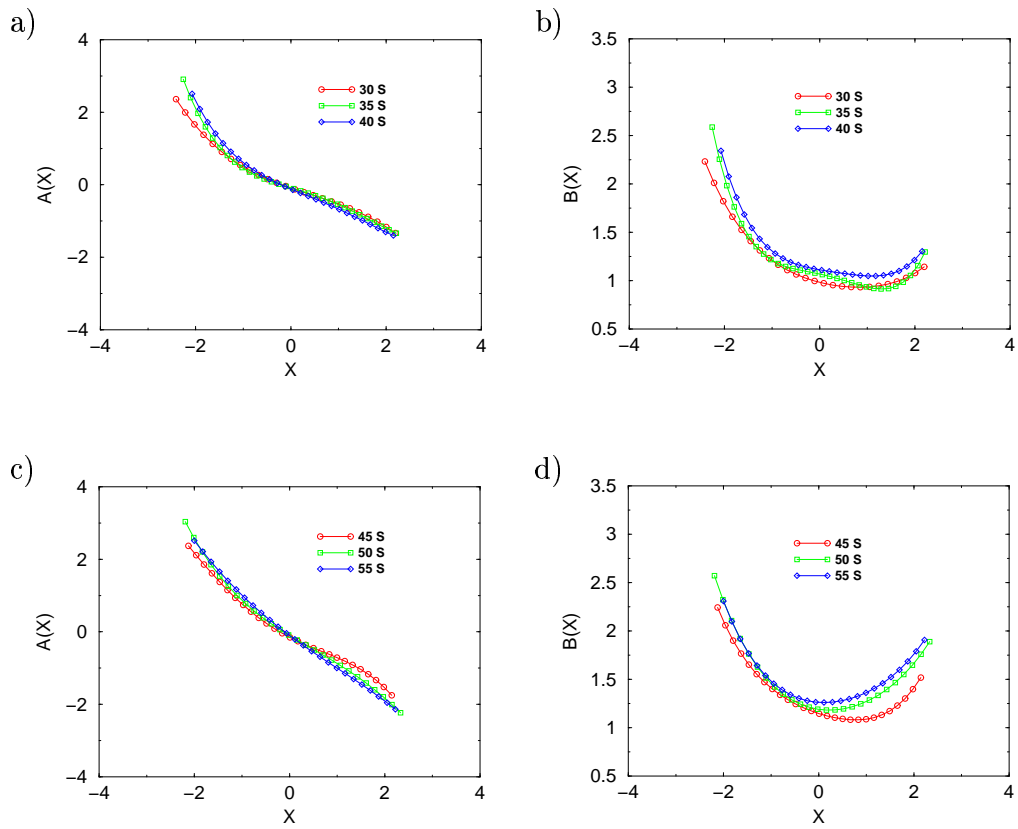


Figure 9: The estimated deterministic drift  $A(x)$  and the estimated noise  $B(x)$  for the meridional wind over the Pacific Ocean. a) and b): between 30°S and 40°S. c) and d): between 45°S and 55°S.

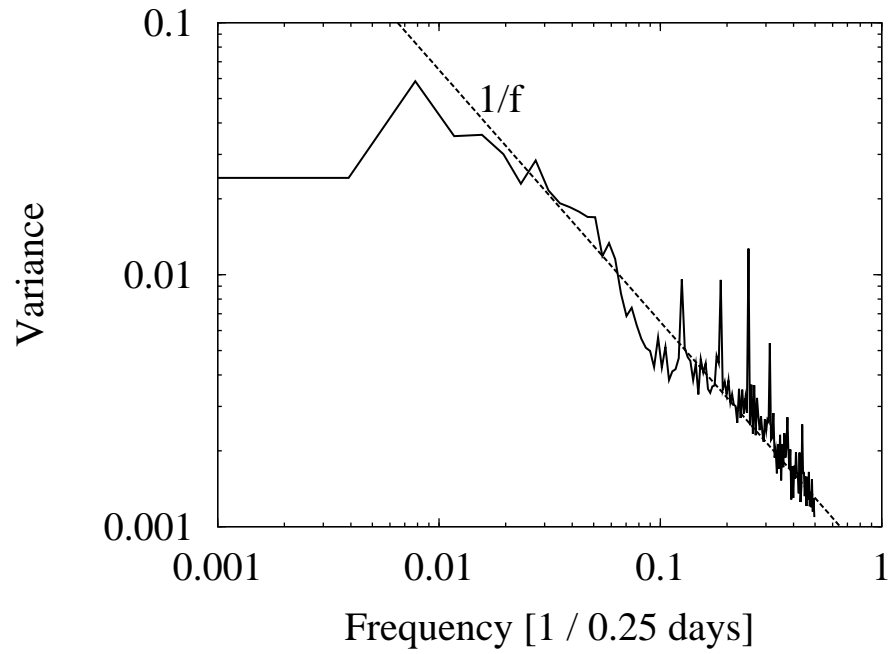


Figure 10: Spectrum (solid line) of the zonal wind at the equatorial Pacific. The meridional wind has the same spectral dependence. The  $1/f$  slope is indicated by the dashed line.

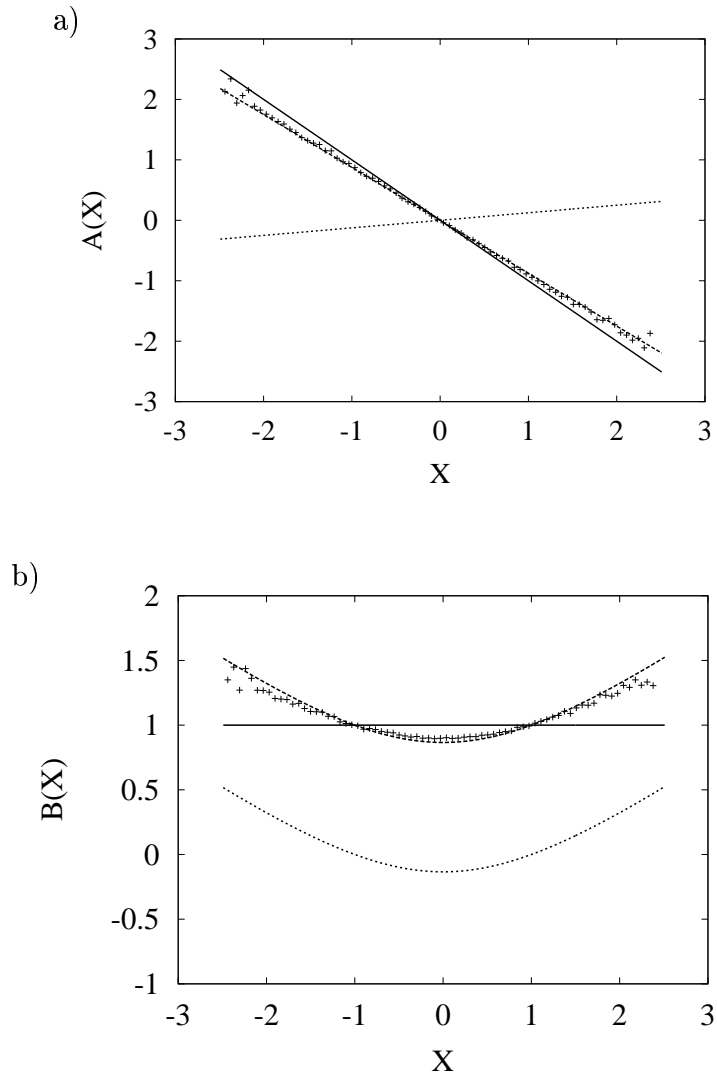


Figure 11: Error estimates of the finite-difference ( $\Delta t = 0.25$ ) approximations a)  $\tilde{A}(x)$  and b)  $\tilde{B}(x)$  in the case of  $A = -x$  and  $B = 1$ .  $A(x)$ ,  $B(x)$ : solid lines;  $\tilde{A}(x)$ ,  $\tilde{B}(x)$ : dashed lines;  $\tilde{A}(x) - A(x)$ ,  $\tilde{B}(x) - B(x)$ : dotted lines. The corresponding numerical estimates are indicated by the  $+$ -signs.

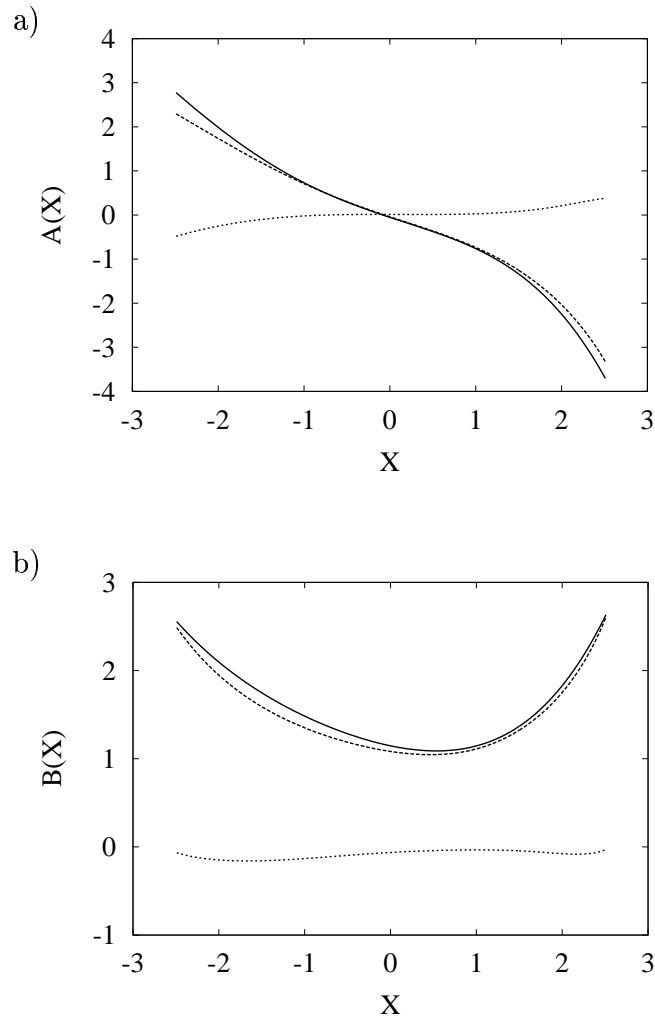


Figure 12: Consistency check of the finite-difference ( $\Delta t = 0.25$ ) approximations in the case of observed data (Southern Ocean zonal winds at  $50^\circ\text{S}$ , as in Fig. 1): The “correct” functions a)  $A(x)$  and b)  $B(x)$  are indicated by solid lines. The theoretically predicted functions  $\tilde{A}(x)$  and  $\tilde{B}(x)$  are indicated by the dashed lines. The errors  $\tilde{A}(x) - A(x)$  and  $\tilde{B}(x) - B(x)$  are indicated by the dotted lines.