Nonlinear modes of North American winter climate variability derived from a general circulation model simulation

Aiming Wu and **William W. Hsieh**
Dept. of Earth and Ocean Sciences, University of British Columbia
Vancouver, B.C., V6T 1Z4, Canada
Tel: (604) 822-3911, Fax: (604) 822-6088
Email: awu@eos.ubc.ca

Francis W. Zwiers
Canadian Center for Climate Modelling and Analysis
Meteorological Service of Canada
c/o University of Victoria
PO Box 1700, STN CSC
Victoria, B.C., V8W 2Y2, Canada

Submitted to *Journal of Climate*

December 18, 2002

Abstract

Nonlinear principal component analysis (NLPCA), via a neural network (NN) approach, was applied to an ensemble of six 47-year simulations conducted by the Canadian Centre for Climate Modelling and Analysis (CCCma) second-generation atmospheric general circulation model (AGCM2). Each simulation was forced with the observed sea surface temperature (from the GISST dataset) from January 1948 to November 1994. The NLPCA modes reveal nonlinear structures in both the winter 500-mb geopotential height (Z500) anomalies and surface air temperature (SAT) anomalies over North America, with asymmetric spatial anomaly patterns during the opposite phases of an NLPCA mode. Only during its negative phase is the first NLPCA mode related to the El Niño-Southern Oscillation (ENSO); the positive phase is related to a weakened jet stream. Spatial patterns of the NLPCA mode for the Z500 anomalies generally agree with those for the SAT anomalies.

Nonlinear canonical correlation analysis (NLCCA), also via a NN approach, was then applied to the mid-latitude winter GCM data and the observed SST of the tropical Pacific. Nonlinearity was detected in both the forcing field (SST) and the response field (Z500 or SAT) at zero time lag. The leading NLCCA mode for the SST anomalies is a nonlinear ENSO mode, with a 30-40° eastward shift of the positive SST anomalies during El Niño relative to the negative SST anomalies during La Niña. The leading NLCCA mode for the Z500 anomaly field is a nonlinear Pacific-North American (PNA) teleconnection pattern. The ray path of the Rossby waves induced during El Niño is 10-15° east of that induced during La Niña. The nonlinear atmospheric response to ENSO is also found in the leading NLCCA mode for the SAT anomalies.

1 Introduction

Substantial progress has been made during the past two decades towards understanding the wintertime extratropical atmospheric responses to the tropical forcings associated with El Niño sea surface temperature (SST) anomalies (Trenberth et al. 1998). The most prominent teleconnection pattern is the Pacific-North America (PNA) pattern (Horel and Wallace 1981), which is thought to link the changes in the extratropical circulation to the tropical SST through Rossby wave dynamics (Wallace and Gutzler 1981; Hoskins and Karoly 1981).

The simplest view of the atmospheric climate signal associated with the El Niño-Southern Oscillation (ENSO) phenomenon is that the atmosphere responds linearly, with anomalies during the El Niño phase being the reverse of those during the La Niña phase. However, recent evidence shows that the responses to warm and cold events are not exactly opposite. Sittel (1994) found the marginal probabilities of extreme rainfall and temperature over the southeastern United States to be highly nonlinear functions of the phase of the Southern Oscillation (SO). For example, it was found that, since 1946, the warm event enhanced rainfall signal has been much larger than the cold event suppressed signal in the southeastern United States. Richman and Montroy (1996) examined the composite January temperature and precipitation patterns over the United States and parts of Canada associated with El Niño and La Niña events. Their results suggest that El Niño and La Niña have their own unique characteristics in terms of temperature and precipitation. Asymmetric spatial patterns of Canadian surface air temperature and precipitation associated with the Southern Oscillation (SO) were also detected by Shabbar and Khandekar (1996) and Shabbar et al. (1997). Further evidence for a nonlinear response of the North America climate to ENSO was provided by Hoerling et al. (1997) with composite analysis and numerical experiments using an idealized atmospheric general circulation model, where a shift in the equatorial positions of the maximum rain responses and a phase shift of teleconnection patterns in the upper troposphere were found. The inherent nonlinearity in the tropical rain response may itself be responsible for the phase shift in the extratropical teleconnection patterns.

The robustness of the nonlinear climate response to ENSO's extreme phases has also been investigated with four GCM simulations (Hoerling et al. 2001), which were all found to have a 500-mb height response to extreme warm tropical Pacific SST that was twice as strong as the response to extreme cold SST. The longitudinal phase of the GCMs' teleconnections also shifted eastward during warm events as compared with cold events, though this displacement is smaller than that observed. A nonlinear identification of the atmospheric response to ENSO was also addressed by Hannachi (2001) using general circulation models.

Nonlinear methods are required to investigate the nonlinear behavior of the North American climate variability and its relation to ENSO. Standard multivariate statistical techniques such as principal component analysis (PCA) [also known as empirical orthogonal function (EOF) analysis], and canonical correlation analysis (CCA) are linear methods. Composite analysis does not assume linearity, but is restricted to the analysis of differences between specific phases of the SO or equatorial SST indices. Recently, neural networks (NN) (Hsieh and Tang 1998) have been used for nonlinear PCA (NLPCA) (Kramer 1991) and nonlinear CCA (NLCCA) (Hsieh 2000). NLPCA has been applied to the Lorenz 3-component chaotic system (Monahan 2000), tropical Pacific SST and sea level pressure fields (Monahan 2001). NLPCA was recently used to represent the quasi-biennial oscillation (QBO) in the equatorial stratospheric wind (Hamilton and Hsieh 2002), and to explore the nonlinear characteristics of Canadian surface air temperature (Wu et al. 2002). NLCCA has been applied to study the nonlinear relation between the tropical Pacific sea level pressure (SLP) and SST fields (Hsieh 2001a), as well as between the wind stress and SST fields (Wu and Hsieh 2002). Also, NLCCA has been applied to forecasting the tropical Pacific sea surface temperatures in the Experimental Long Lead Forecast Bulletin (www.iges.org/ellfb).

In this paper, the NLPCA (Hsieh 2001b) and NLCCA (Hsieh 2001a) models will be applied to study the wintertime climate variability over North America and its relation to ENSO as simulated in an ensemble of six 47-year simulations produced with the Canadian Centre for Climate Modelling and Analysis (CCCma) second generation atmospheric general circulation model (AGCM2). This paper is organized as follows: In Section 2, the data and the two methods (NLPCA and NLCCA) are briefly introduced. Nonlinear modes of the winter 500-mb geopotential height and surface air temperature that were extracted by the NLPCA are described in Section 3. Correlated nonlinear modes of atmospheric and tropical Pacific SST variability that were extracted by the NLCCA are described in Section 4. Section 5 presents a summary and discussion.

2 Methodology and data

2.1 NLPCA

A variable $\mathbf{x}$, which consists of l spatial stations and n observations in time, can be expressed in the form $\mathbf{x}(t) = [x_1(t), \ldots, x_l(t)]$, where for each i $(i = 1, \ldots, l)$, $x_i(t)$ $(t = 1, \ldots, n)$ is a time series of length n. PCA is used to find a scalar variable u and an associated vector $\mathbf{a}$, with

$$u(t) = \mathbf{a} \cdot \mathbf{x}(t)\,, \tag{1}$$

so that

$$\langle \|\mathbf{x}(t) - \mathbf{a}u(t)\|^2 \rangle \ \text{ is minimized,} \tag{2}$$

where $\langle \cdots \rangle$ denotes a sample or time mean. Here u, called the first principal component (PC), is a time series resulting from a linear combination of the original variables x_i, while $\mathbf{a}$, the first eigenvector of the data covariance matrix (the first EOF), often describes a spatial pattern. The second PC can similarly be extracted from the residual $\mathbf{x} - \mathbf{a}u$, and so on for the higher modes. In practice, the common algorithms for PCA extract all modes simultaneously by calculating the eigenvalues and eigenvectors of the data covariance matrix.

The fundamental difference between NLPCA and PCA is that NLPCA allows a nonlinear continuous mapping from $\mathbf{x}$ to u whereas PCA only allows a linear mapping. NLPCA is performed with a NN, such as that displayed in Fig. 1, that contains 3 'hidden' layers of variables (or 'neurons') between the input and output layers. These layers are called the encoding, bottleneck, and decoding layers, respectively. Four transfer functions f_1, f_2, f_3, f_4 are successively used to map from the input layer to the output layer ($\mathbf{x} \longrightarrow \mathbf{h}^{(x)} \longrightarrow u \longrightarrow \mathbf{h}^{(u)} \longrightarrow \mathbf{x}'$), where $\mathbf{x}'$ is the least square approximation of $\mathbf{x}$. The mappings are:

$$h_k^{(x)} = f_1((\mathbf{W}^{(x)}\mathbf{x} + \mathbf{b}^{(x)})_k) \,, \tag{3}$$

$$u = f_2(\mathbf{w}^{(x)} \cdot \mathbf{h}^{(x)} + \overline{b}^{(x)}) \,, \tag{4}$$

$$h_k^{(u)} = f_3((\mathbf{w}^{(u)}u + \mathbf{b}^{(u)})_k) \,, \tag{5}$$

$$x_i' = f_4((\mathbf{W}^{(u)}\mathbf{h}^{(u)} + \overline{\mathbf{b}}^{(u)})_i) \,, \tag{6}$$

where the upper case bold font is reserved for matrices and the lower case bold font for vectors, and $\mathbf{x}$ is the input column vector of length l. The encoding layer $\mathbf{h}^{(x)}$ is described by a column vector of length m (m is the number of the hidden neurons in the encoding layer). The parameters that control the transformation to this layer are $\mathbf{W}^{(x)}$, which is an $m \times l$ weight matrix, and $\mathbf{b}^{(x)}$, a column vector of length m containing the bias parameters. Index $k \in [1, m]$. The bottleneck layer contains a single neuron, which represents the nonlinear principal component u. The transformation between encoding and bottleneck layers is controlled by an m-element weight vector $\mathbf{w}^{(x)}$ and a bias parameter $\overline{b}^{(x)}$. The decoding layer contains the same number of neurons m as the encoding layer, and the output layer is again a column vector of length l. The transformations between these layers are controlled by m-element weight and bias vectors $\mathbf{w}^{(u)}$ and $\mathbf{b}^{(u)}$ (decoding layer), and an $l \times m$ weight matrix $\mathbf{W}^{(u)}$ and an l-element bias vector $\overline{\mathbf{b}}^{(u)}$ (output layer). Transfer functions f_1 and f_3 are generally nonlinear (e.g. the hyperbolic tangent function), while f_2 and f_4 are taken to be the identity function. See Hsieh (2001b) for more details.

Optimal values of the weight parameters for each mapping layer are found by minimizing the cost function J (see Eq (7)), rendering the outputs to be as close to the inputs as possible within constraints imposed by a weight penalty term (discussed below). Data compression is achieved at the bottleneck, with the bottleneck neuron representing a single degree of freedom, namely u, the nonlinear principal component (NLPC).

The nonlinear optimization was carried out with a quasi-Newton method. To avoid the local minima problem (Hsieh and Tang 1998, p. 1859), an ensemble of 30 NNs with random initial weights and bias parameters was run. Also, 20% of the data was randomly selected as testing data and withheld from the training of the NNs. For the NLPCA, runs where the mean square error (MSE) for the testing dataset was 10% larger than that for the training dataset were rejected to avoid overfitted solutions. Then the NN with the smallest MSE was selected as the desired solution.

As noted above, the cost function used to identify the NLPCA model in this study has an extra weight penalty term,

$$J = \langle \|\mathbf{x} - \mathbf{x}'\|^2 \rangle + p \sum_{ki} (W_{ki}^{(x)})^2 \, , \tag{7}$$

where $p \geq 0$ is called the weight penalty parameter. Increasing p increases the concavity of the cost function, thereby pushing the weights $\mathbf{W}^{(x)}$ to be smaller in magnitude and consequently yielding smoother and less nonlinear solutions than when p is small or zero. With a large enough p, the danger of overfitting is greatly reduced (Hsieh 2001b). The NLPCA was run repeatedly with various values of the penalty parameter (ranging from 0.001 to 0.02), and the solution with the smallest MSE was chosen as the final solution.

After the first NLPCA mode has been subtracted from the data, the residual is again input into the NLPCA network to extract the second NLPCA mode. Because of the noisier conditions, the penalty parameter is increased to the range 0.01 to 0.2 for the second mode.

The weight penalty parameters in NLPCA (also in the NLCCA described below) are selected by means of a search. Future research will hopefully provide a more objective way to select these parameters.

2.2 NLCCA

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

$$u = \mathbf{a} \cdot \mathbf{x} \quad \text{and} \quad v = \mathbf{b} \cdot \mathbf{y} \tag{8}$$

where the canonical variates u and v have maximum correlation, i.e. the coefficient vectors $\mathbf{a}$ and $\mathbf{b}$ are chosen such that the Pearson correlation coefficient between u and v, is maximized.

In NLCCA, we follow the same procedure as in CCA, except that the linear mappings in Eq.(8) are replaced by nonlinear mapping functions using 2-layer feed-forward NNs. The mappings from $\mathbf{x}$ to u and $\mathbf{y}$ to v are represented by the double-barreled NN on the left hand side of Fig. 2. By minimizing the cost function $J = -\mathrm{cor}(u, v)$, one finds the parameters that maximize the correlation $\mathrm{cor}(u,\ v)$. After the forward mapping with the double-barreled NN has been solved, inverse mappings from the canonical variates u and v to the original variables, as represented by the two standard feed-forward NNs on the right side of Fig. 2, are to be solved, where the cost function J_1 is the MSE of the output $\mathbf{x}'$ relative to $\mathbf{x}$ (MSE$\mathbf{x}$), and the cost function J_2, the MSE of the output $\mathbf{y}'$ relative to $\mathbf{y}$ (MSE$\mathbf{y}$) are separately minimized to find the optimal parameters for these two NNs (see Hsieh, 2001a for details).

An ensemble approach is also used for the NLCCA— runs where -cor(u,v), the MSE$\mathbf{x}$ or the MSE$\mathbf{y}$ for the testing dataset were 10% larger than those for the training dataset were rejected. The NNs with the highest cor(u,v), and smallest MSE$\mathbf{x}$ and MSE$\mathbf{y}$ were selected as the desired solution. A weight penalty was again used (Hsieh, 2001a) as in NLPCA.

In brief, the major advantage of NLPCA and NLCCA is that both are free from linear constraint. This allows us to identify principal components or canonical variates that vary along empirically derived curves through the data space instead of only straight lines. Consequently, a larger fraction of the variance can be explained (relative to PCA) or higher canonical correlation can be reached (relative to CCA), thus raising the possibility of improving the interpretation and forecasting climate using NLCCA model. The main difficulty in the application of NLPCA or NLCCA is that the cost function often has multiple local minima (Hsieh and Tang 1998). It takes considerable computation to obtain a robust solution that approximates the "global" minimum by running the optimization many times from random initial parameters. Another drawback is overfitting, i.e. fitting to the noise in the data. Using a weight penalty and reserving part of the data as validation data can alleviate overfitting.

2.3 Data

The monthly mean 500-mb geopotential height (Z500) and surface air temperature (SAT) data used in this study were produced by the CCCma AGCM2, a spectral model with T32 resolution in the horizontal, 10 levels in the vertical, semi-implicit time stepping, and a full physics package (McFarlane et al. 1992; Boer et al. 1992). An ensemble of six 47-year runs of AGCM2 were carried out, in which each integration was started from different initial conditions and forced by SSTs from the GISST dataset (version 2.2, Rayner et al. 1996). Simulations were performed from January 1948 to November 1994, and were initially reported by Zwiers et al. (2000). For each run, anomalies were calculated by subtracting the monthly climatology based on the whole period. Monthly anomalies were then smoothed by taking a 3-month running mean and removing linear trends. Only the smoothed data for December, January and February are analyzed, thus the total number of months used from the 6 AGCM2 runs is 840 $[(47 \times 3 - 1) \times 6]$. The domain of interest is 150°E-50°W, 20°N-76°N, covering the North Pacific and North America. For the SAT, only the data over land grids were used.

Monthly SST used in this study was from the reconstructed global historical SST datasets by Smith et al. (1996) for the period 1950-2000 with a resolution of 2° by 2°. As the AGCM2 runs end in 1994, the SST data was actually used up to November 1994 in the NLCCA (the total number of months used in NLCCA is 804 $[(45 \times 3 - 1) \times 6]$). Similar data processing was done for the SST data, as for the Z500 and SAT data. The area of interest for the SST is restricted to the tropical Pacific (123°E-71°W, 21°S-21°N).

Prior to the NLPCA and NLCCA, ordinary PCA (i.e. EOF) analysis was conducted on the Z500, SAT and the SST anomalies to compress the data into manageable dimensions and to insure that the estimated variance-covariance matrices that enter into CCA calculation can be inverted (Barnett and Preisendorfer 1987). EOFs of the 3 leading modes of Z500 and SAT anomalies are shown in Fig. 3. Variance contributions from these 3 modes of the Z500 anomalies are 32.8%, 17.8%, and 12.7%, respectively, and for the SAT anomalies, 26.8%, 16.1% and 8.9%, respectively. The SAT anomaly pattern for each mode generally reflects the anomalous circulation implied by the corresponding Z500 mode, although the signs for EOF2 are opposite (Fig. 3b and 3e). The accumulated variance contribution of the 5 leading modes is 77.0% for the Z500 anomalies, 61.9% for the SAT anomalies, and 83.7% for the tropical Pacific SST anomalies.

The 5 leading principal components (PCs) (i.e. the EOF time series) of the winter Z500 and SAT anomalies from 1948 to 1994 were used as the inputs to the NLPCA. For the NLCCA, only data after 1950 were used since the SST data were available from January 1950. Also for the NLCCA, SST anomalies (DJF) were repeated 6 times to pair with the atmospheric data (Z500 and SAT) from the 6 GCM ensemble runs. The 5 leading SST PCs, and the 5 leading Z500 (or SAT) PCs were used as the inputs to the NLCCA.

3 Nonlinear modes extracted by NLPCA

3.1 Mode 1

Figs. 4a and b, which show the first NLPCA mode for the Z500 and SAT anomalies, respectively, reveal nonlinear structures in both datasets. The Z500 NLPCA mode 1 explains 35.7% of the total variance, versus 32.8% explained by PCA mode 1 (straight line in Fig. 4a). For SAT, the NLPCA mode 1 explains 30.6% of the variance, versus 26.8% by the linear mode. A measure of the degree of nonlinearity is the ratio (r) between the MSE (mean square error) of the NLPCA mode 1 and that of the corresponding PCA mode. Smaller r means stronger nonlinearity. When r=1, the nonlinear mode is reduced to the linear mode. Here r is 0.921 for the Z500, and 0.901 for the SAT, suggesting moderate nonlinearity in both datasets, with SAT being somewhat more nonlinear.

Changing the value of a PC (i.e., selecting a point on the straight line in Fig. 4a) has the effect of changing the amplitude but not the spatial structure of the corresponding EOF pattern (Fig. 3a). In contrast, both the structure and amplitude of the spatial pattern of the NLPCA mode change smoothly as the NLPC u (which traces the curve in Fig. 4a) changes value. The NLPCA maps the bottleneck neuron u to the output layer of the NN used to represent the data. This produces values for the first 5 PCs which in turn can be combined with the corresponding PCA spatial patterns (i.e. the EOFs) to yield the spatial pattern corresponding to u. When u takes on its minimum value, 3 large anomalies appear in the Z500 anomaly field over the North Pacific and the North American continent, resembling a negative PNA pattern (Fig. 5a), with negative height anomalies over Canada and the United States (US) except western Alaska and the southeastern USA, where there are positive height anomalies. The SAT anomaly pattern associated with the minimum u is roughly in agreement with the Z500 anomaly pattern, with positive anomalies (1-2°C) over the southeastern US and negative anomalies over the rest of North America (Fig. 5c).

When u takes on its maximum value, the Z500 anomalies (Fig. 5b) show no closed contours north of 50°N, with the USA covered by negative height anomalies, and Canada by positive anomalies, implying a weakened jet stream because the geopotential gradient is reduced in the region that is usually occupied by the jet stream. With maximum u, the SAT anomalies (Fig. 5d) show that the USA and southern Canada are cooler, while other areas of Canada and Alaska are warmer.

The NLPCA Z500 anomaly pattern for maximum u (Fig. 5b) resembles EOF1 pattern (Fig. 3a) while the pattern for minimum u (Fig. 5a) resembles EOF2 (Fig. 3b). This means that the weakened jet stream state associated with maximum u does not have an equally strong negative counterpart (i.e. an enhanced jet stream state). Instead the enhanced jet steam state is overshadowed by the negative PNA state found during minimum u.

Comparing Fig. 5a with 5c, and 5b with 5d, we can see some spatial correspondence between the anomalies of Z500 and SAT, where positive anomalies of Z500 tend to occur roughly together with positive anomalies of SAT, and similarly for the negative anomalies, thereby revealing some consistency in the nonlinear structures found by the NLPCA approach.

3.2 Mode 2

After removing the NLPCA mode 1 from the data, the residuals were again input into the NLPCA network to extract the second NLPCA mode, which was also found to contain notable nonlinearity in both the Z500 and SAT fields. The second Z500 NLPCA mode explains 16.5% of the total variance, which is slightly more than the 15.2% explained by the corresponding linear mode— here the linear mode is not the same as the PCA mode 2, as the latter is extracted from the residual with the PCA mode 1 (not the NLPCA mode 1) subtracted from the original data. The MSE ratio between the second NLPCA mode and the corresponding linear mode for Z500 is 0.953. There is stronger nonlinearity in the second SAT NLPCA mode, which explains 12.5% of the total variance, versus the

9.3% explained by the corresponding linear mode, with a MSE ratio of 0.874.

The second Z500 NLPCA mode reveals mainly east-west differences over North America and the adjacent North Pacific. Positive anomalies lie in the west and negative anomalies to the east at minimum u (Fig. 6a). The anomalies reverse sign and shift southwestward for maximum u (Fig. 6b). The second SAT NLPCA mode shows even larger changes in the anomaly pattern when comparing minimum u (Fig. 6c) and maximum u (Fig. 6d).

The Z500 and SAT anomaly patterns corresponding to minimum u (Fig. 6a and 6c) resemble EOF3 (Figs. 3c and 3f), with positive height and SAT anomalies centered over western Canada. Spatial patterns corresponding to maximum u (Fig. 6b and 6d) resemble EOF2 for both Z500 and SAT (Figs. 3b and 3e) with positive height and SAT anomalies over much of USA and Canada. We can see some similarity between the spatial structure of the second Z500 and SAT NLPCA modes, even though NLPCA was performed separately on the two datasets.

To investigate the relations between the NLPCA modes and ENSO, composites of the winter (DJF) Z500 and SAT anomalies during warm (El Niño) event years (1958, 1966, 1969, 1973, 1983, 1987, 1988, 1992) and cold (La Niña) event years (1950, 1951, 1955, 1956, 1965, 1971, 1974 1976, 1989) were computed. The years used for the composites are the same as those used by Hoerling et al. (1997). The Z500 composites during El Niño years and La Niña years show positive and negative PNA patterns respectively (Fig. 7b and 7a). When El Niño takes place, positive height anomalies dominate over North America except over the southeastern USA. The corresponding composite SAT anomaly field has negative SAT anomalies over the southeastern USA and positive SAT anomalies over the rest of North America (Fig. 7d). The composite Z500 and SAT anomalies during La Niña years are basically opposite to those during El Niño years. Thus the asymmetries between El Niño and La Niña in the mid latitudes simulated by the CCCma AGCM are weaker than observed (Hoerling et al. 1997). Patterns displayed in Figs. 7a and 7c are similar to those shown in Figs. 5a and 5c, but with weaker amplitudes. However, patterns in Figs. 7b and 7d are quite different from those shown in Figs. 5b and 5d. Apparently, NLPCA mode 1 is related to ENSO only when u is negative. To relate the tropical Pacific SST to the North American Z500 and SAT, we turn to the NLCCA approach.

4 Nonlinear modes extracted by NLCCA

4.1 NLCCA of SST and Z500

NLCCA between winter tropical Pacific SST anomalies and simultaneous Z500 anomalies (Fig. 8) reveals nonlinearity when compared to the linear CCA solution which is shown as a straight line. For SST (Fig. 8a), nonlinearity in the the PC_1-PC_2 plane is manifested by a curve which links La Niña states at the left end to El Niño states at the right end. For Z500 (Fig. 8b), stronger nonlinearity appears in the PC_1-PC_3 and PC_2-PC_3 planes than in the PC_1-PC_2 plane. The first NLCCA mode for SST explains 64.6% of SST variance, versus 62.3% explained by the first CCA mode. The MSE ratio r is 0.856. The corresponding NLCCA mode for Z500 explains 23.6% of the variance, versus 21.5% explained by the first CCA mode, with a r value of 0.953. The correlation between the canonical variates (u and v) is 0.702 for the nonlinear mode and 0.675 for the linear mode.

As in NLPCA, one can map values of canonical variate u and v onto SST and Z500 anomaly patterns respectively. Here, minimum and maximum u are chosen to present the La Niña and El Niño states respectively and Z500 anomaly patterns are considered for the values of v that correspond to minimum and maximum u. The SST field that corresponds to minimum u presents a La Niña with negative anomalies (about -2.0°C) over the central-western equatorial Pacific. The corresponding Z500 field has a negative PNA pattern with a positive anomaly center over the North Pacific, a negative center over western Canada and a positive center over the eastern USA (Fig. 9a).

When u takes on its maximum value, the SST field presents a fairly strong El Niño with positive anomalies (about 2.5-3.0°C) over the central-eastern Pacific (Fig. 9b). The SST warming center shifts eastward by 30-40° longitude relative to the cooling center in Fig. 9a. This asymmetric SST variation between El Niño and La Niña states has also been found by Hsieh (2001a) and Wu and Hsieh (2002). Note that the warming in Fig. 9b does not display a maximum off Peru, in contrast to the first NLPCA mode of SST (Hsieh, 2001b, Fig. 10d). This difference between the first NLPCA and NLCCA modes suggests that warming confined to the eastern equatorial Pacific does not have a strong mid-latitude atmospheric response, in agreement with Hamilton (1988). The Z500 response field contains a PNA pattern (Fig. 9b), that is roughly opposite to that shown in Fig. 9a, but with a notable eastward shift. The zero contour surrounding the North Pacific anomaly lies close to the western coastline of North America during El Niño (Fig. 9b), while it is about 10-15° further west during La Niña (Fig. 9a). The positive anomaly over eastern Canada and USA in Fig. 9a becomes a negative anomaly shifted southeastward in Fig. 9b. Evidently, the Rossby wave train linking the extratropical atmospheric response to the tropical source changes between La Niña and El Niño events in the CCCma model. As the equatorial SST anomalies shift eastward from La Niña to El Niño, the Rossby wave train also shifts eastward. However, the atmosphere's eastward shift is not as large as that in the SST anomalies. This is probably because the atmospheric response is a direct result of tropical heating anomalies (the response to tropical convection) instead of the SST anomalies, although the former is caused by the latter. The amplitude of the Z500 anomaly over the North Pacific is stronger during El Niño than La Niña, but the anomaly over western Canada and USA is weaker during El Niño than La Niña (Figs. 9a and b).

For comparison, CCA mode 1 describes a pattern of atmospheric response to La Niña that, apart from magnitude, is opposite to the response pattern for El Niño (Fig. 9c and 9d). Note, however, that the El Niño response does have somewhat stronger amplitudes. The SST anomaly patterns extracted with CCA are also completely symmetric between the two extremes. The anomaly centers are located at about the average positions between those shown in Fig. 9a and in Fig. 9b.

4.2 NLCCA of SST and SAT

NLCCA was also applied to analyse the covariability of SST and SAT. The first NLCCA mode for SST (shown in Fig. 10a) is very similar to that shown in Fig. 8a. For SAT, considerable nonlinearity occurs between the PC_1 and PC_2 (Fig. 10b). The first NLCCA mode for SST explains 65.5% of the total variance, versus 62.3% explained by the first CCA mode, with r being 0.850. The first NLCCA mode for SAT explains 23.2% of the total variance, versus 22.1% by the first CCA mode, with an MSE ratio r of 0.967. The canonical correlation is 0.605 for NLCCA and 0.588 for CCA.

Figs. 11a,b show the spatial anomaly patterns for both SST and SAT associated with La Niña and El Niño, respectively. When u takes on its minimum value, positive SAT anomalies (about 1°C) appear over the southeastern USA, while much of Canada and northwestern USA are dominated by negative SAT anomalies. The maximum cooling center (−4°C) is located over northwestern Canada and Alaska (Fig. 11a). When u takes on its maximum value (Fig. 11b), the warming center (3°C) is shifted to the southeast of the cooling center in Fig. 11a, with warming over almost all of North America except the southeastern USA. The SAT anomaly patterns here are roughly consistent with the composite analysis from observed data by Hoerling et al. (1997). The first CCA mode for SAT and SST is shown for reference in Figs. 11c,d.

We note that the nonlinear SAT response is more strongly captured by NLCCA than by composite analysis (compare Figs. 9a,b with Figs. 7a,b and Figs. 11a,b with Figs. 7c,d). This is not surprising because the averaging in the composite method mixes the strong and weak events, thereby producing weaker features than NLCCA. Another disadvantage with the composite approach is that one must decide, a priori, which events to include in the composites.

5 Summary and discussions

Two nonlinear multivariate statistical techniques, NLPCA and NLCCA, developed from neural networks, were applied to investigate the nonlinear behaviour of the North American winter climate as simulated by the CCCma AGCM2. NLPCA was used to find the nonlinear modes which could account for the maximum variance in a dataset, while NLCCA was used to detect the most strongly correlated nonlinear modes relating the atmospheric (Z500 or SAT) and tropical Pacific SST variability. These two methods provided valuable nonlinear diagnostics of a GCM simulation.

Nonlinearity is detected by NLPCA in both the winter Z500 and SAT anomaly fields, which displays asymmetric spatial patterns of variability during the opposite extremes of the nonlinear principal component (NLPC) u. The first NLPCA mode for Z500 describes a negative PNA pattern during the minimum u, and a weakened jet stream pattern during maximum u. Although NLPCA was conducted on the Z500 and SAT anomalies separately, the derived NLPCA modes have similar spatial patterns.

The first NLCCA mode shows there is nonlinearity in both the SST anomaly field and the related atmospheric response fields (Z500 and SAT). The leading NLCCA mode for SST is a nonlinear ENSO mode, showing asymmetry between the warm El Niño states and the cool La Niña states with a 30-40° longitude eastward shift of the SST anomalies during El Niño relative to the anomalies during La Niña. The leading NLCCA mode for Z500 is a nonlinear PNA teleconnection pattern, in which the positive PNA pattern associated with El Niño is situated about 10-15° east of the negative PNA pattern associated with La Niña. As with NLPCA, the spatial anomaly pattern of the first NLCCA mode for the SAT generally corroborates with the corresponding Z500 anomaly pattern. The NLCCA was able to extract a stronger nonlinear atmospheric response to ENSO than the composite method.

Even with NLCCA, the nonlinearity of the atmospheric responses to ENSO reported in this paper is weaker than that found by Hoerling et al. (1997), where an appreciable 30° longitude phase shift between the warm and cold event circulation composites was detected from observational data. However, we should note that, the atmospheric sample available here from an ensemble of 6 runs is much bigger than is available for an observational study, so the opportunities for overfitting and consequently finding apparent strong nonlinearity are somewhat reduced. It might therefore be more appropriate to compare our results with other models. Hoerling et al. (2001) surveyed the response to ENSO in four GCMs. Nonlinearity was found in all four models, existing in both the strength of the midlatitude response and its spatial phase. Similar to our results, the spatial phase displacements of the four GCMs' teleconnections during warm events and cold events are smaller than observed. The degree of nonlinear atmospheric responses to ENSO is model dependent.

Hoerling et al. (2001), used *one-sided linear regression* to analyse the nonlinearity. The procedure involves calculating linear regressions between all occurrences of one sign of an SST index (e.g. the leading EOF PC of tropical Pacific SST anomalies) and the response fields (e.g. Z500 or SAT anomalies). For comparison, this method was also applied to CCCma AGCM2's simulation and the results are shown in Fig. 12. The regression patterns for Z500 and SAT anomalies are generally consistent with the spatial patterns given by NLCCA (Fig. 9a,b and Fig. 11a,b), with the positive PNA teleconnection shifted 10-15° eastward of the negative PNA teleconnection (Figs. 12a,b). The phase difference in the CCCma AGCM2, as indicated by the placement of zero contours, is comparable to that in the NCAR CCM3 and ECHAM-3 models, but weaker than that in the GFDL and MRF models (see Hoerling et al. (2001) for details of the four models). The negative Z500 anomalies over the North Pacific are much stronger during warm events than the positive anomalies that occur during cold events. However, the anomalies over the North American continent have similar amplitudes during warm and cold events (Fig. 12a,b), which can also be seen in the NLCCA results (Fig. 9a,b).

Hence the CCCma AGCM2 is basically capable of simulating the nonlinear responses of North American climate to ENSO, although the nonlinearity is somewhat weaker than that observed. NL-

CCA, which successfully extracted the nonlinear mode, offers a fully nonlinear diagnostic tool to study GCM output.

Acknowledgements

This work was supported by the Natural Sciences and Engineering Research Council of Canada via a strategic grant to Hsieh, Shabbar and Zwiers, and a research grant to Hsieh. This paper was improved by constructive comments on an earlier draft provided by Greg Flato and John Fyfe.

References

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.

Boer, G.J., N.A. McFarlane and M. Lazare, 1992: Greenhouse gas-induced climate change simulated by the CCC second generation circulation model. *J. Climate,* **5,** 1045-1077.

Hamilton, K., 1988: A detailed examination of the extratropical response to tropical El Nino/Southern Oscillation events. *J.Climatology*, **8,** 67-86.

Hamilton, K., and W.W. Hsieh, 2002: Representation of the QBO in the tropical stratospheric wind by nonlinear principal component analysis. *J. Geophys. Res.–Atmos.*, in press.

Hannachi, A., Toward a nonlinear identification of the atmospheric response to ENSO, *J. Climate,* **14,** 2138-2149, 2001.

Hoerling, M.P., A. Kumar, and M. Zhong, 1997: El Niño, La Niña and the nonlinearity of their teleconnections, *J. Climate,* **10,** 1769-1786.

Hoerling, M.P., A. Kumar, and T. Xu, 2001: Robustness of the nonlinear climate response to ENSO's extreme phases, *J. Climate,* **14,** 1277-1293.

Horel, J,D., and J.M. Wallace, 1981: Planetary scale atmospheric phenomena associated with the Southern Oscillation. *Mon. Wea. Rev.,* **109,** 813-929.

Hoskins, B., and D. Karoly, 1981: The steady linear response of a spherical atmosphere to thermal and orographic forcing. *J. Atmos. Sci.,* **38,** 1179-1196.

Hsieh, W.W. and B. Tang, 1998: Applying neural network models to prediction and data analysis in meteorology and oceanography, *Bull. Amer. Meteor. Soc.,* **79,** 1855-1870.

Hsieh, W.W., 2000: Nonlinear canonical correlation analysis by neural networks, *Neural Networks,* **13,** 1095-1105.

Hsieh, W.W., 2001a: Nonlinear canonical correlation analysis of the tropical Pacific climate variability using a neural network approach, *J. Climate, 14,* 2528-2539.

Hsieh, W.W., 2001b: Nonlinear principal component analysis by neural network. *Tellus,* **53A,** (in press).

Kramer, M.A., 1991: Nonlinear principal component analysis using autoassociative neural networks, *AIChE Journal,* **37,** 233-243.

McFarlane, N.A., G.J. Boer, J.-P. Blanchet, and M. Lazare, 1992: The Canadian Climate Center second generation general circulation model and its equilibrium climate. *J. Climate,* **5,** 1013-1044.

Monahan, A.H., 2000: Nonlinear principal component analysis by neural networks: theory and application to the Lorenz system. *J. Climate,* **13,** 821-835.

Monahan, A.H., 2001: Nonlinear principal component analysis: Tropical Indo-Pacific sea surface temperature and sea level pressure. *J. Climate,* **14,** 219-233.

Rayner, N.A., E.B. Horton, D.E. Parker, C.K. Folland and R.B. Hackett, 1996: Version 2.2 of the global sea-ice and sea surface temperature data set, 1903-1994. *Hadley Centre Climate Research Technical Note,* CRTN 74, Hadley Centre, Meteorological Office, Bracknell, U.K., 21 pp.

Richman, M.B., and D.L. Montroy, 1996: Nonlinearities in the signal between El Niño/La Niña events and North American precipitation and temperature. Preprints, *13th Conf. on Probability and Statistics in the Atmospheric Sciences*, San Francisco, CA, Amer. Meteor. So., 90-97.

Shabbar, A., and M. Khandekar, 1996: The impact of El Niño-Southern Oscillation on the temperature field over Canada. *Atmos.-Ocean,* **34,** 401-416.

Shabbar, A., B. Bonsal and M. Khandekar, 1997: Canadian precipitation patterns associated with Southern Oscillation. *J. Climate,* **10,** 3016-3027.

Sittel, M.C., 1994: Marginal probabilities of the extremes of ENSO for temperature and precipitation in the southeastern United States. *FSU Tech. Rep.,* 94-1.

Smith, T.M., R.W. Reynolds, R.E. Livezey, and D.C. Stokes, 1996: Reconstruction of historical sea surface temperatures using empirical orthogonal functions. *J. Climate,* **9,** 1403-1420.

Trenberth, K.E., G.W. Branstator, D. Karoly, A. Kumar, N.-C. Lau, and C. Ropelewski, 1998: Progress during TOGA in understanding and modelling global teleconnections associated with tropical sea surface temperatures. *J. Geophys. Res.,* **103,** 14291-14324.

Wallace, J.M., and D. Gutzler, 1981: Teleconnection in the geopotential height field during the Northern Hemisphere winter. *Mon. Wea. Rev.,* **109,** 784-812.

Wu, A., W.W. Hsieh, and A. Shabbar, 2002: Nonlinear characteristics of the surface air temperature over Canada. *J. Geophys. Res.–Atmos.* (in press).

Wu, A., W.W. Hsieh, 2002: Nonlinear canonical correlation analysis of the tropical Pacific wind stress and sea surface temperature. *Climate Dyn.,* **19,** 713-722.

Zwiers, F.W., X. Wang and J. Sheng, 2000: Effects of specifying bottom boundary conditions in an ensemble of atmospheric GCM simulations. *J. Geophys. Res.,* **105,** 7295-7315.

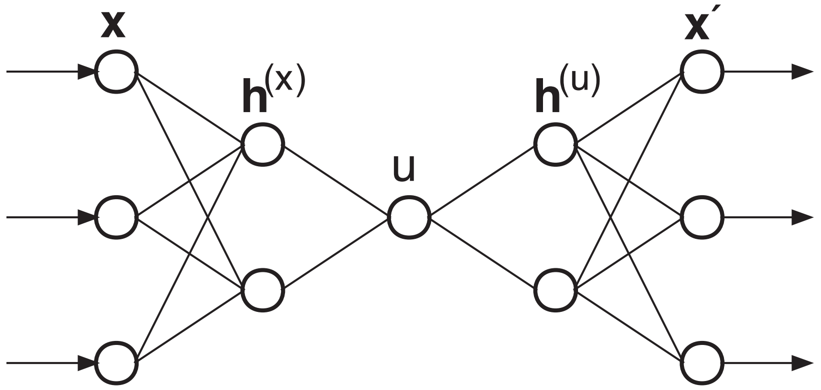

Figure 1: A schematic diagram of the neural network (NN) model for calculating nonlinear PCA (NLPCA). The network is a standard feed-forward NN (i.e. multi-layer perceptron), consisting of several layers of 'neurons' (i.e. variables). The network maps from the input layer on the left to the encoding layer, then to the bottleneck layer (with a single neuron), to the decoding layer, then finally to the output layer on the right. In the applications here, the hidden layers $\mathbf{h}^{(x)}$ and $\mathbf{h}^{(u)}$ both have 3 neurons, while the input $\mathbf{x}$ are the 5 leading principal components.

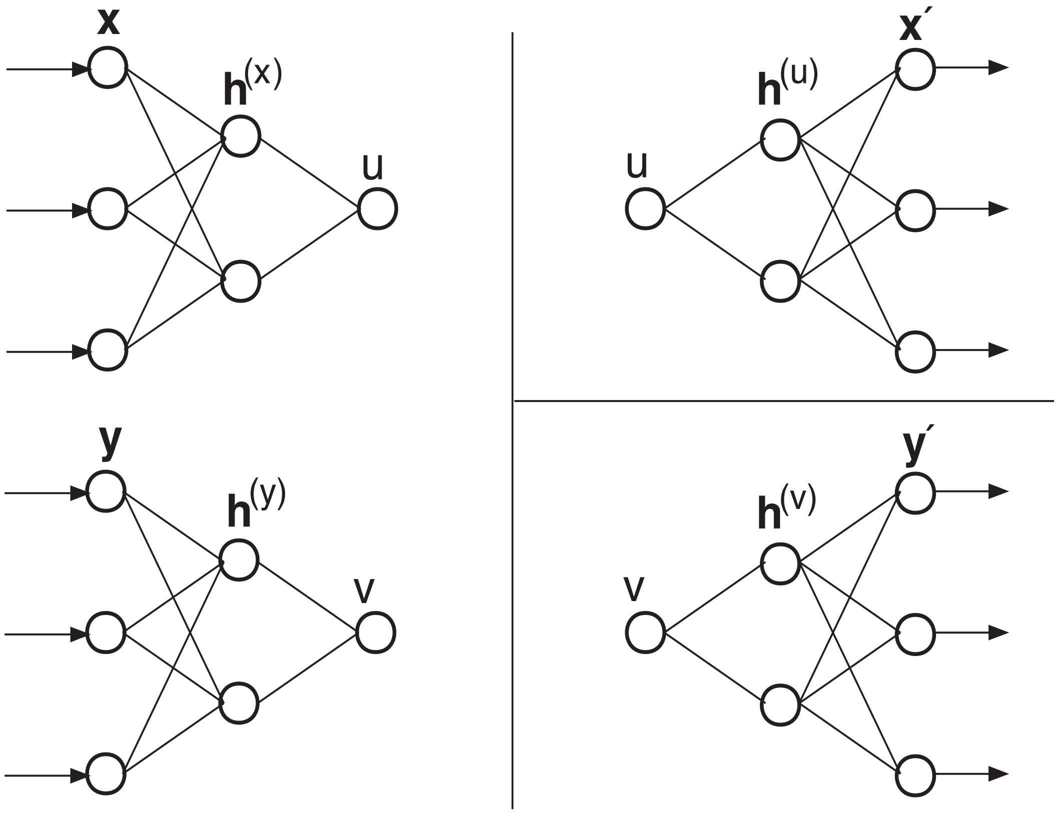

Figure 2: A schematic diagram illustrating the three feed-forward NNs used to perform nonlinear canonical correlation analysis (NLCCA). In the applications here, all the hidden layers $\mathbf{h}^{(x)}$, $\mathbf{h}^{(y)}$, $\mathbf{h}^{(u)}$ and $\mathbf{h}^{(v)}$ have 3 neurons each.

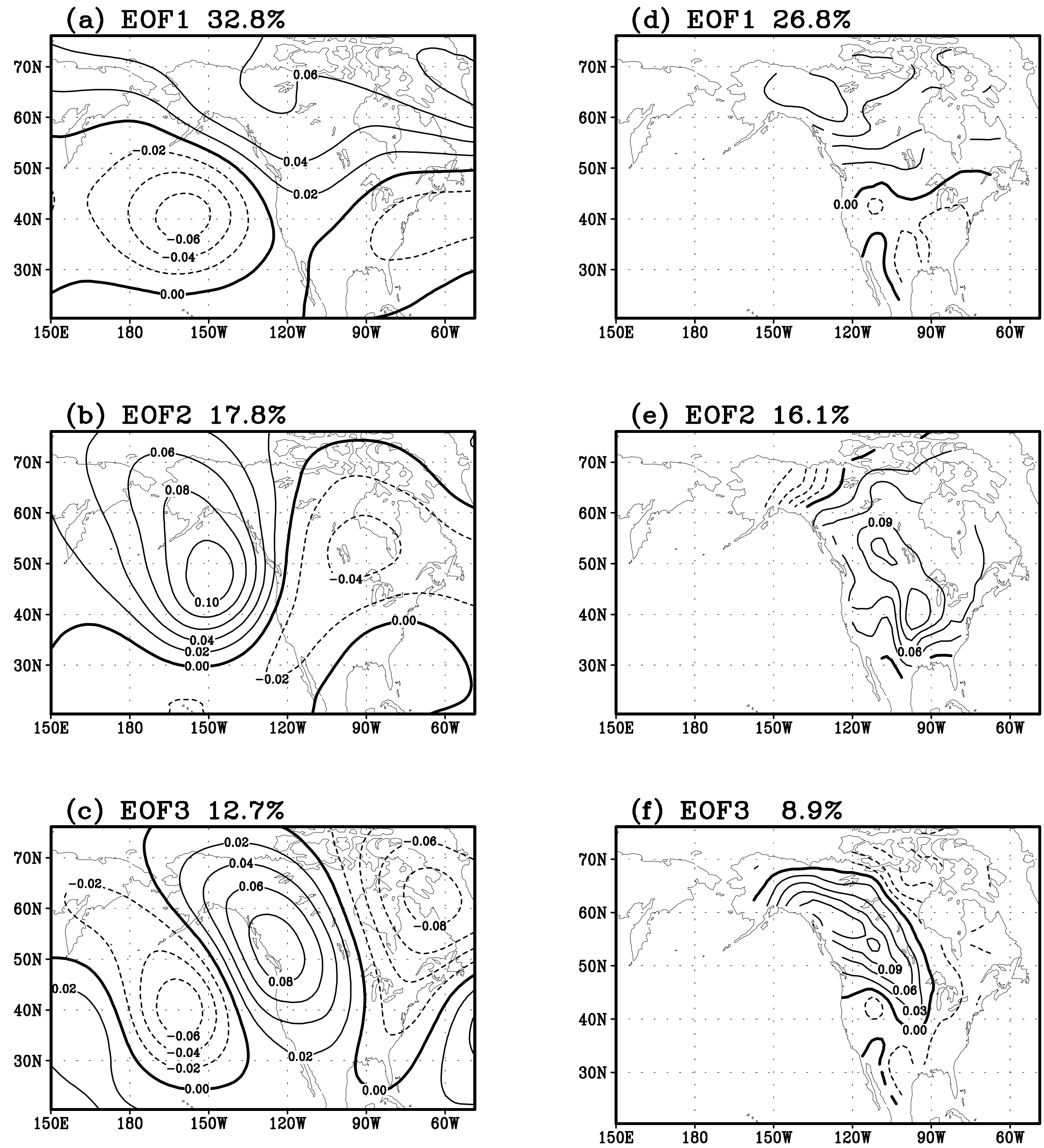

Figure 3: The first 3 EOFs of the winter 500-mb geopotential height (Z500) anomalies (left column) and surface air temperature (SAT) anomalies (right column). Solid curves denote positive contours, dashed curves, negative contours, and thick curves, zero contours. The contour interval is 0.02 in panels (a), (b) and (c), and 0.03 in panels (d), (e) and (f). The EOFs have been normalized to unit norm. If the sign in panel (b) is reversed, then over the North American continent, the patterns in the left column generally agree with those in the right column.

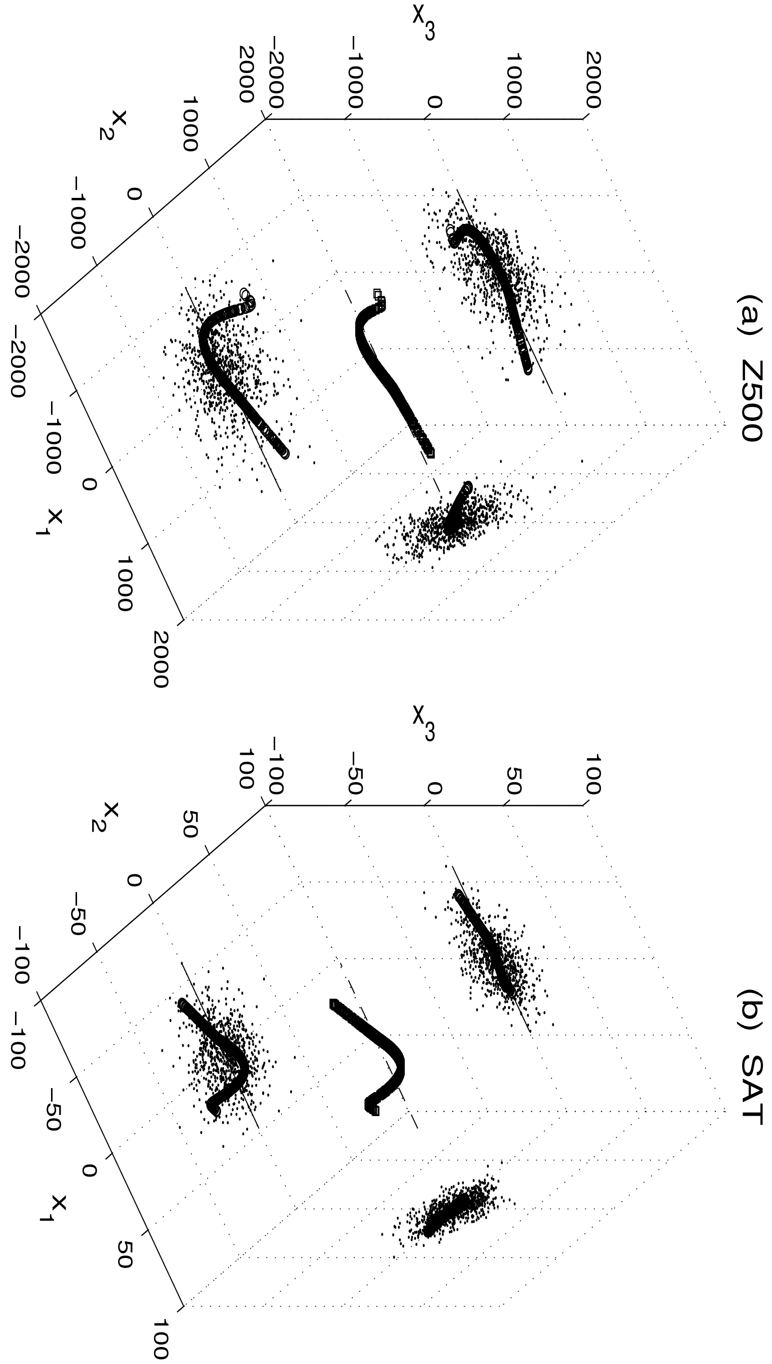

Figure 4: The first NLPCA mode for the winter (a) Z500 anomalies and the (b) SAT anomalies plotted as (densely overlapping) squares (which produce a thick curve) in the PC_1-PC_2-PC_3 3-D space. The linear (PCA) mode is shown as a dashed line. The NLPCA mode and the PCA mode are also projected onto the PC_1-PC_2 plane, the PC_1-PC_3 plane, and the PC_2-PC_3 plane, where the projected NLPCA is indicated by (densely overlapping) circles, and the PCA by thin solid lines, and the projected data points by the scattered dots. The minimum nonlinear principal component (NLPC) u occurs at the left end, and the maximum u at the right end of the curve in both (a) and (b).

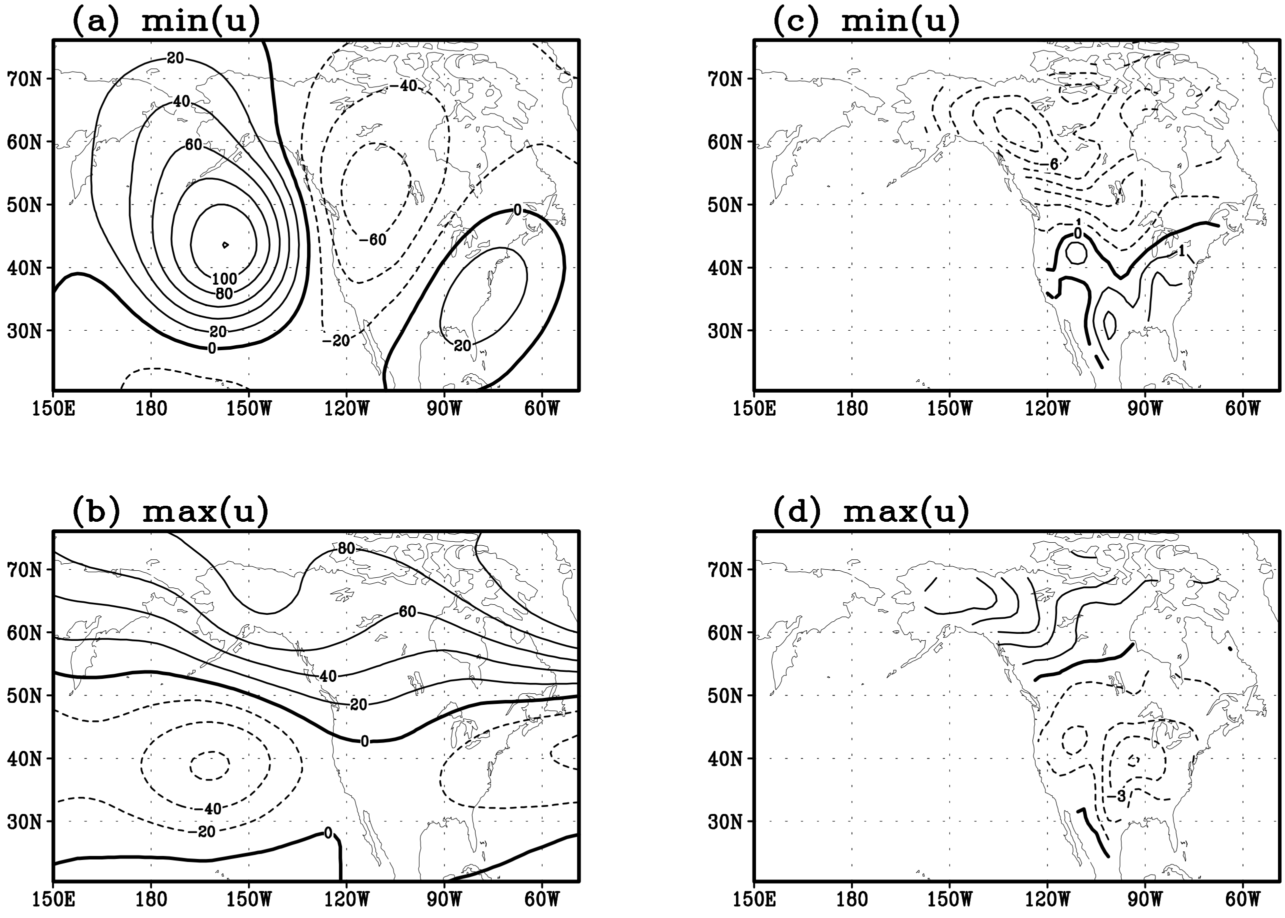

Figure 5: The spatial patterns of the NLPCA mode 1 as the NLPC u takes its minimum and maximum values. The Z500 anomaly patterns for minimum u and maximum u are shown in panels (a) and (b), respectively, with contour intervals of 20 m, and the SAT anomaly patterns, in panels (c) and (d), respectively, with contour intervals of 1°C.

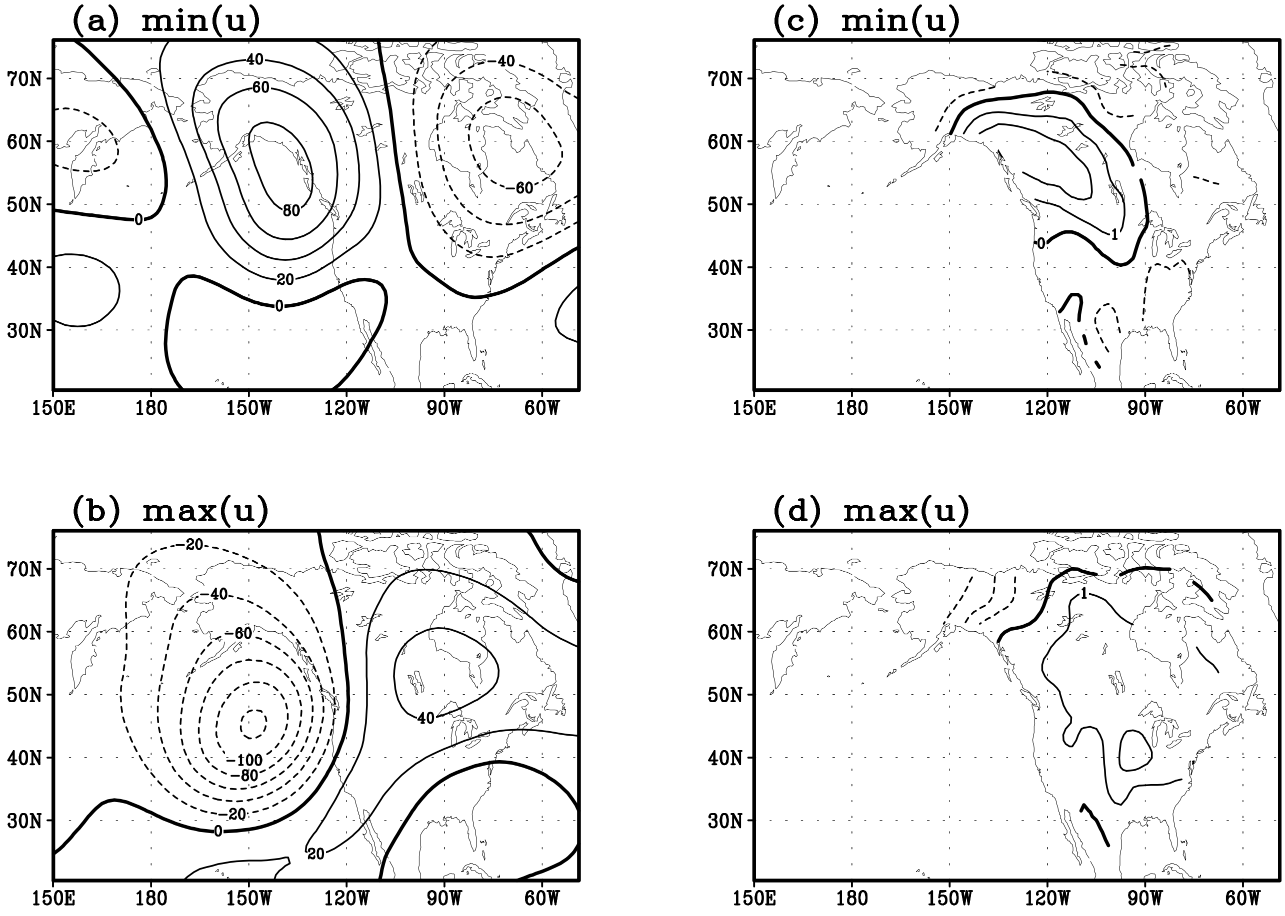

Figure 6: Similar to Fig. 5 but for the NLPCA mode 2.

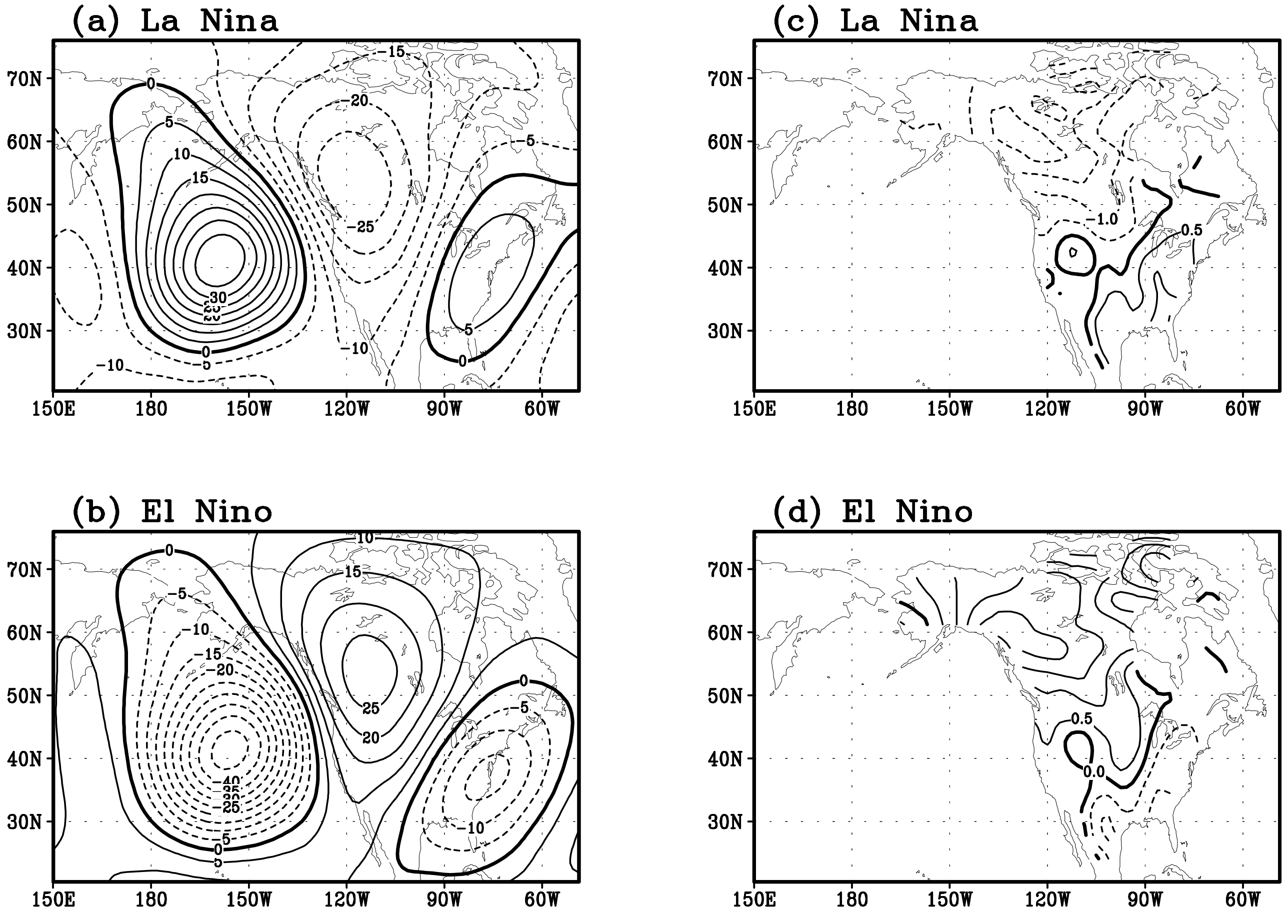

Figure 7: Composites of the winter Z500 and SAT anomalies for La Niña and El Niño states. The Z500 anomalies are shown in panels (a) and (b), respectively, with contour interval of 5 m, and the SAT anomalies, in panels (c) and (d), respectively, with contour interval of 0.5°C.

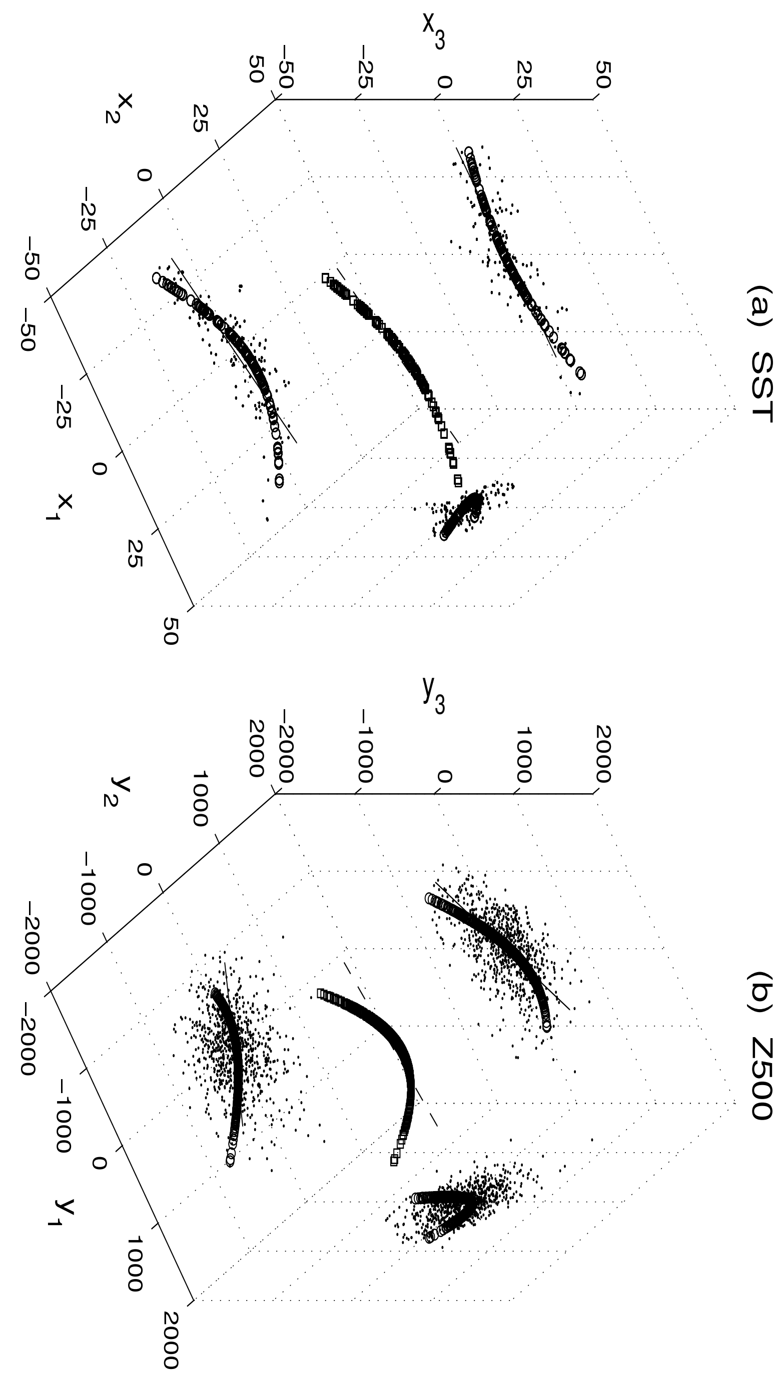

Figure 8: The first NLCCA mode between (a) the winter tropical Pacific SST anomalies and (b) the winter Z500 anomalies, plotted as (overlapping) squares in the PC_1-PC_2-PC_3 3-D space. The linear (CCA) mode is shown as a dashed line. The NLCCA mode and the CCA mode are also projected onto the PC_1-PC_2 plane, the PC_1-PC_3 plane, and the PC_2-PC_3 plane, where the projected NLCCA is indicated by (overlapping) circles, and the CCA by thin solid lines, and the projected data points by the scattered dots. There is no time lag between the SST and the corresponding Z500 data.

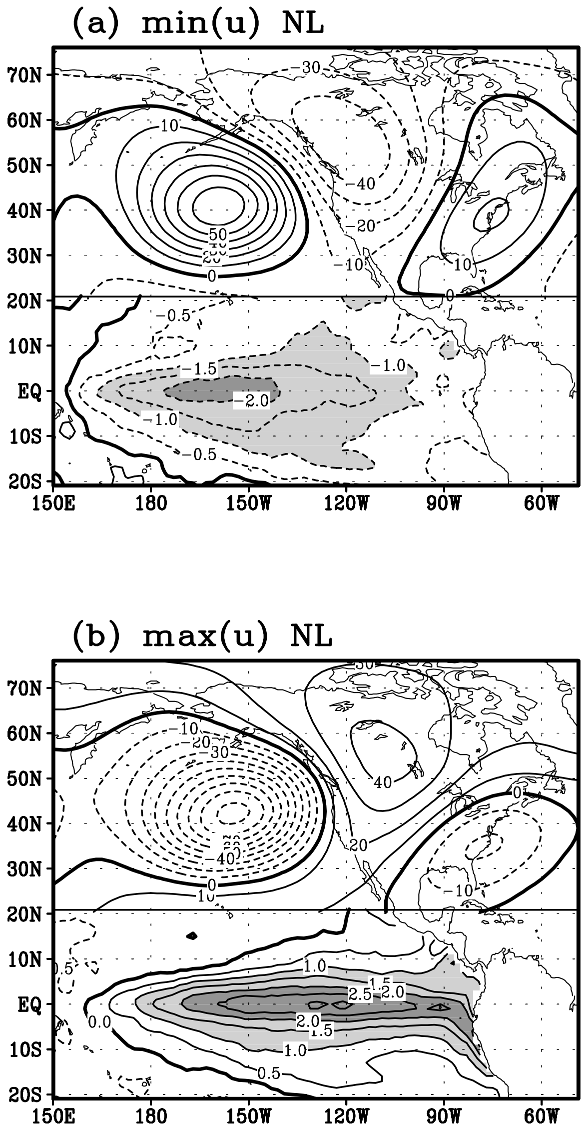

Figure 9: The spatial patterns for the first NLCCA mode between the winter Z500 anomalies and the tropical Pacific SST anomalies as the canonical variate u takes its (a) minimum value and (b) maximum value. The Z500 anomalies with contour intervals of 10m are shown north of 20°N. SST anomalies with contour intervals of 0.5°C are displayed south of 20°N. The SST anomalies greater than +1°C or less than -1°C are shaded, and heavily shaded if greater than +2°C or less than -2°C. The linear CCA mode 1 is shown in panels (c) and (d) for comparison.

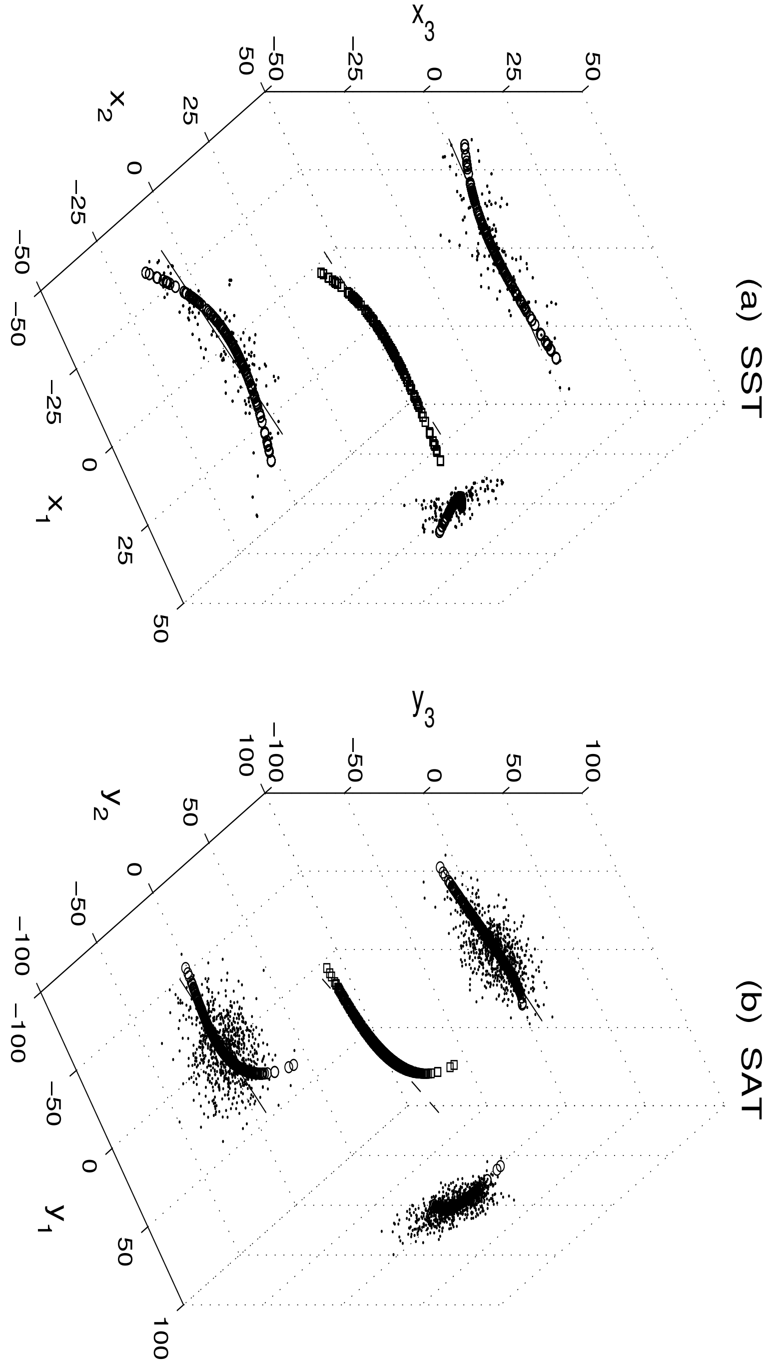

Figure 10: Similar to Fig. 8 but for the NLCCA mode 1 between the SAT anomalies and the tropical SST anomalies.

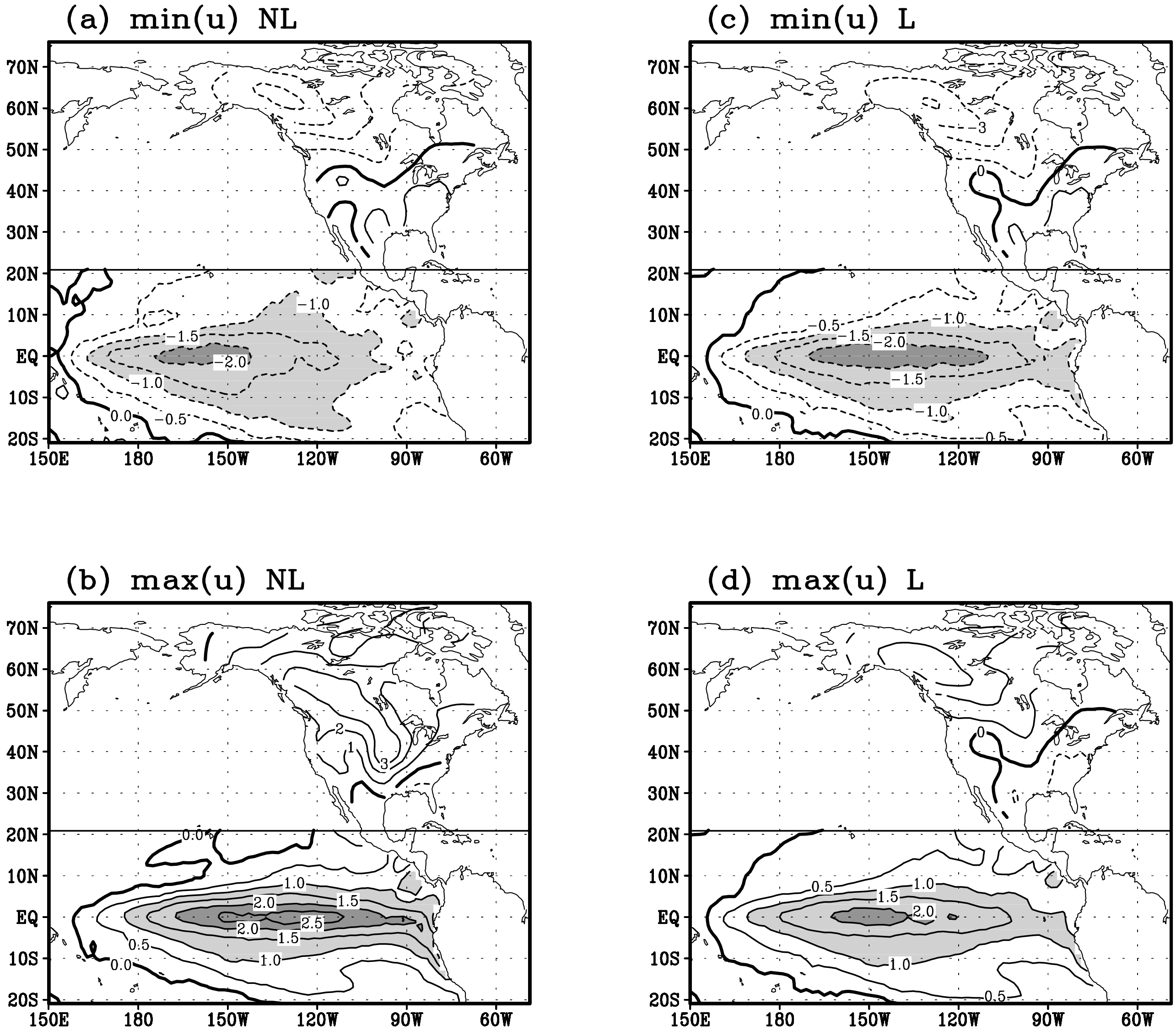

Figure 11: Similar to Fig. 9 but for the spatial patterns for the NLCCA mode 1 between the SAT anomalies and the tropical SST anomalies. The contour interval for the SAT anomalies is 1°C.

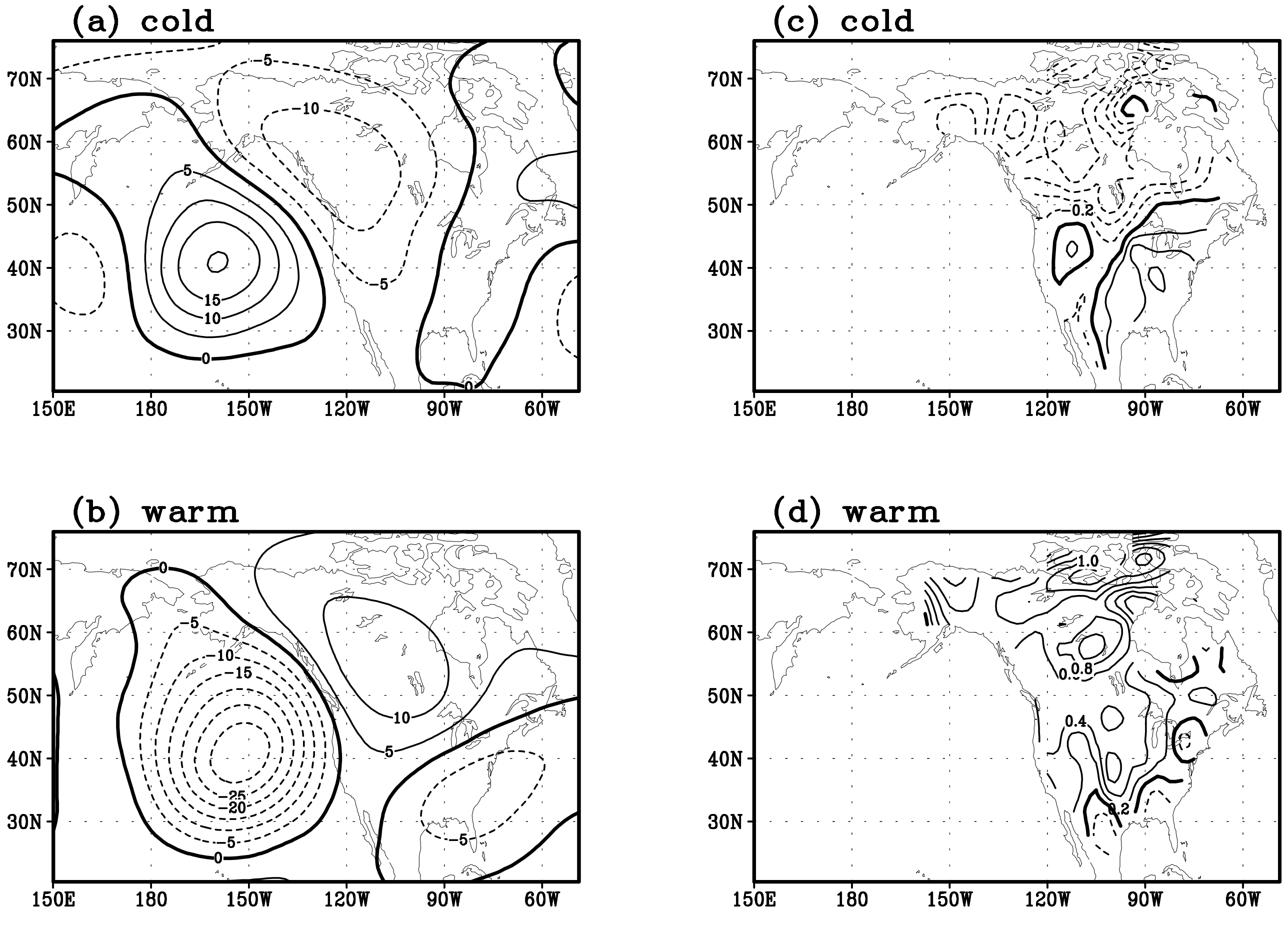

Figure 12: One-sided regressions of the winter (DJF) 500-mb height and SAT for warm and cold phases of the leading EOF of tropical Pacific SSTs. The monthly height and SAT anomalies were regressed onto the positive and negative phases of the first EOF principal component (PC) of the tropical SST anomalies separately. Shown are height and SAT anomalies associated with +1 standard deviation of the EOF PC1. Contour interval is 5m in panels a and b, and 0.2°C in panels c and d.