Nonlinear canonical correlation analysis of the tropical Pacific wind stress and sea surface temperature

Aiming Wu and **William W. Hsieh**
Dept. of Earth and Ocean Sciences, University of British Columbia
Vancouver, B.C., V6T 1Z4, Canada
Tel: (604) 822-5691, Fax: (604) 822-6091
Email: awu@eos.ubc.ca

Submitted to *Climate Dynamics*
Revised Apr. 2002

Abstract

Nonlinear canonical correlation analysis (NLCCA) formulated by a neural network approach was applied to the monthly surface wind stress (WS) and sea surface temperature (SST) in the tropical Pacific. The strength of the nonlinearity varies with the lead/lag time between WS and SST. Compared to the CCA modes, the NLCCA modes explain more variance of the two sets of variables and have higher canonical correlations, particularly, at longer lead/lag times. Unlike the CCA, the NLCCA modes are capable of capturing the asymmetry in the spatial patterns between El Niño and La Niña episodes in both SST and WS fields. With the WS lagging and then leading the SST, the roles of the predictor field and the lagging response field were interchanged— the spatial asymmetry was found to be considerably stronger in the response field than in the predictor field. Hindcasts for SST (using WS as predictor) shows that the NLCCA model is generally slightly better than the CCA model in terms of the correlation skills and the root mean square error (RMSE), mainly in the eastern equatorial regions (e.g. Niño12 and Niño3). Hindcasts for WS (using SST as predictor) shows that despite the relatively rapid decrease of skills with the lead time (compared to the skills of SST prediction from WS), the NLCCA is slightly better than the CCA model, with greater improvement over the western equatorial Pacific.

1 Introduction

Canonical correlation analysis (CCA), a multivariate statistical technique (Hotelling 1936; Anderson 1984), has become widely used in meteorology and oceanography (von Storch and Zwiers 1999). Given two sets of variables $\mathbf{x}$ and $\mathbf{y}$, CCA is used to extract the correlated modes between $\mathbf{x}$ and $\mathbf{y}$ by looking for linear combinations

$$u = \mathbf{a} \cdot \mathbf{x} \quad \text{and} \quad v = \mathbf{b} \cdot \mathbf{y} \tag{1}$$

where the canonical variates u and v have maximum correlation, i.e. the weight vectors $\mathbf{a}$ and $\mathbf{b}$ are chosen such that the Pearson correlation coefficient between u and v, is maximized. CCA is routinely used in seasonal climate prediction (Barnston and Ropelewski 1992; Shabbar and Barnston 1996).

Recently, different approaches to performing nonlinear canonical correlation analysis (NLCCA) have been proposed (Lai and Fyfe 1999, 2000; Hsieh 2000). Hsieh (2000) used a neural network (NN) approach to find general nonlinear continuous functions mapping $\mathbf{x}$ to u, and $\mathbf{y}$ to v. This method has been applied to study the relation between the tropical Pacific sea level pressure (SLP) and sea surface temperature (SST) fields (Hsieh 2001a). The resulting NLCCA modes exhibited some nonlinearity with decadal dependence. The tropical Pacific has also been studied by the nonlinear principal component analysis (NLPCA) method by Monahan (2001) and Hsieh (2001b). Codes for both NLPCA and NLCCA are downloadable from the Web site http://www.ocgy.ubc.ca/projects/clim.pred.

As only simultaneous SLP and SST fields have been studied by Hsieh (2001a), the objective of this paper is to study the nonlinear air-sea interactions between the wind stress (WS) and the SST fields over the tropical Pacific at various lead/lag times up to one year using the NLCCA. This paper is organized as follows: The data used are described in Section 2. The NLCCA mode 1 with WS lagging SST (Section 3) and WS leading SST (Section 4) are then presented, followed by the spatial anomaly patterns of the NLCCA mode 1 (Section 5). NLCCA mode 2 is described in Section 6. Section 7 gives a summary and conclusion, followed by an Appendix, where hindcasts for SST (using WS as predictors), and for WS (using SST as predictors) by the NLCCA and the CCA are compared.

2 Data

The monthly WS came from the Florida State University (FSU) stress analyses (Stricherz et al. 1997). The data period is January 1961 through December 2000 covering the whole tropical Pacific from 124°E to 70°W, 29°S to 29°N with a grid of 2° by 2°. The monthly SST came from the reconstructed historical SST data sets by Smith et al. (1996) covering the period of January 1950 to December 2000 with a resolution of 2° by 2° over the global oceans. Monthly WS and SST anomalies were calculated by subtracting the climatological monthly means. The WS anomalies (for both zonal and meridional components) were then smoothed by a 3-month running average.

Prior to NLCCA, traditional principal component analysis (PCA), also called empirical orthogonal function (EOF) analysis, was performed on the WS and SST anomalies over the tropical Pacific (124°E-70°W, 21°S-21°N) to compress the data. For WS, a combined EOF was applied to the zonal and meridional components of the WS anomalies (τ_x τ_y). This pre-processing to reduce the data to a manageable size is common in rotated PCA (Barnston and Livezey 1987) and CCA (Barnett and Preisendorfer 1987). Variance contributions from the 6 leading modes of WS are 14.83%, 11.13%, 7.84%, 4.98%, 3.54% and 2.88%, respectively, and for SST, 57.24%, 10.93%, 8.57%, 3.48%, 2.93% and 2.38%, respectively. The fractional amount of variance captured by the first 6 WS modes is considerably less than that by the first 6 SST modes, as there is relatively more energy in small-scale fluctuations in the WS. Since we are only interested in large-scale features such as the El Niño-Southern Oscillation, the neglect of the higher WS modes containing smaller scale features does not pose a significant problem.

The 6 leading principal components (PCs, i.e. the EOF time series) of WS and SST from January 1961 to December 2000 are then used as the inputs to the NLCCA model (Fig. 1). Fig. 2 shows the first 3 EOFs, i.e. the spatial patterns, of the WS (here only the zonal component is shown) and the SST. During the positive phase, the WS EOF1 depicts a westerly anomaly patch over the central equatorial Pacific, from 150°E to 120°W, approximately centred at 175°W, 3°S (Fig. 2a). The WS EOF2 is basically uniform in sign over the whole domain (Fig. 2b), while the WS EOF3 displays a northwest-southeast contrast across the equator (Fig. 2c). Correspondingly, the SST EOF1 shows a general El Niño pattern (Fig. 2d), with positive anomalies over the eastern-central equatorial Pacific. More concentrated SST anomalies over the eastern equatorial Pacific is revealed by the EOF2 (Fig. 2e) and an isolated SST anomaly over the central equatorial Pacific by the EOF3 (Fig. 2f). We will see how these patterns or combinations of these patterns will appear in the CCA or NLCCA modes in the following sections.

3 NLCCA mode 1 with WS lagging SST

The first NLCCA mode consists of a curve in the 6-dimensional WS PC-space and another curve in the 6-dimensional SST PC-space. Difficulties in visualizing curves in a 6-dimensional space render it necessary that we focus mainly on the solution projected onto the WS PC_1-PC_2 plane and onto the SST PC_1-PC_2 plane. Fig. 3 shows the NLCCA mode 1 solutions with WS lagging SST by 0, 3, 6, 9 and 12 months. Unlike the 5 curves for the SST modes (Fig. 3b), the 5 curves for the WS modes are dispersed (Fig. 3a), as are the CCA WS modes (Fig. 3c). Compared to the CCA modes, the curvature of the NLCCA modes varies with the lag time. The nonlinearity (i.e. curvature) is moderate at 0 month, very weak at 3 and 6 months, but increases at 9 and 12 months lag. Both the NLCCA and CCA modes rotate about 90° anticlockwise as lag varies from 0 to 12 months, indicating that the oscillations have changed from PC_1 dominated, to ones where PC_2 has an increasing role.

As the 5 curves for the SST modes are close to each other (Fig. 3b), the curvature of the NLCCA modes for the SST does not change much as the lag time varies, indicating that the nonlinearity in the SST is not dependent on the lag. The curves in Fig. 3b link the cool La Niña (also called El Viejo) states on the left to the warm El Niño states in the lower right corner.

Projections onto the PC_1-PC_3 planes are shown in Fig. 4, where the curves of the WS modes shrink with increased curvature (i.e. increased nonlinearity) as the lag increases, indicating that the PC_3 is becoming relatively more important. In contrast, the NLCCA SST modes rotate about 45° clockwise with weakening curvature (nonlinearity) as the lag increases from 0 to 12 months .

4 NLCCA mode 1 with WS leading SST

We interchange the roles of the predictor field and the response field between WS and SST by changing from WS lagging SST to WS leading the SST. Fig. 5 shows the first NLCCA modes with WS leading SST by 0, 3, 6, 9 and 12 months, as projected onto the PC_1-PC_2 planes. The curves of the WS modes at 0, 3 and 6-month lead are close to each other, but different from those for 9 and 12-month lead (Fig. 5a), which can also been seen in the corresponding CCA modes (Fig. 5c). The curves of the SST modes maintain a hump shape, getting increasingly curved as the lead increases from 0 to 12 months (Fig. 5b). The CCA SST modes also slightly rotate anticlockwise with lead time (Fig. 5d).

Projections of the NLCCA modes onto the PC_1-PC_3 planes are shown in Fig. 6. The clockwise rotation of the SST modes with increasing lead reveals that at 12-month lead, the NLCCA oscillation is more along PC_3 than along PC_1.

In summary, at different lag/lead times, the correlated modes of both WS and SST are different, as manifested by the varying orientation of the CCA solutions (Figs. 3c,d and Figs. 5c,d), as well as the changing orientation and curvature of the NLCCA solutions. Generally, the NLCCA modes possess rotations similar to the CCA modes as the lag/lead time varies. The curvature and its variations in the NLCCA solutions indicates that the NLCCA modes for both WS and SST are generally nonlinear, with the degree of nonlinearity changing with the lag/lead time between the two variables.

5 Spatial patterns of the NLCCA mode 1

For a given value of the canonical variate u, one can map from u to the 6 PCs of the WS at the output layer (of the network in the top right corner of Fig. 1). Each of the PCs can be multiplied by its associated eigenvector (i.e. the EOF spatial pattern), and the six modes added together gives the spatial anomaly pattern for that value of u. The zonal WS anomalies corresponding to minimum u and maximum u for the NLCCA modes are shown in Fig. 7. Corresponding to minimum u, negative (easterly) anomalies appears over the central-western equatorial Pacific (Fig. 7a), slightly slanted towards the northwest-southeast direction, resembling the EOF1 pattern (Fig. 2a). Corresponding to maximum u, positive (westerly) anomalies dominates over the central-western equatorial Pacific (Fig. 7b). At lead times of 3-12 months (WS leading SST), the spatial patterns are similar to the EOF1, while at leads of 0 to -12 months, the westerly anomalies shift eastward, with easterly anomalies appearing over the northwestern equatorial Pacific, resembling the pattern of EOF3 (Fig. 2c), or more precisely, a combination of EOF1 and EOF3 (as expected from Fig. 4a). As a result, the spatial patterns of WS anomalies are quite asymmetric on the opposite extremes of u, as the equatorial westerly anomalies (Fig. 7b) are further east than the equatorial easterly anomalies (Fig. 7a). The asymmetry is much more apparent when WS lags SST (the top 4 rows in Fig. 7), i.e., when WS is the response field. The zonal displacement between the positive center and the negative center reaches nearly 30° longitude. The magnitude of the westerly/easterly anomalies decreases rapidly with increasing lag time, but relatively slowly with increasing lead time— probably indicative of the relatively short 'memory' of the wind stress response to SST anomalies.

For comparison, the WS spatial patterns for the CCA modes are shown in Fig. 8, where we can see some features similar to the NLCCA modes, e.g., the eastward shift of the equatorial anomalies as the lead time decreases. All anomaly patterns of the CCA modes for lead time from -6 to 12 months resemble the EOF1 pattern (Fig. 2a), despite the zonal shifts as the lead time varies. Furthermore, the spatial patterns on opposite extremes of u are completely symmetric, i.e. mirror images, though the amplitudes may differ by a constant at each lead time. Compared to the corresponding CCA modes, the westerly anomalies for the NLCCA modes are relatively stronger with their center located further east (Fig. 7b).

Similarly, anomaly patterns of SST associated with some specific values of v are calculated. Here the values of v are chosen at the time when u takes its minimum value or maximum value, For leads not exceeding 9 months, as WS shifts from extreme easterly anomalies to extreme westerly anomalies, the SST field varies from strong La Niña states to strong El Niño states. For the NLCCA modes (Fig. 9), the positive SST anomalies are basically distributed over the eastern equatorial Pacific, especially off the west coast of Southern America, while negative SST anomalies has various locations depending on the lead/lag times. When SST is a forcing field, i.e. WS lags SST in time, the negative SST anomalies appear over the eastern Pacific. As the lag time decreases and lead time increases (SST becomes the response field), the negative SST anomalies move to the central Pacific, illustrating the asymmetry between the La Niña states and the El Niño states. In Fig. 9b, at lead times of 9 and 12 months, the SST anomalies can no longer be accounted for mainly by EOF1, so higher EOF modes are needed. In particular, the EOF3 pattern (Fig. 2f) can be seen in Fig. 9b at the lead time of 12 months. In general, Fig. 9 reveals that the spatial asymmetry between the two extremes of the oscillation is greater when SST is a response field (WS leading SST) than when SST is a predictor field (WS lagging SST).

In Figs. 3b and 5b, the large positive PC_1 values are accompanied by large negative values of PC_2, which gives stronger positive SST anomalies off Peru (Fig. 9b) than the CCA modes (not shown). Meanwhile, the large negative PC_1 values are accompanied by negative values of PC_2 in the NLCCA modes, resulting in weakened negative SST anomalies over the eastern Pacific and strengthened negative SST anomalies over the central Pacific (Fig. 9a), relative to the CCA modes. The CCA modes (not shown) give symmetric SST anomaly patterns for the La Niña and El Niño phases.

Extreme situations have been discussed above. For a more general view of the spatial patterns, WS and SST anomaly fields were reconstructed from the NLCCA mode 1 and CCA mode 1 at zero lead time. Time-longitude cross-sections along the equator (averaged between 5°S and 5°N) from January 1981 to December 2000 were examined. For the CCA mode, both westerly and easterly anomalies are centred at 175°W, and positive and negative SST anomalies, approximately at 120°W. For the NLCCA mode, the easterly anomalies are centered at 175°W, while the westerly anomalies are centered about 10° further east ; the negative SST anomalies are mainly over the central equatorial Pacific (centered at 150°W), while the positive anomalies are over the eastern equatorial Pacific (centered at 90-120°W). Meanwhile, the positive SST anomalies during the 1982-83 and 1997-98 strong El Niño events from the NLCCA are much stronger than those from the CCA, indicating that the NLCCA mode is closer to reality.

The mean square error (MSE) of the first NLCCA mode is less than that of the first CCA mode at all lead times from -12 to 12 months, for both WS (Fig. 10a) and SST (Fig. 10b). Similarly, the explained variance by the first NLCCA mode is higher than that by the CCA mode at all lead times for WS (Fig. 10c) and SST (Fig. 10d). The canonical correlation (i.e. the correlation between the canonical variates u and v) of the NLCCA mode is higher than that of the CCA mode, especially at longer lead/lag times (Fig. 10e). The ratio between the MSE of the NLCCA mode and that of the corresponding CCA mode indicates how far the nonlinear mode deviates from the linear mode (Fig. 10f)— the greater the nonlinearity, the smaller the ratio. This ratio varies with the lead time, and is generally smaller for the SST than for the WS, indicating greater nonlinearity in the SST.

6 NLCCA mode 2

After the first NLCCA mode had been subtracted from the data, the residual was served as input to the NLCCA network again to extract the second mode. Interestingly, the typical El Niño-La Niña patterns which were missing from the NLCCA mode 1 SST patterns at leads of 9 and 12 months (Fig. 9b) are clearly manifested in the NLCCA mode 2 SST at leads of 9 and 12 months (not shown), with the corresponding WS patterns strongly resembling the WS EOF3 pattern (Fig. 2c).

In general, the MSE of the NLCCA mode 2 is less than that of the CCA mode 2; and the variance explained by the NLCCA mode 2 higher than that by the CCA mode 2. For some leads, the variance explained by the NLCCA mode 2 can be higher than that by the NLCCA mode 1. This occurred in the SST at leads of 9 and 12 months, where as mentioned above, the familiar El Niño-La Niña patterns were manifested in NLCCA mode 2 instead of in NLCCA mode 1. Thus at long lead times, the correlation between the WS and SST during the El Niño-La Niña oscillation may not be strong enough to be picked up by the first NLCCA mode.

The canonical correlation of the NLCCA mode 2 was found to be higher than that of the CCA mode 2, especially at leads from -6 to 6 months. That the NLCCA modes 1 and 2 have higher canonical correlations and smaller MSE than the corresponding CCA modes at almost all lead times (from -12 to $+12$ months) raises the possibility of improving ENSO predictions using the NLCCA models. Hindcast results are presented in the Appendix.

7 Summary and conclusion

In this paper, the NLCCA model of Hsieh (2000, 2001a) was applied to the surface wind stress and the sea surface temperature over the tropical Pacific. Nonlinearity can generally be detected in the NLCCA modes for both WS and SST, with the SST tending to be more nonlinear than the WS. The asymmetry between the warm El Niño states and cool La Niña states was well modelled by the first NLCCA mode, where westerly anomalies and positive SST anomalies are located further east than the easterly anomalies and negative SST anomalies, while the first CCA mode is incapable of modelling the spatial asymmetry. With the WS lagging and then leading the SST, we interchanged the roles of the predictor field and the response field between WS and SST. The spatial asymmetry is much more apparent in the response field than in the predictor field. Compared to the CCA modes, the NLCCA modes explain more variance of the two sets of variables and have higher canonical correlations, particularly at longer lead/lag time. The degree of nonlinearity varies with the lead/lag time between WS and SST.

Hindcast skills were then compared in the Appendix between the NLCCA model and CCA model. Predictions for SST (using WS as predictor) show that the NLCCA model slightly outperformed the CCA model at 0-12 months of lead time over the equatorial Pacific, mainly in the east (e.g. Niño12 and Niño3 areas). WS predictions (using SST as predictor) from the NLCCA model were also found to be slightly better than those from the CCA model, with more improvement occuring over the western equatorial Pacific. Both NLCCA modes 1 and 2 contributed to the improved predictions.

In the earlier work of our group, ENSO predictions were based on nonlinear multiple regressions using neural networks (Tangang et al. 1997; 1998a,b; Tang et al. 2000). The next step is to advance the architecture of the forecast model from nonlinear regression to NLCCA. Although the prediction skills shown in this paper are not cross-validated, they are promising enough to warrant further efforts to develop NLCCA into operational ENSO prediction models and to evaluate model prediction skills under cross-validation.

Appendix: Hindcast with the NLCCA model

Once the weights parameters (in the network of Fig. 1) have been determined, an NLCCA or CCA model is built, the standard deviation $\mathrm{std}(u)$ and $\mathrm{std}(v)$ are known, and $\langle u \rangle = \langle v \rangle = 0$. If new $\mathbf{x}$ data become available, then u can be calculated, and v estimated by $u \cdot \mathrm{std}(v)/\mathrm{std}(u)$, which can then be used to predict $\mathbf{y}$. Similarly, $\mathbf{x}$ can be predicted using new $\mathbf{y}$ data. For each lead time, an NLCCA or CCA model is built.

Hindcast for the tropical Pacific SST

We used the 6 leading PCs of the FSU WS anomalies as the predictors, i.e., the inputs $\mathbf{x}$ of the network. From $v = u \cdot \mathrm{std}(v)/\mathrm{std}(u)$, we received the the 6 SST PCs as $\mathbf{y}'$ from the output (Fig. 1). Upon combining the PCs with their associated EOF spatial patterns, we obtained the predicted SST anomaly field. Correlations and RMSE (root mean square error) between the observed and predicted SST anomalies averaged over Niño12 [90°W-80°W, 10°S-0°], Niño3 [150°W-90°W, 5°S-5°N], Niño3.4 [170°W-120°W, 5°S-5°N] and Niño4 [160°E-150°W, 5°S-5°N] areas were calculated.

Prediction skills (correlation) with only mode 1 used are shown in Table 1, where apparent improvement of NLCCA over CCA can be seen in the Niño12 area at $0 - 12$ months of lead time, and other areas at longer lead times (9 and 12 months). The decrease of the RMSE of NLCCA relative to CCA is also more apparent in the Niño12 area, despite the higher RMSE relative to other areas.

When modes 1 and 2 are used, the prediction skills for all 4 areas are enhanced for both the NLCCA and CCA models, particularly, at lead times of 9 and 12 months. The NLCCA model still has higher or comparable prediction skills relative to the CCA model at all lead time over the 4 areas, with the most apparent improvement over the Niño12 area, followed by the Niño3 area. In Niño3.4 and Niño4, the skills are similar between NLCCA and CCA.

Hindcast for the tropical Pacific surface wind stress

Similarly, hindcasts for the WS anomalies over the tropical Pacific were made using the 6 leading PCs of SST as predictors. WS anomalies averaged over the TW1 area [135°E-180°, 5°S-5°N] and the TW2 area [175°W-140°W, 5°S-5°N] were calculated for prediction skill evaluation. We found the correlation skills decrease rapidly beyond 6 months lead, even when modes 1 and 2 are used. Thus, comparisons of RMSE and correlations between the NLCCA and CCA models were made only at 0, 3 and 6 months lead (Table 2), where we see higher correlations and lower RMSE for the NLCCA models than the corresponding CCA models at all lead time over both areas. Interestingly, although the prediction skills for WS are higher over the TW2 area in the central equatorial Pacific, greater improvements of the NLCCA over CCA occur over the TW1 area in the western equatorial Pacific.

Acknowledgements

This work was supported by research and strategic grants to Hsieh from the Natural Sciences and Engineering Research Council of Canada.

References

Anderson CW (1984) An Introduction to Multivariate Statistical Analysis. 2nd Ed. Wiley and Sons 675pp

Barnett TP, Preisendorfer R (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 AG, Livezey RE (1987) Classification, seasonality and persistence of low frequency atmospheric circulation patterns. Mon Wea Rev 115:1083-1126

Barnston AG, Ropelewski CF (1992) Prediction of ENSO episodes using canonical correlation analysis. J Clim 5:1316-1345

Hotelling, H (1936) Relations between two sets of variates. Biometrika 28:321-377

Hsieh WW (2000) Nonlinear canonical correlation analysis by neural networks. Neural Networks 13:1095-1105

Hsieh WW (2001a) Nonlinear canonical correlation analysis of the tropical Pacific climate variability using a neural network approach. J Clim 14:2528-2539

Hsieh WW (2001b) Nonlinear principal component analysis by neural networks. Tellus (in press)

Lai PL, Fyfe C (1999) A neural implementation of canonical correlation analysis. Neural Networks 12:1391-1397.

Lai PL, Fyfe C (2000) Kernel and non-linear canonical correlation analysis. Int J Neural Systems 10:365-377.

Monahan AH (2001) Nonlinear principal component analysis: Tropical Indo-Pacific sea surface temperature and sea level pressure. J Clim 14:219-233

Shabbar A, Barnston AG (1996) Skill of seasonal climate forecasts in Canada using Canonical Correlation Analysis. Mon Wea Rev 124:2370-2385

Smith TM, Reynolds RW, Livezey RE, Stokes DC (1996) Reconstruction of historical sea surface temperatures using empirical orthogonal functions. J Clim 9:1403-1420

Stricherz JN, David ML, O'Brien JJ (1997) TOGA Pseudostress Atlas 1985-1994, Volume II: Pacific Ocean, Florida State University Tallahassee FL 155 pp.

Tang B, Hsieh WW, Monahan AH, Tangang FT (2000) Skill comparisons between neural networks and canonical correlation analysis in predicting the equatorial Pacific sea surface temperatures. J Clim 13:287-293

Tangang FT, Hsieh WW, Tang B (1997) Forecasting the equatorial Pacific sea surface temperatures by neural network models. Clim Dyn 13:135-147

Tangang FT, Hsieh WW, Tang B (1998a) Forecasting the regional sea surface temperatures of the tropical Pacific by neural network models, with wind stress and sea level pressure as predictors. J Geophys Res 103:7511-7522

Tangang FT, Tang B, Monahan AH, Hsieh WW (1998b) Forecasting ENSO events— a neural network-extended EOF approach. J Clim 11:29-41

von Storch H, Zwiers FW (1999) *Statistical Analysis in Climate Research. Cambridge*, Cambridge Univ Pr 484 pp

Table 1: Correlation and RMSE (in °C) of the predicted SST anomalies relative to observations, at lead times (LT) from 0 to 12 months. 'NL' means NLCCA, while 'L' means CCA models were used. 'Mode 1' means only the first mode was used in prediction, while 'modes 1+2', both modes 1 and 2 were used.

		Mode 1				Modes 1+2			
	LT	rmse:L	rmse:NL	cor:L	cor:NL	rmse:L	rmse:NL	cor:L	cor:NL
Niño12	0	0.74	0.71	0.73	0.76	0.73	0.65	0.74	0.80
	3	0.85	0.79	0.61	0.69	0.84	0.73	0.64	0.75
	6	0.98	0.95	0.43	0.51	0.98	0.87	0.44	0.60
	9	1.06	1.04	0.17	0.33	1.05	1.03	0.30	0.50
	12	1.08	1.07	0.09	0.12	1.06	1.04	0.30	0.42
Niño3	0	0.38	0.37	0.90	0.90	0.38	0.35	0.90	0.91
	3	0.51	0.48	0.80	0.82	0.51	0.45	0.80	0.85
	6	0.71	0.71	0.56	0.57	0.69	0.66	0.60	0.66
	9	0.84	0.84	0.06	0.17	0.80	0.79	0.43	0.53
	12	0.85	0.84	0.04	0.08	0.83	0.82	0.33	0.41
Niño3.4	0	0.31	0.31	0.92	0.92	0.31	0.31	0.92	0.92
	3	0.42	0.41	0.86	0.86	0.40	0.37	0.88	0.89
	6	0.65	0.64	0.61	0.60	0.59	0.58	0.70	0.71
	9	0.79	0.79	0.10	0.19	0.74	0.73	0.53	0.56
	12	0.80	0.80	0.04	0.09	0.77	0.76	0.40	0.41
Niño4	0	0.34	0.29	0.81	0.85	0.33	0.30	0.83	0.85
	3	0.32	0.30	0.83	0.86	0.32	0.29	0.84	0.87
	6	0.46	0.44	0.63	0.66	0.43	0.42	0.69	0.71
	9	0.58	0.56	0.09	0.20	0.53	0.52	0.57	0.58
	12	0.59	0.57	0.03	0.09	0.55	0.55	0.41	0.45

Table 2: Table 2 Correlation and RMSE (in m^2/s^2) of the predicted wind stress relative to the FSU wind stress in the TW1 (western equatorial Pacific) and TW2 (central equatorial Pacific) areas.

		Mode 1				Modes 1+2			
	LT	rmse:L	rmse:NL	cor:L	cor:NL	rmse:L	rmse:NL	cor:L	cor:NL
TW1	0	6.1	5.9	0.55	0.60	5.4	5.1	0.69	0.71
	3	7.0	6.7	0.31	0.39	6.8	5.9	0.37	0.59
	6	7.3	6.9	0.18	0.23	7.2	6.9	0.23	0.36
TW2	0	6.4	6.2	0.74	0.75	6.2	6.1	0.76	0.78
	3	7.6	7.5	0.60	0.62	6.9	6.5	0.66	0.69
	6	9.6	9.5	0.21	0.25	9.3	9.1	0.35	0.39

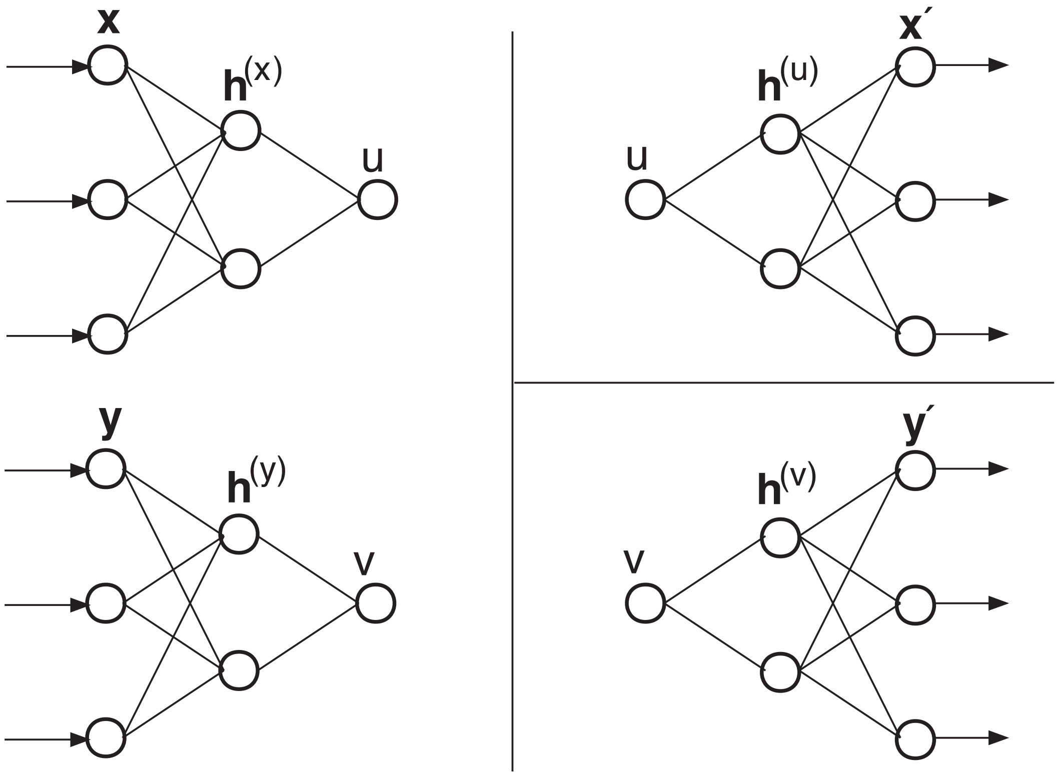

Figure 1: A schematic diagram illustrating the three feed-forward neural networks (NN) used to perform the NLCCA model of Hsieh (2001a). The double-barreled NN on the left maps from the inputs $\mathbf{x}$ and $\mathbf{y}$ to the canonical variates u and v. Starting from the left, there are l_1 input $\mathbf{x}$ variables ('neurons' in NN jargon), denoted by circles. The information is then mapped to the next layer (to the right)— a 'hidden' layer $\mathbf{h}^{(x)}$ (with l_2 neurons). For input $\mathbf{y}$, there are m_1 neurons, followed by a hidden layer $\mathbf{h}^{(y)}$ (with m_2 neurons). The mappings continue onto u and v. The cost function J forces the correlation between u and v to be maximized, and by optimizing J, the weights (i.e. parameters) of the NN are solved. On the right side, the top NN maps from u to a hidden layer $\mathbf{h}^{(u)}$ (with l_2 neurons), followed by the output layer $\mathbf{x}'$ (with l_1 neurons). The cost function J_1 minimizes the MSE($\mathbf{x}$), the mean square error of $\mathbf{x}'$ relative to $\mathbf{x}$. The third NN maps from v to a hidden layer $\mathbf{h}^{(v)}$ (with m_2 neurons), followed by the output layer $\mathbf{y}'$ (with m_1 neurons). The cost function J_2 minimizes the MSE($\mathbf{y}$), the MSE of $\mathbf{y}'$ relative to $\mathbf{y}$. When applied to the tropical Pacific, the $\mathbf{x}$ inputs were the first 6 PCs of the wind stress field, the $\mathbf{y}$ inputs, the first 6 PCs of the SST, and the number of hidden neurons were $l_2 = m_2 = 3$. The weight penalty parameter (Hsieh 2001a) (same for all 3 networks) was varied from 0.01 to 0.25 in steps of 0.02, and for each value of the penalty parameter, an ensemble of 30 trials with random initial weights were run. Among all these runs, the one achieving minimum MSE($\mathbf{x}$) $\cdot$ MSE($\mathbf{y}$) was selected as the solution.

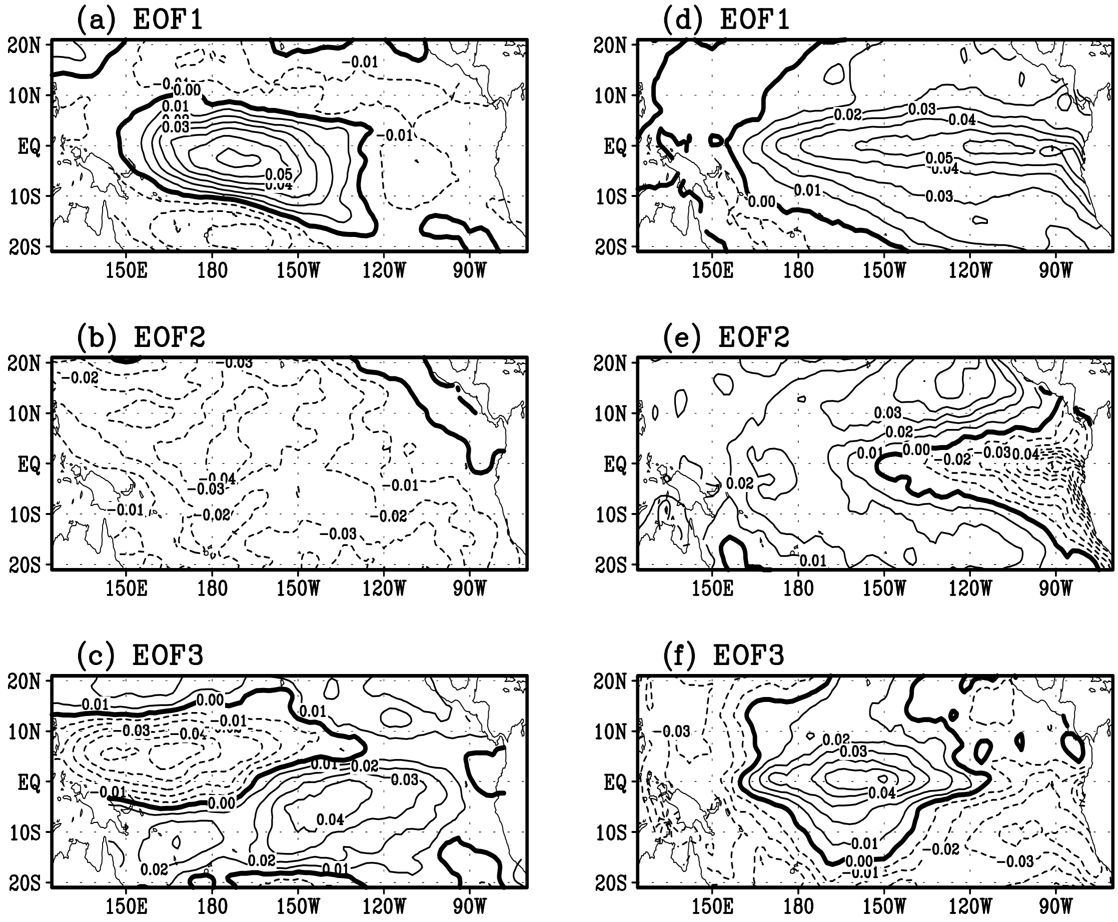

Figure 2: The first 3 EOFs of the wind stress anomalies (left column, only zonal component is shown) and SST anomalies (right column). Solid curves denote positive contours, dashed curves, negative contours, and thick curves, zero contours. The contour interval is 0.01. The EOFs have been normalized to unit norm.

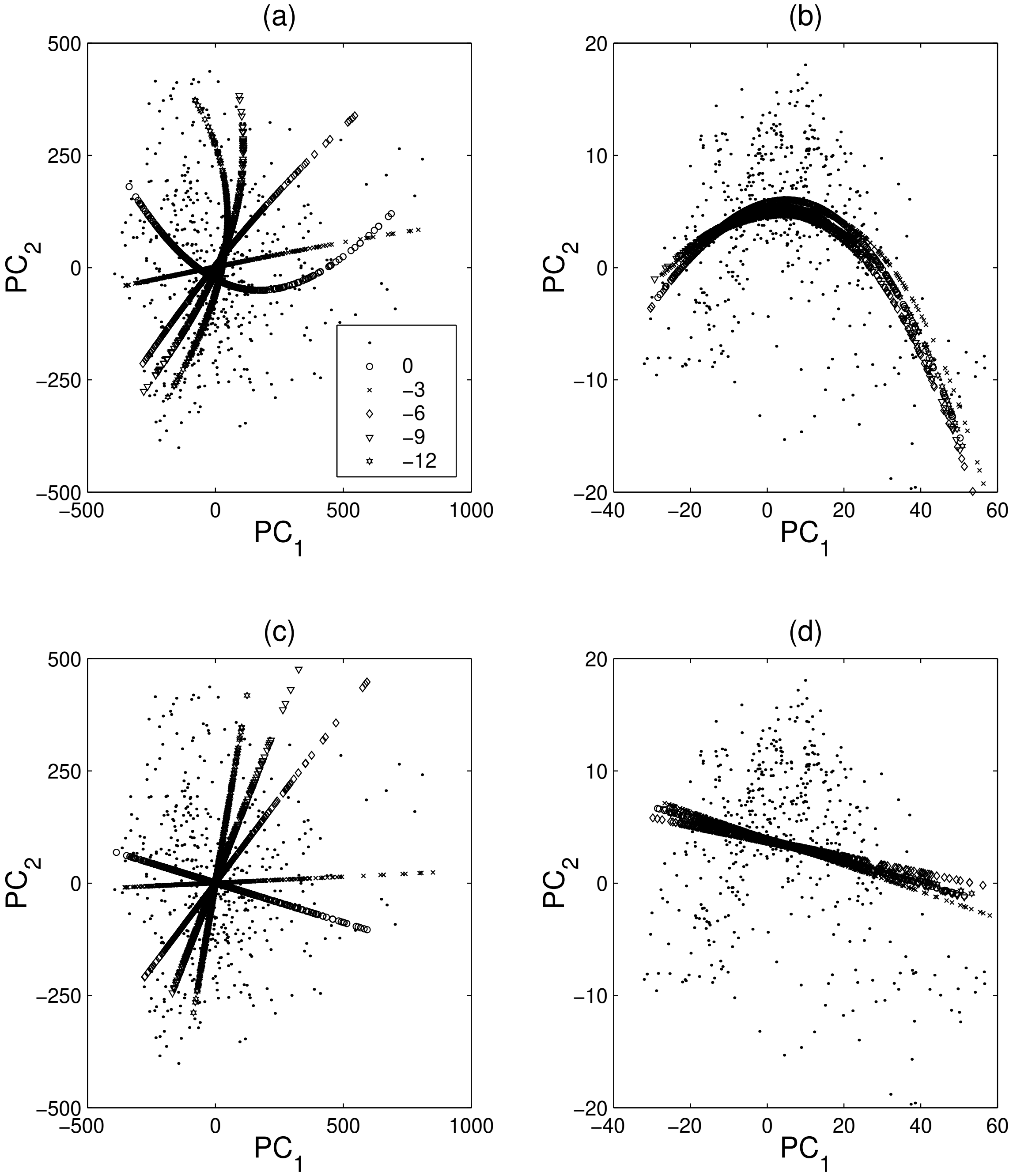

Figure 3: The first NLCCA mode of the wind stress (WS) and SST anomalies projected onto the PC_1-PC_2 planes, as shown in panels (a) and (b), respectively. The data are denoted by dots, and the projected NLCCA solutions with WS lagging SST by 0, 3, 6, 9 and 12 months are denoted by the circle, cross, diamond, triangle and star, respectively. The corresponding CCA modes are shown in panels (c) and (d) for the WS and SST, respectively.

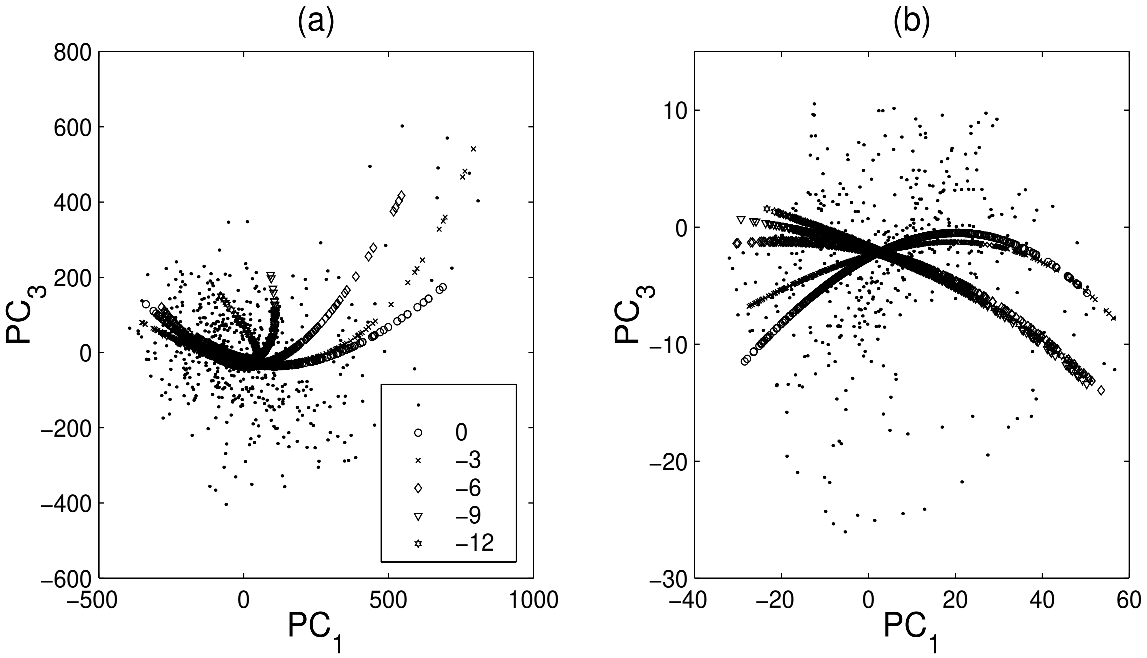

Figure 4: The first NLCCA mode of the WS and SST anomalies projected onto the PC_1-PC_3 planes, as shown in panel (a) and (b), respectively, with WS lagging SST by 0, 3, 6, 9 and 12 months.

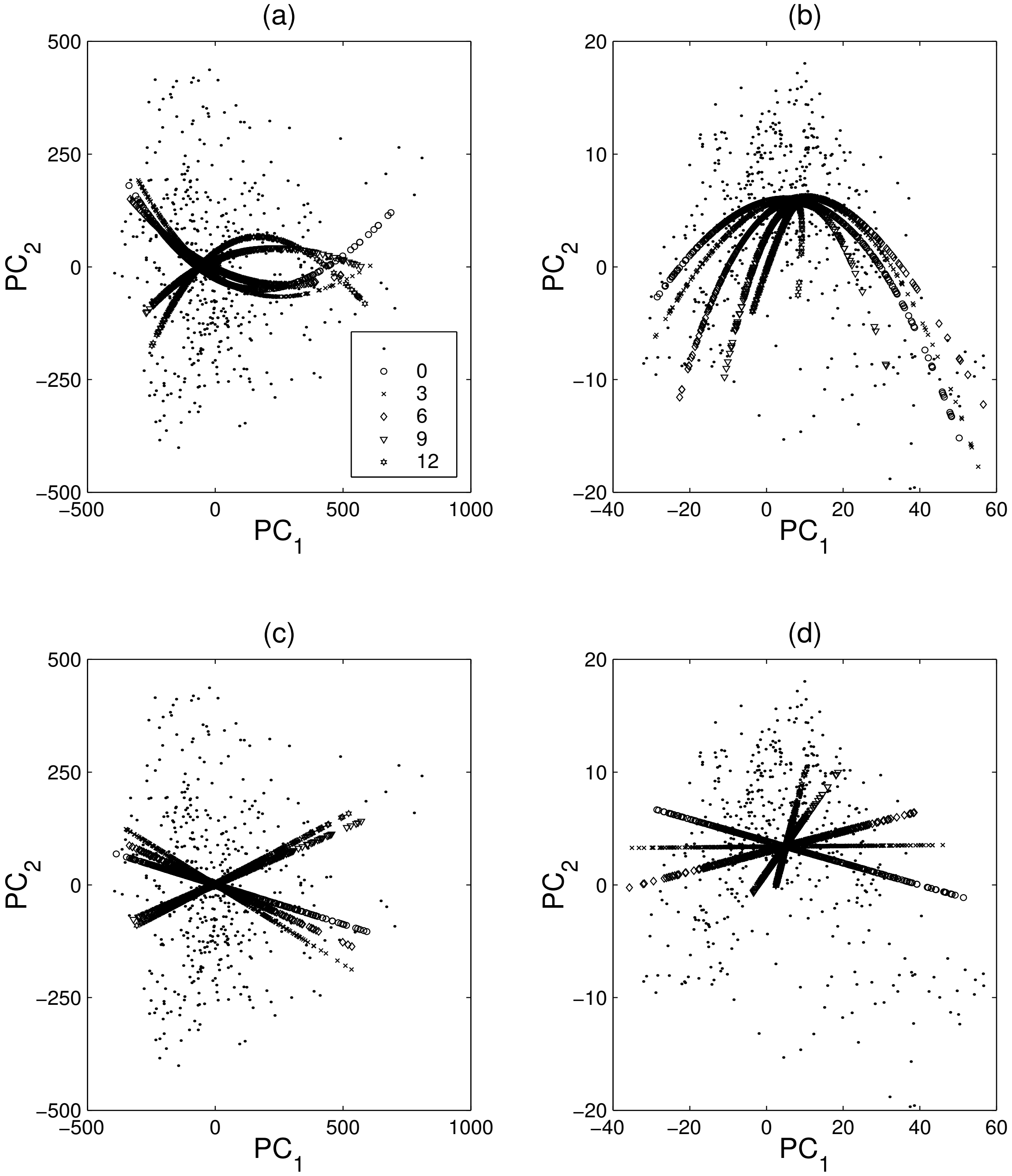

Figure 5: Similar to Fig. 3 but with WS leading SST by 0, 3, 6, 9, and 12 months.

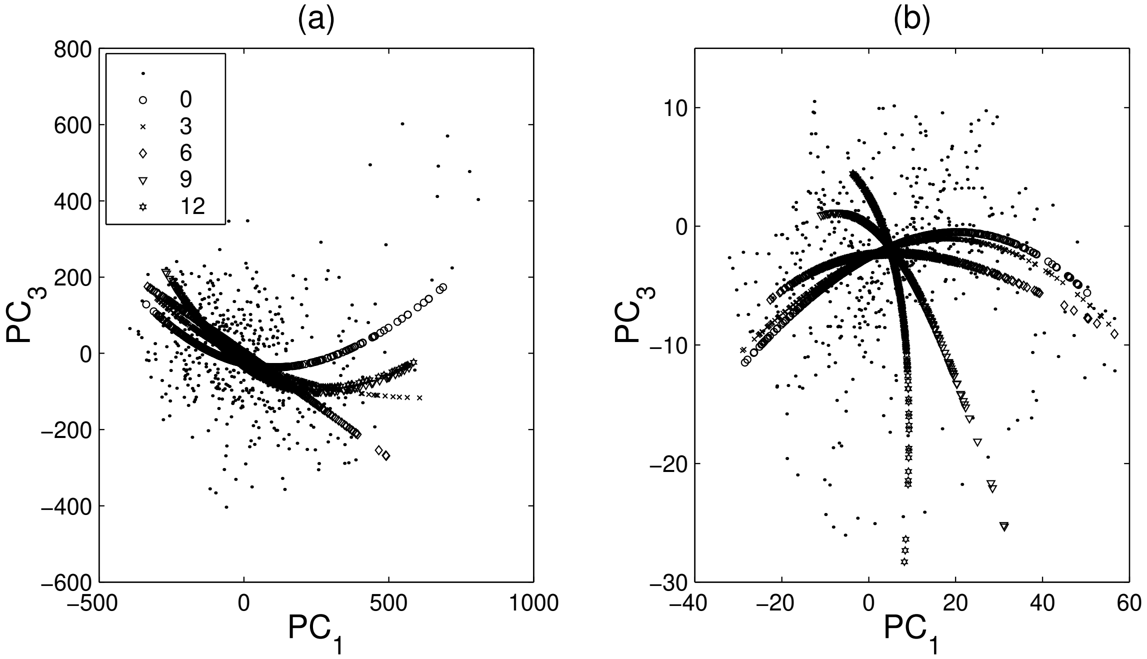

Figure 6: Similar to Fig. 4 but with WS leading SST.

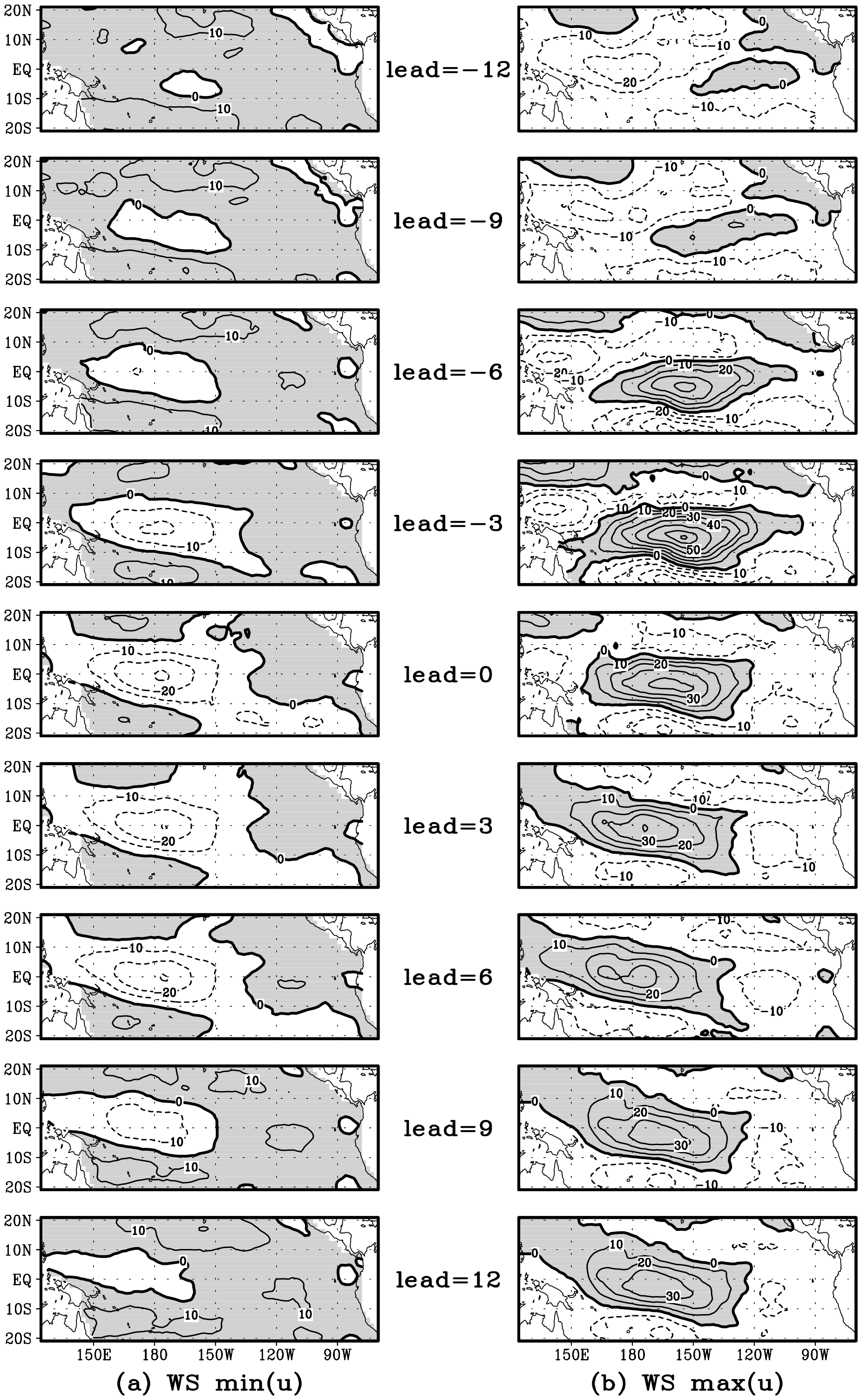

Figure 7: The zonal WS anomaly patterns from the NLCCA mode 1 with the canonical variate u taking its minimum (left column) and maximum values (right column). From top to bottom, we have WS leading SST by $-12, -9, -6, -3, 0, 3, 6, 9$ and 12 months, with a negative lead time denoting a positive lag time. The contour interval is 10 $\mathrm{m}^2/\mathrm{s}^2$. Areas with positive values (westerly anomalies) are shaded.

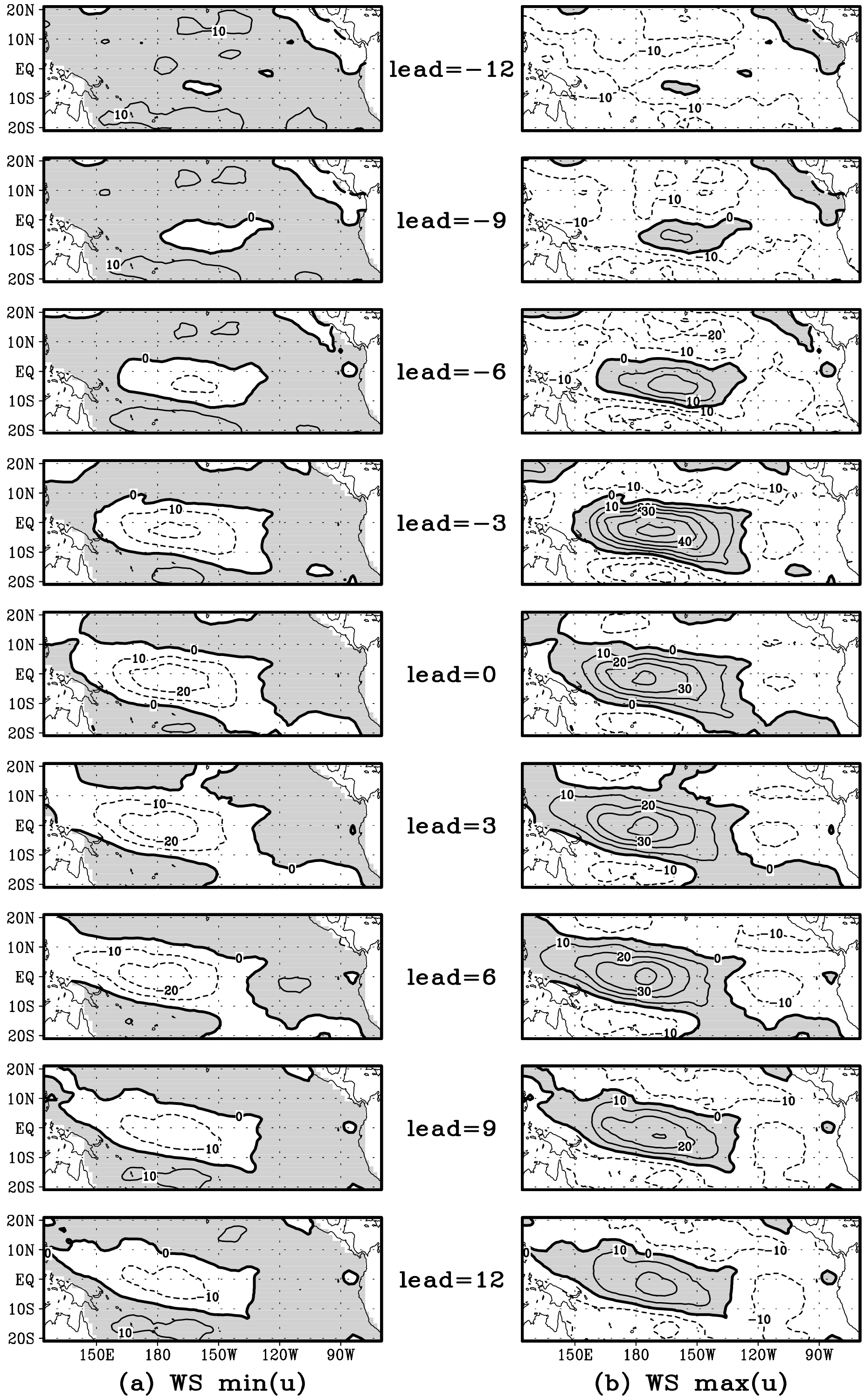

Figure 8: Similar to Fig. 7 but for the CCA mode 1.

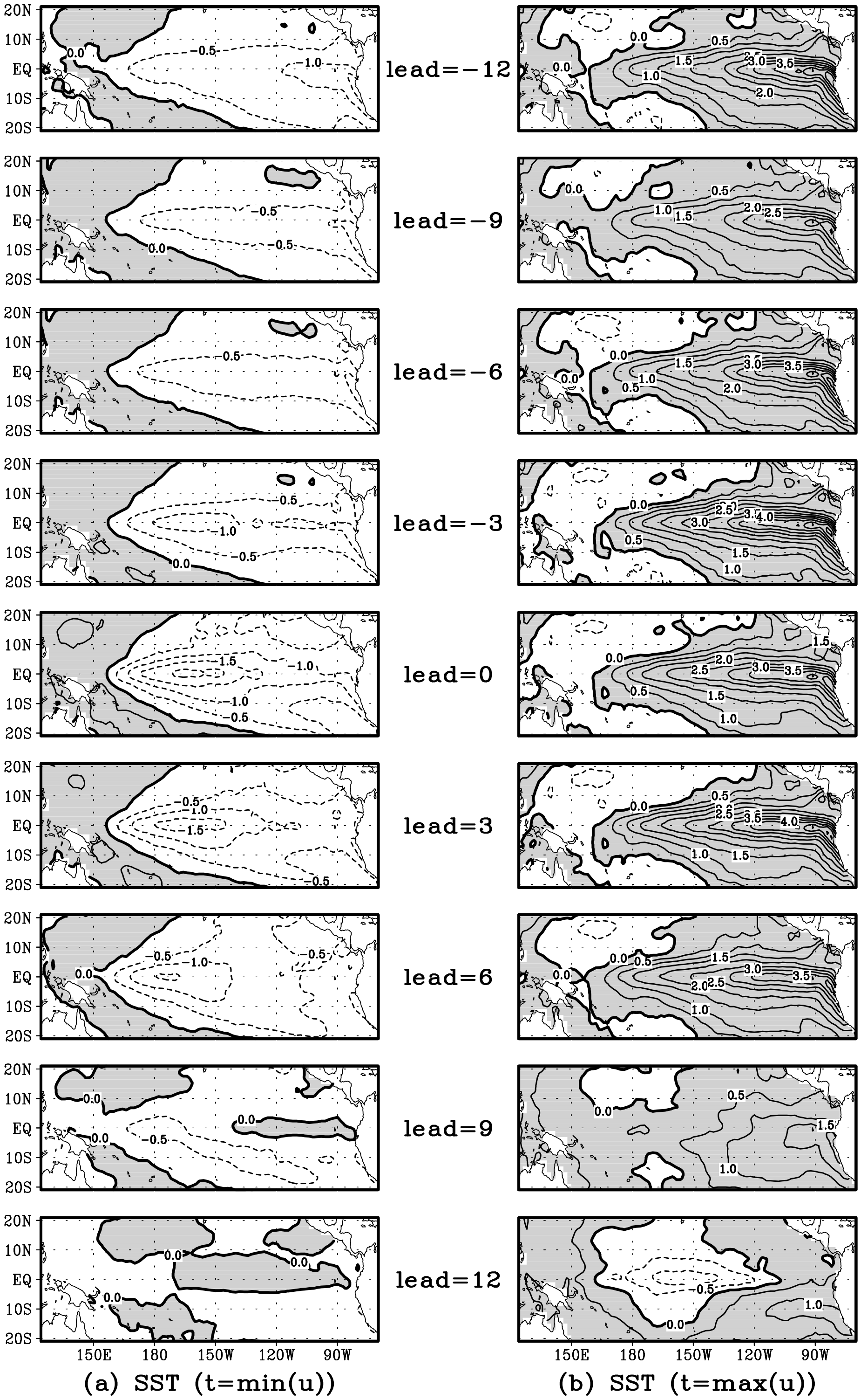

Figure 9: Similar to Fig. 7, but for the spatial patterns of SST anomalies at the time when u takes its minimum and maximum values. Areas with positive anomalies are shaded. The contour interval is 0.5°C.

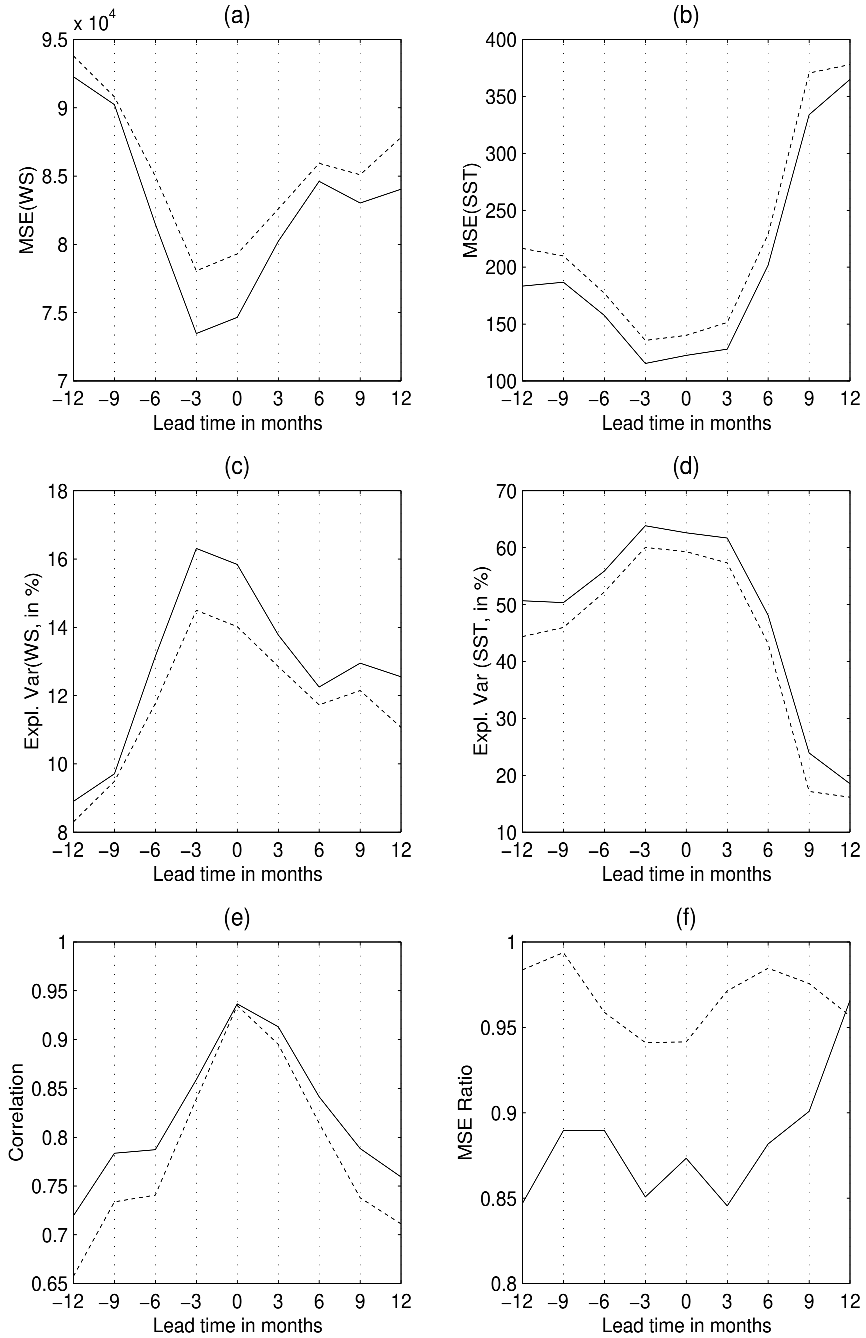

Figure 10: Comparisons between the NLCCA mode 1 and the CCA mode 1 in terms of the mean square error (MSE) for (a) the WS, and (b), the SST; the explained variance (in %), for (c) the WS, and (d) the SST; and (e) the canonical correlations. In panels (a)-(e), the solid lines represent the NLCCA mode, and the dashed lines, the CCA mode. Ratios of MSE between the NLCCA mode and CCA mode are shown in panel (f), with solid line for the SST, and dashed line, the WS.