

Predictive Skills of Seasonal to Annual Rainfall Variations in the U.S. Affiliated Pacific Islands: Canonical Correlation Analysis and Multivariate Principal Component Regression Approaches*

ZHI-PING YU

Pacific ENSO Applications Center, University of Hawaii at Manoa, Honolulu, Hawaii

PAO-SHIN CHU AND THOMAS SCHROEDER

Department of Meteorology, School of Ocean and Earth Science and Technology, University of Hawaii at Manoa, Honolulu, Hawaii

(Manuscript received 22 October 1996, in final form 3 March 1997)

ABSTRACT

Drought and flooding are recurrent and serious problems in the U.S. Affiliated Pacific Islands (USAPI). Given the agricultural and water-dependent characteristics of the USAPI economies, accurate forecasts of seasonal to interseasonal rainfall variations have the potential to provide important information for decision makers involved in resource management issues and response strategies related to drought and flood events.

Climatology of rainfall and outgoing longwave radiation (OLR) cycle in the USAPI and the response of OLR to the El Niño–Southern Oscillation (ENSO) are addressed. Boxplot and harmonic analyses indicate that the annual cycles in rainfall and OLR are generally strong in USAPI except those stations close to the equator. Northern USAPI have positive (negative) OLR anomalies during El Niño (La Niña) winters.

Two statistical models, canonical correlation analysis (CCA) and a relatively new method called multivariate Principal Component Regression (PCR), are employed to forecast rainfall variations in 10 USAPI stations. Sea surface temperatures (SSTs) in the Pacific Ocean are used as predictors for both models. The results of this study indicate that both models are potentially useful in predicting seasonal rainfall variations in the USAPI region, especially in winter (DJF) and spring (MAM). CCA cross validation shows that at one and two seasons lead JFM is the most accurately forecast period in the northern USAPI stations, with average skills of 0.53 and 0.41, respectively. However, the authors' analysis indicates a problem of lower predictive skill in summer (JJA) and fall (SON). One reason might be associated with the so-called spring barrier in predictive skill in the tropical ocean–atmosphere system. Another reason might be associated with the tropical cyclone activity during these seasons. Predictions using the PCR model yield similar predictive skill. Though simpler than He and Barnston's model in term of the number of predictor variables used, the authors' CCA and PCR provide comparable skills.

1. Introduction

The U.S. Affiliated Pacific Islands (USAPI) comprise former territories administered under a United Nations trusteeship [now the Federated States of Micronesia (FSM), Republic of Palau, the Commonwealth of Northern Mariana Islands, and the Republic of the Marshall Islands] as well as older U.S. territories (the Territory of Guam, Wake Island, and American Samoa). All but American Samoa lie in the Northern Hemisphere (Fig. 1).

Drought and flooding are recurrent and serious problems in the Pacific islands. Deficient rainfall can lead

to water resource management problems, such as strict water rationing and expensive importation of potable water, as well as deleterious impacts on rain-fed agriculture and increased risk of wildfire hazard. On the other hand, flooding can damage civil infrastructure such as bridges, roads, and housing, as well as destroy agricultural crops and fields. Accurate forecasts of seasonal rainfall variation with sufficient lead times have the potential to provide important information for decision makers involved in resource management issues and response strategies related to drought and flood events in the Pacific islands.

A Pacific ENSO Applications Center (PEAC) was established in early 1994 as a partnership among the University of Guam, the University of Hawaii, NOAA's National Weather Service and Office of Global Programs, and the U.S. affiliated island governments of the Pacific region, through the Pacific Basin Development Council (PBDC). The main purpose of PEAC is to develop and provide tools to assist government officials

* SOEST Contribution Number 4504.

Corresponding author address: Mr. Zhi-Ping Yu, Jr. Researcher, Pacific ENSO Applications Center, University of Hawaii at Manoa, 2525 Correa Rd., HIG 331, Honolulu, HI 96822.
E-mail: yu@soest.hawaii.edu

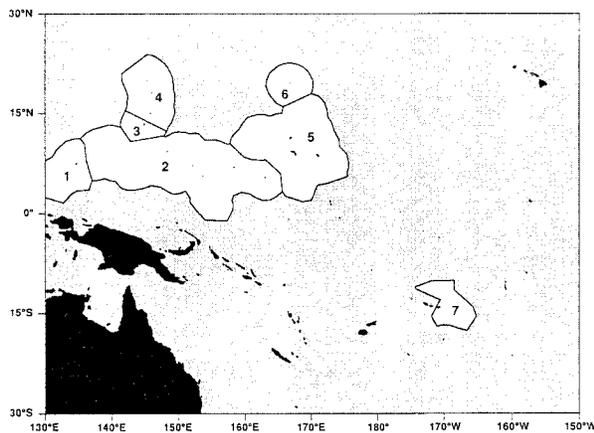


FIG. 1. Geographical map of Exclusive Economic Zone (EEZ) for the United States Affiliated Pacific Islands (USAPI): 1—Republic of Palau; 2—Federated States of Micronesia (FSM); 3—Territory of Guam; 4—Commonwealth of the Northern Mariana Islands (CNMI); 5—Republic of the Marshall Islands; 6—Wake; 7—American Samoa.

and other parties interested in changes in local climate and impacts within the Pacific region arising from the El Niño–Southern Oscillation (ENSO) cycle. Tools developed by the center support decision-making processes undertaken in the Pacific region for water resource management, fisheries management, agriculture, natural disaster mitigation strategies, power utilities, coastal zone management, and other climate-sensitive sectors. We determined that understanding the nature of rainfall and its predictability is essential to serving our clients.

Studies over the past 30 yr have demonstrated that ENSO has a significant impact on the climate variability in the Pacific islands. Bjerknes's (1966, 1969) pioneering studies indicated that tropical climate was strongly influenced by ENSO episodes. Later studies (Meisner 1976; Wright 1979; Ropelewski and Halpert 1987; Chu 1995) have supported the results of Bjerknes. Lau's (1985) global climate model experiments indicate that much of the atmospheric response to ENSO is associated with the changes in SSTs in the Pacific. Pacific SSTs can thus be used to forecast regional climate fluctuations, especially in the tropical Pacific area.

We decided to use two statistical techniques for our application. The first technique is canonical correlation analysis (CCA), a multivariate statistical model that computes linear combinations of a set of predictors that maximize relationships, in a least squares error sense, to similarly calculated linear combinations of a set of predictands. In other words, CCA is used to find linear combinations of two datasets that are most highly correlated. Given this characteristic of CCA and the high correlations between climate variations and ENSO (or SSTs), CCA modeling using Pacific SSTs as predictors certainly provides good potential in predicting seasonal to perhaps interannual climate variations (e.g., rainfall fluctuations) for the USAPI. The second technique that

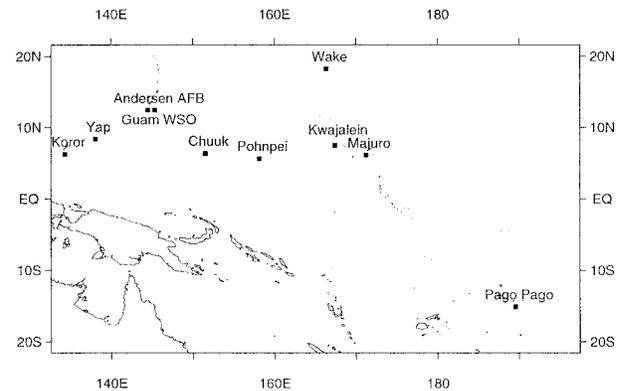


FIG. 2. Locations of selected 10 USAPI stations (Andersen AFB/Guam, Guam WSO, Kwajalein, Koror, Majuro, Pago Pago, Chuuk, Wake, and Yap).

we use is multivariate Principal Component Regression (PCR). It is mathematically simpler than CCA. One of our interests is in comparing performance of the two techniques.

Recently, He and Barnston (1996) used the CCA method to forecast rainfall in the tropical Pacific islands at various lead times. He and Barnston (hereafter referred to as HB) noted a moderate skill for certain islands in the Northwestern Pacific, particularly during late northern winter. Our current study complements the HB analysis by further investigating the seasonality of rainfall predictability in the USAPI. Climatology of the rainfall cycle in the USAPI and the response of rainfall and outgoing longwave radiation (OLR) to ENSO, which are not addressed by HB, will be examined. Furthermore, whereas HB used quasi-global SSTs, Northern Hemisphere 700-mb height, and the past history of rainfall itself in forecasting island rainfall, we have focused on SSTs as the sole predictors. We base this approach on the close proximity of the USAPI to the equatorial center of atmospheric response to ENSO (Ropelewski and Halpert 1987).

Section 2 describes the data, and section 3 discusses climatology and rainfall/OLR variations related to ENSO. The analysis methods and CCA and PCR models are described in section 4. Section 5 discusses the results of cross-validation skills, along with an examination of seasonality of rainfall predictability. Finally, section 6 summarizes the study.

2. Data

Monthly rainfall data for 10 major USAPI stations (Koror, Yap, Guam WSO, Andersen Air Force Base in Guam, Chuuk, Pohnpei, Wake, Kwajalein, Majuro, and Pago Pago) in the Pacific islands are selected as predictands. Locations of selected long-term USAPI rainfall stations are shown in Fig. 2 and a table of geographical details for selected stations is given in Table 1. The rainfall data are for a period of 38 yr (1957–94).

TABLE 1. Geographical details (latitude, longitude, and associated governments) of each rainfall station.

Stations	Latitude	Longitude	Governments
Koror	7°20'N	134°29'E	Palau
Yap	9°29'N	138°05'E	FSM–Yap
Guam WSO	13°33'N	144°50'E	Guam
Andersen AFB Guam	13°35'N	144°55'E	Guam
Chuuk	7°27'N	151°50'E	FSM–Chuuk
Pohnpei	6°58'N	158°14'E	FSM–Pohnpei
Wake	19°17'N	166°39'E	United States
Kwajalein	8°44'N	167°43'E	Marshalls
Majuro	7°05'N	171°23'E	Marshalls
Pago Pago	14°20'S	170°43'W	American Samoa

Data were obtained from the Western Regional Climate Center (WRCC) in Reno, Nevada.

The SST data used in this analysis are a combination of two datasets. The Global Ocean and Global Atmosphere (GOGA) data (4.5° lat × 7.5° long), derived from the Comprehensive Ocean–Atmosphere Data Set (COADS; Slutz et al. 1985), are used in the first part from 1956 to 1988. For the later part of SST data (1989–94), we use the Climate Prediction Center (CPC) blended SST data, which were derived from a blend of in situ data, AVHRR satellite data, and ice data (Reynolds 1988). The SST data cover most of the Pacific Ocean from 100°E to 72°W, and 55°S to 55°N. The SST data are then averaged into 10° lat × 24° long boxes, as in Joseph et al. (1991). There is a total of 80 boxes, or predictor elements, for the forecasts because 8 boxes do not have data.

We also used OLR, which is a commonly used proxy for rainfall in the oceanic Tropics where stations are scarce and the rainfall is primarily convective (Lau and Chan 1983a,b; Weickmann 1983; Murakami and Nakazawa 1985; Yoo and Carton 1990; Lyons 1991). Although Morrissey (1986) and Lyons (1991) demonstrated that OLR can be misleading in certain climate regimes, the USAPI is generally free of that problem. An inverse relation between OLR and rainfall in the Tropics generally holds. This can be further seen in the high negative correlation coefficient between the USAPI monthly rainfall series and OLR series (at nearest corresponding box) in each USAPI station (Table 2). As indicated in Table 2, a correlation analysis of 19-yr USAPI rainfall and OLR monthly series show strong correlation coefficients ranging from -0.60 to -0.78 . All correlation coefficients are significant at the 99% level when the serial correlation is taken into account.

3. Climatology

The Northern Hemisphere USAPI lie within the monsoon region as defined by Ramage (1971). The rainy seasons coincide with the annual northward march and then southward retreat of the surface monsoon trough (Sadler et al. 1987). The cloudiness and precipitation maximum associated with the westerlies equatorward of the trough constitute the intertropical convergence zone (ITCZ) often described in the literature (Sadler 1975). An additional warm season feature is the Tropical Upper-Tropospheric Trough (TUTT), which serves as a primary source of warm season disturbances and some typhoons near Wake Island (Sadler 1978). Wake Island lies well northeast of the maximum northward position of the monsoon trough and experiences no ITCZ rains. During the winter monsoon, dry northeasterly winds blow at all Northern Hemisphere USAPI; cold fronts reach Wake Island.

American Samoa lies eastward of the Southern Hemisphere monsoon region as defined by Ramage (1971) and Webster (1987). Climatologically, the monsoon trough during austral summer is located near 12°S, just to the north of Australia, extending eastward from about 150°E to 170°E (Sadler et al. 1987). However, eastward advances of the monsoon trough can sometimes envelop Samoa, a feature evident during the 1982–83 ENSO episode (Sadler 1983). This trough, together with the South Pacific convergence zone (SPCZ), which is aligned northwest to southeast from about 170°E to 150°W (Vincent 1994), is the source of the cloudiness. In the austral cool season a cloudiness maximum remains near Samoa. This maximum forms in the convergence between two high pressure cells, one in the extreme southeast Pacific and a second immediately east of Australia.

Figure 3 shows boxplots of monthly rainfall data in selected USAPI stations. The annual cycle appears strong in the Northern Hemisphere USAPI stations with maximum median rainfall in ASO (August–October) and minimum in FMA (February–April). The annual cycles for individual stations vary slightly and maximum rainfalls (medians) occur at times ranging from early June to October, depending on the locations of individual stations relative to the seasonal, meridional movement of the monsoon trough. Thus, rainfall maxima occur in July/August in low latitude stations such as Koror, Yap, Pohnpei, Chuuk, and Majuro while they occur in September/October in higher latitude stations such as Guam WSO and Kwajalein. In Wake, however,

TABLE 2. A table of correlation coefficient (CC) between monthly rainfall in the USAPI stations and OLR times series at the nearest box, with * indicating the significance at 99% level (note: Guam rainfall data are rainfall in Guam WSO). Nineteen-year (1975–93) data are used and 10-month OLR missing data in 1978 are replaced by median values.

Station	Koror	Yap	Guam	Chuuk	Pohn	Wake	Kwaj	Maju	Pago
CC	-0.67^*	-0.77^*	-0.78^*	-0.69^*	-0.63^*	-0.66^*	-0.64^*	-0.63^*	-0.60^*

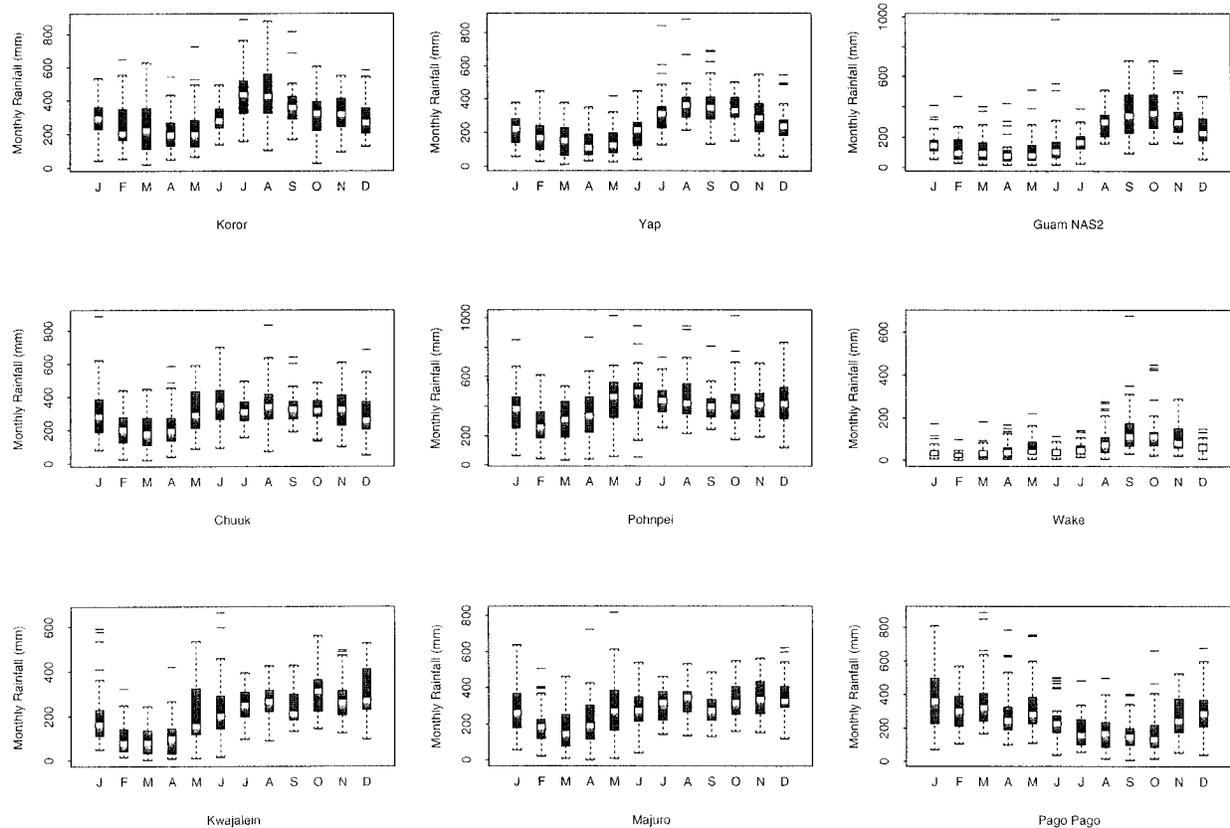


FIG. 3. Boxplots of monthly rainfall totals from 1957 to 94 in selected stations. The box shows the interquartile range extending from the lower quartile ($p = 25\%$) to the upper quartile ($p = 75\%$). White spots inside the boxes are middle halves of the data (medians), the vertical lines (“whiskers”) extending from the box reach to the most extreme nonoutlier ($p = 97.5\%$ and $p = 2.5\%$); outlying points are plotted individually.

maximum rainfall is usually associated with the TUTT as discussed earlier. Low rainfall values in spring are associated with the dry east Asian winter monsoon. In Pago Pago, maximum rainfall occurs in January when the SPCZ is very strong, and the minimum is in September/October when the SPCZ is weak.

The OLR boxplots at locations (Fig. 4) show similar characteristics. Most Northern Hemisphere stations appear to have minimum OLR (or maximum convection) in ASO and maximum OLR (or minimum convection) in FMA.

To evaluate the importance of the annual cycle, harmonic analysis is performed on the long-term monthly mean rainfall and OLR series. The first harmonic, which represents the annual cycle, explains a substantial variance of the rainfall/OLR variability in the monsoonal regions (Fig. 5 and Table 3). Maximum rainfalls (or minimum OLRs) occur from August to October for northern USAPI stations (Table 3). The first harmonic explains most of the variance (at least 53% of total variance for rainfall and 73% for OLR), especially for Guam (92%, 90%), Pago Pago (90%, 99%), and Yap (85%, 87%). The annual cycle is relatively weak (though still explaining over 53% of the variance for

rainfall) at stations near the equator such as Koror, Chuuk, and Pohnpei. These stations receive monsoon trough rains twice a year as the trough crosses on its northward advance and southward retreat. A second harmonic (Table 4), which represents the semiannual cycle, adds significantly to the variance explained at Koror, Pohnpei, and Wake.

The OLR data are utilized to demonstrate ENSO effects in the Tropical Pacific region, as in Chu (1995). Due to the relatively short record of OLR data, we selected five winters (DJF) of warm ENSO events (1976/77, 1982/83, 1986/87, 1991/92, 1992/93) and two winters of cold events (1975/76, 1988/89) for composite analyses. In a normal Northern Hemisphere winter, the northern USAPI are in the high OLR (low rainfall) zone (Fig. 6a). The SPCZ, as defined by a persistent OLR minimum area with values less than 230 W m^{-2} , extends southeastward from New Guinea to American Samoa and beyond.

The OLR composites are shown in Figs. 6b and 6c. For El Niño winters (Fig. 6b), strong negative OLR anomalies (associated with stronger convection) are located in the equatorial central Pacific while strong positive anomalies (associated with weaker convection) are

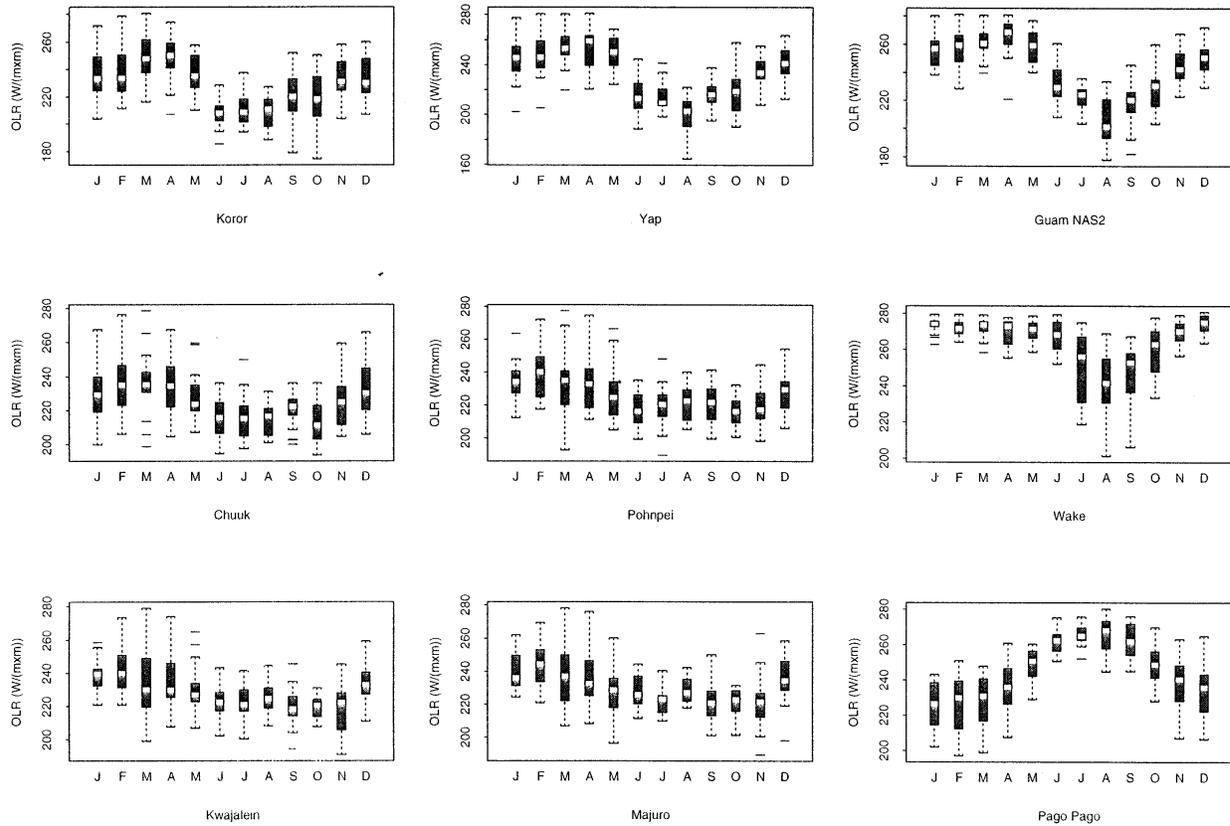


FIG. 4. Same as Fig. 3 but for OLR time series (1975–93) at nearest locations.

found in the equatorial western Pacific. Two relatively weak positive anomalies are found in the subtropical areas of the central Pacific. Pago Pago is in the boundary of the negative anomaly center while most northern Pacific stations are located in the positive anomaly area. Thus, Pago Pago receives more than normal rainfall while most northern Pacific stations become drier during El Niño. The pattern of OLR anomalies during cold events (Fig. 6c) is approximately opposite to warm events: relatively large, negative OLR anomalies are found in the northern tropical Pacific while positive anomalies are located in the equatorial central Pacific.

4. Analysis methods and CCA and PCR model building

a. Predictors

SSTs in the Pacific are used as predictors in the CCA model. Studies by Hastenrath and Heller (1977), Moura and Shukla (1981), and Shukla and Misra (1977) indicated that SSTs play an important role in regulating precipitation on regional or global scales. Lau (1985) also found that much of the atmospheric response to ENSO was associated with the changes in sea surface temperatures in the Pacific. SSTs have long-term records and thus provide a good sample size for statistical models. Indeed, Barnston and

He (1996) used a CCA model to predict rainfall in Hawaii and Alaska and their results suggested that SSTs contribute the most to forecast accuracy.

b. Data prefiltering by empirical orthogonal function (EOF) analysis

It is highly advisable to apply EOF analysis prior to the CCA. This can be accomplished by projecting the data onto EOFs and then retaining only a limited number of principal components in the analysis. There are several reasons for doing this. First of all, the large number of spatial points can cause difficulty in inverting the matrices and in the eigenvalue problem. The preprocessing procedure transforms the large spatial dimension into a smaller number of retained EOF modes and this value is quite modest even for large fields. The computation of canonical modes is thus simplified after preprocessing. In addition, the small-scale noise is filtered out after EOF analysis since the use of EOFs allows the analysis to focus on the dominant modes of variability within each input dataset. The procedure for computing eigenvectors from a matrix of data has been described extensively in the literature and will not be discussed here.

The EOF analysis is thus performed onto the box-

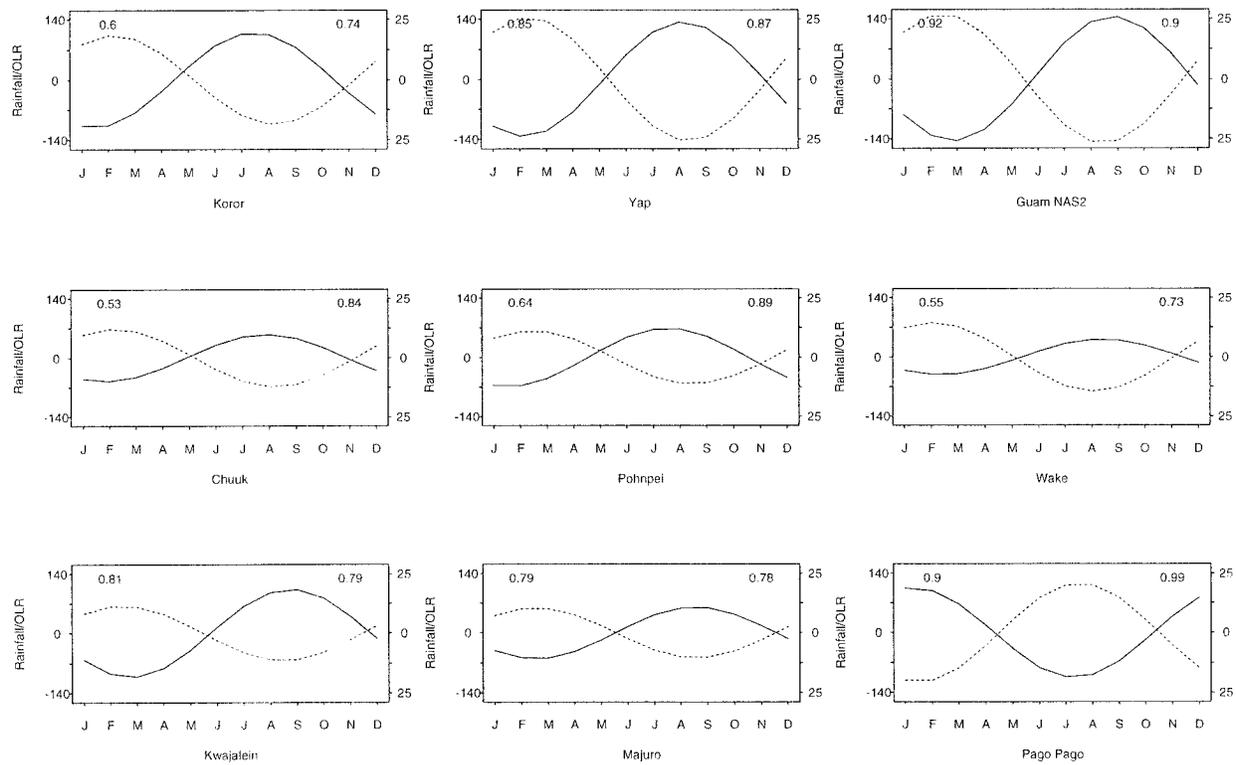


FIG. 5. First harmonic of rainfall and OLR. Solid lines denote rainfall in individual stations; dashed lines denote OLR values at corresponding locations. Values in the upper-left (right) corner are variances explained by the first harmonics of rainfall (OLR).

averaged seasonal mean SSTs for the period of 1956 to 1994 and seasonal rainfall totals for the period of 1957 to 1994. The data have not been detrended or standardized. The EOF analysis is thus performed using a covariance matrix. The advantage of using the covariance matrix is that the strongest variations can be identified or isolated in a dataset.

There is no universally agreed upon procedure for determining how many EOF modes should be retained. Morrison (1976) suggests that the retained components

should explain perhaps 75% or more of the variances. Therefore, the leading eight modes of SSTs and the four leading modes of rainfall are selected since they account for approximately 75% or higher of the total variance, respectively (Tables 5 and 6). Scree graphs (Wilks 1995) are also used and results are consistent with the above selection rule (not shown).

TABLE 3. Harmonic analysis of annual cycles of USAPI rainfall and OLR series (at nearest corresponding box). Here, R^2 denotes the proportion of the variance accounted for by the first harmonic, and T denotes the phase angle of the first harmonic or the time, in month (e.g., 1 means January, 2 means February, . . .) of maximum (or minimum for OLR) of the first harmonic function. Nineteen-year (1975–93) rainfall and OLR data are used.

Stations	Rainfall		OLR at nearest box	
	R^2	T (month)	R^2	T (month)
Koror	0.60	8.46	0.74	9.23
Yap	0.85	9.17	0.87	9.35
Guam WSO	0.92	9.83	0.90	9.46
Chuuk	0.53	8.87	0.84	9.21
Pohnpei	0.64	8.55	0.89	9.46
Wake	0.55	9.39	0.73	9.07
Kwajalein	0.81	9.76	0.79	9.48
Majuro	0.79	9.54	0.78	9.55
Pago Pago	0.90	2.30	0.99	2.51

TABLE 4. Same as Table 3 except for the second harmonic.

Stations	Rainfall		OLR at nearest box	
	R^2	T (months)	R^2	T (months)
Koror	0.28	1.83	0.16	2.57
		7.83		8.57
Yap	0.08	3.18	0.06	3.75
		9.18		9.75
Guam WSO	0.06	5.93	0.07	4.49
		11.93		10.49
Chuuk	0.06	5.53	0.03	2.57
		11.53		8.57
Pohnpei	0.30	5.38	0.06	5.80
		11.38		11.80
Wake	0.25	6.37	0.23	5.64
		12.37		11.64
Kwajalein	0.09	4.51	0.08	4.48
		10.51		10.48
Majuro	0.07	6.04	0.12	4.42
		12.04		10.42
Pago Pago	0.01	4.86	0.01	2.65
		10.86		8.65

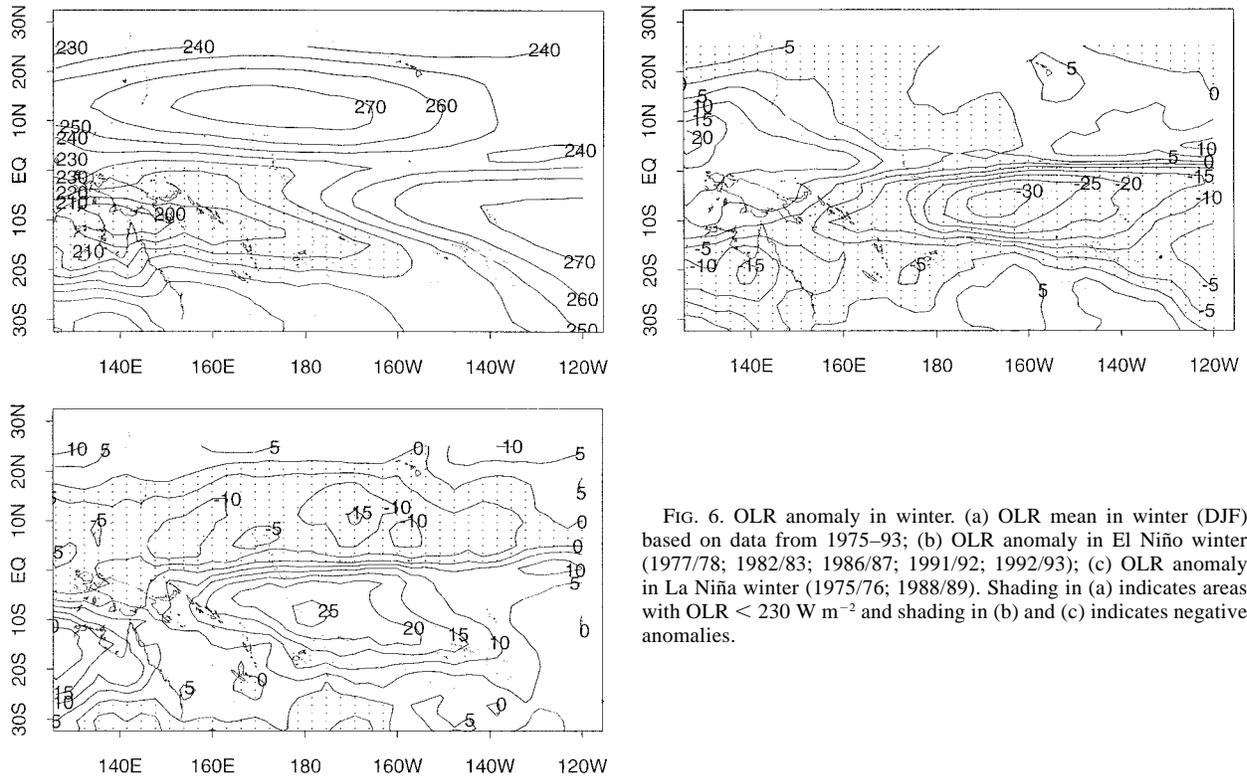


FIG. 6. OLR anomaly in winter. (a) OLR mean in winter (DJF) based on data from 1975–93; (b) OLR anomaly in El Niño winter (1977/78; 1982/83; 1986/87; 1991/92; 1992/93); (c) OLR anomaly in La Niña winter (1975/76; 1988/89). Shading in (a) indicates areas with OLR < 230 W m⁻² and shading in (b) and (c) indicates negative anomalies.

c. CCA forecast model

The CCA technique was introduced by Hotelling (1935), and a useful discussion of the method, with examples of its application to meteorological fields, was presented by Glahn (1968). Since the introduction of CCA, prediction of climate variations using the method has received wide attention and has had some success; for example, CCA was employed to forecast short-term climate fluctuations (Barnett and Preisendorfer 1987; Graham et al. 1987a,b; Barnett et al. 1988) and seasonal rainfall fluctuations (Barnston 1994; Chu and He 1994; Yu 1994; Barnston and He 1996; Barnston and Smith 1996).

Assuming that $X_{s,t}$ denotes the predictor matrix and $Y_{r,t}$ denotes the predictand matrix, where the subscripts s and r represent space, and subscript t represents time, namely,

$$s = 1, \dots, m$$

$$r = 1, \dots, n$$

$$t = 1, \dots, k.$$

Both $X_{s,t}$ and $Y_{r,t}$ are monthly mean (or seasonal mean) removed matrices. Performing EOF analysis on $X_{s,t}$ and $Y_{r,t}$ leads to

$$X_{s,t} = E_{s,s} T_{s,t} \tag{1}$$

$$Y_{r,t} = E_{r,r} T_{r,t}, \tag{2}$$

where $E_{s,s}$ and $E_{r,r}$ represent EOF spatial modes of the predictor and predictand matrices, respectively, and $T_{s,t}$ and $T_{r,t}$ are their attendant time coefficients.

Assuming that we take the first i EOF modes of the predictor time series ($T_{i,t}$, $i < m$) and the first j EOF modes of the predictand series ($T_{j,t}$, $j < n$) as inputs to CCA analysis, we can determine canonical vectors (u ,

TABLE 5. Percentage of variance explained by eigenvectors. (Values in parentheses are cumulative variance explained by the k largest eigenvalues.)

Ek/SST	DJF SSTs	MAM SSTs	JJA SSTs	SON SSTs
E1	35.5 (35.5)	26.7 (26.7)	22.1 (22.1)	33.7 (33.7)
E2	9.9 (45.4)	12.1 (38.8)	14.7 (36.8)	12.7 (46.4)
E3	8.0 (53.4)	9.5 (48.3)	11.3 (48.1)	8.2 (54.6)
E4	7.0 (60.4)	8.0 (56.3)	7.2 (55.3)	5.7 (60.3)
E5	5.2 (65.6)	6.2 (62.5)	5.8 (61.1)	5.2 (65.5)
E6	4.3 (69.9)	5.0 (67.5)	5.4 (66.5)	4.5 (70.0)
E7	3.9 (73.8)	4.2 (71.7)	4.3 (71.8)	4.1 (74.1)
E8	3.3 (77.1)	3.6 (75.3)	3.8 (74.6)	3.4 (77.5)

TABLE 6. Percentage of variance explained by eigenvectors. (Values in parentheses are cumulative variance explained by the k largest eigenvalues.)

Ek/rainfall	DJF rainfall	MAM rainfall	JJA rainfall	SON rainfall
E1	54.9 (54.9)	58.7 (58.7)	34.6 (34.6)	33.1 (33.1)
E2	13.8 (68.7)	17.6 (76.3)	21.5 (56.1)	22.8 (55.9)
E3	10.6 (79.3)	9.7 (86.0)	6.9 (73.0)	14.5 (70.3)
E4	7.6 (86.9)	4.6 (90.6)	11.4 (84.4)	9.5 (79.8)

v) and linear combinations of $Z = u' T_{i,t}$ and $W = v' T_{j,t}$. The following matrices are defined:

$$\mathbf{U}' = [u'_1, u'_2, \dots, u'_q]' \quad (3)$$

$$\mathbf{V}' = [v'_1, v'_2, \dots, v'_q]' \quad (4)$$

$$\mathbf{A}_{q,q} = \begin{bmatrix} a_1 & & & \\ & a_2 & & 0 \\ & & \cdot & \\ 0 & & & a_q \end{bmatrix}, \quad (5)$$

where a_1, \dots, a_q are ‘‘canonical correlations’’ between Z and W , and $a_1 \geq a_2 \geq \dots \geq a_q$ and q is equal to i or j , whichever is smaller.

Prediction equation at time t for lead time l for the first j modes of predictand can then be constructed as follows (Glahn 1968; Chu and He 1994):

$$\hat{T}_{j,t+l} = (\mathbf{V}')^{-1} \mathbf{A}_{q,q} \mathbf{U}' T_{i,t}. \quad (6)$$

Using the orthogonal characteristics of EOF modes, truncated equations of (1) and (2) can be rewritten as

$$T_{i,t} = (\mathbf{E}_{s,i})' X_{s,t} \quad (7)$$

$$Y_{r,t} = \mathbf{E}_{r,j} T_{j,t}. \quad (8)$$

Using Eqs. (7) and (8), $T_{j,t+l}$ can thus be transformed back to rainfall values as

$$\hat{Y}_{r,t+l} = \mathbf{E}_{r,j} (\mathbf{V}')^{-1} \mathbf{A}_{q,q} \mathbf{U}' (\mathbf{E}_{s,i})' X_{s,t}. \quad (9)$$

d. Principal Component Regression (PCR) model

In the CCA model, we assumed that the predictor field is a matrix of known constants. If that matrix is random, then the analysis is carried out conditional on it so that it is still treated as if it were fixed. The initial EOF truncations also cause a problem. In CCA, it is implicitly assumed that the first four modes of the predictand field are the ones that are most highly correlated with the first eight modes of the predictor fields. However, there is no guarantee that one of the higher modes of variations in the predictor set (e.g., higher than the eighth mode) will not be strongly associated with the predictand set.

To overcome these problems, we use a relatively new approach, principal component regression (Draper and Smith 1981). Using the same notation as in section 4c, we can write the Principal Component Regression (PCR) model as follows:

$$\mathbf{Y}_{r,t} = \mathbf{B}_{r,i} \mathbf{T}_{i,t} + \mathbf{Y}_{r,t}' \quad (10)$$

where $\mathbf{T}_{i,t}$ is a matrix of the truncated EOF time coefficients and $\mathbf{B}_{r,i}$ is a matrix of regression coefficients (obtained from least squares estimates). The error matrix $\mathbf{Y}_{r,t}'$ has rows that are assumed to be multivariate normal with means zero and some appropriate covariance structure. Note that there is no EOF truncation on the pre-

dictand dataset so that the full information contained in the rainfall dataset is used. By doing so, there is no need to assume that the leading truncated modes in the predictor dataset are those that are most highly related to the leading modes in the predictand dataset. Once fitted equations are obtained in terms of the selected modes, they can be transformed back into a function of the original predictor data using Eq. 7.

We still use the first eight modes of SSTs (which explain more than 75% of the total variance as shown in Table 6) as predictors to construct a PCR model. To ensure that we do not miss modes with strong correlation to predictands, the correlations between the time coefficients of the eigenvectors associated with other non-zero eigenvalues and the predictand values are examined, and no particularly strong correlation is found. The selection of the first eight principal components in PCR also make it easier to compare the skills between PCR and CCA model results directly.

5. Model results

a. CCA maps

The loading patterns of mode one for winter (DJF) SSTs and the following spring (MAM) rainfall are shown in Figs. 7a and 7b. Strong negative (positive) SST anomalies in the equatorial central/eastern Pacific are associated with positive (negative) rainfall anomalies throughout the northern USAPI region. The rainfall series in USAPI is highly correlated to the Pacific SST series for this mode, which accounts for 39% of the total variance. In Fig. 7c, the two canonical component time series have a high correlation coefficient (0.80). Higher than normal winter SSTs in the equatorial central Pacific lead to lower than normal spring rainfall in northern USAPI and higher than normal spring rainfall in Pago Pago. Specifically, a maximum anomaly of +0.8°C in the equatorial central Pacific leads to -136 mm and +34 mm of spring rainfall anomalies at Andersen AFB and Pago Pago, respectively. Given the importance of the first mode, the above results suggest that the inter-annual rainfall variations in the tropical Pacific islands are strongly influenced by the SST field and further support the use of the Pacific SSTs as predictors in the model.

The pattern of the observed SST anomalies at the height of the 1982–83 ENSO (Fig. 7d) is similar to the SST loading pattern (Fig. 7a) although the sign is reversed. By combining the canonical component time series of SSTs (solid line in Fig. 7c) and the SST loading pattern (Fig. 7a), we independently obtain a positive anomaly in the equatorial central–eastern Pacific during DJF of 1982/83. Therefore, the first mode is an ENSO-related mode, implying that ENSO plays an important role in the inter-annual rainfall variation.

b. CCA cross-validation results and seasonality of rainfall predictability

Cross validation is a generalization of the common technique of repeatedly omitting a few observations from the

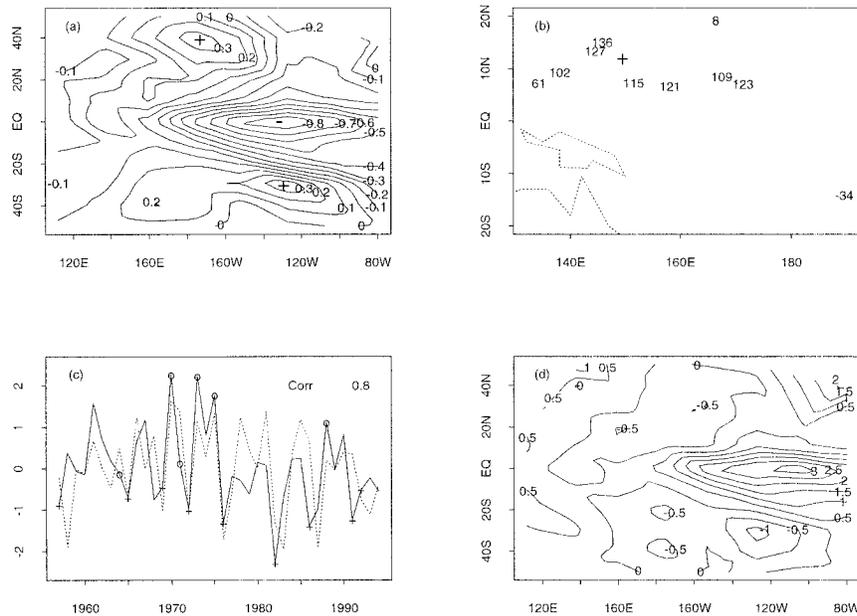


FIG. 7. The principal loading patterns and canonical component time series of mode 1 for winter (DJF) SSTs and the following spring (MAM) precipitation. (a) SST loading pattern. The SST and rainfall loadings have been adjusted to their raw units, namely $^{\circ}\text{C}$ and mm, respectively; (b) rainfall loading pattern; (c) canonical component time series in normalized values. Solid line and dashed line denote SSTs and rainfall, respectively; + and \circ denote El Niño and La Niña years, respectively; (d) observed DJF 1982/83 SST anomalies in $^{\circ}\text{C}$ (based on SST climatology from 1957 to 1994).

data, reconstructing the model, and then making forecasts for the omitted cases (Stone 1974; Chu and He 1994). Cross validation is conducted to evaluate the overall forecasting skill of the CCA model. The cross validation is nonparametric and provides an unbiased estimate of forecast skill. The approach goes as follows (Chu 1989): The predictor and predictand data of N time points are divided into L segments. A model is then developed using the data of $L - 1$ segments. This model is then used to predict the variable in the remaining segment. This process is successively repeated by changing the segment that has been excluded from the model development. By doing this, we obtain N predictions. These predicted values can be correlated with N observations and the overall forecast skill can be determined.

In this study, we remove only one observation at a time for each case. This is justified as interannual autocorrelations in the data are small (e.g., the average absolute autocorrelations of the rainfalls and SSTs is 0.10 and 0.17, respectively). Therefore we use all data available except for the season for which we want to make a prediction. For example, to forecast the summer (JJA) rainfall of 1980 with one-season lead time, we use a 37-yr summer rainfall time series (1957–79, 1981–94) and a 37-yr spring SST series (1957–79, 1981–94) to build a CCA model (redo everything each time, including the pre-EOFs). Then this resulting CCA model is used to forecast rainfall values in summer 1980 using SST values in spring 1980. We use the moving average season of three consecutive months

in order to identify the season with best predictability, yielding 12 target seasons (DJF, JFM, ..., OND, NDJ).

The cross-validation skills for JFM, AMJ, JAS, and OND are shown in Table 7. As indicated, different islands show different levels of predictive skill. Overall, rainfall forecasts for Yap, Guam WSO, Andersen, Chuuk, and Pohnpei are relatively well predicted with a mean skill of 0.33 or higher as demonstrated in the bottom line of Table 7. Rainfall forecasts are most skillful in Guam WSO (0.39) and Andersen (0.36). The CCA model generates very poor rainfall prediction for Pago Pago. One possible reason may be that Pago Pago is not located near any of the other stations in the study and the variance of this station may get chopped off after the pre-EOF truncation because it does not have any “partners” to align with.

The cross validation also indicates a seasonality of predictive skill. In general, forecasts for JFM and AMJ are better than the corresponding ones for JAS and OND. When all stations are considered, the model yields the best forecast skill (0.49) for one-season lead when the target season is JFM. At some stations (Yap and Pohnpei), one-season lead forecasts have skills of 0.70. For some stations (Yap, Andersen, and Wake for JFM; and Yap, Guam WSO, and Andersen for AMJ), the model can even demonstrate moderate predictive skill (≥ 0.30) at four-season lead time when the target seasons are JFM and AMJ.

Forecast skills for JAS rainfall are generally not as good as those for JFM and AMJ. However, moderate skills are found in Guam WSO and Pohnpei. OND is the most

TABLE 7. CCA cross-validation skill (correlation between seasonal rainfall forecasts and observations, negative correlations are set to zero) for all 10 stations in the USAPI. For example, predictions for OND (target season) rainfall using SSTs in JAS and AMJ as predictors are considered as one season and two season leads, respectively.

Season (target)	Lead time (season)	Kor	Yap	Gua	And	Chu	Poh	Wak	Kwa	Maj	Pag	Mean
JFM	1	0.46	0.70	0.52	0.59	0.46	0.70	0.37	0.52	0.53	0.00	0.49
	2	0.44	0.60	0.52	0.61	0.29	0.39	0.38	0.36	0.12	0.42	0.41
	3	0.00	0.22	0.31	0.38	0.00	0.01	0.37	0.14	0.00	0.00	0.14
	4	0.15	0.39	0.25	0.39	0.16	0.12	0.34	0.18	0.01	0.10	0.21
AMJ	1	0.00	0.32	0.44	0.41	0.25	0.00	0.00	0.49	0.43	0.00	0.24
	2	0.14	0.24	0.55	0.51	0.32	0.02	0.24	0.29	0.08	0.00	0.24
	3	0.07	0.31	0.30	0.31	0.37	0.32	0.06	0.23	0.17	0.00	0.21
	4	0.51	0.45	0.31	0.33	0.21	0.09	0.00	0.28	0.18	0.00	0.24
JAS	1	0.00	0.00	0.39	0.25	0.08	0.61	0.12	0.30	0.00	0.05	0.18
	2	0.00	0.00	0.35	0.26	0.09	0.40	0.00	0.11	0.00	0.00	0.12
	3	0.00	0.00	0.31	0.14	0.29	0.39	0.00	0.00	0.00	0.00	0.11
	4	0.00	0.00	0.26	0.09	0.06	0.35	0.00	0.00	0.00	0.00	0.08
OND	1	0.40	0.31	0.18	0.00	0.40	0.12	0.00	0.00	0.00	0.20	0.16
	2	0.00	0.18	0.00	0.00	0.20	0.41	0.00	0.18	0.18	0.16	0.11
	3	0.30	0.21	0.00	0.00	0.40	0.19	0.00	0.00	0.02	0.11	0.12
	4	0.00	0.01	0.00	0.00	0.32	0.00	0.00	0.00	0.03	0.08	0.04
Mean skill		0.18	0.33	0.39	0.36	0.33	0.34	0.16	0.28	0.15	0.09	0.26

difficult season to accurately predict rainfall in the islands. Only one station (Chuuk) shows moderate forecast skill. One reason why JFM and AMJ have better predictability is probably because ENSO responses are most pronounced during boreal winter as the Pacific SSTs and Southern Oscillation index anomalies reach their peaks during boreal winter/spring. Table 8 gives more detailed cross-validation skills for the JFM target season at varying lead times.

Figure 8 provides plots of cross-validation skills for 12 moving seasons (from DJF to NDJ). We can see that at one-season lead JFM is the most skillful forecast period in the northern USAPI, with average skill of 0.53. JFM is also the most skillful forecast period at two seasons lead, with average skill of 0.41. The maximum cross validation skills in JFM and FMA and the minimum skills in ASO and SON yield a seasonality of rainfall predictive skill.

A target season-lead time cross-section plot is a useful way to present the CCA cross-validation skills. Such a plot is given in Fig. 9. Cross-validation skill varies strongly as a function of target season. Skills are high during boreal winter and spring when the contour line of 0.30 extends to 8 months lead. Low skills are found during boreal fall (SON). Therefore, a strong seasonality in empirical rainfall predictive skill is seen. Figure 9 also indicates that cross-validation skill tends to decrease as lead time increases.

As mentioned earlier, one of our objectives is to determine how much predictability can be achieved when only the Pacific SSTs are used as predictors. Compared to HB's Fig. 2a and Fig. 3a, Figs. 8 and 9 imply that our model provides comparable skills (note that HB uses 14 stations, including Johnston and 4 Hawaiian stations, and that the lead time in their study is defined to be 3 months less than in our study such that OND-to-JFM is called a zero lead time). This is intriguing since our model is simpler relative

TABLE 8. Cross-validation skill for JFM (target season) rainfall in the USAPI. Note that the first two rows show partly specification relationships.

Predictor period	Kor	Yap	Gua	And	Chu	Poh	Wak	Kwa	Maj	Pag	Mean
DJF SST	0.49	0.73	0.54	0.55	0.48	0.66	0.40	0.50	0.48	0.08	0.49
NDJ SST	0.42	0.72	0.54	0.59	0.44	0.64	0.38	0.49	0.43	0.00	0.47
OND SST	0.46	0.70	0.52	0.59	0.46	0.70	0.37	0.52	0.53	0.00	0.49
SON SST	0.44	0.72	0.52	0.62	0.44	0.71	0.42	0.47	0.47	0.10	0.49
ASO SST	0.35	0.69	0.55	0.66	0.33	0.62	0.43	0.37	0.33	0.35	0.47
JAS SST	0.44	0.60	0.52	0.61	0.29	0.39	0.38	0.36	0.12	0.42	0.41
JJA SST	0.24	0.42	0.42	0.53	0.11	0.22	0.36	0.30	0.06	0.32	0.30
MJJ SST	0.12	0.30	0.36	0.42	0.06	0.10	0.39	0.24	0.00	0.10	0.21
AMJ SST	0.00	0.22	0.31	0.38	0.00	0.01	0.37	0.14	0.00	0.00	0.14
MAM SST	0.06	0.40	0.37	0.46	0.00	0.14	0.37	0.20	0.00	0.11	0.21
FMA SST	0.07	0.31	0.33	0.42	0.05	0.09	0.39	0.22	0.00	0.20	0.21
JFM SST	0.15	0.39	0.25	0.39	0.16	0.12	0.34	0.18	0.01	0.10	0.21
Mean skill	0.27	0.52	0.44	0.52	0.24	0.34	0.37	0.33	0.20	0.15	0.34

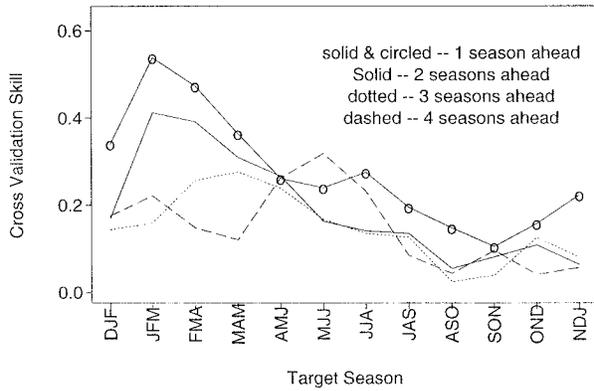


FIG. 8. CCA cross-validation skills (average) for nine stations in the Northern Hemisphere. Solid and circled line, solid line, dotted line, and dashed line denote one, two, three, and four seasons ahead, respectively.

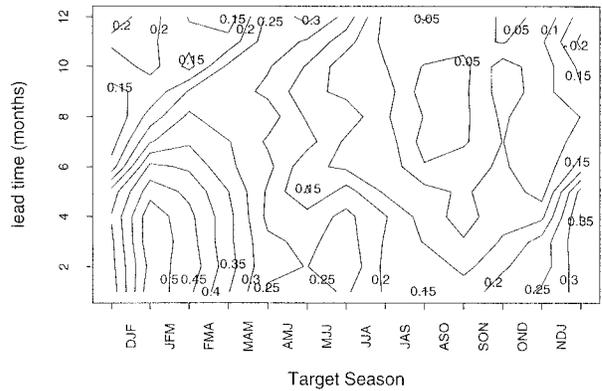


FIG. 9. Lead time-target season cross section of CCA cross-validation skill for northern USAPI rainfall. The horizontal axis denotes the target season, and the vertical axis the lead time in months.

to HB and suggests that the Pacific SSTs alone may be enough to make a moderately skillful rainfall prediction for the USAPI. Given the well-known ENSO's impact on USAPI rainfall, the Pacific SSTs may thus have advantages in predicting Pacific island rainfall and there is no need to include SSTs from other ocean basins and 700-mb height into the prediction scheme.

Since SSTs are the only predictors in the CCA model, the predictability of rainfall variations will depend largely on the characteristics of the ocean surface. Latif and Graham (1991) noticed that there appeared to exist a "predictability barrier" around the time of boreal spring. They found that the correlations between observations and predictions of the SSTs of the coupled system of the Pacific Ocean decrease rapidly between April and June. Similar decreases across the spring period appear in the results of other coupled models (e.g., Cane 1991). Focusing on the rapid decline in forecast skills during boreal spring, Webster and Yang (1992) indicated that the summer monsoon circulation develops fastest from April to May, the time when the Walker circulation is the weakest and most susceptible to external noise. Thus, the typical ocean-atmosphere interaction may be least robust during boreal spring and thus subject to larger error growth. This idea was supported by Xue et al. (1994) who found that the transient initial error grows fastest starting from spring and slowest starting from late summer; they attributed the rapid decline in forecast skill in boreal spring (the "spring barrier") to the smallness of the signal to be forecast.

In our CCA model, no useful skill can be produced for summer and fall rainfall prediction if variations in SSTs are uncertain in the antecedent boreal spring. As shown in Fig. 10, rainfall forecasts using AMJ SSTs as predictors generate the worst skill for 4–6 months ahead forecasts and, as a result, the average skill (0.17) is the lowest among the four seasons. JAS also has very low skill from 1 to 4 months ahead. Forecasts initiated in OND provide the best overall skill (0.37). The model's low rainfall forecast skills in summer and fall (using AMJ SSTs as predictors) and

the good rainfall forecast skills in winter and spring (using OND SSTs as predictors) suggest that a spring barrier occurs in the CCA when SSTs are used as predictors. These results are consistent with previous studies (e.g., Cane et al. 1986; Webster and Yang 1992; Xue et al. 1994).

Another possible explanation of the low rainfall forecast skills in summer and fall relates to tropical cyclone activity in the corresponding period. It is known that the western North Pacific is the favored breeding ground for tropical cyclones, which generally form in late summer and early fall. These intense convection systems can produce substantial rainfall over limited areas and can cause significant local variability in seasonal rainfall at an individual island.

We calculated tropical cyclone frequency in the area of the North Pacific islands. A box covering a domain of 0°–20°N and 130°E–180° is selected. Tropical cyclone frequency is counted only when a tropical cyclone is located within the domain of the selected box and with a maximum sustained wind speed of 17 m s⁻¹ or more. The relationship

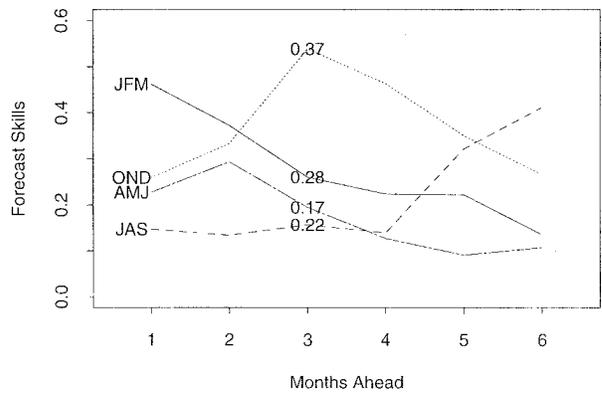


FIG. 10. The correlations between observations and predictions (initiated in four different seasons) of the northern USAPI rainfall as a function of lead times. JFM, AMJ, JAS, and OND indicate rainfall forecasts using JFM, AMJ, JAS, and OND SSTs as predictors, respectively. The average forecast skill values are shown in the middle of individual lines.

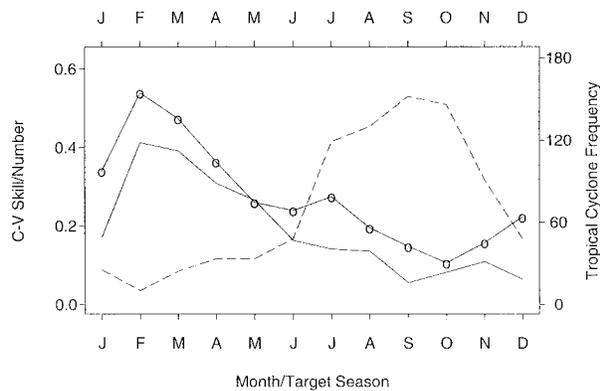


FIG. 11. CCA cross-validation skills and tropical cyclone frequency. Solid line with open circle and solid line denote one season ahead and two season ahead forecasts, respectively. Monthly tropical cyclone frequency (total number of tropical cyclones affecting the selected box within 0°–20°N and 130°E–180° from 1949 to 1994) is plotted as dashed line and its scale is shown in the right vertical axis.

between cross-validation skills and tropical cyclone frequency is plotted in Fig. 11. One-season and two-season lead cross-validation skills are shown in this figure. The predictive skill of rainfall is negatively proportional to the tropical cyclone frequency in the northwest Pacific. The minimum tropical cyclone frequency in February is associated with the highest cross-validation skill in JFM season. After February, the cross-validation skills decline rapidly as the spring barrier is penetrated, and then reach minimum values in September and October when maximum tropical cyclones form in this region.

c. PCR results

We employ the same cross-validation scheme (described section 5b) to assess our PCR model validity and adequacy. In addition to cross-validation skill (correlation), we compare the results of the PCR and CCA models by computing what is known as the PRESS statistic (the sum of the squared differences between the actual observations and the predictions from the leave one out cross-validation

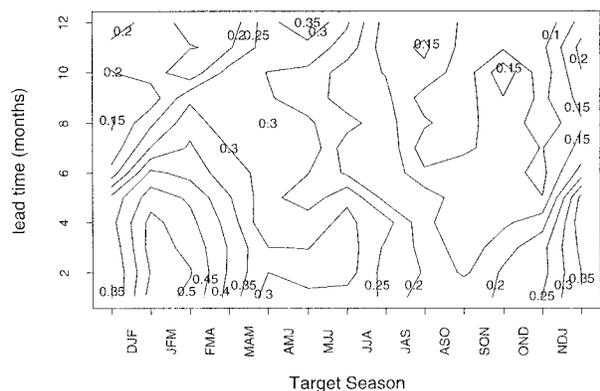


FIG. 12. Same as Fig. 9 but for PCR.

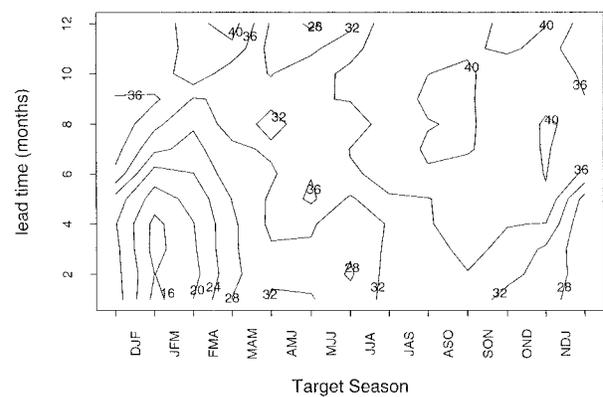


FIG. 13. Lead time-target season cross section of PCR PRESS statistic (units: 1000 mm²) for the northern USAPI rainfall. The horizontal axis denotes the target season, and the vertical axis the lead time in months.

results). The PRESS selection procedure was proposed by Allen (1971). It is a combination of all possible regressions, residual analysis, and validation techniques. The PRESS statistic (Raymond 1990) can be expressed as

$$PRESS = \sum_{t=1}^k (Y_t - \hat{Y}_{t,-t})^2, \quad (11)$$

where Y_t denotes the t th observation and $\hat{Y}_{t,-t}$ is the leave one out prediction for the t th value. In general, a model with the smallest PRESS is preferred.

Figure 12 shows lead time-target season cross section of PCR cross-validation skill for northern USAPI rainfall. It has an almost identical pattern with Fig. 9. The PCR model shows high prediction skills during boreal winter and spring, and again JFM is also the most skillful forecast period. A further comparison of cross-validation skills is made by computing the PRESS statistic for both PCR and CCA models (shown in Figs. 13 and 14, respectively). The PCR model results are generally as good as the CCA results. In both PCR and CCA models, small PRESS values (useful predictive skills) are found during boreal winter and

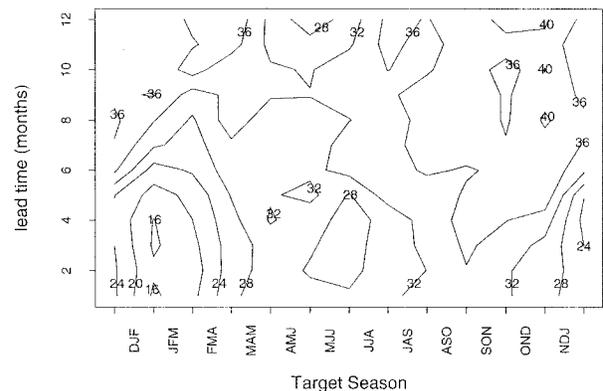


FIG. 14. Same as Fig. 13 but for CCA.

spring, and large PRESS values (weak skills) in summer and fall.

6. Summary

Following an overview of the climatology of the USAPI stations and the effects of ENSO, CCA and PCR are employed to forecast rainfall and study rainfall predictability for 10 stations in the USAPI. Box-averaged SSTs in the Pacific Ocean are used as predictors. To filter the noise and reduce sample size, EOF analysis is conducted as a preprocessing method. The following conclusions are drawn:

- 1) Boxplots and harmonic analysis of rainfall series indicate that the annual cycle in rainfall is generally strong for most USAPI stations with maximum rainfall in August/September and minimum in February/March. However, the annual cycle is relatively weak at stations close to the equator. This is evidenced by these equatorial stations (e.g., Koror, Pohnpei), which encounter monsoon trough rains twice each year as the trough crosses on its northward advance and southward retreat. The results are supported by a similar study of OLR.
- 2) ENSO has a strong impact on the climate of USAPI. For example, most stations have positive OLR anomalies (less convection) during El Niño winters and negative anomalies (more convection) during La Niña winters. The ENSO impact in Pago Pago in the Southern Hemisphere is generally opposite.
- 3) The CCA model provides useful skill in predicting rainfall in the Pacific Islands. Rainfall at most USAPI stations (e.g., Yap, Guam WSO, Andersen, Chuuk, and Pohnpei) shows moderate overall forecast skill with mean cross-validated correlation skill of 0.33 or higher. This result is intriguing in that it is a simpler (one predictor variable: SST) model relative to He and Barnston (1996), who used three predictor variables, and it further suggests that the Pacific SSTs alone might be enough to make a moderate-skill rainfall prediction for the USAPI. This might be expected in view of the important role of SSTs in regulating precipitation and the strong ENSO impact on USAPI rainfall.
- 4) The CCA cross-validated predictive skills show a strong annual cycle (seasonality). JFM is the most accurately forecast period in the northern USAPI at one and two seasons lead time with average correlation skill of 0.53 and 0.41, respectively. Generally speaking, high predictability in rainfall is found in boreal winter and spring while relatively low predictability is found in summer and fall.
- 5) The poor forecast skills resulting from AMJ SSTs as predictors suggest that the so-called spring-barrier effect may exist for SST-based predictions in linear statistical models. It is very difficult to generate accurate forecasts using spring SSTs as predictors if variations in SSTs are uncertain during boreal spring. The moderately useful skill using OND SSTs to predict forthcoming seasonal rainfall further supports this claim. The spring-barrier effect may contribute, at least partially, to the seasonality of rainfall predictions.
- 6) Our study also indicates that there might be a relationship between the predictability of seasonal rainfall in the USAPI region and tropical cyclone activity in the area. The low CCA skills in boreal summer and fall may be attributed in part to the strong tropical cyclone activity during these periods. On the other hand, the high forecast skills in winter and early spring are associated with the minimum tropical activity in the same period.
- 7) The relatively high predictability in boreal winter and early spring might be attributed to the least SST barrier in fall, the low tropical cyclone activity, and the pronounced ENSO responses in winter and spring.
- 8) A relatively new multivariate Principal Component Regression model is also employed to predict USAPI rainfall. Comparison of both PCR and CCA model results indicates that PCR provides comparable skills in predicting USAPI rainfall variation relative to CCA. Because of its simplicity, PCR offers a potential as a new tool for statistical climate prediction research.

Acknowledgments. Deep appreciation is extended to the Office of Global Programs (OGP) of the National Oceanic and Atmospheric Administration (NOAA) for supporting this research under Research Grant NA37RJ0199 and NA67RJ0154. The views expressed herein are those of the authors and do not necessarily reflect the views of NOAA or any of its subagencies. We are grateful to Alan Hilton LT/NOAA, who make some helpful contributions to this research and to Dr. Fei-Fei Jin for his comments. We are also thankful to Tony Barnston for his careful reading and reviewing of the paper. The addition of principal component regression was recommended by an anonymous reviewer and it improved the manuscript.

REFERENCES

- Allen, D. M., 1971: The prediction sum of squares as a criterion for selecting predictor variables, Department of Statistics, University of Kentucky Tech. Rep. 23, 32 pp. [Available from Dept. of Statistics, University of Kentucky, Lexington, KY 40506.]
- Barnett, T. P., and R. Preisendorfer, 1987: Origins and levels of monthly and seasonal forecast skill for United States surface air temperatures determined by canonical correlation analysis. *Mon. Wea. Rev.*, **115**, 1825–1850.
- , N. Graham, M. Cane, S. Zebiak, S. Dolan, J. O'Brien, and D. Legler, 1988: Prediction of the El Niño of 1986–87. *Science*, **241**, 192–196.
- Barnston, A. G., 1994: Linear statistical short-term climate predictive skills in the Northern Hemisphere. *J. Climate*, **7**, 1513–1564.

- , and Y. He, 1996: Skill of CCA forecasts of 3-month mean surface climate in Hawaii and Alaska. *J. Climate*, **9**, 2020–2035.
- , and T. M. Smith, 1996: Specification and prediction of global surface temperature and precipitation from global SST using CCA. *J. Climate*, **9**, 2660–2697.
- Bjerknes, J., 1966: A possible response of the atmospheric Hadley circulation to equatorial anomalies of ocean temperature. *Tellus*, **18**, 820–829.
- , 1969: Atmospheric teleconnections from the equatorial Pacific. *Mon. Wea. Rev.*, **97**, 163–172.
- Cane, M. A., 1991: Forecasting El Niño with a geophysical model. *Teleconnections Linking World-Wide Climate Anomalies*, M. Glantz, R. Katz, and N. Nicholls, Eds., Cambridge University Press, 345–369.
- , S. E. Zebiak, and S. C. Dolan, 1986: Experimental forecasts of El Niño. *Nature*, **321**, 827–832.
- Chu, P.-S., 1989: Hawaiian drought and the Southern Oscillation. *Int. J. Climatol.*, **9**, 619–631.
- , 1995: Hawaii rainfall anomalies and El Niño. *J. Climate*, **8**, 1697–1703.
- , and Y. He, 1994: Long-range prediction of Hawaiian winter rainfall using canonical correlation analysis. *Int. J. Climatol.*, **14**, 659–669.
- Draper, N. R., and H. Smith, 1981: *Applied Regression Analysis*. 2d ed. Wiley and Sons, 709 pp.
- Glahn, H. R., 1968: Canonical correlation analysis and its relationship to discriminant analysis and multiple regression. *J. Atmos. Sci.*, **25**, 23–31.
- Graham, N. E., J. Michaelsen, and T. P. Barnett, 1987a: An investigation of the El Niño–Southern Oscillation cycle with statistical models. Predictor field characteristics. *J. Geophys. Res.*, **92**, 14 251–14 270.
- , —, and —, 1987b: An investigation of the El Niño–Southern Oscillation cycle with statistical models. 2. Model results. *J. Geophys. Res.*, **92**, 14 271–14 289.
- Hastenrath, S., and L. Heller, 1977: Dynamics of climatic hazards in northeast Brazil. *Quart. J. Roy. Meteor. Soc.*, **103**, 77–92.
- He, Y., and A. G. Barnston, 1996: Long-lead forecasts of seasonal precipitation in the tropical Pacific Islands using CCA. *J. Climate*, **9**, 2020–2035.
- Hotelling, H., 1935: The most predictable criterion. *J. Educ. Psychol.*, **26**, 139–142.
- Joseph, P. V., B. Liebmann, and H. H. Hendon, 1991: Interannual variability of the Australian summer monsoon onset: Possible influence of Indian summer monsoon and El Niño. *J. Climate*, **4**, 529–538.
- Latif, M., and N. E. Graham, 1991: How much predictive skill is contained in the thermal structure of an OGCM? *TOGA Notes*, **2**, 608.
- Lau, K.-M., and P. H. Chan, 1983a: Short-term climate variability and atmospheric teleconnection from satellite-observed outgoing longwave radiation, Part I: Simultaneous relationships. *J. Atmos. Sci.*, **40**, 2735–2750.
- , and —, 1983b: Short-term climate variability and atmospheric teleconnection from satellite-observed outgoing longwave radiation, Part II: Lagged correlations. *J. Atmos. Sci.*, **40**, 2751–2767.
- Lau, N. C., 1985: Modeling the seasonal dependence of the atmospheric response to observed El Niño. *Mon. Wea. Rev.*, **113**, 1970–1996.
- Lyons, S. W., 1991: Origins of convective variability over equatorial southern Africa during austral summer. *J. Climate*, **4**, 23–29.
- Meisner, B. N., 1976: A study of Hawaiian and line islands rainfall. Department of Meteorology, University of Hawaii, UHMET 76-04, 82 pp. [Available from Dept. of Meteorology, UHM, 2525 Correa Rd., HIG 331, Honolulu, HI 96822.]
- Morrisey, M. L., 1986: A statistical analysis of the relationships among rainfall, outgoing longwave radiation and the moisture budget during January–March 1979. *Mon. Wea. Rev.*, **114**, 931–942.
- Morrison, D. F., 1976: *Multivariate Statistical Methods*. 2d ed. McGraw-Hill, 273 pp.
- Moura, A. D., and J. Shukla, 1981: On the dynamics of droughts in Northeast Brazil: Observations, theory and numerical experiments with a general circulation model. *J. Atmos. Sci.*, **38**, 2653–2675.
- Murakami, T., and T. Nakazawa, 1985: Tropical 40–50-day oscillations during the 1979 Northern Hemisphere summer. *J. Atmos. Sci.*, **42**, 1107–1122.
- Ramage, C. S., 1971: *Monsoon Meteorology*. Academic Press, 296 pp.
- Raymond, H. M., 1990: *Classical and Modern Regression with Applications*. 2d ed. PWS-Kent Publishing, 488 pp.
- Reynolds, R. W., 1988: A real-time global sea surface temperature analysis. *J. Climate*, **1**, 75–86.
- Ropelewski, C. F., and M. S. Halpert, 1987: Global and regional scale precipitation patterns associated with the El Niño–Southern Oscillation. *Mon. Wea. Rev.*, **115**, 1606–1626.
- Sadler, J. C., 1975: The upper tropospheric circulation over the global tropics. Dept. of Meteorology University of Hawaii, UHMET 75-05, 35 pp. [Available from Dept. of Meteorology, UHM, 2525 Correa Rd., HIG 331, Honolulu, HI 96822.]
- , 1978: Mid-season typhoon development and intensity changes and the tropical upper tropospheric trough. *Mon. Wea. Rev.*, **106**, 1137–1152.
- , 1983: Tropical Pacific atmospheric anomalies during 1982–83. *1982–83 El Niño/Southern Oscillation Data Display Workshop*, NOAA/AOML, 1–10.
- , M. A. Lander, A. M. Hori, and L. K. Oda, 1987: Tropical marine climatic atlas, Vol II: Pacific Ocean. Dept. of Meteorology University of Hawaii, UHMET 87-01, 28 pp. [Available from Dept. of Meteorology, UHM, 2525 Correa Rd., HIG 331, Honolulu, HI 96822.]
- Shukla, J., and B. N. Misra, 1977: Relation between SST and wind speed over the central Arabian Sea and monsoon rainfall over India. *Mon. Wea. Rev.*, **105**, 998–1002.
- Slutz, R., S. J. Lubler, J. D. Hiscox, S. D. Woodruff, R. J. Jenne, D. H. Joseph, P. M. Steurer, and J. D. Elius, 1985: *Comprehensive Ocean Atmosphere Data Set*. National Oceanic and Atmospheric Administration, Boulder, CO, 268 pp.
- Stone, M., 1974: Cross-validated choice and assessment of statistical predictors. *J. Roy. Stat. Soc.*, **B36**, 111–147.
- Vincent, D. G., 1994: The South Pacific Convergence Zone (SPCZ): A review. *Mon. Wea. Rev.*, **122**, 1949–1970.
- Webster, P. J., 1987: The elementary monsoon. *Monsoons*, J. S. Fein and P. L. Stephens, Eds., Academic Press, 3–32.
- , and S. Yang, 1992: Monsoon and ENSO: Selectively interactive systems. *Quart. J. Roy. Meteor. Soc.*, **118**, 877–926.
- Weickmann, K. M., 1983: Intraseasonal circulation and outgoing longwave radiation modes during Northern Hemisphere winter. *Mon. Wea. Rev.*, **111**, 572–577.
- Wilks, D. S., 1995: *Statistical Methods in the Atmospheric Sciences*. Academic Press, 467 pp.
- Wright, P. B., 1979: Hawaiian winter rainfall related to Pacific sea surface temperature. *Mon. Wea. Rev.*, **107**, 492–495.
- Xue, Y., M. A. Cane, S. E. Zebiak, and M. B. Blumenthal, 1994: On the prediction of ENSO: A study with a low-order Markov model. *Tellus*, **46A**, 512–528.
- Yoo, J.-M., and J. A. Carton, 1990: Annual and interannual variation of the freshwater budget in the tropical Atlantic Ocean and the Caribbean Sea. *J. Phys. Oceanogr.*, **20**, 831–845.
- Yu, Z.-P., 1994: On the prediction of rainfall variability in the Hawaiian Island. IRICP Final Rep. International Research Institute on Climate Predictions, Lamont-Doherty Earth Observatory, Columbia University, 33 pp. [Available from IRICP, Lamont-Doherty Earth Observatory, Columbia University, Rt. 9W, Palisades, NY 10964.]