Complex Singular Spectrum Analysis and Multivariate Adaptive Regression Splines Applied to Forecasting the Southern Oscillation

Christian Keppenne[1] and Upmanu Lall[2]

[1]clk@jpl.nasa.govQhttp://yabloko jpl.nasa.gov/clk.html
[2]ulall@kernel.uwrl.usu.eduQhttp://grumpy.usu.edu/~FALALL/ulall.html

[1]Jet Propulsion Laboratory, Pasadena, California 91109
[2]Utah Water Research Laboratory, Utah State University, Logan, Utah 84322

A few years ago, Keppenne and Ghil (1992a,b; see also previous issues of this Bulletin) introduced a methodology to forecast the Southern Oscillation Index (SOI) by applying the maximum entropy method (MEM) to produce autoregressive forecasts of a set of adaptively filtered time series resulting from the application of singular spectrum analysis (SSA) to the raw monthly mean SOI. The success of this methodology has led to the development of a multivariate prediction scheme based on the same concepts, but with the substitution of multivariate SSA for univariate SSA (Keppenne and Ghil 1993, Jiang et al. 1995). The technique described herein introduces the following improvements to the linear prediction scheme used to issue the SSA/MEM predictions presented in earlier issues of this Bulletin.

First, the data base used to compute the forecasts has been extended backward from June 1945 to August 1881. Our earlier work had excluded the pre-World War II data, mainly because of numerous gaps in the Tahiti SLP. Rather than doing so here, we have developed a variation of SSA capable of handling missing values. Most data adaptive statistical prediction methods are best understood in terms of an "analog forecast" (e.g. Toth 1991, Huang et al. 1993, Livezey et al. 1994). Consequently, the extension of the data base increases the likelihood of identifying a suitable "analog" that will influence the determination of the forecast's basis functions. Figure 1 illustrates this principle by showing an adaptively filtered SOI indicator resulting from the complex SSA (CSSA) of the last 114 years (as of February 1996) of the Darwin and Tahiti SLP. The sequence of events between 1910 and 1915 presents some similarities with the early 1990s: a positive excursion of the SOI (La Nina event) is followed by two brief mild negative excursions. A strong La Nina event follows, in 1917-18. The series of circles on the righthand side of the curve shows the result of forecasting the real and imaginary parts of the SOI's leading four complex principal components (CPCs) using a variation of multivariate adaptive regression splines (MARS: Friedman 1991, Lewis and Stevens 1991, Lall et al. 1996), a nonlinear data-adaptive statistical method whose application to the SOI is discussed below. The forecast is remarkably similar to its "analog" in the 1910s and thus testifies of MARS' ability to model the dynamics of rarely occurring events. The prediction shown in Fig. 1 is markedly different from predictions obtained from either the application of linear methods to the entire data base, or from the application of MARS to the post-World War II data. Such predictions forecast near-normal conditions in the late 1990s (e.g. the Jiang et al. articles in the March and September 1995 issues of this Bulletin).

Second, in contrast with our earlier work (Keppenne and Ghil 1992a,b) in which SSA was applied to the difference between the Tahiti and Darwin normalized SLP time series, we apply CSSA to the complex time series whose real and imaginary parts consist in the Darwin and Tahiti SLP, forecast the real and imaginary parts of the resulting CPCs separately, and then take their differences to construct a forecast for the filtered SOI. This seemingly innocent procedural modification results in significant enhancements of the objective forecast skill, because taking the difference between two noisy time series increases the noise-to-signal ratio. The application of CSSA to the Darwin and Tahiti SLP followed by the subtraction of the filtered real parts of the resulting CPCs from the corresponding filtered imaginary parts circumvents this problem, thereby leading to the improved forecast skill.

Third, we have replaced the linear autoregressive MEM predictions by the nonlinear MARS methodology,

MARS has advantages that significantly increase forecast skill. Among these are the ability to propagate a periodic oscillation without damping the underlying signal, and the data-adaptive capability discussed above. The latter advantage provides MARS with the capability of "analog" forecast schemes--such as radial basis functions (Casdagli 1989), nearest- neighbor bootstrap schemes (Lall and Sharma 1996) and local polynomials (Abarbanel and Lall 1996)--of reconstructing the dynamics of rarely occurring events (i.e "recording" and "reconstructing" the character-istics of sparsely populated regions of phase space). To enhance this property, we have developed a variation of MARS in which appropriate "neighbors" of the prevailing climate conditions are identified in the phase space. The regression-splines model used to issue the predictions is then fitted to represent the mapping of each selected "neighbor" to the corresponding successor in the phase-space trajectory. More details about this specific procedure are provided in Keppenne and Lall (1995, 1996).

We use the following approach to objectively evaluate our algorithm's forecast skill. Starting with 1200 complex values in our data base, we apply CSSA with a 60-month wide time window to the data, and embed the real and imaginary parts of each CPC in 60-dimensional phase space using lagged versions of those time series as phase-space coordinates. The embedding phase spaces are then searched for the nearest two hundred neighbors of the time series' last points and MARS models are fitted using the phase-space coordinates of the neighbors as predictor variables and their temporal successors as predictands. A 60-month forecast is then issued for each real and imaginary part. The corresponding forecast for the SOI is obtained by convoluting the extended (as a result of the forecast's issuance) CPCs with the corresponding CEOFs, and subtracting the real part of the resulting time series from its imaginary part. The scheme is then repeated with one more complex number in the data base representing the following monthly mean SLP values at Darwin and Tahiti, and a new 60-month forecast is issued. This procedure is repeated until the data base is exhausted and the resulting 168 sets of eight forecasts (one for each real and imaginary part of the leading four CPCs) are used to objectively measure the procedure's predictive ability. Note that this is a "retroactive real-time" simulation, in that future "analogs" are not used.

Figure 2 illustrates the differences in skills between various forms of MARS models. In it, the average error of applying a 60-month forecast with either one of the following methods is compared to the average error of a persistence forecast: (a) MEM as in Keppenne and Ghil (1992a,b), (b) MARS with interaction level one, (c) MARS with interaction level two, and (d) MARS with interaction level three. MARS employs multivariate cubic spline basis functions for regression. The interaction level determines the types of terms that are considered in forming a tensor product across variables or coordinates. Inclusion of higher- order interaction terms indicates the presence of an increasing amount of nonlinearity in the underlying dynamics.

The forecasts in Fig. 2 are for the time series reconstructions involving the leading four CPCs rather than for the monthly mean SOI itself or for its five- month running mean. Note that all three types of MARS models dramatically outperform the MEM models at short (<30-month) leads. All forms of MARS models have comparable skills for most lead times, although the higher-interaction-level models slightly outperform the lower-interaction-level ones.

Figure 3 shows eight 60-month lead SOI forecasts issued at intervals of 24 months between August 1981 and August 1993, including the relatively recent (but not most current) forecast based on data up to November 1995. The solid line in Fig. 3 denotes the last 15 years of the five-month running-mean SOI. Each series of connected circles is a 60-month lead forecast. Note how well the 82-83 and 86-87 El Nino events could have been forecasted at leads of several years. The forecast skill corresponding to the prediction of the 1985 and 1988 La Nina events is also impressive. However, the skill is much lower in the early 1990s, where all our forecasts miss the doubly recurring mild El Nino event. This fact is not surprising, since the interannual variability from the mid 1960s to the late 1980s has been highly regular (Fig. 1). Indeed, one has to go back almost 80 years in the data base to encounter an event reminiscent of the recent conditions (Fig. 1). As discussed above, the strong La Nina event predicted for the late 1990s (Fig. 1) is a result of our variation of MARS' ability to produce "analog" type forecasts, a capability not present in the SSA-MEM approach of Keppenne and Ghil (1992a,b).

Compared with the forecast issued 3 months ago in the December 1995 issue of this Bulletin, the present

forecast is reasonably similar. The strength of the La Nina is now predicted to be slightly less than before (but still substanual), and to peak somewhat earlier--in early to middle 1997

Abarbanel, H.D. and U. Lall, 1996. Nonlinear dynamics of the Great Salt Lake: system identification and prediction. *Clim. Dyn.*, in press.

Casdagli, M , 1989: Nonlinear prediction of chaotic time series. *Physica D, 35*, 335-356.

Friedman, J.H., 1991: Multivariate adaptive regression splines *Ann Stat, 19*, 1-50.

Huang, J.P., Y.H. Yi, S.W. Wang and J.F. Chou, 1993 An analog-dynamic long-range numerical weather prediction system incorporating historical evolution. *Q J R Met. Soc., 119*, 547-565.

Jiang, N., M. Ghil and D. Neelin, 1995: Forecasts of equatorial Pacific SST using an autoregressive process using singular spectrum analysis. *Exp. Long-Lead Forcst. Bull., 4*, No. 1, 24-27.

Keppenne, C.L. and M. Ghil, 1992a: Forecasting extreme weather events. *Nature, 358*, 547.

Keppenne, C.L. and M. Ghil, 1992b: Adaptive Spectral Analysis and Prediction of the Southern Oscillation Index. *J. Geophys. Res., 97*, 20449-20554.

Keppenne, C.L. and M. Ghil, 1993: Adaptive filtering and prediction of noisy multi-variate signals: an application to atmospheric angular momentum. *Intl. J. Bifurcations and Chaos, 3*, 625-634.

Keppenne, C.L. and U. Lall, 1995: A new methodology to forecast paleoclimate time series with application to the Southern Oscillation index. EOS Trans AGU. 1995 Fall Meeting Supplement, 76, F327.

Keppenne, C.L. and U. Lall, 1996: Complex singular spectrum analysis and multivariate adaptive regression splines applied to forecasting the Southern Oscillation. *J. Clim.*, 9, submitted.

Lall, U. and A. Sharma, 1996: A nearest-neighbor bootstrap for resampling hydrologic time series. *Water Resources Res.*, in press.

Lall, U., T. Sangoyomi and H.D. Abarbanel, 1996: Nonlinear dynamics of the Great Salt Lake: nonparametric short term forecasting. *Water Resources Res.*, in press.

Lewis, P.A.W. and J.G. Stevens, 1991: Nonlinear modeling of time series using multivariate adaptive regression splines (MARS). *J. Amer. Stat. Assoc., 86*, 864-877.

Livezey, R.E., A.G. Barnston, G.V. Gruza and E.Y. Rankova, 1994: Comparative skill of 2 analog seasonal temperature prediction systems: Objective selection of predictors. *J. Clim., 7*, 608-615.

Toth, Z., 1991: Estimation of atmospheric predictability by circulation analogs. *Mon. Wea. Rev., 119*, 65-72.

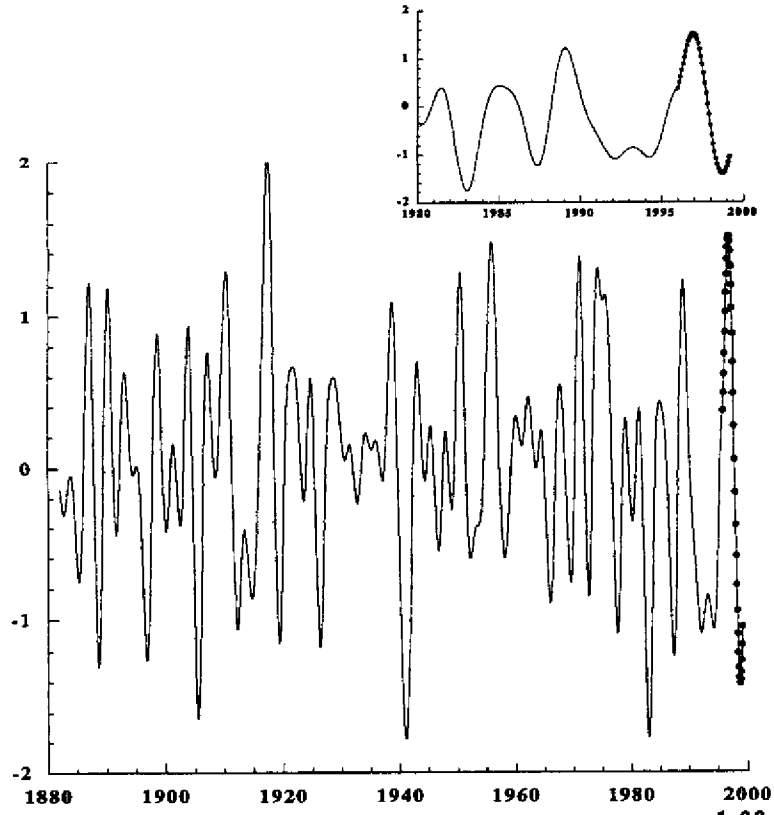


Fig. 1. *Adaptively filtered Southern Oscillation Index (SOI) time series resulting from the complex singular spectrum analysis (CSSA) of the monthly mean Darwin and Tahiti sea–level pressure (SLP) data through February 1996 (solid). Note the similarity between the two brief negative excursions of the filtered SOI following the strong La Nina event in the early 1910s and the recent conditions. The application of a variant of multivariate adaptive regression splines (MARS) to the real and imaginary parts of the leading four complex principal components (CPCs) resulting from the CSSA yields a forecast (circles on right side of curve) reminiscent of the conditions that dominated in the late 1910s (a strong La Nina) and illustrates MARS' capability to model the conditions of rare events.*

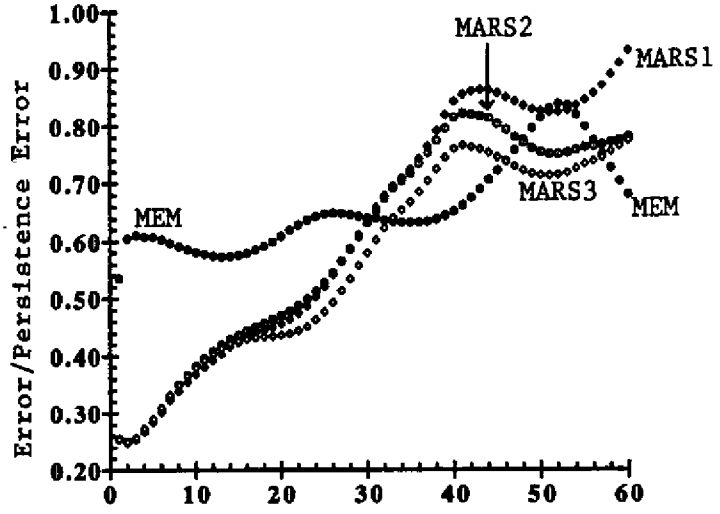


Fig. 2. *Ratio of the average forecast error of 60–month forecasts issued with either MEM (full circles), interaction–level–one MARS models (full diamonds), interaction–level–two MARS models (open circles) and interaction–level–three MARS models (open diamonds), to the average error of a same–lead persistence forecast. Shown is the forecast error for the adaptively filtered time series obtained by convoluting each of the leading four CPCs with the corresponding complex empirical orthogonal function (CEOF). For example, the average error of interaction–level–three MARS forecasts grows from about 0.25 times that of a persistence forecast at one–month lead to about 0.8 times it at 60–month lead.*

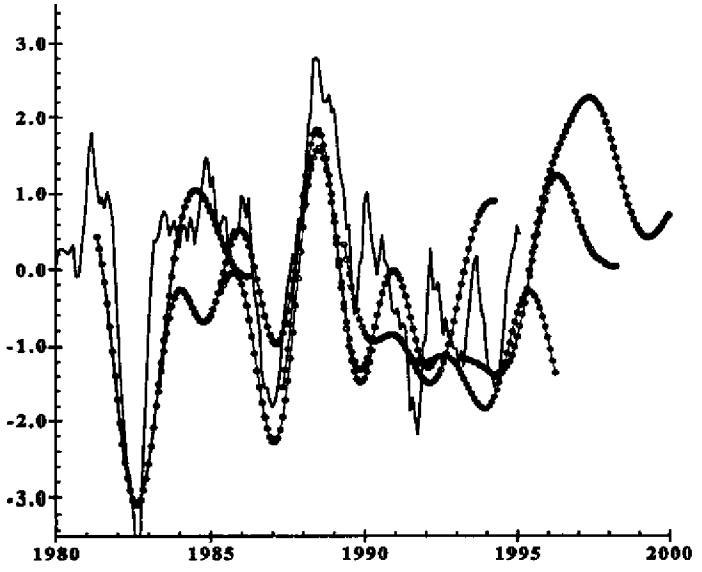


Fig. 3. *Five–month running–mean SOI (solid) and series of eight 60–month lead forecasts (series of connected circles) obtained by combining the forecasts resulting from the application of cubic MARS models to the real and imaginary parts of the leading four CPCs resulting from the Darwin and Tahiti data's CSSA. See text.*

Forecasts of Equatorial Pacific SST Anomalies Based on Singular Spectrum Analysis Combined with the Maximum Entropy Method

Ning Jiang, Michael Ghil and David Neelin

Department of Atmospheric Sciences and Institute of Geophysics and Planetary Physics
University of California, Los Angeles, California

Singular spectrum analysis (SSA: Vautard and Ghil 1989; Ghil and Vautard 1991; Plaut et al. 1995) and the maximum entropy method (MEM: Burg 1968; Penland et al. 1991) are used here for long-lead forecasts of the sea-surface temperature (SST) anomalies averaged over the Niño 3 area and the Southern Oscillation Index (SOI). The forecast is for up to one year ahead, based on the last 45 years of observed data. More detailed information on the forecast method based on single-channel SSA combined with MEM is given by Keppenne and Ghil (1992), while multi-channel SSA (M-SSA: Kimoto et al. 1991; Keppenne and Ghil 1993; Plaut and Vautard 1994) combined with MEM is documented in the March 1995 issue of this Bulletin (Jiang et al. 1995). Briefly, the time series is filtered first by SSA (if univariate) or M-SSA (if multivariate), so that the statistically significant components are retained, specifically the quasi-quadrennial (QQ) and the quasi-biennial (QB) components of ENSO variability (Rasmusson et al. 1990; Keppenne and Ghil 1992; Jiang et al. 1995). Then MEM is applied to advance these components in time.

Figure 1 shows area-averaged Niño 3 SSTAs, forecast and observed, since 1990, using the SSA and MSSA-MEM schemes for a 3-, 6-, 9- and 12-month lead. The last forecast, for the next 1-4 seasons, using data through January 1996, is shown in Fig. 2. The vertical bars are one standard deviation in length, based on forecast verification over the 1984-93 time span. The forecasts indicate that the presently slightly cooler than normal or near-normal conditions in Niño 3 may continue through 1996. Figure 3 shows the SSA-MEM forecast for the SOI from February 1996 through January 1997. The SOI is expected to remain close to its mean over this year. The present SOI forecast agrees with the Niño 3 SSTA forecast, although the anomalously weak (anti)correlation between these two ENSO signals, which we pointed out in the September 1995 issue of this Bulletin, continues.

Burg, J.P., 1968: Maximum entropy spectral analysis. *Modern Spectrum Analysis*, 34-48. IEEE Press.

Ghil, M. and R. Vautard, 1991: Interdecadal oscillations and the warming trend in global temperature time series. *Nature*, **350**, 324-327.

Jiang, N., D. Neelin and M. Ghil, 1995: Quasi-quadrennial and quasi-biennial variability in the equatorial Pacific. *Clim. Dyn.*, **12**, 101-112.

Jianag, N., M. Ghil and D. Neelin, 1995: Forecasts of Equatorial Pacific SST anomalies using an autoregressive process using singular spectrum analysis. *Experimental Long-Lead Forecast Bulletin*, **4**, **1**, 24-27.

Keppenne, C.L. and M. Ghil, 1992: Adaptive filtering and prediction of the Southern Oscillation Index. *J. Geophys. Res.*, **97**: 20449-20454.

Keppenne, C.L. and M. Ghil, 1993: Adaptive filtering and prediction of noisy multivariate signals Adaptive filtering and prediction of noisy multivariate signals: An application to subannual variability in atmospheric angular momentum. *Intl. J. Bif. & Chaos*, **3**, 625-634.

Kimoto, M., M. Ghil and K.C. Mo, 1991: Spatial structure of the 40-day oscillation in the Northern Hemisphere extratropics. *Proc. 8th Conf. Atmos. & Oceanic Waves & Stability*. Amer. Met. Soc., Boston, 115-116.

Penland, C., M. Ghil and K.M. Weickmann, 1991: Adaptive filtering and maximum entropy spectra, with application to changes in atmospheric angular momentum. *J. Geophys. Res.*, **96**, 22,659-22,671.

Plaut, G.R., M. Ghil and R. Vautard, 1995: Interannual and interdecadal variability in 335 years of central England temperature. *Science*, **268**, 710-713.

Rasmusson E.M., X. Wang and C. F. Ropelewski, 1990: The biennial component of ENSO variability. *J. Mar. Sys.*, **1**, 71-96.

Vautard R., and M. Ghil, 1989: Singular spectrum analysis in nonlinear dynamics with applications to paleoclimatic time series. *Physica D*, **35**, 395-424.

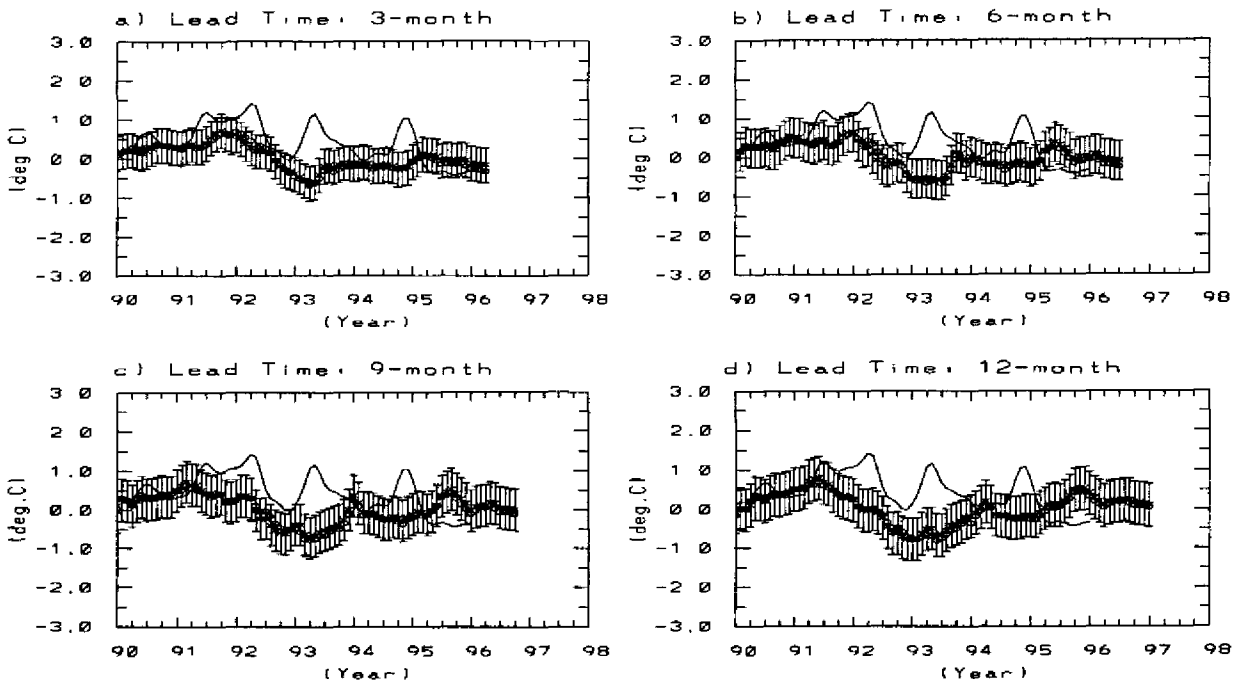


Fig. 1. *Forecasts of the area-averaged Niño 3 SST anomalies (SSTAs) by the SSA–MEM (star) and MSSA–MEM (circle) schemes. The solid line indicates the observed Niño 3 SSTAs. The latest forecast starts from January 1996, shown for lead times of (a) 3 months, (b) 6 months, (c) 9 months and (d) 12 months.*

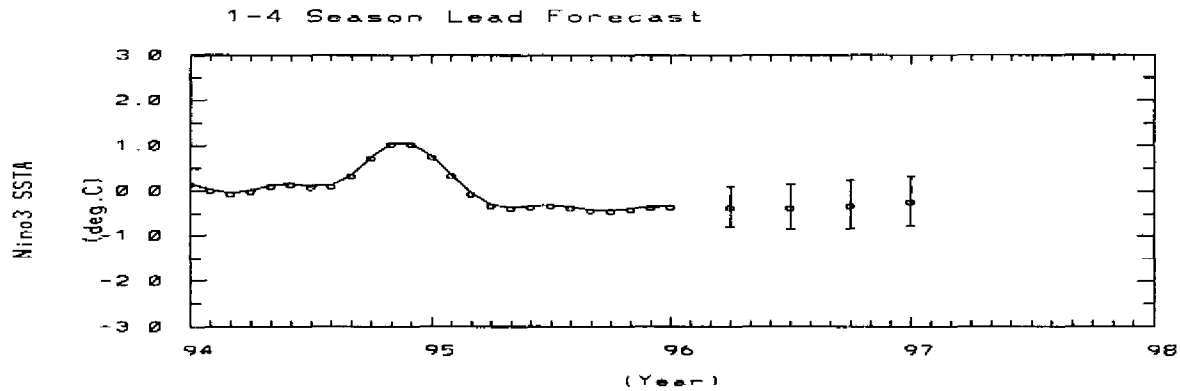


Fig. 2. *SSA–MEM forecasts of Nino–3 SSTAs for the upcoming 4 seasons, with a January 1996 start time. The solid line indicates the observed Nino–3 SSTAs.*

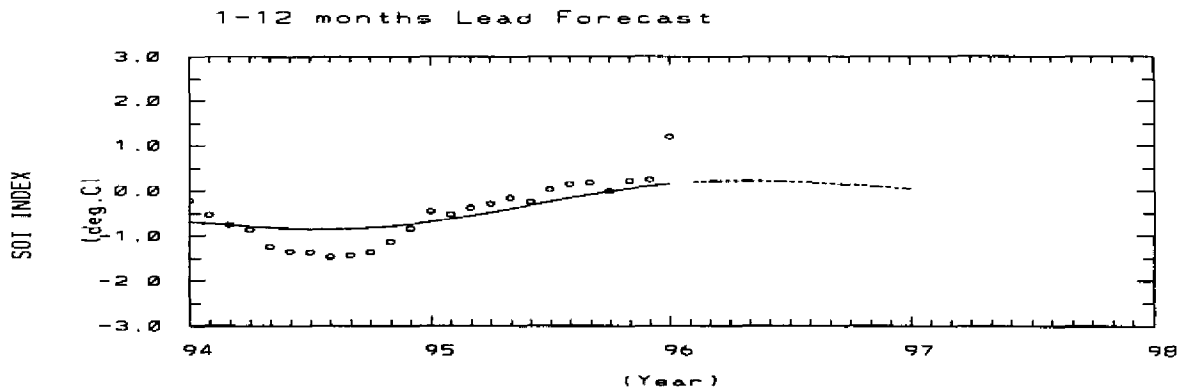


Fig. 3. *SSA–MEM forecast for the SOI for February 1996 through January 1997. The circles represent the raw monthly SOI values without the seasonal cycle, and the solid line the SSA–filtered SOI. The dashed line indicates the forecast for the coming 12 months.*

Canonical Correlation Analysis (CCA) Forecasts of Canadian Temperature and Precipitation -- Apr-May-Jun 1996

Contributed by Amir Shabbar

Atmospheric Environment Service, Environment Canada, Downsview, Ontario, Canada
ashabbar@ccrdp03.dow.on.doe.ca

In the last two issues of this Bulletin, forecasts of Canadian temperature and precipitation using the multivariate statistical technique of canonical correlation analysis (CCA) were presented. For Canada, we have developed the predictive relationships between evolving large scale patterns of quasi-global sea surface temperature, Northern Hemisphere 500 mb circulation, and the subsequent Canadian surface temperature and precipitation. In this issue we present the forecasts for Apr-Jun 1996 using the predictor fields through February 1995. These forecasts are made with a lead time of 4 months, where lead time is defined as the time between the end of the latest predictor season and the end of the predictand season. Further detail about the Canadian CCA-based seasonal climate prediction is found in Shabbar (1996a, 1996b) and Shabbar and Barnston (1996).

Figure 1(a) shows the CCA-based temperature forecast for the 3 month period of April-June 1996 expressed as standardized anomaly. Table 1 shows the value of the standard deviation in °C at selected stations. The mean skill over all 51 stations is given in the caption beneath each forecast map. The field significance is also shown, reflecting the probability of randomly obtaining overall map skill equal to or higher than that which actually occurred. Field significance is evaluated using a Monte Carlo procedure in which the forecast versus observation correspondences are shuffled randomly 1000 times. The field of cross-validated historical skill (correlation) for the forecast time period is shown in Figure 1(b). The forecast has a modest expected skill: a mean national score of 0.14 and a field significance of 0.082. The skill of the temperature forecast drops off considerably in spring in Canada (see the September 1995 issue of this Bulletin, page 28). Local skills are highest over the northern Canadian Prairies, and modest skill is found on the west coast of Canada. A large area of the country from the Rockies to Hudson Bay is expected to experience a negative temperature anomaly; positive temperature anomalies are forecast for both coasts.

Figure 2(a) shows the CCA-based precipitation forecast for the 3 month period of April-June 1996 expressed as a standardized anomaly. Table 1 shows the value of the standard deviation (mm) at selected stations. Cross-validated historical skill (correlation) for this time period is shown in Figure 2(b). The forecast has moderate expected skill: a mean national score of 0.14 and a field significance of 0.050. Local skills are highest over sections of the Prairies, over southwestern British Columbia and the east coast. Large areas of Canada extending from Saskatchewan to Quebec are expected to have a deficit in April-June precipitation. The largest deficit is forecast in central Saskatchewan. An excess in spring precipitation is expected over central regions of British Columbia and the east coast of Canada.

Following the normal evolution of the current cold ENSO episode, some models are projecting a return to normal conditions in the central and eastern equatorial Pacific by the middle of 1996. The April-June '96 forecast recognizes the near future demise of the current cold ENSO episode which started in the middle of last year, and also reflects the decreasing influence of ENSO in the warm half of the year as compared with winter and the first half of spring. Nonetheless, the April-June forecast reflects a strong component of persistence from the winter circulation pattern.

Table 1. Standard deviation of temperature (Temp) and precipitation (Prcp) for the 3 month period April through June at selected Canadian stations.

Station	Temp (°C)	Prcp (mm)
Whitehorse	1.6	13.2
Fort Smith	2.5	19.5
Innujjuak	1.9	18.2
Eureka	2.6	3.5
Vancouver	1.3	26.6
Edmonton	1.7	26.3
Regina	2.1	30.9
Winnipeg	2.2	37.4
Churchill	2.1	24.6
Moosonee	1.9	27.6
Toronto	1.6	30.0
Quebec City	1.3	35.3
Halifax	1.2	42.7
St. John's	1.6	46.6

Shabbar, A., 1996a: Seasonal prediction of Canadian surface temperature and precipitation by canonical correlation analysis. *Proceedings of the 20th Annual Climate Diagnostics Workshop*, Seattle, Washington, October 23-27, 1995, in press.

Shabbar, A., 1996b: Seasonal forecast of Canadian surface temperature by canonical correlation analysis. *13th Conference on Porbability and Statistics in the Atmospheric Sciences*. American Meteorological Society, San Francisco, California, February 21-23, 339-342.

Shabbar, A. and A.G. Barnston, 1996: Prediction of Canadian seasonal temperature and precipitation using canonical correlation analysis. *Mon. Wea. Rev., 124*, accepted.

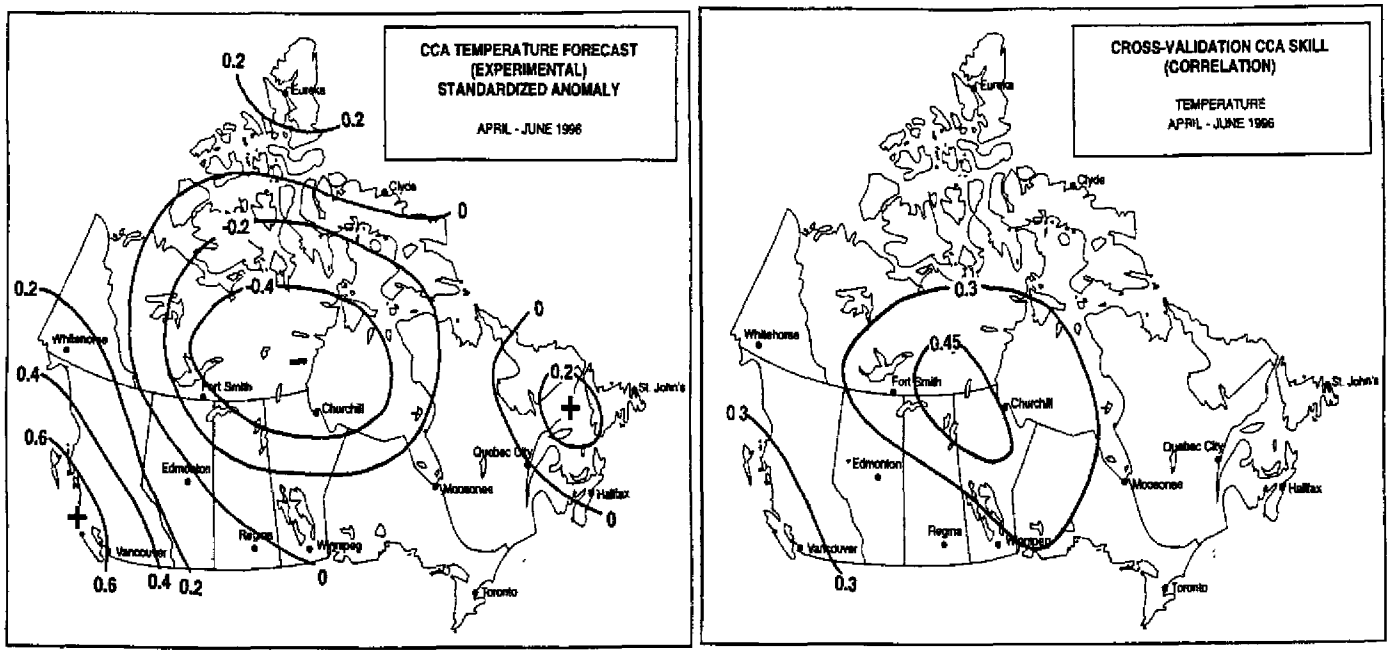


Fig. 1. *Panel (a): CCA–basedtemperature forecast for the 3 month mean period of Apr–Jun1996. Forecasts are represented as standardized anomalies. Panel (b): Geographical distribution of cross–validated historical skill for the forecast shown in (a), calculated as a temporal correlation coefficient between forecasts and observations. Areas having forecast skill of 0.30 or higher are considered to have utility. The mean score over 51 stations is 0.14. Field significance is 0.08.*

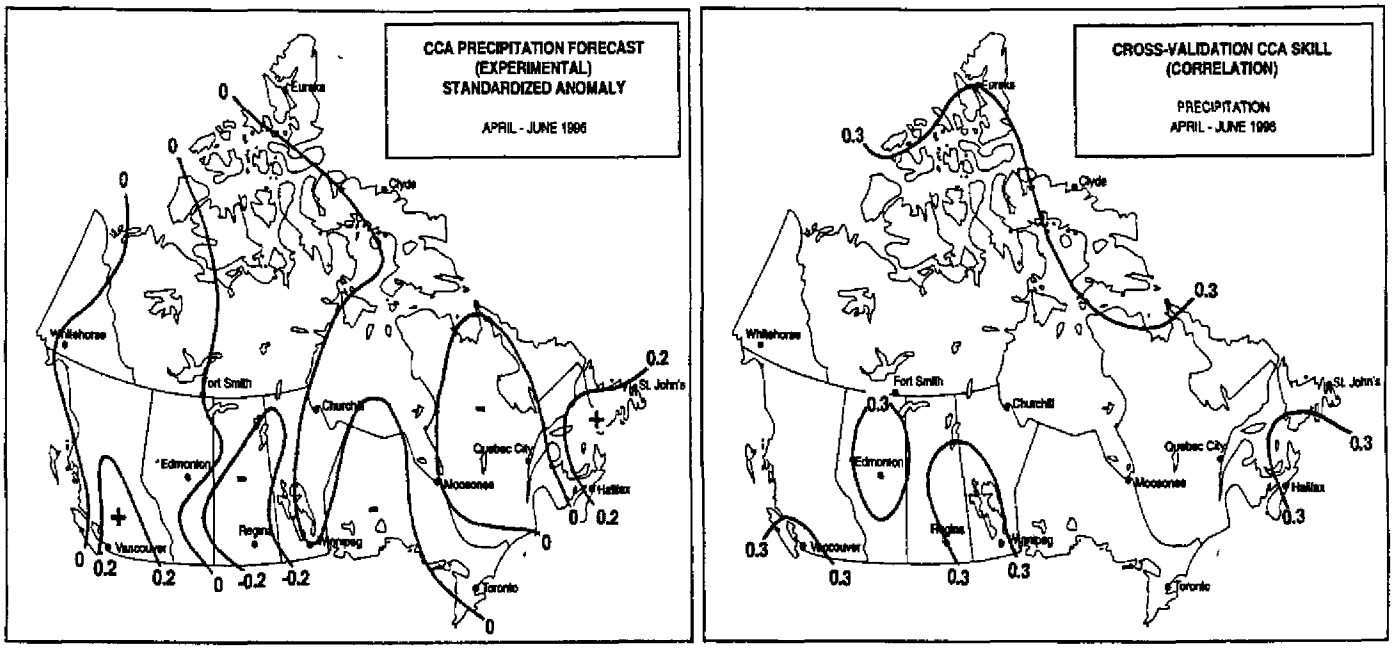


Fig. 2. *Panel (a): As in Fig. 1a (CCA anomaly forecast), except for Apr–Jun1996 precipitation. Panel (b): As in Fig. 1b (skill for the forecast shown in [a]), except for precipitation. The mean score over 69 stations is 0.14. Field significance is 0.05 (see text).*

Precipitation Forecasts for the Tropical Pacific Islands Using Canonical Correlation Analysis (CCA)

contributed by Yuxiang He and Anthony Barnston

Climate Prediction Center, NOAA, Camp Springs, Maryland

In canonical correlation analysis (CCA), relationships between multicomponent predictors and multicomponent predictands are linearly modeled. These typically take the form of pattern-to-pattern relationships in space and/or time. CCA is designed to minimize squared error in hindcasting linear combinations of predictand elements from linear combinations of the predictor elements.

CCA has been used in the social sciences for many decades, but only in the last 10 years has it begun being used in the atmospheric sciences. For example, Barnett and Preisendorfer (1987) applied CCA to monthly and seasonal prediction of U.S. temperature. Graham et al. (1987a,b) and Barnston and Ropelewski (1992) applied it to predicting aspects of the ENSO phenomenon, and Barnston (1994) forecasted short-term climate anomalies in the Northern Hemisphere. Recently, Barnston and He (1996) explored CCA as a tool for seasonal temperature and precipitation forecasts for Hawaii and Alaska. The skills resulting from the latter two studies, while generally modest, were high enough for the U.S. National Weather Service to use the forecasts on a real-time, operational basis.

Here, CCA is used to predict 3-month total precipitation anomalies in the Pacific Islands out to a year in advance, as described in He and Barnston (1996). It is known from past work that rainfall in the tropical and subtropical Pacific is strongly related to ENSO (Ropelewski and Halpert 1987). Therefore it is worthwhile to set up a seasonal prediction system that produces real-time forecasts on a monthly basis for the benefit of agricultural and commercial interests in the Pacific Islands.

The predictor fields used for the forecasts include quasi-global sea surface temperature (SST), Northern Hemisphere 700 mb geopotential height, and the precipitation itself over the 33 stations used as the predictand. Experiments with different subsets of predictors and predictor field weights showed that the most valuable predictor field is SST, with 700 mb heights and prior precipitation somewhat helpful. The SST predictors are therefore given double their natural weight. Further details about the skills, the underlying relationships, and the need to weight the SST double are provided in He and Barnston (1996). The set of predictors is configured as four consecutive 3-month periods prior to the time of the forecast, followed by a variable lead time, and then a single 3-month predictand period. The predictand includes 3-month total rainfall at 33 Pacific Island stations within 25°N-30°S, including 4 Hawaiian stations. The lead time is defined as the time between the end of the final (fourth) predictor period (i.e., the time of the forecast) and the *beginning* of the 3-month predictand period. This strict definition contrasts with that in which the shortest lead forecast would be called 3-month lead instead of zero lead.

The expected skill of the forecasts was estimated using cross-validation, in which each year in turn was held out of the model development sample and used as the forecast target. These skill estimates indicated that at 1 month lead time the highest correlation skill across the Pacific Islands occurs in Jan-Feb-Mar at 0.44 (0.29) averaged over all stations north (south) of the equator, and the lowest occurs from September through December at about 0.15 (0.30) for stations north (south) of the equator. At four months lead skills are only slightly lower except for the Jan-Feb-Mar average skill north of the equator which drops significantly to 0.26.

Figure 1a shows standardized precipitation anomaly forecasts for 33 Pacific Island stations for Jul-Aug-Sep 1996 made using data through Feb 1996 (3 months lead). The geographical distribution of expected skill for this forecast, based on cross-validation, is shown in part (b) in terms of a correlation between forecasts and observations. The forecasts are fairly weak in amplitude. However, a tendency toward positive rainfall anomalies is noted north of 10°N. While this response agrees with the findings of Ropelewski and Halpert (1987) for the cold phase of ENSO, that response was found to be limited to the cold half of the Northern Hemisphere year. However, Barnston and He (1996) showed that the expected effects from either phase of ENSO may continue for several additional seasons in Hawaii. This delay may be caused partly by lingering SST anomalies off the equator at higher tropical latitudes. Presently the east-west band of negative SST anomalies along the equator in the central and eastern Pacific has not expanded north of 10°N, except near Central America.

More detailed forecasts for 9 U.S -affiliated Pacific Island stations, located as shown in Fig. 2, are provided in Fig. 3. In the latter figure, long-lead rainfall forecasts from 1 to 13 seasons lead are shown (solid bars), along with their expected skills (lines). The horizontal axis reflects the lead time, whose corresponding actual target period for this forecast is indicated in the legend along the top of the figure (e.g. 1=AMJ 1996). The same ordinate scale is used for both forecasts and skills (standardized anomaly and correlation, respectively).

The skill curve applies to the target season for the associated lead time of the present forecast. Sometimes a "return of skill" occurs as the lead is increased because a more forecastable target season has been reached. The forecasts and their skills differ as a result of both location differences within the Pacific basin and differences in orientation with respect to the local orography (if any) and subsequent exposure to the prevailing low-level wind flow. We note that at most stations no substantial anomalies in either direction are predicted in the next few months. It should also be noted that the expected skill for the boreal warm half of the year is generally relatively low. However, enhanced rainfall is predicted with modest but usable skill at Johnston Island this summer. At longer lead, a tendency for dryness is noted for boreal winter 1996-97 at Wake, Yap and Johnston Islands. While the associated skills are not high enough to react with concern at this point, expected skill will slowly rise as the lead time decreases.

The CCA modes (not shown) emphasize ENSO as the leading influence on tropical Pacific climate, but most strongly during the months of Nov-Dec-Jan-Feb-Mar-Apr-May (and even earlier than Nov along the immediate equator near and somewhat east of the dateline). Mild to moderate cold episode conditions have now prevailed for about 9 months. Their effect on the forecasts has begun overshadowing that of the long warmish period that ended in spring 1995, although the forecast magnitudes are weak. Another important mode is a long-term trend related to a warming of the global tropical SST. This mode can cause the forecasts to repeat from one year to the next at given times of the year, and may govern the forecasts by a higher proportion in the northern summer when ENSO's influence is diminished at many of the off-equator stations.

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.

Barnston, A.G., 1994: Linear statistical short-term climate predictive skill in the Northern Hemisphere. *J. Climate*, **5**, 1514-1564.

Barnston, A.G. and C.F. Ropelewski, 1992: Prediction of ENSO episodes using canonical correlation analysis. *J. Climate*, **7**, 1316-1345.

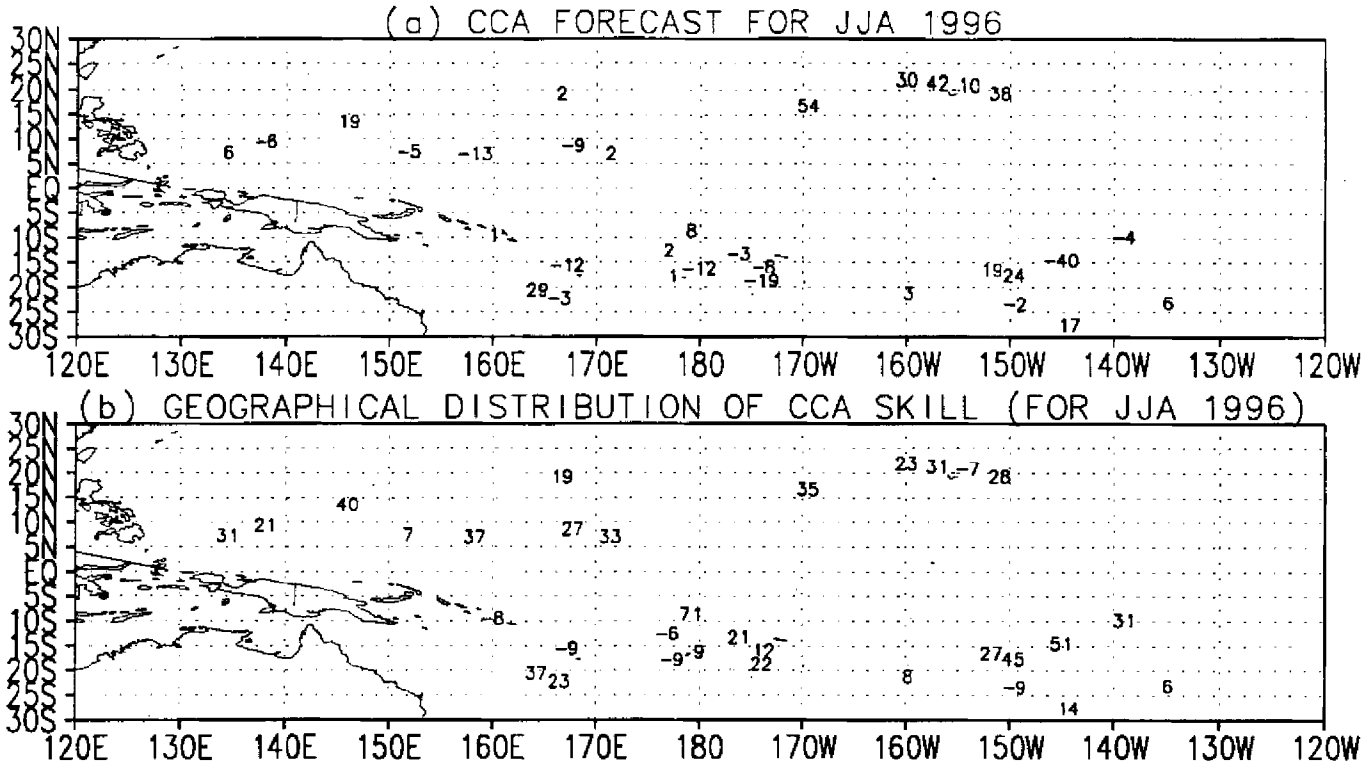
Barnston, A.G. and Y. He, 1996: Skill of CCA forecasts of 3-month mean surface climate in Hawaii and Alaska. *J. Climate*, **9**, submitted..

Graham, N.E., J. Michaelsen and T. Barnett, 1987a: An investigation of the El Nino-Southern Oscillation cycle with statistical models. 1. Predictor field characteristics. *J. Geophys. Res.*, **92**, 14251-14270.

Graham, N.E., J. Machaelsen and T. Barnett, 1987b: An investigation of the El Nino-Southern Oscillation cycle with statistical models. 2. Model results. *J. Geophys. Res.*, **92**, 14271-14289.

He, Y. and A.G. Barnston, 1996: Long-lead forecasts of seasonal precipitation in the tropical Pacific islands Using CCA. *J. Climate*, **9**, in press.

Ropelewski, C.F. and M.S. Halpert, 1987: Global and regional scale precipitation patterns associated with the El Nino/Southern Oscillation. *Mon. Wea. Rev.*, **115**, 1606-1626.



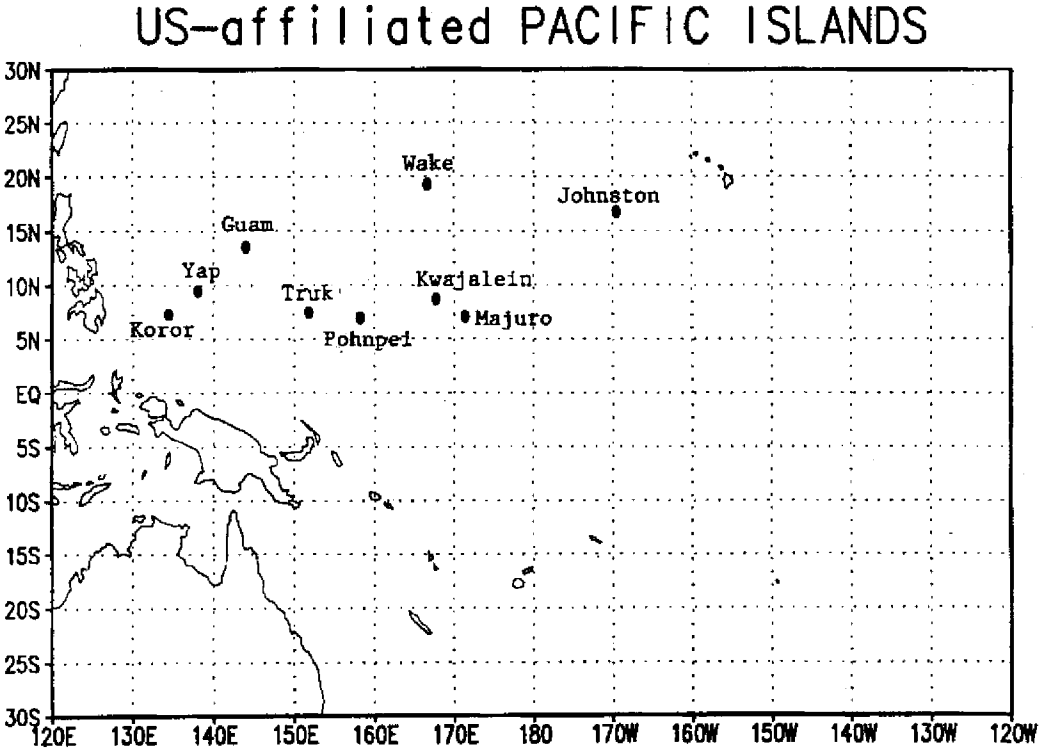


Fig. 1 (above). *(a): CCA-derivedprecipitation standardized anomaly forecast (X100) for 33 Pacific Islands stations for Jun-Jul-Aug1996 made at 3 months lead (latest data February 1996). (b): The cross-validated skill expected for the forecast shown in (a), expressed as a correlation X100.*

Fig. 2. *Locations of the 9 U.S.-affiliated Pacific Island stations whose long-lead precipitation forecasts are shown in detail in Fig. 3.*

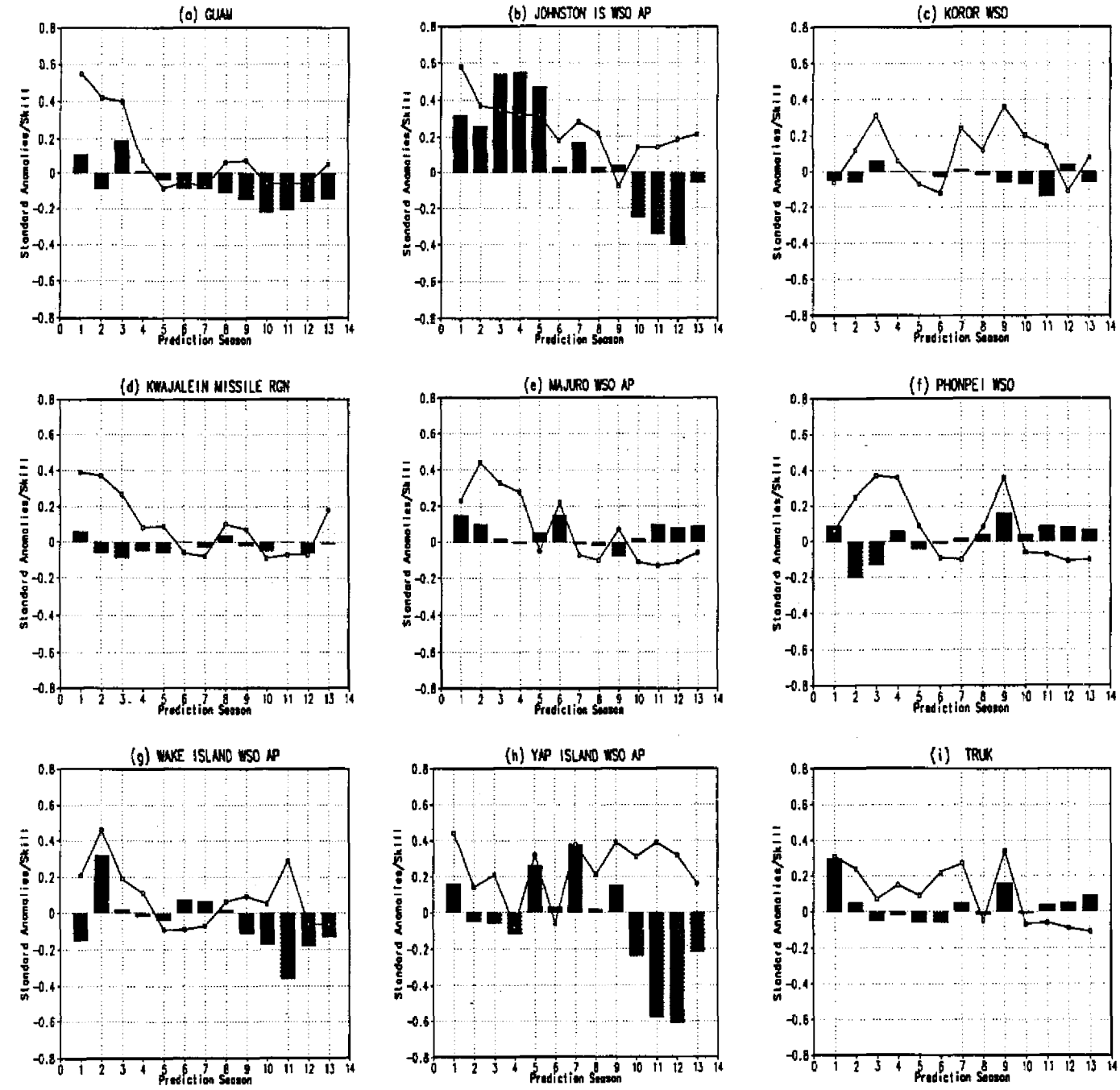


Fig. 3. *Time series of CCA-based long-lead precipitation anomaly forecasts, and their expected skills, out to one year into the future for 9 U.S.-affiliated Pacific Island stations (see Fig. 2). The bars indicate the forecast values (as standardized anomalies) and the lines indicate the associated skills (as correlation coefficients). Both forecasts and skills use the same ordinate scale. The target season is indicated on the abscissa, ranging from 1 (Apr-May-Jun 1996) through 13 (Apr-May-Jun 1997); see the legend at top.*