Low-Frequency Variability of the Equatorial Pacific Ocean Using a New Pseudostress Dataset: 1930–1989

JAY F. SHRIVER and JAMES J. O'BRIEN

Center for Ocean–Atmospheric Prediction Studies, The Florida State University, Tallahassee, Florida

(Manuscript received 14 January 1994, in final form 25 April 1995)

ABSTRACT

Interannual and interdecadal variability of the equatorial Pacific are examined using a new pseudostress dataset. The monthly mean pseudostress fields (1930–89) are derived from Comprehensive Ocean–Atmosphere Data Set (COADS) pseudostresses using climate basis functions obtained from the Florida State University pseudostress product (1966–90). To validate the new wind fields, a two-tier validation scheme was used. The new wind fields were first examined to see if they exhibited characteristics that have been shown to be important in terms of exciting El Niño events. Next, the new wind fields are used to force an ocean model, thereby obtaining model estimates of tropical Pacific currents and model upper-layer thickness (ULT). Observed sea level and spatially averaged SST anomalies are used to validate the hindcasts. The new wind fields were found to have a significant El Niño mode (accounting for 41% of the variance), which possesses features consistent with those that theory and numerical simulations dictate. Interannual fluctuations in modeled and observed sea level fluctuations are compared at Galapagos and Truk, yielding correlation r values of 0.73 and 0.71, respectively. The comparison of the interannual fluctuations in modeled ULT and observed SST anomalies, which are both spatially averaged over a subdomain in the eastern Pacific basin, results in an r value of 0.64. Interdecadal fluctuations in eastern Pacific model ULT are found to be qualitatively consistent with those in the spatially averaged observed SST anomalies. Stronger El Niño events are observed to occur during periods of higher Pacific ULT (for decadal and longer timescales). Comparison of interdecadal fluctuations in global mean land air temperature and eastern Pacific ULT suggests a connection between eastern Pacific SST and tropical mean land air-temperature warming for interdecadal timescales.

1. Introduction

El Niño is currently the most studied feature of global interannual variability. By definition, El Niño is the occurrence and persistence of anomalously high SST along the coast of Ecuador and Peru. The term El Niño is also used to represent anomalously high sea level and thermocline depression in the eastern Pacific, which is consistent and coincident with SST change (Busalacchi et al. 1983). El Niño events are also thought to decrease hurricane activity in the Atlantic, increase both droughts in southeast Asia and rainfall in parts of the Americas, and cause many other anomalous weather conditions (Enfield 1992).

One key to our understanding of this phenomena was documented by Wyrtki (1975). El Niño was hypothesized to be a response of the equatorial Pacific Ocean to atmospheric forcing. In the tropical Pacific prior to the onset of El Niño, strong trades exist over the central Pacific. These winds drive water toward the western Pacific, increasing the east–west sea level slope. Subsequent, relaxation of the trades then results in a decrease of the east–west sea level slope, accomplished through an equatorial Kelvin wave. Numerical and analytical models by Hurlburt et al. (1976) and McCreary (1976) demonstrated that the downwelling associated with this Kelvin wave yields both a depressed thermocline and an increased sea level in the eastern Pacific. This proposed dynamic mechanism for El Niño was given further credulity when a statistical analysis linked the variability of the zonal component of the southeast trades west of the date line with sea level changes across the Pacific basin (Barnett 1977).

However, to numerically validate Wyrtki’s hypothesis using observed winds, wind fields of sufficient resolution to force an ocean model were necessary. Wyrtki and Meyers (1976) produced the first maps of both time-varying winds and wind stress over the Pacific Ocean for 1947–72. This dataset was on a coarse 2° longitudinal by 10° latitude grid that contained many data-void regions as well as some obvious errors.

The Mesoscale Air–Sea Interaction Group at The Florida State University (FSU) chose to reanalyze the Wyrtki and Meyers wind data (Goldenberg and O'Brien 1981). The analysis was performed on pseudostress (defined as the wind-velocity vector multiplied...

---

Corresponding author address: Dr. Jay F. Shriner, Naval Research Laboratory, Code 7323, Bldg. 1103, Stennis Space Center, MS 39529.
E-mail: shriner@nrlssc.navy.mil

© 1995 American Meteorological Society
by its magnitude), rather than wind speed and direction, in order to produce wind stress fields for forcing ocean models. The three reasons for the reanalysis were

1) to eliminate obvious errors in the Wyrtki and Meyers dataset,
2) to yield a finer resolution dataset, suitable for forcing ocean circulation models, and
3) to estimate data in regions void of observations.

These Pacific wind analyses continue today (Stricherz et al. 1992) using National Climatic Data Center (NCDC) global marine surface observations.

Busalacchi and O'Brien (1981) and Busalacchi et al. (1983) were the first to verify Wyrtki’s hypothesis numerically using observed wind fields. They used a shallow water, reduced gravity, linear model on an equatorial \( \beta \) plane. The forcing used was an estimate of Pacific wind stress (Stricherz et al. 1992). The model simulations covered the period January 1961 through December 1978. Model pycnocline depth variations compared with observed sea level at Galápagos Island showed excellent agreement, both in timing and amplitude. This successful simulation of the El Niño events of 1963, 1965, 1969, 1972, and 1976 confirmed Wyrtki’s hypothesis.

Subsequent El Niño models have been developed by Barnett (1984), Inoue and O'Brien (1984), and Zebiak and Cane (1987) for use in forecasting El Niño occurrences. The basic characteristics of these models are different, even though all three use equatorial Pacific wind information. The models documented in Inoue and O'Brien and Zebiak and Cane are dynamic in nature, while Barnett uses a statistical approach.

Previous modeling efforts have focused on the relatively data-rich period from the 1960s to the present (e.g., Busalacchi and O'Brien 1981; Cane et al. 1986; Kubota and O'Brien 1988; Bigg and Inoue 1992), however, ventured into the data-poor period, investigating wind-forced variability in 1935–46. They were able to reasonably replicate observed eastern Pacific...
sea level trends. No one else has attempted to examine wind-forced variability over a time period longer than about 30 yr, the approximate length of the relatively data-rich period from the 1960s to the present.

A pseudostress dataset, which encompasses a longer period than the FSU dataset, would prove invaluable in the examination of tropical Pacific variability. Model simulations of wind-forced equatorial Pacific variability could be conducted over a longer time period than that spanned by the FSU dataset (1966–present), thereby increasing the ability to develop some insight into the low-frequency variability of the ocean–atmosphere system.

In this study, a pseudostress dataset is compiled for the period 1930–89. To validate the new wind fields, a two-tier validation scheme was used. The new wind fields were first examined to see if they exhibited characteristics that have been shown to be important in terms of exciting El Niño events. Next, using an equatorial Pacific model (Inoue and O'Brien 1984), model estimates of tropical Pacific current and model upper-layer thickness (ULT) variability are then obtained for the period of interest. Historical documentation of El Niño events (Quinn et al. 1987), spatially averaged sea surface temperature (SST), and equatorial Pacific sea level data are then used to validate the model hindcasts. Interannual and interdecadal fluctuations in model ULT and the north equatorial countercurrent (NECC) are also examined and compared with observations.

Other uses for this new pseudostress dataset include deriving equatorial Pacific sea surface temperature distributions through the use of an equatorial Pacific model with thermodynamics and examining midlatitude Pacific variability caused by coastal Kelvin waves.
(e.g., Johnson and O'Brien 1990; Shrider et al. 1991). Using data assimilation techniques, it would be possible to assimilate observed ocean temperatures into the hindcasted SST field to actually improve the wind set. The next section discusses the SST, sea level, and pseudostress data used in this research. Section 3 will present the model used. Section 4 will present and discuss the results. The last section is devoted to our conclusions.

2. Data

a. Pseudostress datasets

1) Comprehensive Ocean–Atmosphere Data Set

The Comprehensive Ocean–Atmosphere Data Set (COADS) (Slutz et al. 1985) covers the period 1854 to the present and has global spatial coverage. The wind observations in this dataset are compiled from ship data and buoys. Some basic quality control is performed so that duplicates and statistical outliers are removed. The threshold used to determine statistical outliers was 3.5 standard deviations from the median. Monthly mean values are formed and then gridded on a $2^\circ \times 2^\circ$ mesh. The domain of interest covers $29^\circ$N–$29^\circ$S, $124^\circ$E–$70^\circ$W, which includes 2155 grid elements at $2^\circ$ resolution. A plot of available observations over time is shown in Fig. 1. Between 70% and 80% of the $2^\circ \times 2^\circ$ grid elements in the domain have at least one observation from 1960 to 1990, tailing off to about 50% by 1950. Prior to 1950, the number of grid elements with observations continues to decrease, with the exception of two surges in the number of observations in the mid-1930s and 1940s, reaching a value of approximately 20% by 1930.
The Florida State University Dataset

The FSU pseudostress dataset consists of monthly fields that span the period 1966 to the present, with spatial coverage from 124°E to 70°W and 29°S to 29°N. The analysis procedure used involves manual contouring and digitization (Stricherz et al. 1992). The data are then gridded on a $2^\circ \times 2^\circ$ spatial grid. The analyses for the period 1966–80 are based on individual ship reports obtained from COADS. The analyses from 1981 to the present are based on NCDC global marine surface observations.

Historical

This dataset consists of monthly wind fields that span the period 1930–89, with spatial coverage from 124°E to 70°W and 29°S to 29°N. The data are gridded on a $2^\circ \times 2^\circ$ spatial grid. The analysis procedure (Shriver 1993) used to form this historical data set use a vector EOF analysis to derive dominant spatial patterns from the FSU Pacific wind stress product. These spatial patterns are then used as basis functions to derive pseudostress fields for the period of interest from COADS. The reader is referred to appendix A for a detailed discussion of the technique used to form this dataset.

Ramage (1987) notes a positive post-WWII trend in the magnitude of the COADS winds, which he attributes to the change in estimating procedure and the growing proportion of ships equipped with anemometers. COADS pseudostress anomalies averaged over the area 19°N–19°S and 142°E–88°W exhibit a decreasing trend in the zonal anomaly (Fig. 2a). This decreasing trend in the zonal anomaly results in an increasing trend in the strength of the prevailing easterlies. However, the spatially averaged meridional component of the COADS pseudostress anomalies (Fig. 2b) does not exhibit a pronounced trend. A secular trend such as that noted by Ramage would act to increase the magnitude of the wind observations over time, but the sense of direction would be preserved. Hence, in a region dominated by easterlies such as the tropical Pacific, the observed trend is predominantly zonal. To avoid having this spurious trend distort the low-frequency variability that this wind dataset would excite in an ocean model, the new wind fields are detrended prior to use (see appendix B for a discussion of the technique used).

The detrended EOF-based wind dataset and COADS pseudostress anomalies averaged over the area 19°N–19°S and 142°E–88°W are shown in Figs. 2a,b. The analyzed anomalies preserve the amplitude and timing of the El Niño/El Viejo events indicated in the COADS anomalies. The analyzed anomalies are somewhat smoother than the COADS anomalies and there is minimal degradation of the amplitude of the observed anomaly signal. Section 4 will present results from the validation of this wind dataset.

Sea level

The sea level data used are monthly averages from Galapagos (1960–90) and Truk (1947–84). Galapagos is located at approximately 0.5°S, 90°W and Truk is located at 152°E, 7.5°N. These two positions are chosen due to their locations on opposite sides of the equatorial Pacific basin. The Galapagos sea level data were provided by K. Wyrkki, and the Truk sea level data were obtained from the Permanent Service for Mean Sea Level (PSMSL) dataset.

Sea surface temperature (SST)

Using available COADS observations, SST values averaged over the area 4°N–4°S, 150°W–90°W are formed for the period 1930–89. The definition of El Niño used in this study is that of the Japan Meteorological Agency (JMA), which defines El Niño as a period over which the 5-month running mean of SST anomalies averaged over the previously mentioned area is at least +0.5°C for a minimum of 6 consecutive months (Marine Department, JMA 1991). The 5-month running mean of SST anomalies is implemented to smooth possible intraseasonal variations.

The years of El Niño events as documented in Quinn et al. (1987) are 1930–31, 1932, 1939, 1940–41, 1943, 1951, 1953, 1957–58, 1965, 1972–73, 1976, 1982–83, and 1987. The El Niño events indicated by the JMA SST anomaly criteria contain all the preceding El Niño events with the exception of 1932 and 1953. Inspection of the spatially averaged SST anomalies (Fig. 3), however, reveals that there was anomalous warming in the eastern equatorial Pacific in 1932 and 1953, but they were not of sufficient duration and amplitude to satisfy the JMA criteria. Since this dataset is a good indicator of the timing and magnitude of El Niño events, a version of it filtered with a 12-month moving average will be used as a validation dataset in section 4. The 12-month moving average is used to highlight the interannual variability in the SST data.

Model

A nonlinear reduced gravity model in spherical coordinates is used to simulate the variability in the equatorial Pacific Ocean due to wind stress forcing. This type of model has been used successfully in previous studies of tropical Pacific variability (e.g., Busalacchi and O'Brien 1981; Kubota and O'Brien 1988; Bigg and Inoue 1992). The model consists of one dynamically active layer of a density $\rho$ and depth $h$ on top of an infinitely deep layer of slightly higher density $\rho + \Delta \rho$. The interface between these two layers is a proxy of the ocean pycnocline. Spherical coordinates are used due to the latitudinal extent of the model (20°S–25°N). The equations defining the model are
Fig. 3. Spatially averaged SST anomalies (monthly mean removed) for the period 1930–89. The SST anomalies were averaged over the area 4°N–4°S and 150°W–90°W. A 5-month running mean is applied to smooth out possible intraseasonal variations. Periods where the SST anomaly is at least +0.5°C for a minimum of 6 consecutive months is classified as El Niño by the Japan Meteorological Agency (JMA).

\[
\frac{\partial U}{\partial t} + \frac{1}{a \cos \theta} \frac{\partial}{\partial \varphi} \left( \frac{U^2}{h} \right) + \frac{1}{a} \frac{\partial}{\partial \theta} \left( \frac{UV}{h} \right)
\]

\[- (2\Omega \sin \theta)V = - \frac{g'}{2a \cos \theta} \frac{\partial h^2}{\partial \varphi} + \frac{\tau_{\varphi}}{\rho} + A\nabla^2 U, \tag{1a}\]

\[
\frac{\partial V}{\partial t} + \frac{1}{a \cos \theta} \frac{\partial}{\partial \varphi} \left( \frac{UV}{h} \right) + \frac{1}{a} \frac{\partial}{\partial \theta} \left( \frac{V^2}{h} \right)
\]

\[+ (2\Omega \sin \theta)U = - \frac{g'}{2a} \frac{\partial h^2}{\partial \theta} + \frac{\tau_{\theta}}{\rho} + A\nabla^2 V, \tag{1b}\]

\[
\frac{\partial h}{\partial t} + \frac{1}{a \cos \theta} \left( \frac{\partial U}{\partial \varphi} + \frac{\partial}{\partial \theta} (V \cos \theta) \right) = 0, \tag{1c}\]

where \(\varphi\) and \(\theta\) are the longitude and latitude, respectively; \(U\) and \(V\) are the transport in the east–west and north–south directions, respectively; \(H\) is the initial depth of the upper layer; \(g' = (\Delta \rho / \rho)g\) is the reduced gravity; \(A\) is a horizontal eddy viscosity coefficient; \(\tau = \rho_s C_D W |W|\) is the wind stress; \(\rho_s\) is the air density; \(C_D\) is a constant drag coefficient; \(W\) and \(|W|\) are the wind velocity vector and its magnitude, respectively; \(a\) is the radius of the earth; and \(\Omega\) is the angular velocity of rotation of the earth. The terms \(A\nabla^2 U\) and \(A\nabla^2 V\) parameterize the horizontal transfer of momentum by turbulent eddy processes. Model parameter values used are listed in Table 1.

Equations (1a–c) were integrated over the domain 20°S–25°N and 124°E–76°W. The equations were discretized using a staggered grid (Arakawa C-grid). At the coastlines, kinematic and no-slip boundary conditions are used. The north and south boundaries are
open, employing a numerically implemented Sommerfeld radiation condition (see Camerlengo and O'Brien 1980).

<table>
<thead>
<tr>
<th>Model parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reduced gravity ($g'$)</td>
<td>0.02 m s$^{-2}$</td>
</tr>
<tr>
<td>Earth's rotation rate ($\Omega$)</td>
<td>$7.292 \times 10^{-5}$ s$^{-1}$</td>
</tr>
<tr>
<td>Initial depth ($H_{init}$)</td>
<td>200 m</td>
</tr>
<tr>
<td>Air density ($\rho_{air}$)</td>
<td>1.2 kg m$^{-3}$</td>
</tr>
<tr>
<td>Sea water density ($\rho_{sea}$)</td>
<td>1025 kg m$^{-3}$</td>
</tr>
<tr>
<td>Eddy viscosity ($A$)</td>
<td>750 m$^2$ s$^{-1}$</td>
</tr>
<tr>
<td>Drag coefficient ($C_d$)</td>
<td>$1.5 \times 10^{-3}$</td>
</tr>
<tr>
<td>Time step ($\Delta T$)</td>
<td>1800 s</td>
</tr>
<tr>
<td>Zonal resolution ($\Delta \varphi$)</td>
<td>$1^\circ$</td>
</tr>
<tr>
<td>Meridional resolution ($\Delta \theta$)</td>
<td>$1^\circ$</td>
</tr>
</tbody>
</table>

4. Results and discussion

a. Sample analysis month

To test the ability of the technique described in appendix A to reliably reconstruct complete pseudostress fields from a limited number of observations, data of several months from the data-rich period (post-1960) were subsampled to simulate months with poor data coverage and then analyzed. The observed pseudostress anomalies, which corresponded to the original and subsampled observed fields, were then checked for consistency.

December 1962 is presented as an example. In Fig. 4a the monthly mean COADS pseudostress anomalies (defined to be departures from the FSU pseudostress dataset's monthly climatology) are shown. There were a total of 1545 $2^\circ$ grid elements in the domain of interest that had at least one observation. The monthly mean pseudostress anomalies for December 1962 were then subsampled to resemble the observation spatial distri-

![Fig. 4.](http://journals.ametsoc.org/jcli/article-pdf/8/11/2762/4699707/1520-0442(1995)008_2762_lfvote_2_0_co_2.pdf)
bution for December 1934 (Fig. 4b). In this example, there were a total of 639 positions with at least one observation in the domain of interest.

The pseudostress anomalies yielded via the analysis procedure described in appendix A for the original and subsampled observed fields for December 1962 are shown in Figs. 5a,b. Root-mean-square differences are calculated between these two fields using the following formula:

\[
\text{RMS}(m) = \left\{ \frac{\sum_{s} \left[ |A(s, m)| - |O(s, m)| \right]^2 \text{obs}(s, m)}{\sum_{s} \text{obs}(s, m)} \right\}^{1/2},
\]

where \(|A(s, m)|\) and \(|O(s, m)|\) denote the magnitude of the analyzed and observed pseudostress for month \(m = 1, 2, \ldots, 720\) (January 1930, February 1930, \ldots, December 1989) and spatial position \(s = 1, 2, \ldots, 2155\); \(\text{obs}(s, m)\) denotes the number of observations; and \(\Psi\) denotes a sum over all spatial positions. Note since \(\text{obs}(s, m) = 0\) at a point with no observations, RMS \((m)\) is only calculated considering points with observations. The number of observations were incorporated into the rms statistic to account for the nonuniform distribution of observations in space and time.

The root-mean-square difference between the COADS anomalies and the analyzed values (Figs. 4a and 5a), respectively, for this month is 2.7 m² s⁻². The difference between the analyzed fields of the original and subsampled observations is shown in Fig. 5c. Note the regions of large differences coincide with areas where there were observations in 1962, but not in 1934. The large anomaly band south of the equator in Fig. 5a is significantly weakened and a spurious anomaly is created in the southeast corner of the domain. This example then demonstrates the capability of the analysis procedure used to produce anomalies that are consistent with the observed COADS anomalies.

The monthly root-mean-square discrepancy (between analysis and observations) for the period 1930–89 is shown in Fig. 6. The monthly mean discrepancy shows a downward trend throughout the period, coinciding with an increase in the quantity of observations. The large peak in the discrepancy near 1942 does not have a corresponding large decrease in the percent possible observations during this period (see Fig. 1). This large discrepancy during this period may be related to a breakdown of the assumptions made to form this new wind dataset during this time. The upturn in the discrepancy during 1980–89 corresponds to a period where the number of gridpoints with at least one observation decreases by approximately 10%.

The mean discrepancy in Fig. 6 is 3.13 m² s⁻². The mean wind speed over the tropical Pacific is approximately 7 m s⁻¹, or 49 m² s⁻² in pseudostress units. This mean discrepancy would correspond to a percent "error" of roughly 6%. The error in anemometer measurements of wind speed from moored buoys is approximately 10% (Hamilton 1990). The error associated with sea state estimates of wind speed and direction are even larger. The COADS wind observations rely heavily on Beaufort estimates of wind speed, particularly toward the beginning of our period of interest (Ramage 1987). We then find that the wind fields produced by the technique we used agree with the observations (in an rms sense) as well or better than the percent error associated with the measurements themselves.

b. Interannual variability

To determine how well the analysis scheme reconstructed the overall wind fields, a two-tier validation scheme was used:

1) The new wind fields were examined directly. That is, the new wind product was examined to see if it exhibited certain characteristics that have been shown to be important in terms of exciting El Niño events.

2) The new pseudostress dataset was then used to force an ocean model. Interannual fluctuations in ULT and currents were then examined and compared with observations.

In section 4b(1) the new pseudostress dataset will be examined to see if it contains features that have been shown to be important in terms of exciting El Niño events. In sections 4b(2) and 4b(3), fluctuations in ULT are investigated at several positions in the domain of interest. In section 4b(4) fluctuations in the NECC are examined. In all cases, for validation purposes, model output is compared with either observed island sea level or spatially averaged SST anomalies.

1) EOF ANALYSIS OF THE NEW PSEUDOSTRESS DATASET

A first approach to validating that the new pseudostress dataset contains sufficient physics to excite an El Niño response consistent with observations is to examine the pseudostress dataset directly. Through numerical simulations (e.g., Busalacchi and O'Brien 1981; Busalacchi et al. 1983; Cane 1984), it has been shown that the most important aspect of the equatorial Pacific wind field in terms of exciting El Niño events is the zonal component in the equatorial, west-central part of the basin. If an El Niño mode could be isolated within the new pseudostress dataset, we should expect to see a large spatial contribution somewhere near the date line and El Niño-type variations in time. If this El Niño model is not present in the new wind dataset, then the new wind fields would be lacking the dynamics necessary to excite an oceanic El Niño response consistent with observations.
An EOF analysis was performed on the zonal pseudo-stress anomaly along the equator averaged between 5°N and 5°S. The first mode accounts for 41.29% of the variance. The spatial pattern (Fig. 7a) shows a large positive contribution centered near the date line, with an area of negative contribution in the
far eastern basin. This pattern is similar to one found by Wang (1992). The time series that modulates Fig. 7a shows oscillations that are consistent with those in the average SST anomalies (Fig. 7b). The correlation coefficient value for the first mode time series and the average SST anomalies has a maximum value of 0.71, with the first mode time series leading the SST anomalies by 1 month. The peaks in the first mode time series correspond to an increase in the zonal anomaly in the western basin, which would then result in a relaxation or reversal of the zonal component of the trades. This connection between the periods where the trades relax or reverse and the warm SST's is exactly the El Niño mechanism put forth by Wyrski (1975). It can then be concluded that there is a significant El Niño mode within the new dataset that has features consistent with those that theoretical and numerical simulations suggest we should see.

2) Sea Level Comparisons

The next approach used to validate the new pseudostress dataset was to use it to force an ocean model and then compare the model results with observations. A traditional dataset used in the validation of a reduced-gravity model is observed sea level (e.g., Busalacchi and O'Brien 1981; Busalacchi and Cane 1985; Kubota and O'Brien 1988). Examination of the sea level data available in the PMSL dataset shows that the equatorial Pacific is extremely lacking in observed sea level data prior to about 1960. The lack of observed sea level prior to 1960 therefore prevents a thorough validation in space and time.

We chose to examine interannual fluctuations in the oceanic response near the equator by considering two locations on opposite ends of the equatorial Pacific basin. The two locations used are Galapagos and Truk. The model sea level anomalies were formed by multi-
Fig. 7. (a) The first eigenmode’s spatial pattern obtained from the EOF decomposition of the new pseudostress product and (b) its corresponding time series. (The average SST anomalies are plotted for comparison.) The first mode accounts for 41.29% of the variance.

Applying the model ULT anomalies by $(\Delta \rho / \rho)$ to get model sea level.

Model sea level fluctuations at a location corresponding to Truk are shown in Fig. 8. The model sea level fluctuations and observed sea level (plotted for comparison) have had their monthly means removed and were then filtered using a 12-month running mean to highlight the interannual variability. The y axes in Fig. 8 have been reversed so that peaks in the curve correspond to El Niño years. Model sea level falls corresponding to the El Niño events of 1957–58, 1963, 1965, 1969, 1972, 1976, and 1982–83 are evident. The model reproduces the observed sea level fluctuations at Truk well, with the exception of some phase disagreement around 1967 and missing the positive sea level anomaly around 1962. The correlation coefficient for these two time series is 0.71.

Model sea level anomalies at a location corresponding to Galapagos are shown in Fig. 9. The model sea level fluctuations and observed sea level (plotted for comparison) have had their monthly means removed and were then filtered using a 12-month running mean to highlight the interannual variability. The model clearly identifies El Niño events (positive anomalies) in 1963, 1965, 1969, 1972, 1976, 1982, and 1986. The timing and relative amplitudes of these events are consistent with the El Niño occurrences and strengths documented by Quinn et al. (1987). The fluctuations in model sea level compare well with those in the observed Galapagos sea level anomalies ($r = 0.73$). However, the model sea level data exhibits a “climatic shift” in the mid-1970s, which is not present in the observed sea level data. This shift in the model sea level data is evident by noting the sharp increase in the mean model sea level anomaly between the periods 1966–76 and 1976–86. This climatic shift that the model sea level indicates was also observed in North Pacific sea level pressure and 500-mb heights (Trenberth 1990).

Note also that the cold events (negative anomalies) are also resolved well in terms of amplitude and timing.
To assess how well this new wind dataset was doing exciting El Niño events, a separate model run was done using the FSU dataset as the wind forcing. The idea behind this run was to see how well the new dataset did exciting eastern Pacific variability as compared to a case using the FSU dataset as the wind forcing. The FSU dataset was chosen for comparison due to it being the longest (1966–present), most consistently analyzed wind dataset for the tropical Pacific, which is suitable for model forcing.

The model ULT response compared with observed sea level from Galapagos using the FSU dataset as the wind forcing is shown in Fig. 10a. The two Galapagos model sea level time series for the cases where the FSU dataset and the new wind dataset were used as the forcing are plotted against each other in Fig. 10b. The observed interannual fluctuations are replicated well by the model. The model, however, overestimated the warm event in 1965 and generally overestimated the cold events. The correlation coefficient for these two time series is 0.71. Since the results using the two different wind forcings account for the same amount of variance (approximately 50%), we can conclude that the new wind dataset does as good a job as the FSU dataset in exciting eastern Pacific variability which is consistent with observations.

3) Comparisons with spatially averaged SST anomalies

We next examined model ULT anomalies averaged over the area 4°N–4°S and 150°W–90°W (Fig. 11).
Fig. 8. The model sea level anomalies at a location corresponding to Truk and observed Truk sea level anomalies. The model sea level anomalies were formed by multiplying the model ULT anomalies by $\Delta \rho / \rho$.

The area $4^\circ N - 4^\circ S$ and $150^\circ W - 90^\circ W$ is the same area over which the spatially averaged SST anomalies discussed in section 2c were computed. Comparison of the model ULT time series with the spatially averaged SST anomalies from section 2c demonstrates that the model resolved the warm events (positive anomalies) and the cold events (negative anomalies) well, with the exception of the period 1945–49. The disagreement in the period 1945–49 can be explained by the following. As discussed in section 4b(1), the region of largest influence for forcing eastern equatorial Pacific sea level variations is found close to the equator in the area around the date line. While looking at the observation distributions over the equatorial Pacific, note the good density of observations in the vicinity of the date line for the period 1940–44 (Fig. 12a) and 1950–54 (Fig. 12c). There are almost no observations, however, along the equator for the 5-yr period 1945–49 (Fig. 12b). The disagreement during the 1945–49 period can then be attributed to extremely poor data coverage in the west-central basin during these 5 yr. The data coverage during this period was so poor in the region of largest influence that the analysis scheme was not able to produce pseudostress fields that contained the necessary physics to excite eastern Pacific variability that was consistent with observations.

Two features of note in Fig. 11 are the general warming trend in the SST anomalies between 1950 and 1989 and the climate shift of the mid-1970s. The warming trend is evident by noting the increase in the mean av-
Figure 9. The model sea level anomalies at a location corresponding to Galapagos and observed Galapagos sea level anomalies.

Average SST anomaly between the periods 1950–59 and 1980–89. The high correlation ($r = 0.63$) between the SST and Ult anomalies suggests that this warming trend in the SST anomalies is due to a trend in the wind forcing. The climatic shift of the mid-1970s was previously noted in North Pacific sea level pressure and 500-mb heights and eastern Pacific model sea level anomalies. The shift of the mid-1970s is also evident in the spatially averaged SST and Ult anomalies depicted in Fig. 11.

4) North Equatorial Countercurrent

The NECC, an eastward flowing current between the westward flowing north and south equatorial currents, owes its existence to variations in the trades from north to south across the equator. In the model simulation, the NECC shows significant variability about its mean position of approximately $6^\circ$N (see Fig. 13). This type of meridional variability in the position of the NECC is consistent with observations (e.g., Metzger et al. 1992). To make a single time series to examine fluctuations in the modeled NECC, the zonal transport ($m^3 s^{-1}$) component at this longitude is averaged between $3^\circ$ and $12^\circ$N.

Interannual fluctuations in the modeled NECC and the spatially averaged SST anomalies from section 2c for 1950–89 are shown in Fig. 14. The monthly climatology (1966–89) has been removed and a 12-month moving average has been applied to highlight the interannual variability. There are several interesting features evident in this figure. The two curves
have a maximum correlation value of $r = 0.70$, where the NECC anomalies lead the SST anomalies by 5 months. The NECC shows a strengthening during El Niño events, where the relative magnitudes of the NECC fluctuations are indicative of the strength of the event. The enhanced NECC acts to transport warm water eastward during an El Niño event and it relaxes during cold events. Wyrtki (1974) inferred variations in the strength of the NECC by considering observed sea level differences across the current. The modeled NECC shows fluctuations that are consistent, in terms of amplitude and timing, with the results of Wyrtki.

c. Interdecadal variability

A fifth-order Butterworth filter, with a cutoff frequency corresponding to a period of 10 yr, was used to low pass the ULT and SST data in Fig. 11. The Butterworth filter was chosen due to its magnitude response being maximally flat in the passband and monotonic overall. The filtered curves are shown in Fig. 15. The low-passed model ULT exhibits several very interesting features. The model ULT anomalies and the SST anomalies show good qualitative agreement. There are positive anomaly peaks around 1940, 1950, 1958, and 1982, with a prolonged negative anomaly
connecting the positive peaks in 1958 and 1982. However, the model ULT misses the large cooling around 1946 indicated in the SST anomalies [as discussed in section 4b(3)].

Global mean land air temperature anomalies (Oce-anic Interdecadal Climate Variability 1992) exhibit interdecadal fluctuations that are consistent with those in the eastern Pacific ULT and SST anomalies. This qualitative agreement would suggest that there is a connection between interdecadal fluctuations in both eastern Pacific ULT and global mean land air temperature. This connection is only robust in the Tropics however, as noted in Diaz and Kiladis (1992). The tropical warming projects onto global mean land air temperature because the Tropics compose approximately one half of the earth's surface area. A connection between El Niño warming in the tropical Pacific and tropical mean land temperature has been previously demonstrated (e.g., Diaz and Kiladis 1992). The results in this study would suggest that this connection is also valid for interdeca-dal timescales and not just for the El Niño—period band (approximately 4–7 yr).

Stronger El Niño events are observed to occur during periods of higher Pacific ULT (for decadal and longer timescales). The large positive peak near 1940 corresponds to the anomalously warm climate in the tropical Pacific during the late 1930s and early 1940s. This was a highly active period, with three El Niño events happening in this time period (as documented by Quinn et al. 1987). Also, the positive peaks in the late 1950s and early 1980s coincide with the severe El Niño events of 1957–1958 and 1982–1983, respectively.
5. Summary and conclusions

Poor data coverage over the world’s oceans has greatly limited the period over which wind-forced ocean models can be used in climate studies. Due to intense interest recently in low-frequency climatic variability, a question arises: how far back in time can equatorial Pacific wind fields be compiled that are suitable for model forcing, whereby accurate simulations of the equatorial Pacific can be carried out? Previous modeling efforts have focused on the relatively data-rich period from the 1960s to the present (e.g., Busalacchi and O’Brien 1981; Cane et al. 1986; Kubota and O’Brien 1988). During this period, wind analyses produced at The Florida State University have been shown to excite equatorial Pacific Ocean variability consistent with observations. In this research, a new pseudostress dataset spanning the period 1930–89 was formed, effectively doubling the period over which wind-forced oceanic variability can be examined.

The new pseudostress dataset was derived from observed COADS pseudostresses. The monthly mean COADS data has 70%–80% coverage during the period 1960–90, but decreases to approximately 20% by 1930 (see Fig. 1). The sparse COADS coverage, particularly before 1960, necessitated an objective analysis scheme that incorporated physical constraints to overcome the analysis problems that very sparse data coverage poses.

Dominant spatial patterns for the period 1966–90 were determined through an EOF analysis of the FSU pseudostress product. It was then assumed that these
dominant spatial patterns (termed climate basis functions) were representative of the spatial variability of the equatorial Pacific wind field for the period of interest (1930–89). An equivalent statement of this assumption is that the set of climate basis functions derived from the FSU dataset are representative of the climate basis functions that would be obtained for the period 1930–89 if the data were adequate. Using the
climate basis function set, COADS pseudostresses, and a technique discussed in appendix A, monthly mean pseudostress fields are obtained for the period of interest (1930–89).

After compiling the new dataset, the next step was to determine how well the analysis scheme reconstructed the overall wind field. This question was answered in two ways: direct examination of the new wind fields and using them to force an ocean model. In both cases, observed data, sea level, and spatially averaged SST anomalies were used for comparison purposes.

Direct examination of the new dataset was done via an EOF analysis. The EOF analysis was done on equatorial zonal pseudostress averaged between 5°N and 5°S. The first mode was found to account for 41.29% of the variance. The spatial pattern for this mode showed a large contribution centered near the date line. The time series that modulates the first mode spatial pattern displayed variability that was consistent with that exhibited by the spatially averaged SST anomalies. It was concluded that this mode was an El Niño mode, possessing features consistent with those which theory and numerical simulations dictate.

A shallow water, reduced gravity, nonlinear model was then forced with this new wind dataset and the model response was compared with observations. Observed sea level at Galapagos and Truk were compared with model sea level values corresponding to the same locations, yielding correlation $r$ values of 0.73 and 0.71, respectively. To validate the eastern Pacific response, model ULT anomalies and observed SST anomalies averaged over the area 4°N–4°S, 150°W–90°W were compared, yielding a correlation coefficient $r$ value of 0.63. The period of 1945–49, however, yields poor agreement.
The poor agreement during 1945–49 is due to the extremely poor data coverage in the west-central basin during this period. The data coverage was so poor in the region of largest influence during this period that the analysis scheme was not able to produce pseudostress fields that contained the necessary physics to excite eastern Pacific variability which was consistent with observations.

Next, interdecadal variability in eastern Pacific ULT and SST were examined. The low-frequency fluctuations in the ULT and SST data showed good qualitative agreement. Examination of global mean land air temperature showed low-frequency fluctuations that were consistent with those in the ULT data. This qualitative agreement would suggest that there is a connection between interdecadal fluctuations in both eastern Pacific ULT and global mean land air temperature. This connection is only robust in the Tropics however, as noted in Diaz and Kiladis (1992). It has been previously demonstrated that there is a clear relationship between El Niño warming in the eastern Pacific and an increase in mean land air temperature in the Tropics. Our results suggest that the connection between eastern Pacific Ocean and mean land air temperature warming in the Tropics also holds for interdecadal timescales.

Stronger El Niño events are observed to occur during periods of higher Pacific ULT (for decadal and longer timescales). The peaks in low-passed eastern Pacific ULT and SST in 1940, the late 1950s, and the early 1980s correspond to the strong El Niño events of the
Fig. 15. Low-passed ULT and SST anomalies averaged over the area 4°N–4°S, 150°W–90°W.
To filter periods less than 10 yr, a Butterworth filter was used.

late 1930s and early 1940s, 1957–58, and 1982–83, respectively.

Inspection of the model-observed data time series shows noticeably less agreement prior to 1950. This raises the question of how useful the wind data is prior to 1950. The ability of the technique used to estimate data in regions void of observations is related to the amount of data missing, that is, the more data that is missing, the less reliable the technique’s estimates in these regions are. To expect agreement during the period of 1930–50 comparable to that after 1960 is not reasonable, since the data coverage during 1930–50 is significantly less than during the more recent part of the dataset. Considering the poor density of observed data prior to 1950 (approximately 30% coverage), good qualitative and reasonable quantitative agreement is still found during this period. The agreement during the period 1930–50 certainly is not ideal, but with the exception of the period 1945–49, the warm and cold events are resolved well. The wind fields derived during the period 1930–50 are still usable to examine low-frequency oceanic variability, provided the user understands the limitations of the data due to the low density in space and time of the observations.

Acknowledgments. This research was supported by NASA (Grant NGT50653) and the Physical Oceanography Division of the Office of Naval Research. The additional support and data received from the TOGA Project at the Office of Global Programs at NOAA
proved to be a valuable aid. We would like to thank the members of the Center for Ocean–Atmosphere Prediction Studies for their help and support during this research. We would also like to acknowledge the insight and encouragement of Joe Fletcher who inspired this work. Finally, we would like to thank Harley Hurlburt, Brian Kelly, and the anonymous reviewers whose comments and suggestions significantly improved the final manuscript. The undetrended and detrended wind fields produced through this research are available via anonymous ftp: please e-mail winds@coaps.fsu.edu for details.

APPENDIX A

Derivation of New Pseudostress Fields

Dominant spatial patterns for the period 1966–90 (termed climate basis functions) can be determined through a complex EOF analysis of the FSU Pacific wind stress product. It is assumed that the spatial variability contained within these climate basis functions is sufficient to reproduce the El Niño events for 1930–89. An equivalent statement of this assumption is that the set of climate basis functions derived from the FSU dataset are representative of the climate basis functions that would be obtained for the period 1930–89 if the data were adequate. Desired wind fields for the period 1930–89 can then be represented as the superposition of this set of climate basis functions (each modulated by a certain time coefficient). The method used to solve for these time coefficients will be developed later in this section.

The FSU Pacific wind stress product (1966–90) is analyzed using complex EOF analysis. Let \( \mathbf{w} = \mathbf{T}_x + i\mathbf{T}_y \), where \( \mathbf{w} \), \( \mathbf{T}_x \), and \( \mathbf{T}_y \) are each 2155 (spatial positions) by 300 (months) matrices, and \( i = \sqrt{-1} \). The matrices \( \mathbf{T}_x \) and \( \mathbf{T}_y \) represent the eastward and northward components of the FSU pseudostress product, respectively. Before EOF processing, the climatological monthly means from \( \mathbf{T}_x \) and \( \mathbf{T}_y \) are removed. The EOF decomposition of \( \mathbf{w} \) yields 300 pairs of space–time vectors \( Z \) and \( M \).

We next wanted to isolate the statistically significant spatial patterns yielded by the FSU vector EOF analysis. The significant patterns are determined using a statistical test (Overland and Preisendorfer 1982). Spatial patterns are determined to not be significant if the eigenvalue corresponding to that spatial pattern is less than one generated from a random dataset. The set of \( K \) statistically significant spatial patterns (\( K = 40 \)) will then be referred to as the climate basis functions set. The percent variances accounted for by the first 40 eigenmodes are shown in Table A1. The percent variances range from a high of 9.18\% for mode 1 to 0.51\% for mode 40. In an EOF analysis it is desirable to have a large portion of the variance accounted for by the first few eigenmodes, thereby increasing the likelihood that the most significant eigenmodes are physically meaningful. In this case, the first three eigenmodes account for less than 24\% of the variance and there are not large differences in the variance accounted for between adjacent eigenmodes. Since it is difficult to distinguish physically the meaning of a given eigenmode in cases where the percent variances are close together, spatial patterns and time series will not be shown.

The climate basis functions set, along with COADS winds, can now be used to determine complete wind fields for the years prior to 1966. Let the complex matrix \( \mathbf{T} \) represent the COADS data. Since the climate basis functions were derived from wind fields where the monthly climatology was removed, the monthly climatology must also be removed from \( \mathbf{T} \).

The calculation of a monthly climatology from \( \mathbf{T} \) for the period 1930–89 would produce a product where the representativeness would vary due to spatial and

<table>
<thead>
<tr>
<th>Mode</th>
<th>Percent variance</th>
<th>Cumulative variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>9.18</td>
<td>9.18</td>
</tr>
<tr>
<td>2</td>
<td>7.41</td>
<td>16.59</td>
</tr>
<tr>
<td>3</td>
<td>6.98</td>
<td>23.58</td>
</tr>
<tr>
<td>4</td>
<td>4.45</td>
<td>28.02</td>
</tr>
<tr>
<td>5</td>
<td>3.66</td>
<td>31.69</td>
</tr>
<tr>
<td>6</td>
<td>3.43</td>
<td>35.12</td>
</tr>
<tr>
<td>7</td>
<td>3.25</td>
<td>38.36</td>
</tr>
<tr>
<td>8</td>
<td>2.76</td>
<td>41.13</td>
</tr>
<tr>
<td>9</td>
<td>2.46</td>
<td>43.58</td>
</tr>
<tr>
<td>10</td>
<td>2.28</td>
<td>45.87</td>
</tr>
<tr>
<td>11</td>
<td>2.11</td>
<td>47.98</td>
</tr>
<tr>
<td>12</td>
<td>1.94</td>
<td>49.92</td>
</tr>
<tr>
<td>13</td>
<td>1.77</td>
<td>51.69</td>
</tr>
<tr>
<td>14</td>
<td>1.64</td>
<td>53.32</td>
</tr>
<tr>
<td>15</td>
<td>1.42</td>
<td>54.75</td>
</tr>
<tr>
<td>16</td>
<td>1.36</td>
<td>56.11</td>
</tr>
<tr>
<td>17</td>
<td>1.33</td>
<td>57.44</td>
</tr>
<tr>
<td>18</td>
<td>1.23</td>
<td>58.67</td>
</tr>
<tr>
<td>19</td>
<td>1.18</td>
<td>59.84</td>
</tr>
<tr>
<td>20</td>
<td>1.07</td>
<td>60.92</td>
</tr>
<tr>
<td>21</td>
<td>1.04</td>
<td>61.95</td>
</tr>
<tr>
<td>22</td>
<td>0.99</td>
<td>62.95</td>
</tr>
<tr>
<td>23</td>
<td>0.95</td>
<td>63.89</td>
</tr>
<tr>
<td>24</td>
<td>0.93</td>
<td>64.82</td>
</tr>
<tr>
<td>25</td>
<td>0.87</td>
<td>65.70</td>
</tr>
<tr>
<td>26</td>
<td>0.85</td>
<td>66.55</td>
</tr>
<tr>
<td>27</td>
<td>0.82</td>
<td>67.36</td>
</tr>
<tr>
<td>28</td>
<td>0.79</td>
<td>68.15</td>
</tr>
<tr>
<td>29</td>
<td>0.73</td>
<td>68.89</td>
</tr>
<tr>
<td>30</td>
<td>0.73</td>
<td>69.62</td>
</tr>
<tr>
<td>31</td>
<td>0.70</td>
<td>70.32</td>
</tr>
<tr>
<td>32</td>
<td>0.68</td>
<td>71.01</td>
</tr>
<tr>
<td>33</td>
<td>0.65</td>
<td>71.66</td>
</tr>
<tr>
<td>34</td>
<td>0.62</td>
<td>72.28</td>
</tr>
<tr>
<td>35</td>
<td>0.60</td>
<td>72.88</td>
</tr>
<tr>
<td>36</td>
<td>0.56</td>
<td>73.44</td>
</tr>
<tr>
<td>37</td>
<td>0.55</td>
<td>74.00</td>
</tr>
<tr>
<td>38</td>
<td>0.54</td>
<td>74.54</td>
</tr>
<tr>
<td>39</td>
<td>0.52</td>
<td>75.06</td>
</tr>
<tr>
<td>40</td>
<td>0.51</td>
<td>75.57</td>
</tr>
</tbody>
</table>
temporal variations in the distribution of observations. The FSU monthly climatology is another dataset that could be used to demean \( T \). However, the use of the FSU climatology to demean the COADS data is equivalent to stating that the climatology has not changed significantly over the equatorial Pacific.

The idea that climatology over the equatorial Pacific hasn’t changed significantly over the period 1930–89 is certainly a nontrivial assumption. A two-tailed \( t \) test was used, using 95% confidence limits, to see if the null hypothesis that the FSU climatology was the correct climatology for 1930–89 could be rejected. If the FSU climatology is the proper climatology for the period of interest, then an alternative statement of the null hypothesis would be that the monthly mean COADS departures from the FSU climatology are zero. The alternative null hypothesis is what was actually tested. Approximately 85% of the time the null hypothesis could not be rejected. The FSU climatology is then used to demean the COADS data.

A variation of Cressman’s (1959) objective analysis scheme is first applied to the COADS anomalies to smooth the values at positions with data and to interpolate values (if possible) at positions without data. As originally put forth, the weights for this scheme are radially symmetric. In the equatorial Pacific, east–west length scales are longer than north–south scales (e.g., Kubota and O’Brien 1988); therefore it is expected that the weight contours should resemble ellipses, not circles.

The weights are determined as follows: using the FSU dataset, where the monthly means have been removed, vector autocorrelations \( r^2 \) are calculated first for each north–south transect for each month, yielding an autocorrelation matrix \( \text{acf}1(x, \text{ylag}, t) \), where \( x \) is the longitudinal grid position, \( \text{ylag} \) is the north–south lag, and \( t \) is the time index. The same is then done for each east–west transect yielding a matrix \( \text{acf}2(y, \text{xlag}, t) \), where \( \text{xlag} \) is the east–west lag and \( y \) corresponds to the latitudinal grid position. The functions \( \text{acf}1 \) and \( \text{acf}2 \) are then averaged to yield monthly average autocorrelation matrices \( \text{acf}1(x, y, \text{lag}, m) \) and \( \text{acf}2(y, x, \text{lag}, m) \), where \( m = 1, 2, \ldots, 12 \) (January, February, ... , December). We then define a weight matrix \( \text{acf} \) as follows:

\[
\text{acf}(x, y, m; x_0, y_0) = \exp \left[ - \frac{(x - x_0)^2}{\gamma(y, m)^2} - \frac{(y - y_0)^2}{\Delta(x, m)^2} \right],
\]

where \( x_0 \) and \( y_0 \) correspond to the position where either an existing value is being smoothed or a value is being interpolated, and \( x \) and \( y \) correspond to some neighboring position. The \( e \)-folding matrix \( \Delta(x, m)^2 \) is defined as the \( (y - y_0)^2 \) distance in which \( \text{acf}1 \), for fixed \( x \) and \( m \), decreases to \( e^{-1} \). The \( e \)-folding matrix \( \gamma(y, m)^2 \) is defined as the \( (x - x_0)^2 \) distance in which \( \text{acf}2 \), for fixed \( y \) and \( m \), decreases to \( e^{-1} \). Any acf values less than \( e^{-2} \) are set to zero.

The algorithm used for the scan is as follows: at first we calculate

\[
\hat{T}(x_0, y_0, t) = \sum_x \sum_y \text{acf}(x, y, m; x_0, y_0) \times \text{nobs}(x, y, t)T(x, y, t),
\]

where \( T(x, y, t) \) is the COADS observation and \( \text{nobs}(x, y, t) \) is the number of observations at position \((x, y)\) and time \( t \). After one pass of the entire data set, \( \hat{T} \) is then smoothed in time at each spatial position using one \( 1/4 - 1/2 - 1/4 \) weighting Hanning pass.

The desired pseudostress product can be represented as a superposition of the climate basis functions by the following:

\[
Y = \mathbf{Z}a,
\]

where \( Y \) is a complex column vector of length 2155 (spatial positions), representing a complete wind field. The complex matrix \( \mathbf{Z} \) is a 2155 (spatial position) by \( K \) (number of spatial patterns) matrix, where the column vectors of \( \mathbf{Z} \) constitute the set of climate basis functions. The complex column vector \( a \) contains the \( K \) unknown modulation coefficients. Note that the procedure used is done for a particular month \( t \) and that each monthly wind field \( Y \) is solved for independently.

Next, we need to determine a solution for \( a \). Determining the discrepancy between (A2) and \( W \) is straightforward. The column vector \( W \) is of length 2155 (spatial positions) and has the values yielded by (A1) for a month \( t \) mapped into it. Weights related to the number of observations at a given point can also be incorporated. This leads to the following discrepancy sum, which we would want to minimize:

\[
D = |HW - \mathbf{HZ}a|,
\]

where \( H \) is a square diagonal matrix of order 2155 (spatial positions) with real weights \( \alpha_x \) on the diagonal and \( x \) is an arbitrary solution vector of length \( K \). The weights \( \alpha_x \) are equal to the number of observations at position \( x \) if \( W_x \) (corresponding to an element of \( W \) is a smoothed value, 1 if \( W_x \) is an interpolated value, or 0 if \( W_x \) is missing. The notation \(|j|\) denotes the Euclidean norm of \( j \).

To solve for the optimal solution vector \( a \) that minimizes (A3), it can be shown that the vector \( (HW - \mathbf{HZ}a) \) must be orthogonal to all vectors of the form \( \mathbf{HZ}x \). This implies that \((HW - \mathbf{HZ}a)\) is orthogonal to all column vectors of \( \mathbf{HZ} \), that is,

\[
(HZ)^*T(W - ZH \mathbf{a}) = 0.
\]

Using the definition of transpose, the above equation can be rewritten as

\[
Z^*THW = Z^*TH \mathbf{a},
\]

(\text{A4})
where $\hat{H} = H^*H$. The square matrix $\hat{H}$ is a diagonal matrix with weights $a_j^2$ on the diagonal. The optimal solution vector $a$ can be expressed as

$$a = (Z^*H\bar{Z})^{-1}Z^*H\bar{W}.$$ 

Through the minimization of $(A3)$, we can solve for the time coefficients $a$ that modulate each of the climate basis functions $Z_{jk}$ for the month(s) of interest. Using $(A2)$ we then obtain pseudostress fields. Note that because of the assumptions that were made about the equatorial Pacific wind field, the task of forming complete wind pictures reduces to solving a weighted least squares problem.

A complex form of the Gram–Schmidt method (Strang 1986) is used in order to solve $(A4)$ for the optimal solution vector $a$. This technique is used as an alternative way to solve for $a$ without having to invert a matrix.

Now applying the Gram–Schmidt method to this case, we first transform the matrix $Z$ into a new matrix $Z'$ such that product of any two column vectors $Z_j'$ and $Z_k'$, which are the $j$th and $k$th columns of $Z'$, satisfies the following product rule:

$$\langle Z'_j, Z'_k \rangle = Z'_j^*H\bar{Z}_{jk}, \quad (A5)$$

where the result is equal to 1 if $j = k$ or 0 otherwise. Next, we then take the complex conjugate of both sides of Eq. $(A4)$, yielding

$$Z'^*H\bar{Z}'^*a^* = Z'^*H\bar{W}^* \quad (A6)$$

noting that $H^* = \bar{H}$, since all of the elements of $H$ are real. Recognizing that

$$(Z'^*H\bar{Z}'^*)_j = \langle Z'_j, Z'_k \rangle = \delta_{jk},$$

where the notation $X_{jk}$ refers to the element of array $X$ at position $(j, k)$, $Z'^*H\bar{Z}'^*$ is therefore the identity matrix. We then obtain the following solution for $a$ from $(A6)$ (after taking the conjugate of both sides),

$$a = Z'^*H\bar{W}.$$

So the new wind field $\hat{Y}$ expressed in terms of $Z'$ is

$$\hat{Y} = Z'a. \quad (A7)$$

Finally, to correct for any biases the analysis scheme may have introduced in $\hat{Y}$ and to preserve the amplitude of the COADS observations, the following procedure was used:

1) The ratio of the average amplitude of the COADS anomalies to the analyzed anomalies is first formed as follows:

$$\chi = \frac{\sum_{s=1}^{\infty} W_s y_s}{\sum_{s=1}^{\infty} W_s y_s}$$

where $W_s$ represents the COADS observation at position $s$, $y_s$ is the analyzed anomaly at position $s$ yielded by $(A7)$, and $\chi$ is equal to 1 if there is a COADS observation at that position $s$ or 0 otherwise.

2) The analyzed anomalies are then scaled by $\chi$, that is,

$$\hat{Y} = \chi Y.$$

$\hat{Y}$ then constitutes the new pseudostress anomalies.

**APPENDIX B**

**Detrending of New Pseudostress Fields**

Let the new pseudostress fields be represented by $\mathbf{w} = \mathbf{T}_x + i\mathbf{T}_y$, where $\mathbf{T}_x$ and $\mathbf{T}_y$ are each 2155 (spatial positions) by 720 (months) matrices, and $i = \sqrt{-1}$. The matrices $\mathbf{T}_x$ and $\mathbf{T}_y$ represent the eastward and northward components of the new pseudostress product, respectively.

Spatially averaged values are computed for each stress component over the subdomain 19°N–19°S, 142°E–88°W for January 1945 and subsequent months:

$$D_x(t) = \frac{1}{n} \sum_{\Psi} T_x(s, t),$$

$$D_y(t) = \frac{1}{n} \sum_{\Psi} T_y(s, t),$$

where $\Psi$ denotes the $x$–$y$ subdomain encompassing the region 19°N–19°S and 142°E–88°W, $s$ and $t$ denote space and time indices, $n$ denotes the number of spatial positions in $\Psi$, and $D_x(t)$ and $D_y(t)$ represent the zonal and meridional components of $w$ spatially averaged over $\Psi$. This subdomain was chosen to focus the detrending on the interior tropical Pacific, the key atmospheric forcing region for our area of interest. Note $D_x(t)$ and $D_y(t)$ are functions of time only, spanning the period January 1945 (month 181) to December 1989 (month 720).

Straight lines are least squares fit to $D_x(t)$ and $D_y(t)$, yielding two sets of slopes and $y$ intercepts. The post-1945 trend is then removed from the new pseudostress fields as shown in the following equations:

$$T'_x(s, t) = T_x(s, t) - (m_t + b_t),$$

$$T'_y(s, t) = T_y(s, t) - (m_t + b_t),$$

where $m_t$ and $b_t$ are the slope and $y$ intercept corresponding to $D_x(t)$, $m_y$ and $b_y$ are the slope and $y$ intercept corresponding to $D_y(t)$, $t$ = 181, 182, . . . , 720 (January 1945, February 1945, . . . , December 1989). The slopes for the zonal and meridional stress components are $-4.67 \times 10^{-2}$ m$^2$ s$^{-2}$ month$^{-1}$ and $3.50 \times 10^{-3}$ m$^2$ s$^{-2}$ month$^{-1}$, respectively. The $y$ intercept values for the zonal and meridional stress components are 14.66 m$^2$ s$^{-2}$ and $-1.89$ m$^2$ s$^{-2}$, respectively.
A constant value is subtracted from each of the wind fields prior to 1945 in order to allow a smooth transition between the detrended and undetrended periods of the data set. The constant value subtracted from the wind fields prior to 1945 is the same as that removed from the January 1945 fields, that is,

\[ T'_x(s, t) = T_x(s, t) - (m_x \times 181) + b_x \]

\[ T'_y(s, t) = T_y(s, t) - (m_y \times 181) + b_y, \]

where \( t = 1, 2, \ldots, 180 \) (January 1930, February 1930, \ldots, December 1944).

REFERENCES


