Predictability of Monthly North Pacific Sea Level Pressure from Monthly Sea Surface Temperature for the Period 1933–1976

ANTHONY J. BROCCOLI AND ROBERT P. HARNACK

Department of Meteorology and Physical Oceanography, Cook College and New Jersey Agricultural Experiment Station, Rutgers, The State University of New Jersey, New Brunswick 08903

(Manuscript received 8 September 1980, in final form 1 June 1981)

ABSTRACT

Statistical models were developed to specify and predict the mean monthly sea level pressure (SLP) distribution over the central and eastern North Pacific Ocean from the mean monthly sea surface temperature (SST) distribution for the same area. These models were derived from data for the period 1947–71, with data from two additional periods (1933–41 and 1972–76) retained for independent testing.

The earlier period SST data is taken from a new data set compiled at the National Climatic Center and processed for use in examining large-scale air-sea interactions. This procedure is described.

Empirical orthogonal function (EOF) analysis was used to represent each field (SST and SLP) by a small number of composite variables. Regression analysis was then used in which SST EOF amplitudes were the predictors and SLP EOF amplitudes were the predictands. The analyses were stratified by month, with lags from 0–3 months considered. Of the 84 models developed, 18 were statistically significant at the 10% level. The number of significant relationships was found to decrease with increasing lag, being greatest for SST contemporaneous with SLP. All statistically significant models involved SST's from the period June–January.

Each of the significant models was tested on the independent samples, using the reduction of error (RE) statistic as the measure of skill. An adjustment was made to the 1933–41 SST data to remove a systematic bias, and the RE scores recomputed. Using the adjusted data, RE scores for the 1933–41 period improved, with 10 of 18 models demonstrating skill overall. Most of the skill, especially at longer lags, was associated with models using late autumn and early winter SST as predictors.

Possible reasons for the seasonal distribution of the skillful relationships are discussed.

1. Introduction

The large heat capacity and volume of the oceans cause them to play an important role in the climate system. Large quantities of energy are transferred across the air-sea interface in the form of sensible and latent heat, and variations in these exchanges are believed by many to have an effect on short-period climate fluctuations on monthly and seasonal time scales. Sea surface temperature anomalies (SSTA's) which are associated with variations in energy fluxes across the air-sea interface play an important role in such fluctuations.

Various studies have considered the existence of SSTA's and their relationship to atmospheric circulation. In numerous case studies, Namias (1959, 1966, 1969, 1971) demonstrated an apparent relationship between the tropospheric circulation over the North Pacific and underlying SSTA's. Namias and Born (1972) examined some contemporaneous and lag correlations between fields of sea surface temperature (SST) and sea level pressure (SLP), uncovering some statistically significant relationships. Harnack and Broccoli (1979) found significant linear correlations between anomalous SST gradients and anomalous flow at the 700 mb level on monthly and seasonal time scales.

In addition, long-range forecasters have considered the relatively long "memory" of the ocean and the relationship between SSTA's and tropospheric circulation in the formulation of monthly or seasonal forecasts. Ratcliffe (1970) described the use of North Atlantic SSTA's as input for monthly and seasonal forecasts by the British Meteorological Office, and Farmer (1973) has considered the usefulness of North Pacific SST data for forecasting seasonal circulation. Harnack (1979) has developed statistical models in which winter temperatures in the United States are predicted using North Pacific, tropical Pacific, and North Atlantic SST's as predictors. Significant skill has been demonstrated over several years of independent forecasts made with some of these models.

A series of studies by Davis (1976, 1978) examined the predictability of North Pacific SST and SLP. In the first of these, monthly fields of SSTAs and SLP anomalies were expressed in terms of empirical orthogonal functions (EOF's) and then related using

0027-0644/81/102107-11$06.75

linear statistical estimators. Davis found that while SST was skillful in specifying contemporaneous and previous SLP, it was not skillful in predicting future SST. In his second study, Davis performed a similar analysis using data stratified according to 3-month seasons. The seasonal stratification is significant since the nature of air-sea interactions changes with season due to variations in mixed-layer depth and atmospheric stability. Using the stratified data, Davis found instances in which fall and winter SLP could be predicted from summer and fall SST. He also found that forecasts of winter SLP from November SST are skillful when tested on independent data, and explain 20% of the SLP variance.

Thus, substantial evidence exists for a relationship between North Pacific SST and the overlying tropospheric circulation, with a relationship to the downstream circulation likely through various well-established teleconnections. This implies that information about temperature and precipitation over the United States and the storm climate of the Pacific Ocean may be obtainable up to several months in advance by considering the Pacific SST distribution. The economic implications for agriculture, energy consumption, transportation, and offshore operations make further study of the effects of the SST distribution on tropospheric circulation worthwhile.

2. Objectives

Most investigations of large-scale air-sea interactions in the North Pacific have used only data from the post-World War II period, owing to the better quality and easy availability of SST data during this time. In almost all cases, these are data provided by the Scripps Institution of Oceanography in the form of monthly temperatures at 5° latitude-longitude intersection points. These data originate as averages of individual ship intake observations over 2° latitude-longitude squares. The averages for these 2° × 2° squares were subjectively analyzed and contoured in map form, then digitized to a 5° × 5° grid (Namias and Born, 1972).

This study proposes to extend the use of existing SST data backward in time, and to examine the predictability of *monthly* mean SLP in the North Pacific region. This will be done in two ways: first, by deriving SST-SLP relationships for the period 1947–71 using regression models stratified by month, and second, by testing these relationships on independent data for two periods—the 1972–76 period and the 1933–41 period.

This analysis is similar to those performed by Davis (1976, 1978) and discussed in the previous section. The differences between this analysis and those of Davis involve the time-scale and seasonal stratification. In this study, monthly averages of SST and SLP were used to develop regression models, with the regression analyses stratified by month.

This differs from Davis's earlier study because of the use of seasonally stratified data. As was stated previously, this is necessary since the nature of the atmosphere-ocean interaction changes with season due to variations in oceanic mixed-layer depth and atmospheric stability. It is likely that a failure to stratify the data by season could lead one to conclude that no SST-SLP relationships exist, since the seasonal nature of atmosphere-ocean exchange processes precludes combining cases from different months.

In addition, this study differs from Davis's later study by examining SST-SLP relationships on a monthly rather than a seasonal time scale. Evidence is accumulating that climatic fluctuations on a monthly time scale are of a distinctly different nature than climatic fluctuations on a seasonal time scale (Harnack, 1979; Harnack and Broccoli, 1979). This is further implied by differences between National Weather Service 30-day forecasts and 90-day forecasts with regard to forecast methodology and verification scores (Harnack, 1981).

Finally, previous studies (including those of Davis) used a recent period for independent testing of derived relationships. This study uses an additional period (1933–1941) to further test the relationships obtained. This period has not been examined previously with regard to large-scale air-sea interactions.

3. Data sources and analysis

For the preparation of the North Pacific SST data, two sources were used. SST data for the period 1947–76 were provided by the Scripps Institution of Oceanography. This source, which will be called the Scripps data base in future references, was described in Section 2. For the pre-1947 period, SST data were obtained on magnetic tape from the National Climatic Center. These consisted of monthly averages of individual ship bucket observations in 10° × 10° latitude-longitude squares. This source will be called the "historical" data base in future references. The following information was also provided for each month for each 10° × 10° square: 1) the number of individual observations, 2) the mean latitude and mean longitude of the observations, 3) the number of 1° × 1° latitude-longitude squares with at least one observation, 4) the mean day of the observations, and 5) the number of days with at least one observation.

This additional information was used to assess the representativeness of the SST data and to screen out those monthly averages judged to be unrepresentative. Data for individual months for any

10° × 10° square were dropped if any of the following criteria were not met: 1) at least a total of 10 observations; 2) the centroid of the observations, as defined by the mean latitude and mean longitude, within a 4° × 4° latitude-longitude square centered at the middle of the 10° × 10° square; 3) at least six 1° × 1° latitude-longitude squares with one or more observations; 4) the mean day of the month of the observations within an 11-day period centered on the 15th day of the month; and 5) at least five days with one or more observations.

Through the requirement that all of these criteria be met, several sources of unrepresentativeness and observational error were limited. Requiring a minimum total number of observations decreased the impact of individual observational errors. Checking the location of the centroid of the observations and the number of 1° × 1° squares with observations helped ensure that the observations were centrally located or evenly distributed in the 10° × 10° square, while checking the number of days with observations and the mean day of the month of the observations helped ensure that the observations were centrally located or evenly distributed within a month.

For a 10° × 10° latitude-longitude square that passed these quality checks for an individual month, the SST value was attributed to the center of the square. Data points chosen for use in the analysis were located in an area of the east central North Pacific bounded by 170°E–130°W and 20–50°N. This is the primary region of the previous investigations which were cited in Section 1.

Fig. 1 shows the average number of 10° × 10° squares per month for which data were both available and of sufficient quality to pass the aforementioned requirements for each year in the period 1861–1946. The maximum possible number of 10° × 10° squares available for each month was 18. An inspection of this figure shows a substantial increase in the average frequency of observations (as reflected in the number of 10° × 10° squares) in 1933 and the sharp decrease after 1941. As a result, only the years from 1933 through 1941 were chosen to extend the SST observations, and only the thirteen 10° × 10° boxes with the highest frequency of observations were used. These data points are shown in Fig. 2.

For the 1947–76 period, the Scripps data base values for grid points along the border and at the center of a 10° × 10° square were averaged to produce a value attributable to the center of that square. This made the Scripps data consistent in form with the data from the earlier period. In order to make some statement about the compatibility of the two data sources, SST's were obtained for the period 1947–60 from the "historical" data base using the same procedures that were used in obtaining the 1933–41 SST data. These data were then compared with contemporaneous data obtained from the Scripps data base. The bias, root-mean-square (rms) difference and mean absolute differ-

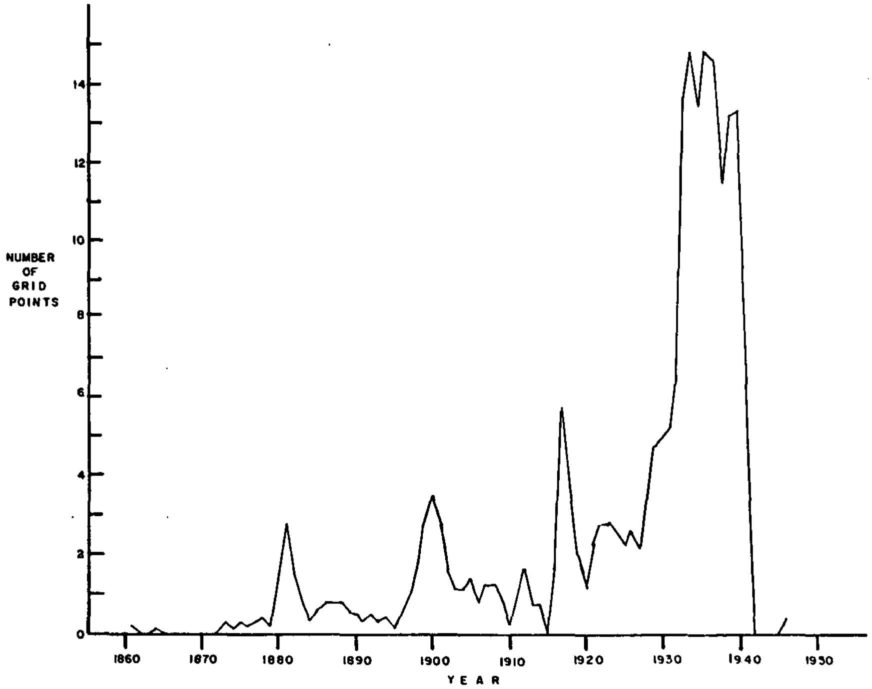

Fig. 1. Average number of grid points with data available as a function of year for the period 1861–1946.

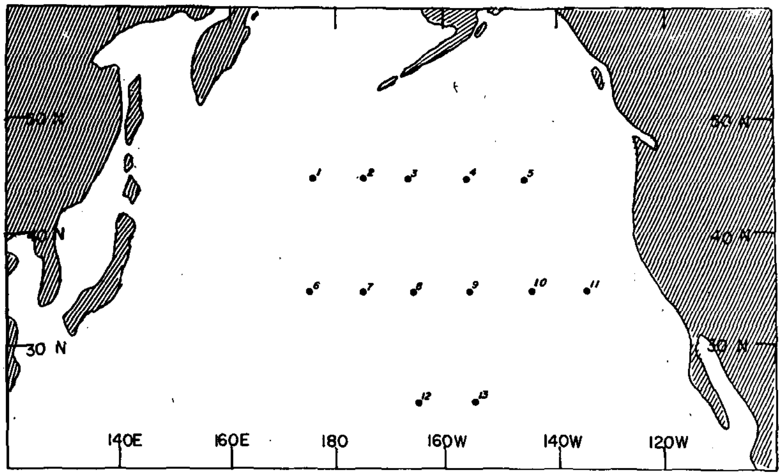

FIG. 2. Sea surface temperature grid points used in the analysis.

ence were computed for each grid point, and are given in Table 1.

A pronounced negative bias can be noted at almost all grid points, meaning that the SST's from the "historical" data base tend to be colder than those from the Scripps data base. The bias is largest in the northwest section of the area being studied, decreasing to the east and to the south.

A probable cause for a small amount of this bias is the difference between a bucket observation of sea surface temperature and an intake measurement. Since all of the "historical" SST observations were bucket observations, the difference in the manner of observation could produce a systematic bias. Roll (1965, pp. 30–33) has noted that differences between bucket and intake temperatures are generally small, exceeding 1.0°C in only 6% of cases. While there is little doubt that the difference in the method of observation is partially responsible for the observed bias, it is unlikely that this difference alone can account for the size of the biases observed.

TABLE 1. Comparison statistics [bias, root-mean-square difference (rmsd), and mean absolute difference (mad)] for "historical" data base sea surface temperatures versus Scripps data base sea surface temperatures during the 1947–60 period. (Refer to Fig. 2 for grid point numbers; units are °C).

Grid point	Bias	rmsd	mad
1	−2.12	2.36	2.12
2	−1.50	1.91	1.62
3	−1.38	1.57	1.40
4	−1.17	1.59	1.44
5	−0.89	1.23	1.02
6	−0.97	1.52	1.19
7	−0.83	1.29	1.05
8	−0.62	1.12	0.88
9	−0.44	1.21	0.89
10	−0.10	0.93	0.77
11	−0.31	0.88	0.67
12	−0.02	0.59	0.47
13	0.08	0.52	0.42

A more significant cause of the bias is the difference in the manner in which the SST values were arrived at in the two data bases. SST values from the Scripps data base closely resemble area-weighted values, since the grid points from which they originally come are evenly distributed within each 10° × 10° square, while the "historical" data base values are the means of ship observations in raw form and are not area-weighted. Under ordinary circumstances with observations randomly distributed, this would not produce a bias. However, SST observations do not occur randomly; they are found most frequently along shipping lanes. Since the center of each 10° × 10° square may not lie along a shipping lane, this could well be the major cause of the observed bias.

As a check of this hypothesis, a "sampling centroid anomaly" was computed for each 10° × 10° square for each month. The sampling centroid anomaly was defined as the vector difference between the centroid of the SST observations and the geometric center of the 10° × 10° square. This vector was then decomposed into north-south and east-west components. SST anomalies also were computed for each 10° × 10° square for each month. The SST anomalies were then cross correlated with each component of the sampling centroid anomaly. The cross-correlation analyses were stratified according to 10° × 10° square and month.

The correlations between the SST anomalies and the east-west component of the sampling centroid anomaly were insignificant. Correlations between SST anomalies and the north-south component of the sampling centroid anomaly were generally significant, with most correlation coefficients ranging from −0.4 to −0.9 for sample sizes of fifteen to thirty cases. The negative correlation implies that negative SST anomalies are associated with sampling centroids which are north of the center of a 10° × 10° square. This is to be expected since SST tends to decrease with latitude in the North Pacific.

Thus, it seems reasonable to attribute much of the bias to systematic displacements of the sampling centroid. While the differences between data bases are of approximately the same magnitude as the standard deviation of monthly SST in the North Pacific, it was originally believed that the data analysis procedure used, empirical orthogonal function analysis, would reduce the impact of these differences. This topic will be discussed more thoroughly in later sections.

Monthly sea level pressure data for both the 1933–41 period and the 1947–76 period were obtained from a data set prepared by the National Center for Atmospheric Research and corrected along the lines described by Trenberth and Paolino (1980). The data used in this study consist of monthly means for 21 grid points at 10° latitude-longitude intersections between 170°E–130°W and 30–50°N, as shown in Fig. 3.

4. Methodology for relating SST and SLP

a. EOF analysis

Empirical orthogonal function (EOF) analysis was used to reduce the number of original variables (21 SLP variables and 13 SST variables) to a smaller number of variables which explain a significant proportion of the variance of the original variables. Reducing the number of variables is desirable in a linear statistical model since it decreases the degree to which "artificial predictability" affects the derived level of forecast skill. Davis (1976) has discussed the problem of artificial predictability in which statistically significant relationships may appear to exist due to the limited number of cases included in the dependent set, which may cause the simple correlations to be inaccurate estimates of the population correlations. Davis has shown that this, on the average, will cause skill to be overestimated.

In this analysis, EOF's were derived from a covariance matrix which was calculated using data from the 1947–71 period. The data were stratified by month; thus a covariance matrix for SST variables and a covariance matrix for SLP variables were computed for each month type. An EOF analysis was performed on each of these matrices, so that a total of 24 analyses were performed.

The number of EOF's retained for further analysis was determined by subjecting them to a test of statistical significance proposed by Preisendorfer and Barnett (1977). Only those EOF's statistically significant at the 2.5% level were retained. Table 2 shows the number of SLP and SST components (EOF's) retained for each month, along with the total variance accounted for by those components.

In general, only 1–3 EOF's were statistically significant for both the SLP and SST fields. Most months had only two SST components that were significant, while the predominant number of significant SLP components was one. The statistically significant EOF's explained anywhere from 55 to 91% of the variance of the original variables. Thus effective data reduction was accomplished, with the thirty-four original variables for each month being reduced to between two and five composite variables.

b. Regression analysis

The amplitudes of the EOF's retained (those shown in Table 2) were used in regression analyses, in which SLP EOF amplitudes were the pre-

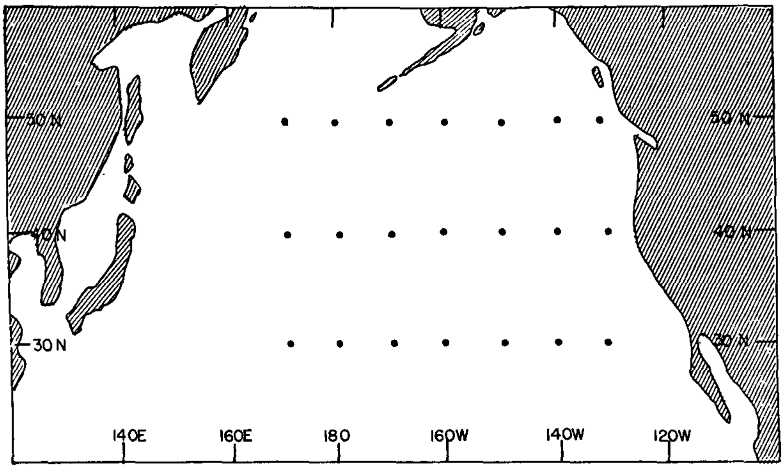

FIG. 3. Sea level pressure grid points used in the analysis.

TABLE 2. Number of statistically significant EOFs and variance explained by each EOF component for each month.

	SST					SLP				
		Variance					Variance			
Month	Number	EOF1	EOF2	EOF3	Total	Number	EOF1	EOF2	EOF3	Total
January	1	65.3			65.3	3	53.8	18.5	17.7	89.9
February	1	55.7			55.7	1	67.2			67.2
March	2	50.5	22.3		72.9	1	54.6			54.6
April	2	46.5	26.1		72.6	1	54.6			54.6
May	3	41.0	23.5	17.8	82.3	1	54.7			54.7
June	2	36.1	24.2		60.3	2	44.9	25.0		69.9
July	2	33.8	25.1		58.9	3	34.1	22.3	17.1	73.5
August	2	38.4	24.5		62.9	2	45.5	23.7		69.3
September	2	38.4	30.2		68.7	1	58.5			58.5
October	2	41.3	29.3		70.5	2	50.7	25.5		76.2
November	2	43.7	33.1		76.9	1	57.2			57.2
December	2	50.0	25.4		75.3	3	56.4	18.9	16.1	91.3

dictands and SST EOF amplitudes were used as predictors. Only data from the period 1947–71 were included in the data set used in the regression analyses. The remaining years (1933–41 and 1972–76) were retained for testing the regression equations on an independent data set.

For a given month and lag, a regression analysis was performed for each retained SLP component, using all of the retained SST component amplitudes from the appropriate month as predictors. Lags of 0, 1, 2 and 3 months were considered, with SST concurrent with or leading the SLP. Thus, for example, June SLP predictability was tested at a 1-month lag by deriving a regression equation for each of the two significant SLP components. Both equations have three predictors—the significant SST component amplitudes for May.

TABLE 3. Statistically significant regression relationships (at the 10% level) with the explained variance for each relationship.

Predictor(s) (number of EOF's)	Predictand (EOF number)	Lag (in months)	R-Squared
January (1)	January (1)	0	0.478
June (2)	June (1)	0	0.307
June (2)	June (2)	0	0.208
July (2)	July (1)	0	0.316
July (2)	August (1)	1	0.257
August (2)	August (1)	0	0.405
August (2)	August (2)	0	0.197
August (2)	October (1)	2	0.295
August (2)	November (1)	3	0.234
September (2)	November (1)	2	0.195
October (2)	October (2)	0	0.254
October (2)	December (2)	2	0.237
November (2)	December (1)	1	0.313
December (2)	December (1)	0	0.323
December (2)	December (3)	0	0.189
October (2)	January (1)	3	0.289
November (2)	January (1)	2	0.291
December (2)	January (1)	1	0.237

Since a regression equation was derived for each of 21 SLP components for each of the four lags, a total of 84 equations resulted. Of these, 15 were statistically significant at the 10% level, using the standard *F*-test for simple or multiple regression, and these are shown in Table 3 along with the explained variance (*R*-squared) for each. The 10% significance level was used to make the probability of rejecting a meaningful relationship (i.e., a type II error) lower than would be the case if the more commonly used 5% level were used. The increased risk of accepting a relationship that is nonmeaningful (i.e., a type I error) is not as important, since independent testing was performed, and any relationship which is not meaningful would probably show no skill in such tests.

An inspection of Table 3 allows several conclusions to be drawn based on the regression analyses. First, predictability generally decreases with increasing lag: 9, 3, 4 and 2 significant relationships were found for lags of 0, 1, 2 and 3 months, respectively. Second, the first SLP EOF (i.e., the one that explains the most variance) shows up most frequently as a predictand. Third, the variance explained by the regression models is modest, with *R*-squared values between 0.19 and 0.48.

Fig. 4 shows the months for which statistically significant relationships exist from which SLP EOF's can be specified or predicted from SST EOFs, depending on lag. Significant relationships seem to be concentrated in the period from June through January, with no significant relationships for February–May. Fig. 5 shows the months for which SST EOF's are statistically significant in specifying or predicting SLP EOF's, depending on lag. Once again, the concentration of significant relationship can be seen to lie in the June–January period. In addition, all significant relationships at longer lags (2 or 3 months) are cases in which SST in the

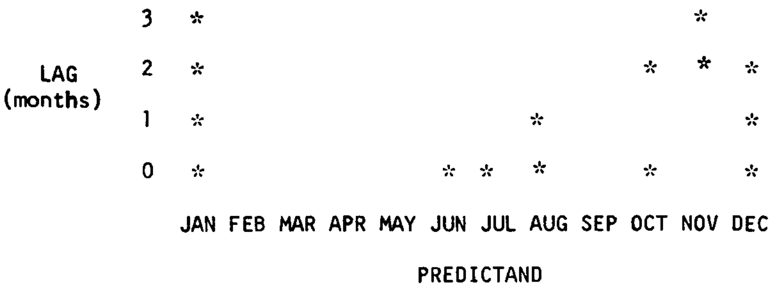

FIG. 4. Statistically significant relationships according to predictand and lag.

August–November period are used as predictors. This would suggest that the effect of SST on SLP is most long-lasting during late summer and autumn. This is consistent with the findings of Davis (1978), in which summer and autumn seasonal SST were found to have value in predicting autumn and winter seasonal SLP; and the results of Harnack and Broccoli (1979), in which the strongest atmosphere-ocean relationships were shown to exist in summer and autumn.

The tendency for strong atmosphere-ocean relationships to favor late summer and autumn seems to have a plausible physical basis. This is the time of year when the water is relatively warm with respect to the air, so that an unstable boundary layer is common in the atmosphere. Instability increases the thermal communication between ocean and atmosphere, producing stronger ocean-atmosphere interactions. While this reasoning applies to contemporaneous relationships, it is more difficult to explain the existence of relationships at lags of 2 or 3 months, since Namias (1972) found the persistence of North Pacific SSTA's (in terms of autocorrelation) to be relatively low during late summer and autumn. One possible explanation for the existence of stronger relationships in this period is related to the negative feedback effects which Namias (1976) found in the North Pacific for the cooling season. This feedback process, which Namias relates to the consequence of cyclonic activity that is enhanced over warm but shallow SSTA's, may be an example of a strong atmosphere-ocean interaction so that unusually strong SST/SLP relationships occur during the cooling season.

5. Independent testing of regression models

To further assess the strength of the regression-derived SST-SLP relationships, independent tests were performed using data from 1933–41 and 1972–76, a total of 14 years. In conducting the independent tests, spatial coefficients derived from EOF analysis of the 1949–71 data were used to compute EOF amplitudes for both periods of independent data. Predicted amplitudes were compared with observed amplitudes, with verification scores computed in terms of the reduction of error statistic (Gilman, 1976). This statistic is defined according to the following formula:

$$\mathrm{RE} = 1 - \sum_{i=1}^{N} (F_i - O_i)^2 / \sum_{i=1}^{N} O_i^2, \qquad (1)$$

where N is the number of forecasts, F_i is the forecast anomaly, and O_i is the observed anomaly. The anomalies are computed by subtracting the 1947–71 mean amplitudes from forecast or observed amplitudes. For a perfect set of forecasts, RE = 1; for normals used as forecasts (i.e., climatology), RE = 0; and for forecasts that are poorer than climatology, RE < 0. When examining RE scores it is necessary to bear in mind that the RE score is a stringent measure of skill, and is sensitive to bias (e.g., forecasts consistently higher or lower than observed values).

Reduction of error scores were calculated sepa-

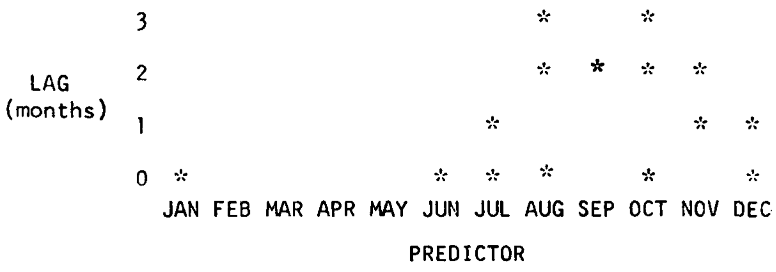

FIG. 5. Statistically significant relationships according to predictor and lag.

TABLE 4. Reduction of error scores (subdivided by period) for statistically significant regression relationships as tested on independent data.

Predictor(s) (number of EOF's)	Predictand (EOF number)	Lag (months)	Reduction of error		
			1933–41	1972–76	Overall
January (1)	January (1)	0	0.156	0.136	0.154
June (2)	June (1)	0	0.185	−0.275	0.112
June (2)	June (2)	0	−0.081	0.243	0.057
July (2)	July (1)	0	−1.331	0.395	−0.347
July (2)	August (1)	1	−0.705	−0.388	−0.601
August (2)	August (1)	0	0.085	−0.151	0.008
August (2)	August (2)	0	−1.041	−0.048	−0.612
August (2)	October (1)	2	−0.885	0.001	−0.655
August (2)	November (1)	3	−1.350	0.142	−0.300
September (2)	November (1)	2	−2.952	−0.298	−1.085
October (2)	October (2)	0	0.229	0.154	0.204
October (2)	December (2)	2	0.242	−0.836	0.075
November (2)	December (1)	1	0.160	0.813	0.278
December (2)	December (1)	0	0.446	0.547	0.464
December (2)	December (3)	0	−0.107	−0.130	−0.116
October (2)	January (1)	3	−0.867	−0.836	−0.865
November (2)	January (1)	2	−0.777	−0.667	−0.770
December (2)	January (1)	1	0.191	−0.029	0.177

rately for the 1933–41 period and the 1972–76 period, and these scores are shown (along with scores for both periods combined) in Table 4. In comparing the values, it can be noted that the RE scores are somewhat higher, in general, for the 1972–76 period than for the earlier period. For 11 of the 18 regression models, the more recent period had the better RE score.

In view of this development, it was decided to reexamine the possibility that the systematic bias between the "historical" data base SST's and the Scripps data base SST's (discussed in Section 3) was responsible for the lower RE scores for the 1933–41 period. This is plausible, since the comparison statistics presented in Table 1 suggest that the differences are quite large.

This development suggested that the systematic bias between the "historical" data base SSTs and the Scripps data base SST's might be too large to be ignored. With the intention of eliminating this bias, linear regression was used to quantify the effects of the sampling centroid anomaly on the SST data. Using seasonally stratified data for each grid point from the years 1861–1960, the "historical" SSTA's were regressed on the latitude and longitude of the sampling centroid anomaly. This analysis yielded equations of the form

$$\mathrm{SSTA_{SCA}} = a_0 + a_1\phi_{\mathrm{SCA}} + a_2\lambda_{\mathrm{SCA}}, \qquad (2)$$

where $\mathrm{SSTA_{SCA}}$ is the component of the SSTA resulting from the sampling centroid anomaly, ϕ_{SCA} is the latitude of the sampling centroid, λ_{SCA} is the longitude of the sampling centroid, and a_0, a_1 and a_2 are regression coefficients. The longitude term was only included in those cases in which it provided a statistically significant contribution to the model.

The historical SST data were then adjusted by subtracting the $\mathrm{SSTA_{SCA}}$ from the new values. These adjusted SST values were used as input for the regression-derived SST-SLP relationships, and RE scores were recalculated for the 1933–41 period. The new RE scores are shown in Table 5. These scores indicate that the disparity between the two time periods has been lessened by adjusting the earlier SST's. Using the adjusted data, the earlier period had the higher RE score for 9 of the 18 cases, as compared to 7 of 18 using the unadjusted data. An overall improvement of the RE scores for the 1933–41 period can be noted by comparing the values in Table 5 to the values in Table 4. For purposes of rough comparison, the arithmetic mean of the RE scores (for the 1933–41 period) using the unadjusted data was −0.467 as compared to −0.078 using the adjusted data.

Two conclusions can be drawn from the independent tests using the adjusted data. First, 10 of the 18 statistically significant relationships (56%) show skill when tested on independent data, with skill being defined as a positive reduction of error (i.e., an improvement over climatology). Second, there is a tendency for skill to decrease with increasing lag. The number of statistically significant relationships judged to have skill when tested on independent data is 7, 2, 1 and 0 for lags of 0, 1, 2 and 3 months, respectively.

Fig. 6 shows the months for which SLP EOFs were skillfully specified or predicted from SST EOFs, depending on lag; and Fig. 7 shows the months for which SST EOFs were skillful in specifying or predicting SLP EOF's, depending on lag, as determined from independent testing. The relationships shown in Figs. 6 and 7 were those which both passed significance tests applied to the

TABLE 5. Reduction of error scores (using adjusted SST data) for statistically significant regression models as tested on independent data.

Predictor(s) (number of EOF's)	Predictand (EOF number)	Lag (months)	Reduction of error		
			1933–41	1972–76	Overall
January (1)	January (1)	0	0.451	0.136	0.421
June (2)	June (1)	0	0.303	−0.275	0.211
June (2)	June (2)	0	0.142	0.243	0.185
July (2)	July (1)	0	0.247	0.395	0.332
July (2)	August (1)	1	−0.396	−0.388	−0.393
August (2)	August (1)	0	0.034	−0.151	−0.027
August (2)	August (2)	0	−1.162	−0.048	−0.113
August (2)	October (1)	2	−0.502	0.001	−0.371
August (2)	November (1)	3	−1.073	0.142	−0.218
September (2)	November (1)	2	−0.985	−0.298	−0.502
October (2)	October (2)	0	0.625	0.154	0.465
October (2)	December (2)	2	0.239	−0.836	0.073
November (2)	December (1)	1	−0.147	0.813	0.026
December (2)	December (1)	0	0.378	0.547	0.408
December (2)	December (3)	0	0.136	−0.130	0.037
October (2)	January (1)	3	−0.528	−0.836	−0.597
November (2)	January (1)	2	−0.401	−0.667	−0.418
December (2)	January (1)	1	0.237	−0.029	0.221

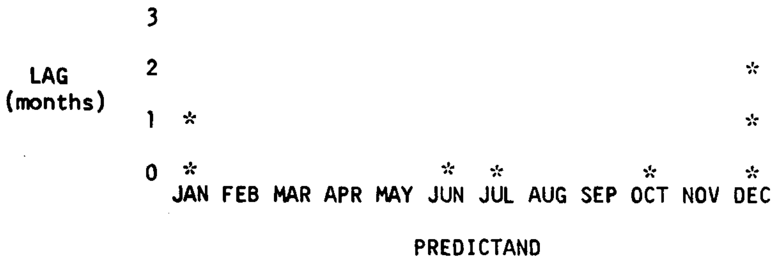

FIG. 6. Temporal distribution of skillful relationships (RE > 0), as tested on independent data, according to predictand and lag.

dependent data set and performed skillfully, as compared with climatology, when applied to two periods of independent data. SLP for January, June, July, August, October and December were specified with skill from contemporaneous SST data. January and December SLP were predicted skillfully from SST data at a one-month lag, while December is the only month for which SLP was predicted with skill from SST data two months in advance. There is some tendency for autumn SST to be most valuable as a predictor, as noted in the previous section.

An examination of the overall RE scores in Table 5 shows that all of the statistically significant SST/SLP relationships in which summer SST's were used as a predictor failed to demonstrate skill when tested on independent data. Most of the relationships in which the predictor month used was in the period from October–January did show skill in independent testing, with the largest RE scores also found in that period.

6. Discussion

The lack of skill for many combinations of months can be attributed to several potential sources. One possibility is to attribute the lack of skill to changes of the atmosphere-ocean interactions over the 44-year period being examined. If such changes occurred, a relationship derived using data from one time period may not be valid during another time period. This is difficult to explain from a physical standpoint, since the laws governing the physical processes involved do not change with time. The problems inherent in sampling a population may make a statistical relationship derived for one time period invalid for another, but this is a statistical effect rather than a physical one.

As a second possibility, the independent sample may not be large enough for reliable RE scores to be computed. While the validity of this premise is difficult to assess objectively, a 14-case independent sample seems, intuitively, to be large enough. Poor spatial coverage of the SST data is another potential cause which is, likewise, difficult to assess objectively.

A third possibility is that SST only leads SLP (i.e., the ocean driving the atmosphere) for certain times of year, while the forcing is in the opposite direction and SLP leads SST (i.e., the atmosphere driving the ocean) at other times of year. Some of the findings of Davis (1978) suggest that this may be the case for the North Pacific.

Finally, using one-month averages of SST and SLP may produce poor results if a monthly time scale is inappropriate for large-scale air-sea interactions of the type hypothesized. Work by Harnack (1979) and Harnack and Broccoli (1979) suggests that this might be the case, since both studies found stronger relationships for three-month periods than for 1-month periods. Nevertheless, some important relationships were found for a monthly time scale, which may be used to advantage for monthly forecasting in selected portions of the year.

7. Conclusions

This study has considered the availability and quality of North Pacific SST data prior to 1947, and the ability of regression models to specify or pre-

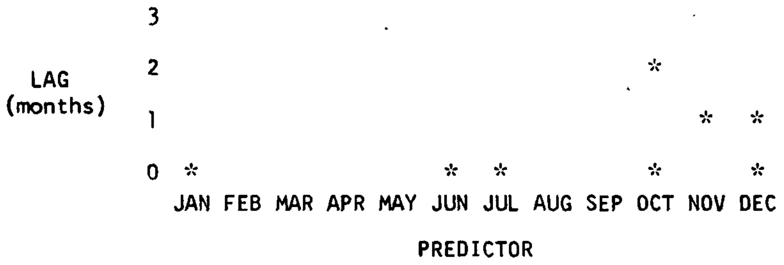

FIG. 7. Temporal distribution of skillful relationships (RE > 0), as tested on independent data, according to predictor and lag.

dict monthly SLP from contemporaneous or previous SST. While the nature of the data and the types of relationships investigated make drawing firm conclusions a difficult task, some statements can be made with sound bases.

First, sufficient SST data exists for the 1933–41 period to further examine large-scale air-sea interactions that heretofore have been investigated using only more recent data. While the "historical" SST data are relatively crude due to poor spatial resolution, it is likely that more useful data could be produced from a similar source. The National Climatic Center Marine Deck contains individual ship observations; these were used to create the "historical" SST data used for this study. Processing these ship observations using methods similar to those used by Scripps for more recent times might produce better SST data.

Second, by developing regression models in which the time-varying amplitudes of statistically significant SLP EOF's are specified or predicted using the amplitudes of statistically significant SST EOF's as predictors, a number of relationships were found for lags of zero to three months that are statistically significant. Most of these relationships were for contemporaneous fields (i.e., zero lag), and the number of significant relationships decreased with increasing lag. The first SLP EOF—the one which explains the most variance for a given month—is the most frequently predictable. All relationships had generally modest *R*-squared values, indicating that they account for only a small amount of the total variance.

An examination of the seasonal distribution of the statistically significant relationships at all lags finds them confined to the period from June to January, while those for longer lags are in the August–November period. This is consistent with previous findings (Davis, 1978; Harnack and Broccoli, 1979) which showed stronger or longer lasting atmosphere-ocean relationships in the late summer and autumn period.

Third, all statistically significant regression models were tested on independent data. The data used in the test included SST's for the 1933–41 period which were adjusted to correct systematic biases in the "historical" SST data base. Verifications of these independent tests were conducted in terms of the reduction of error score, and skill was defined as any improvement over climatology. In these tests, some of the statistically significant regression models failed to show skill when tested on independent data. Most of the models which did show skill involved lags of one month or less, so that there is some indication of decreasing skill with increasing lag.

Several causes for this lack of skill were discussed. Among the more likely causes are the existence of certain times of year in which the atmosphere primarily drives the ocean, rather than the reverse. Also, there exists the likelihood of some incompatibility of monthly averages of oceanic and atmospheric variables with the characteristic time scale of large-scale air-sea interactions. The latter possibility is substantiated by other studies (Harnack, 1979; Harnack and Broccoli, 1979), in which 3-month averages of atmospheric and oceanic variables were found to relate to one another better than one-month averages.

In summary, an attempt to investigate large-scale air-sea interactions over the North Pacific found limited skill in predicting and specifying SLP from SST. The relatively crude nature of the data used to extend the SST's back to 1933 makes it difficult to draw firm conclusions about predictability, but the similarity of skill scores for the 1933–41 verification period and the 1972–76 verification period (once the early data were adjusted), suggests that the relationships found are stationary in time. Continued work to extend high-quality SST data backward beyond 1947 may allow future investigations of these interactions to be more fruitful and conclusive.

Acknowledgments. A thank you is extended to the following individuals who supplied assistance in preparing this work: Ms. Ruth Smith and Ms. Valerie Bowers for typing the manuscript, Ms. Jeremi Harnack for drafting the figures, and Mr. John Lanzante for assisting with some of the computations.

This research was supported by the Climate Dynamics Program, Division of Atmospheric Sciences, National Science Foundation, under Grant ATM-7815434; and by the New Jersey Agricultural Experiment Station under Project 13405.

REFERENCES

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

——, 1978: Predictability of sea level pressure anomalies over the North Pacific Ocean. *J. Phys. Oceanogr.*, **8**, 233–246.

Farmer, S. A., 1973: A note on the long-term effects on the atmosphere of sea-surface temperature anomalies in the North Pacific Ocean. *Weather*, **28**, 102–105.

Gilman, D. S., 1976: Some forecast verifications at NMC. *Weather Forecasting and Weather Forecasts: Models, Systems and Users*, National Center for Atmospheric Research, 797–806.

Harnack, R. P., 1979: A further assessment of winter temperature predictions using objective methods. *Mon. Wea. Rev.*, **107**, 250–267.

——, 1981: Principles and methods of extended period forecasting in the United States. *Nat. Wea. Assoc. Monogr. Ser.*, No. 1, 34 pp.

——, and A. J. Broccoli, 1979: Associations between sea surface temperature gradient and overlying mid-tropospheric circulation in the North Pacific region. *J. Phys. Oceanogr.*, **9,** 1232–1242.

Namias, J., 1959: Recent seasonal interactions between North Pacific waters and the overlying atmospheric circulation. *J. Geophys. Res.*, **64,** 637–646.

——, 1966: Large-scale air-sea interactions as primary causes of fluctuations of prevailing weather. *Trans. N.Y. Acad. Sci.*, **29,** 183–191.

——, 1969: Seasonal interactions between the North Pacific Ocean and the atmosphere during the 1960's. *Mon. Wea. Rev.*, **97,** 173–192.

——, 1971: The 1968–69 winter as an outgrowth of sea and air coupling during antecedent seasons. *J. Phys. Oceanogr.*, **1,** 65–81.

——, 1972: Large-scale and long-term fluctuations in some atmospheric and oceanic variables. *The Changing Chemistry of the Oceans*, Dyrssen, D., and D. Jagner, Eds., John Wiley & Sons, 27–48.

——, 1976: Negative ocean-air feedback systems over the North Pacific in the transition from warm to cold seasons. *Mon. Wea. Rev.*, **104,** 1107–1121.

——, and R. M. Born, 1972: Empirical techniques applied to large-scale and long-period air-sea interactions, a preliminary report. SIO Reference 72-1, Scripps Institution of Oceanography, 47 pp.

Preisendorfer, R. W., and T. P. Barnett, 1977: Significance tests for empirical orthogonal functions. *Preprints Fifth Conf. Probability and Statistics in Atmospheric Sciences,* Amer. Meteor. Soc., 169–172.

Ratcliffe, R. A. S., 1970: Sea surface temperature anomalies and long-range forecasting. *Quart. J. Roy. Meteor. Soc.*, **96,** 337–338.

Roll, H. U., 1965: *Physics of the Marine Atmosphere*. Academic Press, 426 pp.

Trenberth, K. E., and D. A. Paolino, Jr., 1980: The Northern Hemisphere sea-level pressure data set: Trends, errors and discontinuities. *Mon. Wea. Rev.*, **108,** 855–872.