The Quality of Ship Observations in the Equatorial Western Pacific

Mark L. Morrissey
Oklahoma Climatological Survey
The University of Oklahoma
100 East Boyd, Suite 1210
Norman, OK 73019

Abstract

Theoretical and observational studies suggest that the Equatorial Western Pacific plays an important role in the origin and maintenance of the El Niño-Southern Oscillation phenomenon. Historical data within this critical region are sparse except for a scattering of island stations and a merchant shipping lane along 155°E. The usefulness of ship data along this track is assessed utilizing exploratory data analysis and analysis of variance. Systematic biases are revealed in the surface wind, pressure and sea surface temperature measurements. The characteristic spatial scale associated with these variables are identified using spatial correlograms. The noise level in the data is quantified using standard error estimates and signal to noise ratios associated with various time averages. It is shown that zonal wind is capable of detailing synoptic-scale variations. However, meridional wind, surface pressure and sea surface temperature are better suited for estimating lower frequency variations.

1. Introduction

The equatorial western Pacific (EWP) has been the focus of many investigations of the El Niño-Southern Oscillation (ENSO). It has been suggested that synoptic-scale fluctuations in the EWP surface wind flow are involved in the establishment and/or maintenance of the ENSO through oceanic Kelvin wave generation (Eriksen et al., 1983; Harrison and Schopf, 1984; Ramage, 1986). However, monthly averaged data has generally been used to investigate ENSO related phenomena (Wyrtki, 1975; Wyrtki and Meyers, 1976; Barnett, 1977; Goldenberg and O'Brien, 1981; Barnett, 1981; Rasmusson and Carpenter, 1982; Barnett, 1983). Harrison (1987) used monthly averaged surface wind data from 11 equatorial Pacific island stations along 175°E to obtain the anomalous surface wind fields for several ENSO events. His study clearly illustrated the development and maintenance of westerly anomalies in the EAT during ENSO episodes. However, no clear signal was found in the anomalous wind field preceding an ENSO event. As Harrison pointed out, "westerly wind bursts" are one to two week phenomenon and, thus, may not be resolved with monthly data. He also suggested that a precursor signal may exist west of 175°E.

Ship observations are the only source of historical surface data in the EWP west of 175°E (Fig. 1). These data are scant over most of the EWP except for a north-south swath of concentrated ship traffic along 155'E. The data from this track are dense enough to construct a nearly continuous synoptic-scale record and could conceivably supplement the existing island record. However, these data have seldom been used for the study of synoptic-scale phenomena. As noted by Luther et al. (1983) and Wright (1986), researchers have historically questioned the quality of ship data.

Unfortunately, only a few scattered islands and ships regularly report weather conditions in this region. Many problems have been attributed to ship data. These include:

a) variable time and space sampling
b) possible fair weather biases, and
c) uncertainty in time and space averages due to the assemblage of measurements obtained from instruments with various degrees of calibration error and variations in measurement techniques.

Data (especially ship data) are often utilized without proper assessments of data quality, thereby putting study results in question. The purpose of this paper is to summarize the work conducted by Morrissey et al. (1988) and Morrissey (1990). This work provided researchers with an understanding of the limitations and capabilities of EWP ship data to discern natural fluctuations, such as the ENSO phenomenon.

2. General Data Description

Ship data from the EWP ship track were obtained from the Comprehensive Ocean Atmosphere Data set (Woodruff et al., 1987; COADS) and the interim COADS (1957-1987), obtained from Scott Woodruff. Individual ship reports of zonal (u) and meridional (v) wind, sea surface temperature (SST) and surface pressure (p) were extracted.

The density of ship data along the EWP ship track is highly variable, ranging from fewer than 25 observations per 4°x4° box per month during the 1950s to over 350 observations per box per month in the early 1980s (Fig. 2). Variation was least from 1965 to 1979. The interim COADS, in contrast to COADS, contains many more duplicates and outliers. The large decrease in reports after 1982 suggests that many ship reports have not been included. This is now being undertaken.

Lander and Morrissey (1987) discovered a substantial number of duplicate ship reports remaining in COADS. Steurer (1986) indicates that these amount to approximately 1% of the total reports. The number of duplicates remaining in COADS were reduced using the method described in Morrissey et al. (1988).

3. Exploratory Data Analysis

a. Systematic Biases

Ships generally observe at 0000, 0600, 1200 and 1800 UTC. The reports are not evenly distributed with reporting hour with reports most frequent at 0000 UTC (1000 LST) and less frequent at 1800 UTC. Thus, monthly averages will be biased towards the daylight hours unless the diurnal and semi-diurnal cycles are first removed from the data. These cycles (Fig. 3) also vary with season.

The resolution of each variable given in COADS may differ from the actual, or “effective” resolution. This and other systematic biases may be detected through the use of high resolution frequency histograms (Figs. 4 and 5). The resolution of individual reports of wind speed appears to be approximately 1.0 m s^{-1}. The uneven distribution of reports arises from the conversion from knots to meters per second and from a mixture of two measuring procedures. Anemometers are gradually replacing the Beaufort technique which estimates the wind speed from a visual inspection of the sea state. To avoid working with mixed resolution data, the wind components should be rounded to the nearest 1.0 m s^{-1}. In addition to these problems, ship-board anemometers^{-1} are placed at various heights and, for certain wind directions, the ship's superstructure obstructs the air flow. Since ships are moving platforms, their movement should be subtracted from the wind measurements. It is unclear whether this is consistently done.

Depending on the type of compass used, wind direction is reported at various resolutions. A disproportionate amount of reports can be observed in the frequency histogram (Fig. 4) at the cardinal directions (0°, 90°, 180° and 270° especially for easterly wind. This indicates that ship winds along 155°E are biased towards both pure zonal or pure meridional directions. The precise reason for this is unknown.

SST is reported in the COADS at a resolution of a tenth of a degree Celsius. However, as noted in the COADS Users Manual, Slutz et al. (1985), decreased resolution occurred as a result of the conversion from Fahrenheit to Celsius. This is clearly evident in the frequency histogram for SST (Fig. 5), in which there are a predominance of reports at quarter and half degrees. Thus, SST should be rounded to the nearest half degree. Another source of systematic error in SST arises from a mixture of two measurement techniques. Measurements made by the engine intake procedure are approximately 0.5°C higher than those recorded using a bucket drawn from the ocean surface (Ramage, 1984). Unfortunately, COADS indicators identifying the measurement procedure are unreliable (Slutz et al., 1985).

The effective resolution for surface pressure also appears less than the standard resolution (Fig. 5). Errors originating from the conversion of inches of mercury to millibars may be responsible for this. Rounding to the closest 0.25 mb appears to alleviate this problem.

b. Observational Error

To assess how observational error affects the quality of ship data and to examine the spatial characteristics of the various air-sea variables in the EWP, a correlation coefficient was computed for pairs of ship reports at various separation distances. The ship pairs were grouped at 10 km separations from 10 km to 240 km and correlation coefficients were then computed from these grouped pairs. The calculations were performed for each variable. In COADS, ship positions are given to the nearest 10th of a degree latitude/longitude, so two ships reporting the identical location and time could be separated up to 15 km. At the smallest separation distance (0-11 km), the correlation indicates how much of the variance of one observation is explained by another. Assuming minor gradients from 0 to 11 km, this provides a rough estimate of the observational error inherent in the measurement of each variable since measurements from two collocated and properly calibrated instruments should be the same. The characteristic length scale of the various air-sea variables can also be assessed from the variation of correlation coefficient with separation distance.

The correlations indicate that observational errors are relatively small for u and v (Fig. 6).

Between 80% and 90% of the variance in these variables measured at one ship is explained by measurements taken at a collocated ship. In contrast, SST and p measured at one ship explain only 60%-70% of the variance of SST and p taken at a collocated ship.

For SST, the correlation coefficient of 0.77 at the smallest separation distance indicates that the relatively low correlation coefficients for SST beyond 30 km were not due to observational error alone, but to small-scale variability. Perhaps this reflects the pattern of rainfall and/or solar radiation due to mesoscale cloudiness fluctuations.

c. Standard Error and Signal to Noise Calculations

Whether an El Niño signal can be detected depends on the ratio of the signal strength to the noise level in the data. To determine this, averaged standard error estimates associated with various time averaging intervals (1, 3, 5, 14, 30 and 90 day) were computed for a 4° x 4° latitude-longitude box centered on the equator and 155°E. These estimates were then compared to the standard deviations of these intervals (hereafter, the 1, 3, 5, 14 day and 30, 90 day averaging intervals will be referred to as the high and low resolution averages, respectively).

Signal to noise ratios were also calculated. However, as noted by Trenberth (1984a, 1984b) standard signal to noise computations using meteorological data generally contain large errors due to: 1) the general inseparability of signal from noise, 2) the non-independence of the data, 3) the correlation of signal and noise in samples of finite size, 4) large uncertainties in the computed signal to noise ratio and 5) the likely non-stationarity of all moments of the data. Because of these reasons, the signal to noise ratios alone are not conclusive, but were deemed useful for comparative purposes. The degree of inseparability of signal from noise within meteorological data (Hayashi, 1982) may also lead to overestimates of the standard error (Trenberth, 1984b). This constrains these estimates to be interpreted as maximum values. Thus, substantial alterations to the classical standard error formula were required.

Coherence within meteorological data effectively reduces the information content of the data set. Thus, computing statistical quantities, such as the mean standard error of a time average or the correlation coefficient, demands that the number of degrees of freedom used in the test be reduced (Waldo-Lewis and McIntosh, 1952). Allowances were made for missing observations and their distribution with time using the method derived by Parker (1984). The signal to noise ratios were calculated following Trenberth (1984b).

For this analysis, all regularly varying cycles (diurnal, semi-diurnal, semi-annual and annual) have been removed from the data. To reduce the contribution of spatial coherence to our calculations each element was initially spatially averaged into six hour increments utilizing data from the 4° x 4° box.

Since the high resolution averages are not generally used to assess a trend in the data, this variation has been removed for these averages to reduce its effect on the error variance.

The standard error (SE) of a coherent time series including missing values was computed following Morrissey (1990). We have assumed a signal to be present in the data and have, thus, interpreted the F ratio minus one (F-1) as the signal to noise ratio.

The mean standard errors representing the high resolution averages for u, v, p and SST are approximately 1.75 m s^{-1}, 1.3 m s^{-1}, 1.2 mb and 0.5°C, respectively (Figs. 7 and 8). The nearly equal standard error and standard deviation values in SST or v would make synoptic-scale signals hard to detect. However, since the standard error estimates may be overestimated the signal to noise ratios may be underestimated.

For surface pressure, the signal to noise ratios are quite large for the high resolution averages (> 0.5). However, this can be very misleading. An individual ship stays within the 4° x 4° box for approximately 1 day and its observation is often the sole contributor of a day's average. Thus, the large ship-board barometric calibration error reported by Morrissey et al. (1988) contributes to a large interdiurnal rather than diurnal variation in pressure; this causes the synoptic-scale signal to be substantially overestimated. This may also be true for SST as well.

The strong coherence in windspeed at small separation distances (Fig. 6) suggests small calibration error. This indicates that relatively large signal to noise ratios for zonal wind are valid. Since the computed standard error for zonal wind is less than 1.75 m s^{-1}, synoptic-scale fluctuations greater than 3.5 m s^{-1} can be resolved.

During a westerly burst, zonal wind can reach 10 to 15 m s^{-1} (Luther et al., 1983, McPhaden et al., 1987), and in a couple of days, can change by about 15 m s^{-1}. Thus, time and space averaged ship zonal wind probably reflect these fluctuations. The other variables (v, p, SST) may be useful in detecting synoptic-scale fluctuations greater than double their respective mean standard error statistic (i.e. only large fluctuations are detectable). Ship-board SST and p measurements in the EWP appear to be best suited for measuring interannual fluctuations.

Data density variations result in unstable standard error estimates for 1-day averages. However, 3-day averages are considerably more stable, and thus, should be utilized to detect westerly wind burst characteristics.

4. Summary

The results of the exploratory data analysis suggest that ship data obtained from the COADS and Interim COADS for the EWP ship track are quite useful from 1964 to 1987. Prior to 1964, sparse data cause unacceptably high noise levels (Morrissey, 1990).

Several sources of systematic error were revealed. The large difference between the number of nighttime and daytime reports and the rather large semi-diurnal and diurnal cycles in the data suggest that time averages may be significantly biased unless the diurnal cycles are first removed.

Care must be taken in removing. these cycles as the amplitude and phase may be a function of time. Mixed resolution resulted from variations in recording methods and from conversions to standard units. Decreased resolution also occurred from the bias of surface wind towards the cardinal directions. the precise reason for this is unknown. Individual reports of u, v, SST and pressure should be rounded to the nearest 1.0 m s^{-1}, 0.5° C and 0.25 mb, respectively

Results from the analysis of variance indicate that ship-measured zonal wind within the EWP track can be utilized to assess synoptic-scale fluctuations, as the mean standard error suggests that 3- day fluctuations of greater than 4.0 m s^{-1} can, on the average, be resolved. Very little signal in the meridional wind remains after the removal of the annual cycle and, thus, it may be of little use in describing synoptic-scale fluctuations. The degree of uncertainty in these averages resulting from realized and unrealized systematic error is unknown. However, the spatial correlograms indicate that ship wind within the EWP track is spatially more coherent than EWP island wind. In addition, a correlation coefficient of 0.95 (Fig. 6; 95% confidence interval equaled 0.02) was computed from all pairs of zonal wind reports from ships separated by less than 11 kin within the EWP ship track. These results suggest that westerly wind burst statistics may be detailed using zonal wind. It is recommended that 3-day averages of zonal wind be utilized as data density variations render the standard error estimates associated with 1-day averages unstable.

The rather small gradients in SST and p in the EWP, their relatively large standard error estimates, and the uncertainty in the values due to systematic error, suggest that these variables are best suited for estimating low frequency fluctuations. For these variables, synoptic-scale fluctuations greater than twice the respective standard error estimates may be realized.

Overall ship data obtained from the COADS for the EWP ship track appears to be quite useful from 1964 to 1985. The effective recorded length (1964-1987) will allow the investigation of several major ENSO events (1965-66, 1972-73, 1976-77 and 1982-83). Three moderate events also occurred during this period (1963, 1969 and 1987; Rasmusson and Carpenter, 1982; Harrison, 1987; Lander, 1989). Thus, we believe that ship data along the EWP track will provide a useful supplement to the existing EWP island record.

References

Barnett, T. P., 1977: An attempt to verify some theories of El Niño. *Journal of Physical Oceanographer*, **7**, 633-647.

Barnett, T.P., 1981: Statistical relations between ocean/atmospheric fluctuation in the tropical Pacific. *Journal of Physical Oceanographer*, **11**, 1043-1058.

—, 1983: Interaction of the monsoon and Pacific trade wind systems at interannual time scales. Part 1: The equatorial band. *Monthly Weather Review*, **111**, 756-773.

Eriksen, C., M. Blumenthal, S. Hayes and P. Ripa, 1983: Wind-generated equatorial Kelvin waves observed across the Pacific Ocean. *Journal of Physical Oceanographer*, 13, 1622-1640.

Goldenberg, S.B. and J.J. O'Brien, 1981: Time and space variability of tropical Pacific wind stress. *Monthly Weather Review*, **109**, 1190-1207.

Harrison, D.E., 1987: Monthly mean island surface winds in the central
tropical Pacific and El Niño events. Monthly Weather Review, 115, 3133-3145.

— and P.S. Schopf, 1984: Kelvin-wave-induced anomalous advection and the onset of surface warming in El Niño events. *Monthly Weather Review*, **112**, 923-933.

Hayashi, Y., 1982: Confidence intervals of a climate signal. *Journal of Atmospheric Science*, **39**, 1895-1905.

Lander, M. A. and M. L. Morrissey, 1987: Unexpected duplicate ship reports in the Comprehensive Ocean-Atmosphere Data Set (COADS). *Tropical Ocean-Atmosphere Newsletter*, **38**, 13-14.

—, 1989: A comparative analysis of the 1987 ENSO event. Tropical Ocean-Atmosphere Newsletter, 49, 3-6.

Luther, D.S., D.E. Harrison and R.A. Knox, 1983: Zonal winds in the central equatorial Pacific and El Niño. *Science*, **222**, 327-330.

McPhaden, M.J., H.P. Freitag, S. Hayes, B.A. Taft, Z. Chen and K. Wyrtki, 1987: *The response of the equatorial Pacific ocean to a westerly wind burst in May, 1986*. Manuscript, No. 977, NOAA/PMEL. Seattle, Washington.

Morrissey, M.L., M.A. Lander and J.A. Maliekal, 1988: A preliminary evaluation of ship data in the equatorial western Pacific. *Journal of Oceanographic and Atmospheric Technology*, **5**, 251-258.

— 1990: An evaluation of ship data in the equatorial western Pacific. *Journal of Climate*, **3**, 99-112.

Parker, D.E., 1984: The statistical effects of incomplete sampling of coherent data series. *Journal of Climate*, **4**, 445-449.

Ramage, C.S., 1986: El Niño. *Scientific American*, **254**, 76-83.

— 1984: Can shipboard measurements reveal secular changes in tropical air-sea heat fluxes? *Journal of Climate and Applied Meteorology*, **23**, 187-193.

Rasmussen, E.M. and T.H. Carpenter, 1982: Variations in tropical sea surface temperature and surface wind fields associated with the Southern Oscillation/El Niño. *Monthly Weather Review*, **110**, 354-384.

Slutz, R.J., S.J. Lubker, J.D. Hiscox, S.D. Woodruff, R.L. Jenne, D.H. Joseph, P.M. Steurer and J.D. Elms, 1985: *Comprehensive Ocean-Atmosphere Data Set; Release 1*. NOAA Environmental Research Laboratories, Climate Research Program, Boulder, CO, 268 pp.

Steurer, P.M., 1987: Reply: Unexpected duplicate ship reports in the Comprehensive Ocean Atmosphere Data Set (COADS). *Tropical Ocean-Atmosphere Newsletter*, **40**, 12-14.

Trenberth, K.E., 1984a: Some effects of finite sample size and persistence on meteorological statistics. Part I: Autocorrelations. *Monthly Weather Review*, **112**, 2359-2368.

— 9 1984b: Some effects of finite sample size and persistence on meteorological statistics. Part 11: Potential predictability, *Monthly Weather Review*, **112**, 2369-2379.

Waldo-Lewis, R.P. and D.H. McIntosh, 1952: Some effects of the coherence of meteorological time- series. *Meteorology Magazine*, **81**, 242-244.

Woodruff, S.D., R.J. Slutz, R.L. Jenne and P.M. Steurer, 1987: A comprehensive ocean atmosphere data set. *Bulletin of the American Meteorology Society*, **68**, 1239-1250.

Wright, P.B., 1986: Problems in the use of ship observations for the study of interdecadal climate changes. *Monthly Weather Review*, **114**, 1028-1034.

Wyrtki, K., 1975: El Niño—The dynamic response of the equatorial Pacific Ocean to atmospheric forcing. *Journal of Physical Oceanographer*, **5**, 572-584.

— and G. Meyers, 1976: The trade wind field over the Pacific ocean. *Journal of Applied Meteorology*, **15**, 698-704.

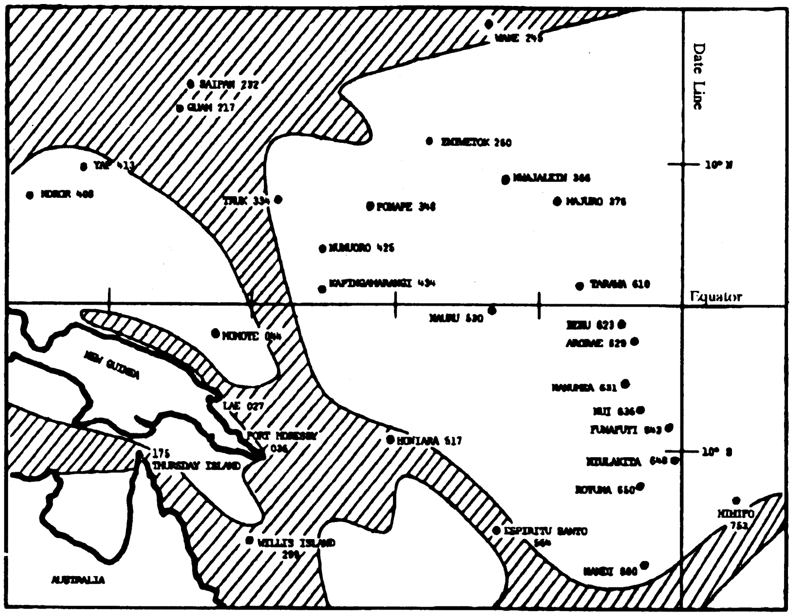

Figure 1. Distribution of COADS ships and location of island stations in the EWP. Shading indicates ship density greater than 10 per month per 2°x2° lat/long box after 1945.

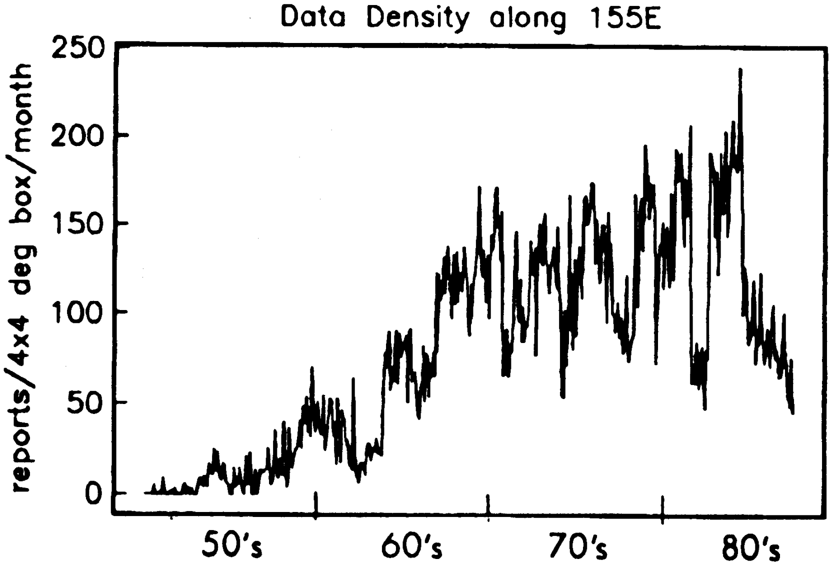

Figure 2. Number of ship reports per 4°x4° lat/long box along the EWP ship track

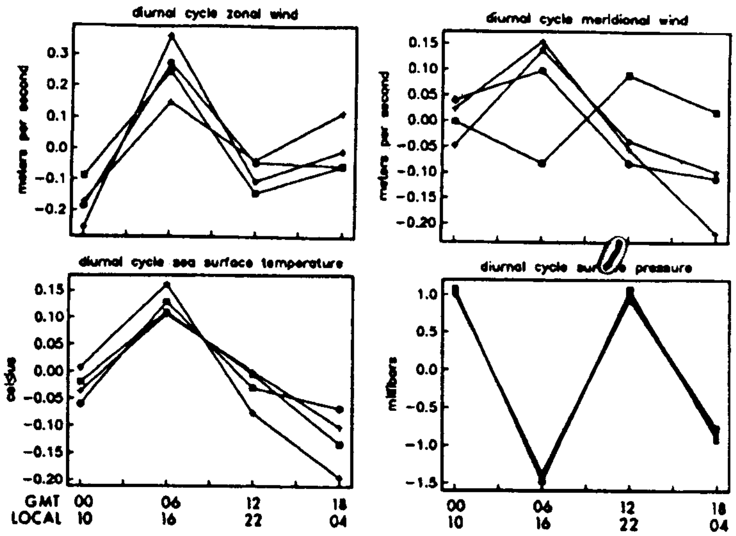

Figure 3. Diurnal variation in the indicated variables. The seasons are denoted by a square: Jan-Mar, a circle: Apr-Jun, a plus: Jul-Sep and a diamond: Oct-Dec.

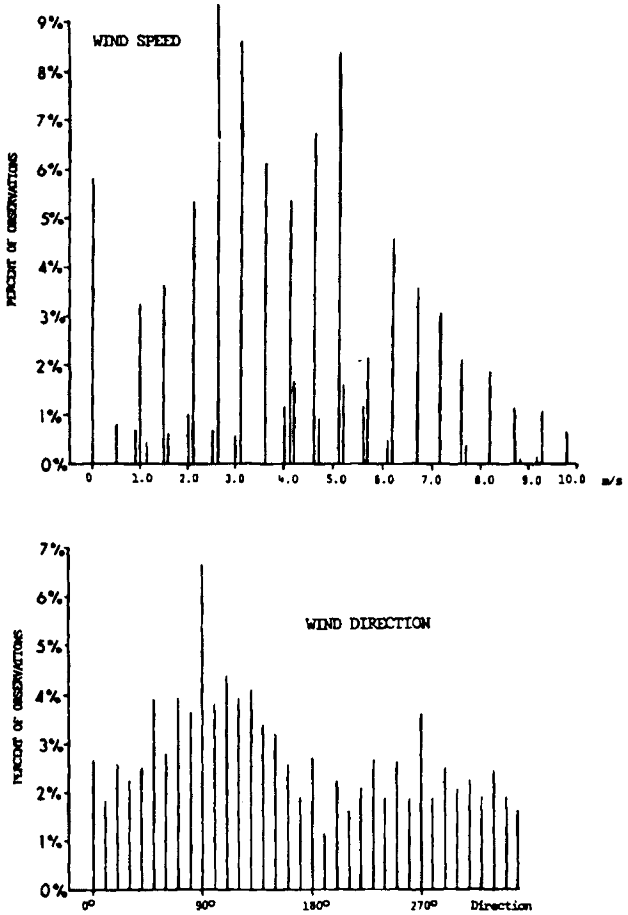

Figure 4. High resolution frequency histograms for wind speed and direction.

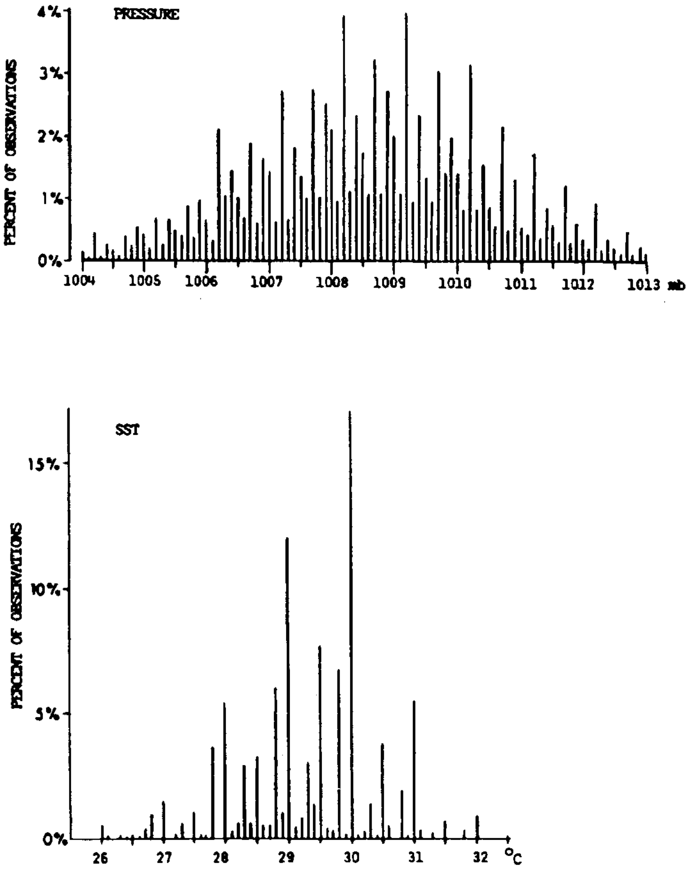

Figure 5. As in Fig. 4, but for Surface pressure and SST.

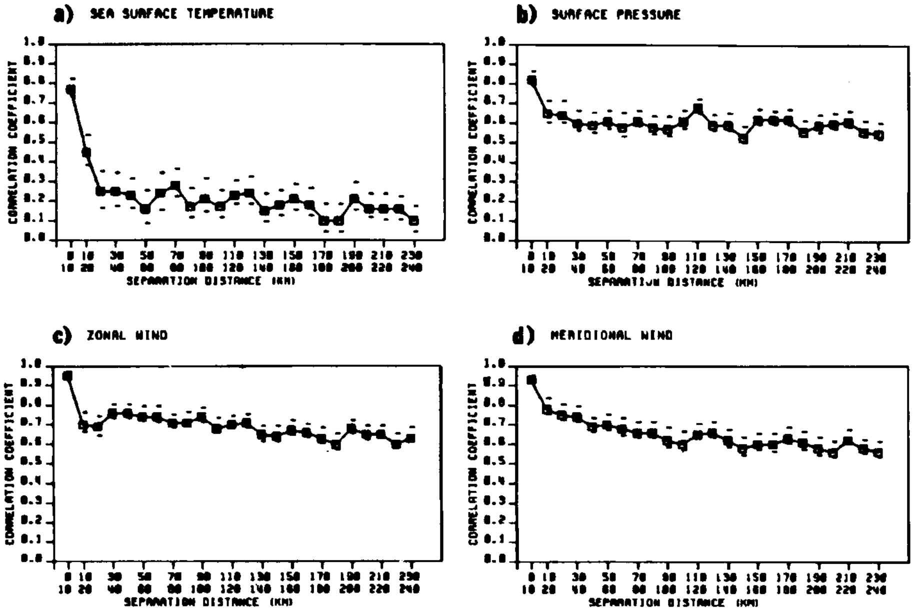

Figure 6. Correlation vs. ship separation distance.

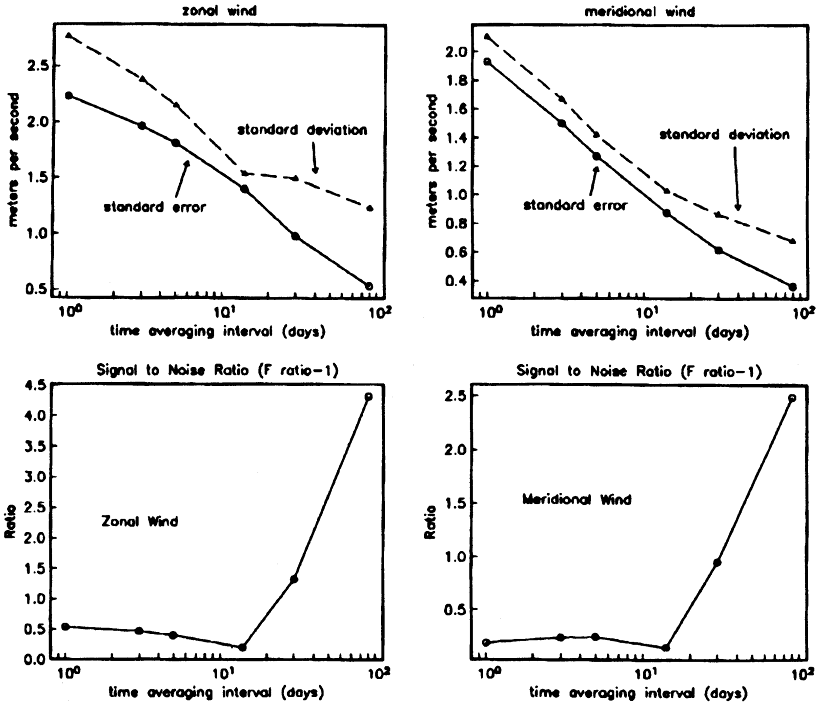

Figure 7. Comparison of standard error and standard deviation for various time averages of data within a 4°x4° box centered on the equator and the EWP ship track. Also shown are the signal to noise ratios.

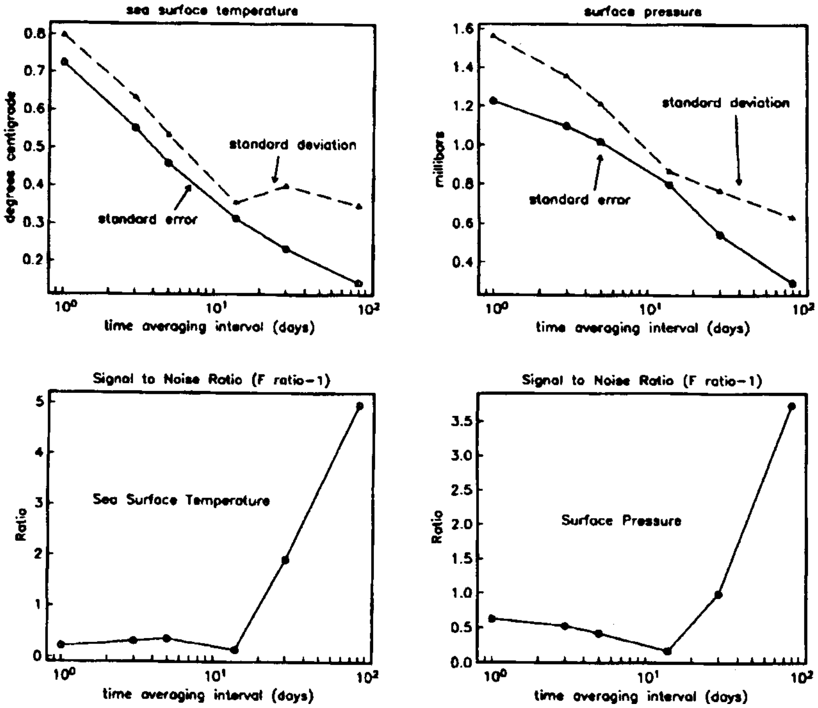

Figure 8. As in Fig. 7, but for SST and surface pressure.