

Hybrid coupled modeling of the tropical Pacific using neural networks

Shuyong Li, William W. Hsieh and Aiming Wu

Dept. of Earth & Ocean Sciences

University of British Columbia

6339 Stores Road

Vancouver, B.C. V6T 1Z4, Canada

Submitted to J. Geophys. Res. (Oceans)

May 28, 2005

Corresponding author: William Hsieh, Dept. of Earth & Ocean Sciences, University of British Columbia,
6339 Stores Road, Vancouver, B.C. V6T 1Z4, Canada

Email: whsieh@eos.ubc.ca, Tel: (604) 822-2821, Fax: (604) 822-6088

Abstract

To investigate the potential for improving hybrid coupled models (HCM) of the tropical Pacific by the use of neural network (NN) methods for nonlinear regression, NN was introduced for the nonlinear parametrization of the subsurface temperature in the Lamont ocean model, and for the nonlinear estimation of the wind stress anomalies (WSA) from the sea surface temperature anomalies (SSTA). For comparison, corresponding linear regression (LR) models were also built. By combining the NN or the LR version of the ocean model and the atmospheric model, four HCMs resulted. For the coupled model Niño3 SSTA spectrum, using NN in the ocean model produced a much broader spectrum than using LR, which gave basically a single narrow spectral peak. Using NN in the atmospheric model in addition to the ocean model further broadened the SSTA spectrum, yielding a spectrum with two main peaks as observed. Principal component analysis (PCA) and nonlinear PCA (NLPCA) were used to analyze the SSTA and WSA. By comparing the NLPCA mode 1 and the PCA mode 1, we found that all the coupled models (including the original Lamont coupled model) were too linear compared to the observations. However, using NN in the ocean model and in the atmospheric model was able to alleviate the weak nonlinearity in the coupled models.

1 Introduction

Numerous models have been developed to study and to forecast the El Niño-Southern Oscillation (ENSO) phenomenon, the most important interannual variability in the tropical Pacific coupled atmosphere-ocean climate system. With complexity lying somewhere between the computationally demanding coupled general circulation models and simple models, “intermediate” coupled models (e.g. *Zebiak and Cane* [1987]) are widely used. Alternatively, ENSO models can be divided into 3 classes: dynamical coupled models, statistical models and hybrid coupled models [*Barnston et al.*, 1994].

A hybrid coupled model (HCM) consists of a dynamical ocean model coupled to a statistical atmospheric model [*Syu et al.*, 1995; *Barnett et al.*, 1993; *Tang and Hsieh*, 2001]. The design of the hybrid coupled model uses the fact that the ocean possesses long-term memory in the coupled atmosphere-ocean system, while the atmosphere can be treated as a fast adjusting component, so a steady-state statistical model for the atmosphere can be used. The hybrid coupled model uses an empirical atmospheric component, based on the assumption that for monthly or longer time scales, contemporaneous correlation between wind stress and oceanic variables such as sea surface temperatures (SST) is associated with the atmosphere’s rapid non-local adjustment to the oceanic anomaly patterns throughout the basin [*Syu et al.*, 1995]. The main merits of a hybrid coupled model are: (1) lower computing cost than a full coupled general circulation model (GCM) [*Blank et al.*, 1997]; (2) The climate drift problem is avoided; (3) and comparable, or even better ENSO simulation and prediction skills relative to a coupled GCM [*Palmer and Anderson*, 1994].

There are two important aspects affecting the HCM performance: One is the construction of the empirical atmospheric model, i.e. the method used to estimate the surface wind stress field from a given ocean state. Most of the empirical atmospheric models used in HCMs so far are linear statistical models and the methods used include correlation [*Latif and Villwork*, 1990], linear regression with EOF

(empirical orthogonal function) modes [Barnett *et al.*, 1993] and SVD (singular value decomposition) [Syu *et al.*, 1995]. Tang *et al.* [2001] tried to improve the empirical atmospheric model by a nonlinear regression approach using artificial neural network (NN) methods. This NN atmosphere was then coupled to a dynamical ocean model for ENSO prediction [Tang and Hsieh, 2002, 2003]. The other aspect affecting the HCM performance is the dynamical ocean model, which in this study was adapted from the ocean component of the Lamont coupled model [Zebiak and Cane, 1987]. The original Lamont ocean model used a simple parameterization scheme for the subsurface temperature T_{sub} , which has been replaced here by an NN nonlinear regression scheme. For comparison, linear regression (LR) was also tested in place of NN.

Upon coupling the oceanic and atmospheric models, a total of 4 HCMs resulted: (1) The Lamont ocean model with a nonlinear NN T_{sub} parameterization coupled to a nonlinear NN atmosphere (henceforth referred to as the NONA HCM), (2) the ocean model with NN coupled to an LR atmospheric model (the NOLA HCM), (3) the ocean model with LR coupled to an NN atmosphere (LONA), and (4) the ocean model with LR coupled to an LR atmosphere (LOLA). We compared these 4 HCMs to see the effects from incorporating nonlinearity in the T_{sub} parameterization and in the atmospheric response to the SST, leading to a better understanding of the role of nonlinearity in intermediate coupled models of ENSO.

This paper is organized as follows: The data and models used are described in Sect. 2. In Sect. 3, the interannual variability of the 4 HCMs over a 167-year period is examined. In Sect. 4, principal component analysis (PCA) and nonlinear principal component analysis (NLPCA) are applied to the SST anomalies (SSTA) simulated by the 4 HCMs, and the results are compared with observations, while in Sect. 5, NLPCA is applied to the wind stress anomalies simulated by the HCMs.

2 Data and models

2.1 Data

The monthly wind stress on a $2^\circ \times 2^\circ$ grid for the period of January 1964 – January 2002 was obtained from Florida State University (FSU). The monthly SST came from the reconstructed historical SST dataset by *Smith et al.* [1996] for the period of January 1950 – December 2001 with a 2° by 2° resolution over the global oceans. The SST were converted to the Lamont ocean model grid using linear interpolation.

2.2 Neural network models

NN is a nonparametric statistical model for extracting nonlinear relations in the data [*Bishop, 1995; Hsieh and Tang, 1998*]. Figure 1a shows a common NN model configuration for nonlinear regression. A “hidden” layer of variables, called “neurons” in NN jargon, is placed between the input and output variables. The j th hidden neuron is assigned the value y_j , given by

$$y_j = \tanh\left(\sum_i w_{ij}x_i + b_j\right),$$

where x_i is the i th input values, w_{ij} and b_j are the weight and bias parameters respectively. The hyperbolic tangent function is used as the transfer function (other forms of the transfer function can also be used, since it only serves as a basis function).

The output neuron z is calculated by a linear combination of the neurons in the hidden layer, i.e.

$$z = \sum_j \tilde{w}_j y_j + \tilde{b}.$$

To construct a NN model for nonlinear regression, the predictor variables are the inputs, and the predictands are the outputs of the network.

The cost function

$$J = \langle (z - z_{\text{obs}})^2 \rangle + p \sum_l W_l^2$$

has a first term measuring the mean square error between the model output z and the observed data z_{obs} , and a second term penalizing the use of excessive weight and bias parameters, where p is a penalty parameter and W_l represents all the weight and bias parameters of the NN. The NN model is trained by finding the optimal parameters w_{ij} , \tilde{w}_j , b_j and \tilde{b} as the cost function is minimised. An ensemble of 200 NN models with random initial parameters were trained, then the 30 NN models with the smallest cost functions were selected, so the final output of the NN model was actually the ensemble average of the 30 individual NN model outputs. Ensembles can alleviate the problem of multiple minima in the cost function, commonly found with NN models [Hsieh and Tang, 1998].

2.3 Ocean model

The ocean model used in this research is the ocean component of the Lamont coupled model (henceforth the Lamont ocean model) but with a new parameterization of the subsurface temperature T_{sub} and two extra terms added to the temperature equation for the surface layer. It is a reduced gravity model consisting of an active upper layer with a fixed depth surface mixed layer, overlying a motionless deep layer, covering the tropical Pacific from 29°S to 29°N. The resolution of ocean dynamics is 2° in longitude and 0.5° in latitude, but that of SST physics is 5.625° in longitude and 2° in latitude. The integration time step is 10 days.

The temperature equation for the surface layer in the Lamont model [Zebiak and Cane, 1987, Eq. A11] has two extra terms added [Boulanger and Menkes, 2001]:

$$\begin{aligned} \frac{\partial T}{\partial t} = & -\mathbf{u}_1 \cdot \nabla(\bar{T} + T) - \bar{\mathbf{u}}_1 \cdot \nabla T - \{M(\bar{w}_s + w_s) - M(\bar{w}_s)\}\bar{T}_z - M(\bar{w}_s + w_s)\frac{T - T_e}{H_1} \\ & -\alpha_s T - K_T \partial_z T + A_H \nabla^2 T, \end{aligned} \quad (1)$$

where T (\bar{T}) is the anomalous (mean) SST, \mathbf{u}_1 ($\bar{\mathbf{u}}_1$) and w_s (\bar{w}_s) the anomalous (mean) horizontal currents and upwelling, respectively, $M(x)$ a function which equals x if x is positive and equals zero otherwise, the entrainment temperature anomaly $T_e = \gamma T_{\text{sub}} + (1 - \gamma)T$, and α_s a damping parameter.

The two extra terms added are the vertical mixing term $-K_T \partial_z T$ and the horizontal diffusion term $A_H \Delta_H T$, where we take $K_T = 2.9 \times 10^{-5} \text{ m s}^{-1}$ and $A_H = 2000 \text{ m}^2 \text{ s}^{-1}$.

In the Lamont ocean model, T_{sub} , the ocean temperature anomaly below the mixed layer, is parameterized in terms of the thermocline depth anomaly h :

$$T_{\text{sub}} = \begin{cases} A_1 \{ \tanh[B_1(\bar{h} + h)] - \tanh(B_1 \bar{h}) \}, & h \geq 0 \\ A_2 \{ \tanh[B_2(\bar{h} - h)] - \tanh(B_2 \bar{h}) \}, & h < 0, \end{cases} \quad (2)$$

where $A_1 = 28^\circ\text{C}$, $B_1 = 0.0125\text{m}^{-1}$, $A_2 = -40^\circ\text{C}$, and $B_2 = 0.03\text{m}^{-1}$, and $\bar{h}(x)$ is specified using an observed equatorial thermocline distribution after *Colin et al.* [1971].

In this paper, we used an NN or LR model to estimate T_{sub} from h . First the Lamont ocean model was forced by the FSU wind stress anomalies from 1964-2001, and the model thermocline depth anomalies and current anomalies were extracted. T_{sub} was then inversely estimated [Zhang et al., 2004] from the SST anomaly equation (1) from 1964-2001 using the simulated current anomalies and the observed monthly SST fields from *Smith et al.* [1996].

Principal component analysis (PCA), i.e. empirical orthogonal function (EOF) analysis, was first applied to the thermocline depth anomaly $h(x, y, t)$ and the subsurface temperature anomaly $T_{\text{sub}}(x, y, t)$:

$$h(x, y, t) = \sum_n a_n(t) f_n(x, y),$$

$$T_{\text{sub}}(x, y, t) = \sum_n c_n(t) e_n(x, y),$$

where n indicates the n th mode. For the thermocline depth anomaly h , the first four PCA modes accounted for 48% of the variance, while for the T_{sub} , the first five modes accounted for 69% of the variance. The NN used has at most four input neurons, namely the first four principal components (PCs) $a_n(t)$ for h , and the single output neuron is one of the five leading PCs $c_n(t)$ for T_{sub} , i.e. a different NN model was used to predict each predictand c_n , using at most 4 a_n as predictors, with no time lag between the predictors and the predictand. Since the NN model is performing nonlinear

regression, we checked the role of nonlinearity by building a corresponding linear regression (LR) model for comparison.

Data from 1964 to 2001 were used to train the NN model or the corresponding LR model. For each T_{sub} PC, we used cross-validation to find out the number of hidden neurons, the number of PC predictors and the weight penalty parameter p for the the best NN result, and the number of PC predictors for the best LR result. Cross-validation was performed as follows: First, the data record was divided into 5 equal segments. One segment was selected to be the test data and the rest, training data. The NN (or LR) model was built using the training data only, and model simulations on the independent test data were obtained. Next, another segment was selected as the test data, and a new model built. This was repeated until the entire data record had been used for independent testing. The model test results on the entire record were compared with the observed T_{sub} PCs. From cross-validation, the optimal NN or LR T_{sub} model was obtained.

2.4 Atmospheric model

The FSU wind stress anomalies were first smoothed by a 3-month running mean, then the ocean models were driven by the FSU wind stress anomalies from 1964-2001. PCA was applied to the model SSTA $T(x, y, t)$, and combined PCA to both components of the FSU wind stress anomalies $\tau(x, y, t)$:

$$T(x, y, t) = \sum_n \tilde{a}_n(t) \tilde{f}_n(x, y),$$

$$\tau(x, y, t) = \sum_n \tilde{c}_n(t) \tilde{\mathbf{e}}_n(x, y).$$

For T , the first four PCA modes accounted for 91.6 % and 91.8 % of the variance respectively for the ocean model with NN and LR T_{sub} parametrization, while for τ , the first 7 modes contained 39.5% of the variance. The NN or LR used has at most four inputs, namely the first four PCs for T , and the single output is one of the first 7 PCs for τ .

Tables 1 shows the cross-validated skills for the first 7 wind stress PCs attained by the NN and

LR models where the predictors were the SSTA PCs from the ocean model with an NN or LR T_{sub} parametrization. In general, the NN atmospheric model predicted the wind stress PCs slightly better than the LR. PC2 was harder to predict than PC3, and PC4 was the most difficult. Henceforth, only if the cross-validated correlation of the wind stress PC is over 0.1, will this PC be included in the atmospheric model, i.e. PC4 will not be used in all the atmospheric models, and some of the higher PCs are also excluded from the LR atmospheric model.

3 Interannual variability from the coupled models

As common with statistical atmospheric models, the variance of the predicted wind stress was lower than that observed, hence the estimated wind stress were scaled up by an adjustable scale factor μ [Barnett *et al.*, 1993; Tang and Hsieh, 2002]. To determine μ , each HCM was repeatedly integrated for 167 years, with μ ranging from 1.10 to 1.30 at increments of 0.01. Among them, the model with overall the most realistic model Nino3 and Nino34 SSTA indices, and La Niña and El Niño SSTA patterns (as determined by the nonlinear principal component analysis method described in the following section) was selected. Hence μ was chosen to be 1.19, 1.26, 1.21 and 1.19 respectively in the NONA, NOLA, LONA and LOLA HCMs.

From construction, the statistical atmospheric models only capture the low frequency relationship between SSTA and the wind stress. But high frequency wind variability is important for the model to exhibit irregular behavior. Neelin *et al.* [1998] pointed out that there is a greater likelihood for the irregularity of ENSO to be due to external uncoupled atmospheric noise as opposed to internal nonlinear dynamics. In order to produce irregular behavior, “atmospheric noise”, i.e. high frequency wind variability with a monthly timescale was added as in Kirtman and Schopf [1998]. The difference between the unsmoothed FSU wind stress from 1964-2001 and its 9 month running mean was taken as the atmospheric noise. At the initial time of the coupled model run, a random date was selected

to begin sampling the noise dataset. Each successive simulation month would sample each successive month of noise, continuing for a period of time until a new random start date was chosen.

In this section, we present results of the 167-year coupling experiments. The ocean model was first forced by the FSU wind stress for 37 years, then coupled to the atmosphere and integrated forward for 167 years (2004 months). Figure 2 shows the 450-month long records of the area-averaged SSTA for the regions Niño3 (5°N - 5°S , 150°W - 90°W) and Niño34 (5°N - 5°S , 170°W - 120°W). The oscillations in the LONA and LOLA models (where the ocean used LR T_{sub} parametrization) appear to be more regular in structure with smaller inter-episode spacing than the oscillations in NONA and NOLA (where the ocean used NN T_{sub} parametrization). That warm episodes tend to have larger amplitude than cool episodes in the more nonlinear models (Figure 2) is consistent with *An and Jin* [2004], where nonlinear dynamical heating terms are shown to strengthen the warm episodes relative to the cool ones.

If we define a strong El Niño as a warm episode where the Niño3 anomaly index is greater than 1.5°C , then over 167 years there were 20, 7, 16, 13 strong El Niño episodes in the NONA, NOLA, LONA and LOLA model runs, respectively. The standard deviation of Niño3 (Niño34) SSTA for the NONA, NOLA, LONA, LOLA models and observation were 0.74 (0.70), 0.63 (0.60), 0.77(0.72), 0.82 (0.76), 0.88 (0.78), respectively. Although among the four HCMs, NONA had more strong warm episodes than others, the standard deviation of Niño3 (Niño34) SSTA for NONA was not the biggest because there was more inter-episode spacing, as the oscillations in NONA were more irregular than in the others.

Figure 3 shows the Fourier spectral analysis of the Niño3 SSTA from the 4 HCMs over 167 years and the observed Niño3 SSTA over the period 1950-2001. Models where the ocean used NN parametrization (NONA and NOLA) had considerably broader spectra than models where the ocean used LR (LONA and LOLA). Using NN instead of LR in the atmosphere also broadened the spectrum (com-

pare NONA with NOLA, and LONA with LOLA). Among the four HCMs, NONA had by far the best agreement with observations, as it exhibited two main spectral peaks at 44-month period ($0.0229 \text{ month}^{-1}$ frequency) and 65-month ($0.0154 \text{ month}^{-1}$), while observations revealed two main peaks at 45-month period ($0.0225 \text{ month}^{-1}$) and 69-month ($0.0144 \text{ month}^{-1}$). The other HCMs showed only a single dominant peak: at 58-month period ($0.0172 \text{ month}^{-1}$) for NOLA, 45-month ($0.0223 \text{ month}^{-1}$) for LONA, and 42-month ($0.0240 \text{ month}^{-1}$) for LOLA.

Figure 4 shows the time-longitude cross sections of the simulated SSTA from the four HCMs along the equator during year 67 to year 96. The oscillations were much more regular in the HCMs with the ocean model using LR parametrization (LONA and LOLA) than the ones with NN parametrization (NONA and NOLA), which displayed quiet inter-episode periods. Figure 5 shows the time-longitude cross sections of the simulated thermocline depth anomaly along the equator from the four HCMs, showing eastward propagation of the thermocline anomalies as the upper ocean heat content in the western equatorial Pacific is discharged.

4 PCA and nonlinear PCA of SSTA

We next compare the 2 leading PCA modes of the model SSTA during the last 100 years from the 4 HCMs. All the spatial patterns of PCA mode 1 from the 4 HCMs (Figures 6a, c, e and g) agreed well with the observed pattern (Figure 6i). The percentage of the SSTA variance accounted for by the first mode is 79.7%, 79.1%, 86.6% and 88.0% for the NONA, NOLA, LONA and LOLA models respectively, versus 56.5% for observations.

For PCA mode 2, the patterns from LONA and LOLA models (Figures 6f and h) did not do as well as those from the NONA and NOLA models (Figure 6b and d) in agreeing with the observed pattern (Figure 6j), which shows a west-east dipole structure. The percentage variance accounted for by mode 2 is 7.5%, 6.1%, 5.4% and 5.6% for the NONA, NOLA, LONA and LOLA models respectively, versus

12.7% for observations. Thus relative to observations, the four HCMs all have too much variance concentrated in the first PCA mode, and not enough in their second mode, the situation being worse for the two HCMs using LR parametrization in their ocean component.

Neural networks have been introduced to nonlinearly generalize the PCA method [Kramer, 1991], in that instead of a straight line, a curve is found to pass through the middle of the data cloud. The nonlinear principal component analysis (NLPCA) code and procedure from Hsieh [2001, 2004] were applied to the SSTA from the four HCMs during the last 100 years. The first 6 PCs of the SSTA were supplied as inputs to the NLPCA network (Figure 1b). The NLPCA model fits a curve to the data in the 6-dimensional PC-space. At one end of this curve, where the nonlinear principal component (NLPC) u assumes its minimum value, one finds the strongest La Niña episodes, while at the other end of the curve, where u assumes its maximum value, one finds the strongest El Niño episodes.

Figure 7 shows the NLPCA mode 1 spatial patterns of SSTA at $\min(u)$ and at $\max(u)$ for the four models and for the observed data. The linear PCA solution has the same standing spatial anomaly pattern (but for a sign change) as the PC flips from its minimum to its maximum value. In contrast, the NLPCA mode 1 of the observed data shows the cool SSTA during strong La Niña (Figure 7i) to be located much further west of the warm anomalies found during strong El Niño (Figure 7j). This asymmetry between La Niña and El Niño SSTA is evident in the NONA, NOLA and LONA HCMs (Figures 7a-b, c-d and e-f), but is least evident in the LOLA model (Figures 7g and h). The enhanced asymmetry between La Niña and El Niño seen in our more nonlinear models demonstrates the value of using the nonlinear NN approach over the earlier use of linear statistical methods in HCMs [Kang and Kug, 2000]. The spatial correlations and root mean square differences (RMSD) between these model SSTA patterns and the observed patterns during strong El Niño and strong La Niña (Table 2) showed that NONA and NOLA did slightly better than LONA and LOLA in simulating the strong El Niño SSTA pattern, but slightly worse in simulating the strong La Niña SSTA pattern.

To gauge the nonlinearity in a dataset, we computed the percentage variance accounted for by the NLPCA mode 1 (p_{NL}) and that by the PCA mode 1 (p_L), and calculated the normalized difference δ ($= (p_{NL} - p_L)/p_L$). If the dataset is completely linear, then the NLPCA mode 1 will retrieve the same straight line approximation of the dataset as the PCA mode 1, and δ will be 0. The larger δ is, the more nonlinear is the dataset and the greater is the asymmetry between the El Niño and La Niña patterns. We found that δ had the values 6.2 % (NONA), 5.2% (NOLA), 0.3% (LONA), 0.03% (LOLA) and 11.7% (observed). In the original Lamont coupled model [Zebiak and Cane, 1987], we found δ to be 2.1%. This implies that none of the models could match the observed nonlinear structure of the SSTA in ENSO. NONA and NOLA managed to improve on the weak nonlinearity in the Lamont coupled model, while LONA and LOLA were even more linear than the original Lamont model. Thus the nonlinearity in the NN parametrization of T_{sub} has helped in giving the HCMs a more asymmetric, nonlinear structure in the ENSO SSTA.

5 NLPCA of wind stress anomalies

The NLPCA was also applied to the model wind stress anomalies (WSA) during the last 100 years and the observed anomalies. The first 7 PCs of the WSA were supplied as inputs to the NLPCA network (Figure 1b). Figure 8 shows the WSA NLPCA mode 1 spatial patterns at $\min(u)$ and at $\max(u)$ for the four HCMs and the observed data. For the strong La Niña WSA pattern (Figures 8a, c, e, g, i), the easterly anomalies in the western equatorial Pacific in the four HCMs were all stronger than the observed anomalies, especially for NOLA and LOLA (Figures 8c and g). Also, strong WSA blowing away from the equator in the region 10°S - 25°S , 165°E - 140°W and the region 5°N - 15°N , 180° - 100°W can be found in NOLA and LOLA (Figures 8c and g), but not in NONA, LONA and the observations (Figures 8a, e and i).

For the strong El Niño WSA pattern (Figures 8b, d, f, h and j), relatively strong wind anomalies

blowing towards the equator in the region 10°S-20°S, 150°E-180° can be found in NOLA and LOLA (Figures 8d and h) but not in NONA, LONA and the observations. The rather strong easterly wind anomalies observed (Figure 8j) in the region 0°N-10°N, 120°E-160°E are only simulated well in the LONA model (Figure 8f) and to a lesser extent in the NONA model (Figure 8b). In sum, the two HCMs with NN atmosphere simulated the observed WSA during strong El Niño and strong La Niña noticeably better than the two HCMs with LR atmosphere. This is confirmed by the spatial correlation and RMSD between the simulated and observed WSA patterns (Table 2), where NONA and LONA did much better than NOLA and LOLA.

To gauge the nonlinearity in the WSA dataset, we compared the NLPCA mode 1 solution to the PCA mode 1 solution, and found that δ had the values 21.9 % (NONA), 0.4% (NOLA), 21.9% (LONA), 0.1% (LOLA) and 34.7% (observed). For the original Lamont coupled model, δ was 5.3%. Again, none of the models quite matched the nonlinearity in the observed ENSO WSA. NONA and LONA managed to dramatically improved on the weak nonlinearity in the Lamont coupled model, while NOLA and LOLA were even more linear than the original Lamont model.

6 Summary and Conclusion

In this paper, we investigated the potential for improving hybrid coupled modeling of the tropical Pacific by the use of nonlinear NN methods. NN was introduced for the nonlinear parametrization of the subsurface temperature in the Lamont ocean model, and for the nonlinear estimation of the WSA from the SSTA. To compare with the nonlinear regression by NN, corresponding linear regression (LR) models were also developed. By combining the NN or the LR version of the ocean model and the atmospheric model, four HCMs resulted — NONA (with NN for both the ocean and atmosphere), NOLA (NN for ocean and LR for atmosphere), LONA (LR ocean and NN atmosphere) and LOLA (LR for both ocean and atmosphere).

For the coupled model Niño3 SSTA spectrum, using NN in the ocean model produced a much broader spectrum than using LR, which gave basically a single narrow spectral peak. Thus the oscillations in LONA and LOLA were all far too regular when compared to the observed oscillations. Using NN in the atmospheric model in addition to the ocean model furthered broadened the SSTA spectrum, yielding a spectrum with two main peaks at periods of 44 and 65 months, in good agreement with the observed spectrum, where there were two peaks at 45 and 69 months (Figure 3).

PCA on the SSTA also showed that using the LR ocean model led to excessive concentration of energy in the first mode and poorer agreement of the mode 2 spatial patterns between the models and observations. Nonlinear PCA (NLPCA) on the SSTA was less conclusive, as using the NN ocean model instead of LR yielded a slightly better spatial pattern during strong El Niño but also a slightly worse pattern during La Niña. From δ (the normalized difference between the percentage variance explained by the NLPCA mode 1 and that by the PCA mode 1), we found that for SSTA all the coupled models were too linear compared to the observations. However, using the NN ocean model was able to alleviate the weak nonlinearity in the original Lamont coupled model.

NLPCA applied to the WSA revealed that using the NN atmospheric model instead of LR led to much better agreement in the spatial anomaly patterns between the coupled model and observations during both strong El Niño and La Niña. For the WSA, δ again revealed the original Lamont coupled model to be too linear, and that using the NN atmospheric model considerably alleviated this problem. Thus, the use of NN in the ocean model and in the atmospheric model improved on the nonlinear behaviour of the coupled model.

Acknowledgements

We are grateful to the Lamont staff, especially Steve Zebiak and Larry Rosen, for providing us the Lamont model. This work was supported by grants to W. W. Hsieh from the Canadian Foundation

for Climate and Atmospheric Sciences, and the Natural Sciences and Engineering Research Council of Canada.

References

- An, S.-I., and F.-F. Jin (2004), Nonlinearity and asymmetry of ENSO, *J. Climate*, *17*, 2399-2412.
- Barnett, T. P., M. Latif, N. E. Graham, M. Flügel, S. Pazan, and W. White (1993), ENSO and ENSO related predictability, part 1: prediction of equatorial sea surface temperature with a hybrid coupled ocean-atmosphere model, *J. Climate*, *6*, 1545-1566, 1993.
- Barnston, A. G., et al. (1994), Long-lead seasonal forecasts—where do we stand? *Bull. Amer. Meteor. Soc.*, *75*, 2097-2114.
- Blank, B., J. D. Neelin and D. Gutzler (1997), Estimating the effect of stochastic wind stress forcing on ENSO irregularity, *J. Climate*, *10*, 1473-1486.
- Bishop, C. M. (1995), Neural Networks for Pattern Recognition, Clarendon Pr., Oxford, pp 482, 1995.
- Boulanger, J-P., and C. Menkes (2001), The Trident Pacific model, part 2: role of long equatorial wave reflection on sea surface temperature anomalies during the 1993-1998 TOPEX/POSEIDON period, *Clim. Dynam.*, *17*, 175-186.
- Colin, C., C. Henin, P. Hisard, and C. Oudot (1971), Le Courant de Cromwell dans le Pacifique central en fevrier, *Cah, ORSTOM Ser Oceanogr.*, *9*, 167-186.
- Hsieh, W. W. (2001), Nonlinear principal component analysis by neural networks, *Tellus*, *53A*, 599-615.
- Hsieh, W. W. (2004), Nonlinear multivariate and time series analysis by neural network methods, *Rev. Geophys.*, *42*, RG1003, doi:10.1029/2002RG000112.
- Hsieh, W. W., and B. Tang (1998), Applying neural network models to prediction and analysis in meteorology and oceanography, *Bull. Am. Meteorol. Soc.*, *79*, 1855-1870.

- Kang, I.-S., and J.-S. Kug (2000), An El-Nino prediction system using an intermediate ocean and a statistical atmosphere, *Geophys. Res. Lett.*, *27*, 1167-1170.
- Kirtman, B. P., and P. S. Schopf (1998), Decadal variability in ENSO predictability and prediction, *J. Climate*, *11*, 2804-2822.
- Kramer, M. A. (1991), Nonlinear principal component analysis using autoassociative neural networks, *AIChE Journal*, *37*, 233-243.
- Latif, M., and A. Villwock (1990), Interannual variability as simulated in coupled ocean-atmosphere model, *J. Mar. Syst.*, *1*, 51-60.
- Neelin, J. D., D. S. Battisti, A. C. Hirst, F.-F. Jin, Y. Wakata, T. Yamagata, and S. E. Zebiak (1998), ENSO theory, *J. Geophys. Res.*, *103*, 14261-14290.
- Palmer, T. N., and D. L. T. Anderson (1994), The prospects for seasonal forecasting— a review paper. *Q. J. R. Meteorol. Sci.*, *7*, 755-793.
- 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.
- Syu, H.-H., J. D. Neelin, and D. Gutzler (1995), Seasonal and interannual variability in a hybrid coupled GCM. *J. Climate*, *9*, 2121-2143.
- Tang, Y., W. W. Hsieh, B. Tang, and K. Haines (2001), A neural network atmospheric model for hybrid coupled modelling, *Clim. Dynam.*, *17*, 445-455.
- Tang, Y., and W. W. Hsieh (2002), Hybrid coupled models of the tropical Pacific — ENSO prediction. *Clim. Dynam.* *19*, 343-353.
- Tang, Y., and W. W. Hsieh (2003), ENSO simulation and prediction in a hybrid coupled model with data assimilation, *J. Meteor. Soc. Japan*, *81*, 1-19.

Zebiak, S. E., and M. A. Cane (1987), A model El Niño-Southern Oscillation, *Mon. Wea. Rev.*, *115*, 2262-2278.

Zhang, R., J. B. Antonio, G. M. Raghuram, C. H. Eric, and B. Joaquim (2004), A new approach to improved SST anomaly simulations using altimeter data: Parameterizing entrainment temperature from sea level, *Geophys. Res. Lett.*, *31*, doi:10.1029/2003GL019237.

Table 1: Cross-validated correlation and root mean square error (RMSE) (in parenthesis) between the predicted wind stress anomaly PCs and the observed PCs. In the first column, NN or LR means the predictors are the SSTA from the ocean model with an NN or LR T_{sub} parameterization, while in the second column, NN or LR indicates the method used to estimate the wind stress anomaly PCs from the SSTA.

SSTA	Atmos.	Correlation (RMSE)						
		PC1	PC2	PC3	PC4	PC5	PC6	PC7
NN	NN	0.836	0.224	0.531	0.085	0.125	0.322	0.207
		(2.04)	(3.26)	(2.71)	(2.60)	(2.12)	(1.91)	(1.95)
NN	LR	0.821	0.144	0.349	0.028	-0.007	0.073	0.274
		(2.12)	(3.29)	(2.99)	(2.58)	(2.15)	(2.01)	(1.80)
LR	NN	0.835	0.276	0.481	0.010	0.099	0.319	0.286
		(2.05)	(3.21)	(2.82)	(2.64)	(2.13)	(1.91)	(1.83)
LR	LR	0.804	0.194	0.414	-0.025	0.101	0.096	0.075
		(2.21)	(3.25)	(2.92)	(2.63)	(2.13)	(2.00)	(1.91)

Table 2: Spatial correlation and RMS difference (RMSD) between the observed and model ENSO patterns, where the spatial patterns of SSTA and wind stress anomaly (WSA) for strong El Niño and strong La Niña were extracted by the NLPCA method.

	Correlation				RMSD			
	NONA	NOLA	LONA	LOLA	NONA	NOLA	LONA	LOLA
La Niña SSTA	0.84	0.86	0.89	0.89	0.17	0.15	0.13	0.13
El Niño SSTA	0.91	0.91	0.88	0.85	0.22	0.19	0.23	0.23
La Niña WSA	0.87	0.68	0.89	0.62	0.06	0.10	0.06	0.14
El Niño WSA	0.92	0.65	0.92	0.73	0.09	0.19	0.08	0.15

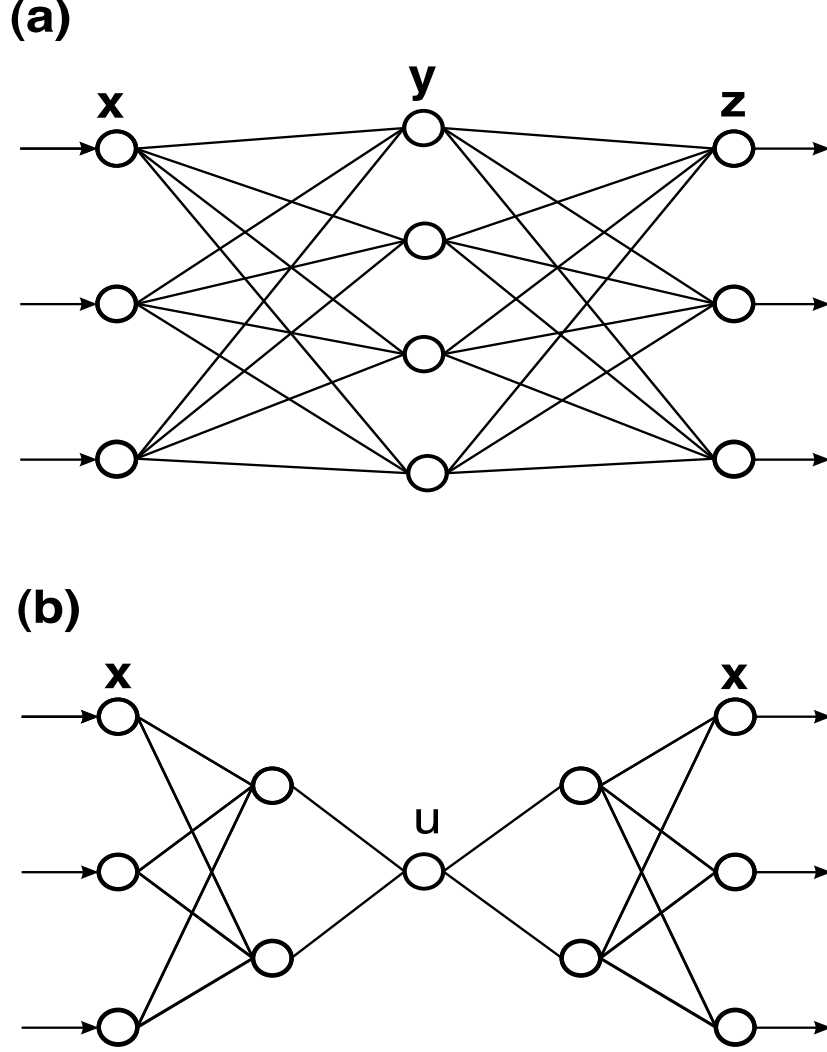


Figure 1: (a) An schematic diagram of a neural network (NN) model for nonlinear regression, with one ‘hidden’ layer of neurons (i.e. variables) (denoted by circles) sandwiched between the input layer \mathbf{x} and the output layer \mathbf{z} . Increasing the number of hidden neurons \mathbf{y} increases the number of model parameters. In this paper, the output layer has only a single neuron. (b) A schematic diagram of the NN model for nonlinear principal component analysis (NLPCA). There are 3 layers of hidden neurons sandwiched between the input layer \mathbf{x} on the left and the output layer \mathbf{x}' on the right. Next to the input layer is the encoding layer, followed by the ‘bottleneck’ layer (with a single neuron u), which is then followed by the decoding layer. Effectively, a nonlinear function $u = F(\mathbf{x})$ maps from the higher dimension input space to the lower dimension bottleneck space, followed by an inverse transform $\mathbf{x}' = \mathbf{G}(u)$ mapping from the bottleneck space back to the original space, as represented by the outputs. To make the outputs as close to the inputs as possible, the cost function $J = \langle \|\mathbf{x} - \mathbf{x}'\|^2 \rangle$ (i.e. the mean square error, MSE) is minimized. Data compression is achieved by the bottleneck, yielding the nonlinear principal component (NLPC) u . See Hsieh (2004) for details. In Sects. 4 and 5, 3 neurons were used in each of the encoding and decoding layers, and there were 6 (7) input and 6 (7) output neurons in Sect.4 (Sect.5).

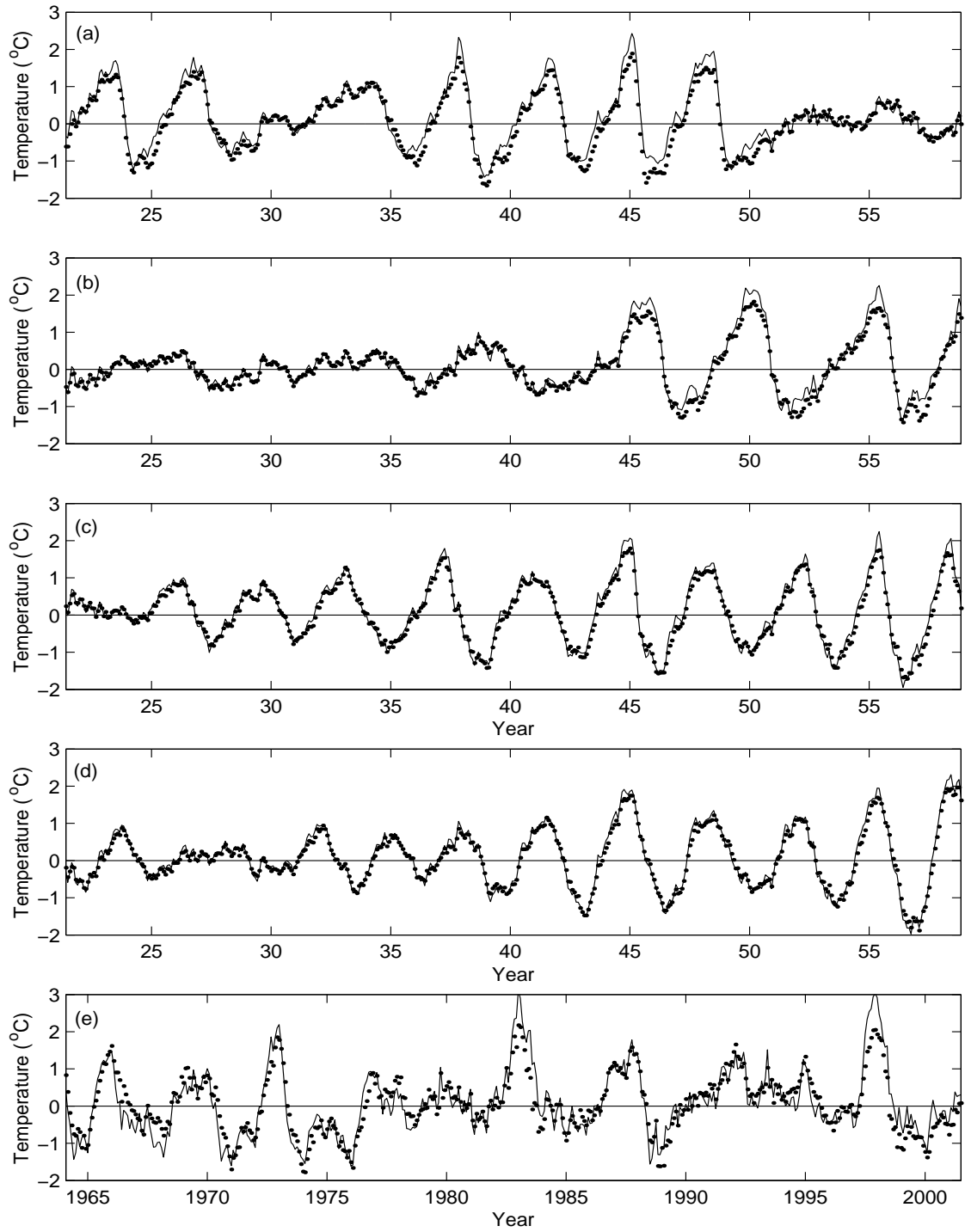


Figure 2: The Niño3 (solid line) and Niño34 (dotted line) SSTA from the (a) NONA, (b) NOLA, (c) LONA and (d) LOLA HCMs, and (e) observations.

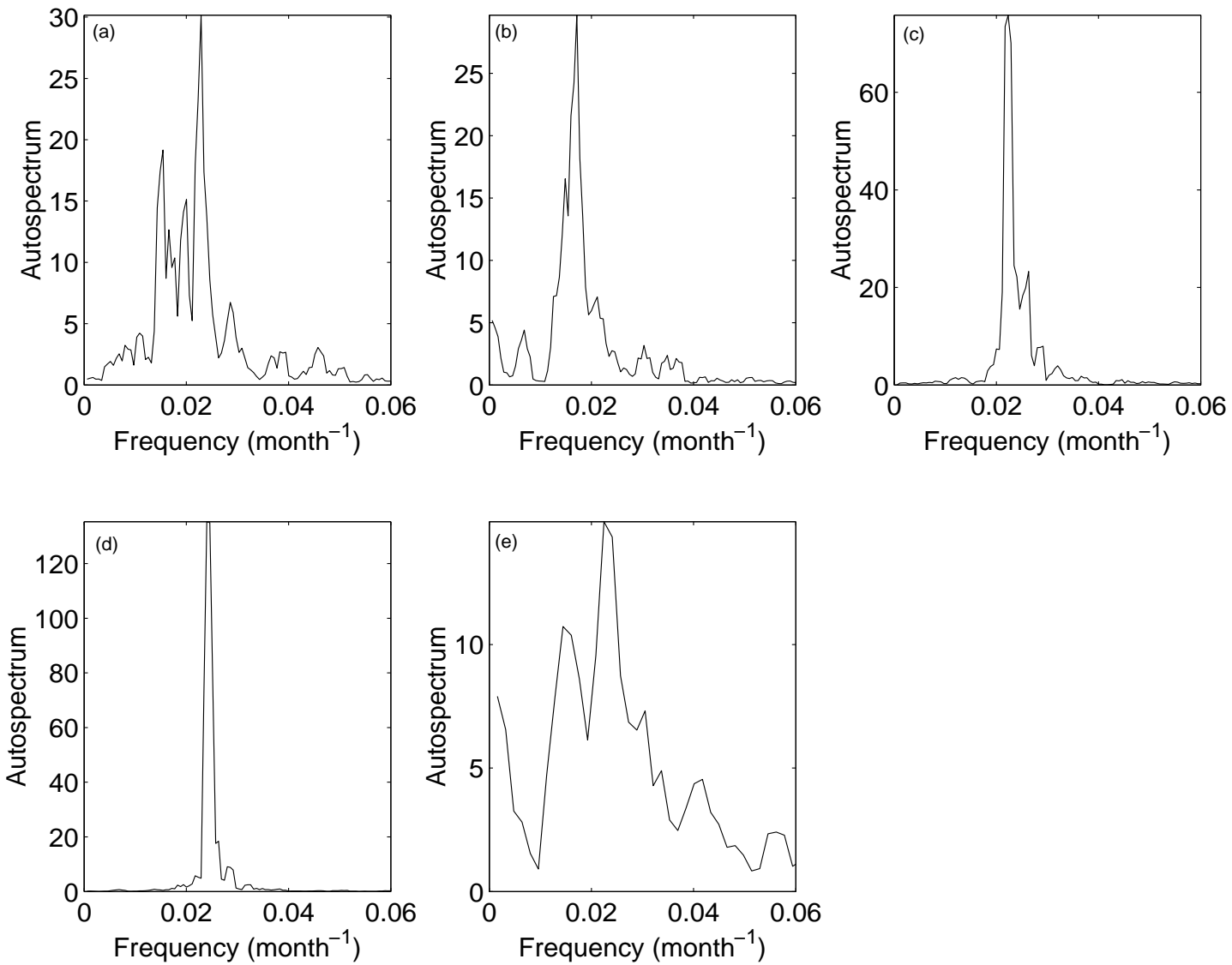


Figure 3: Autospectrum of the Niño3 SSTA time series from the (a) NONA , (b) NOLA, (c) LONA and (d) LOLA HCMs, and (e) observations.

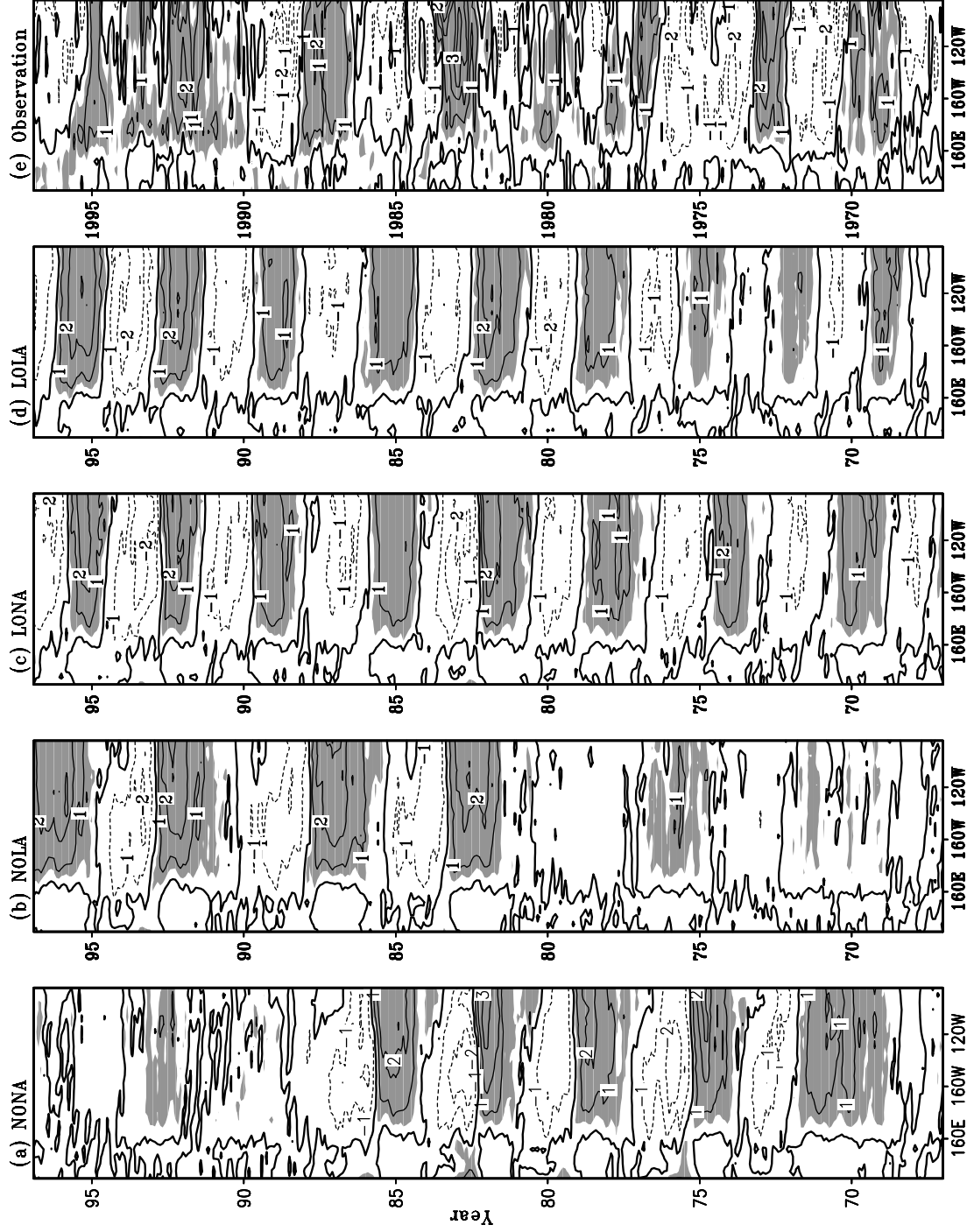


Figure 4: Time-longitude cross section along the equator of SSTA from the (a) NONA , (b) NOLA, (c) LONA and (d) LOLA HCMs, and (e) observations (from 1967-1996). With a contour interval of 1°C, the positive contours are shown as solid curves, the negative contours as dashed curves, the zero contour as a thick solid curve, and positive anomalies above 0.5°C are shaded.

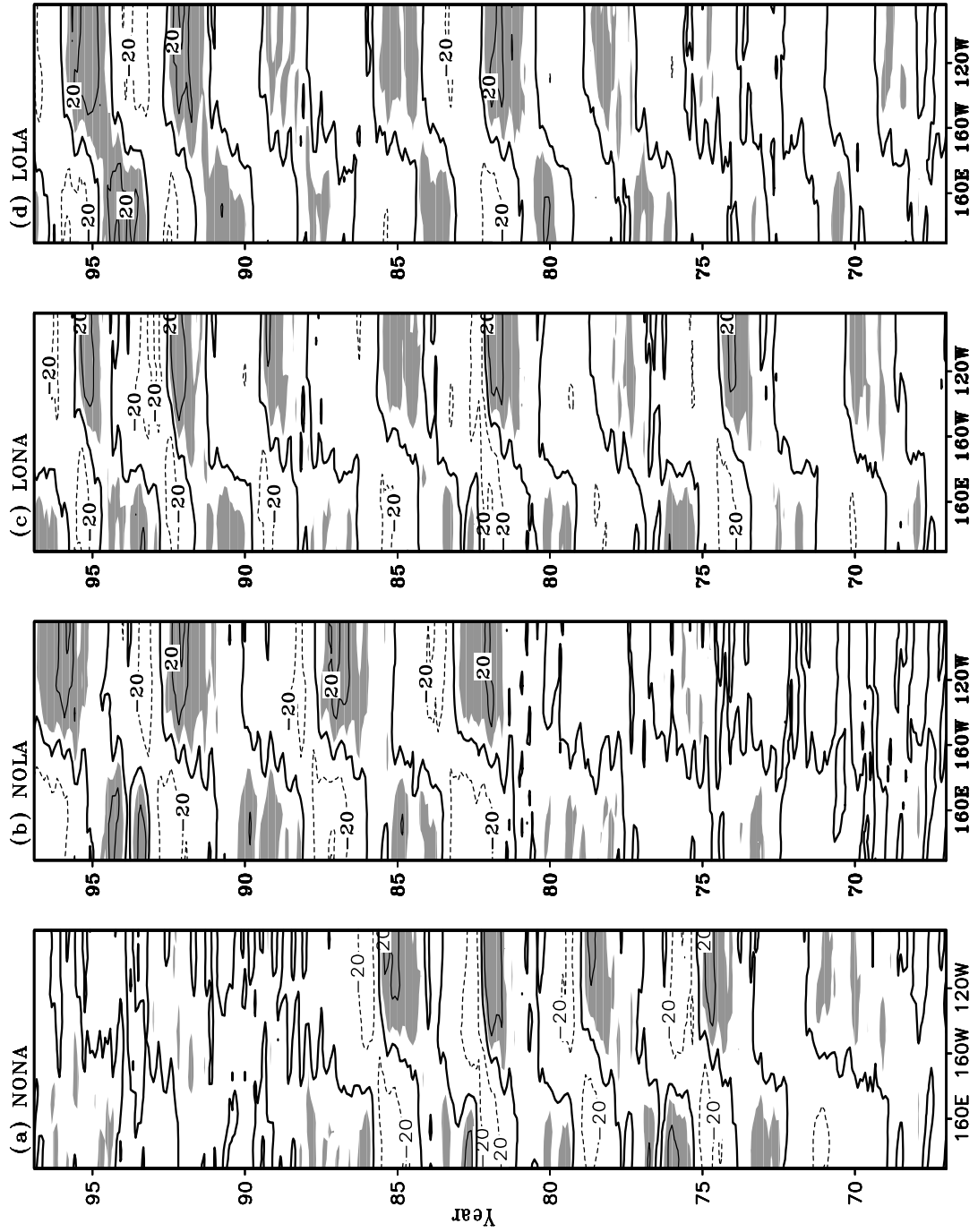


Figure 5: Time-longitude cross section along the equator of the thermocline depth anomaly from the (a) NONA , (b) NOLA, (c) LONA and (d) LOLA HCMs. The plotted anomalies have been smoothed by a 3-month running mean. The contour interval is 20 m, and positive anomalies above 10 m are shaded.

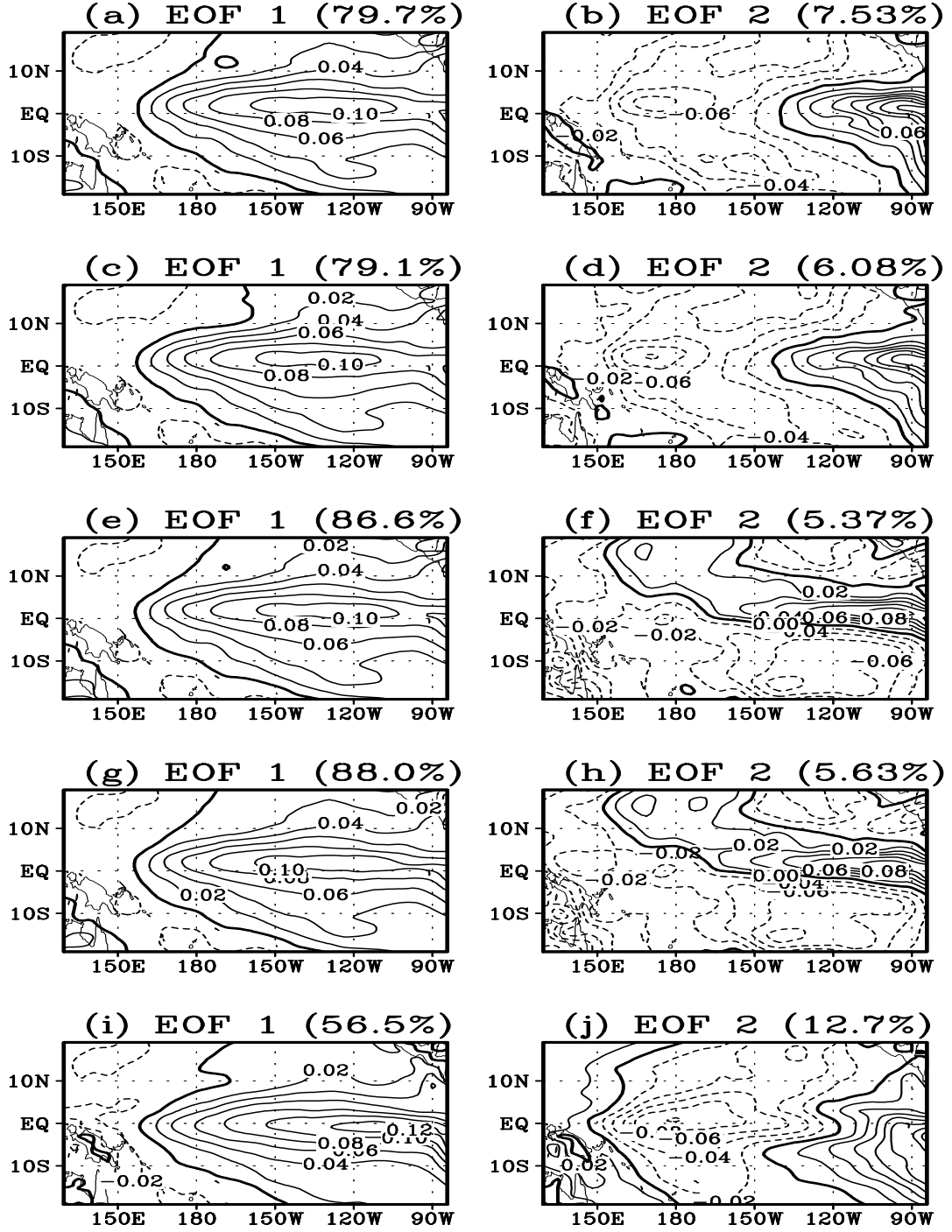


Figure 6: The first two PCA modes of the SSTA. From top to bottom, the five rows present the results from the NONA (a-b), NOLA (c-d), LONA (e-f), LOLA (g-h) models, and observations (i-j). The contour interval is 0.02.

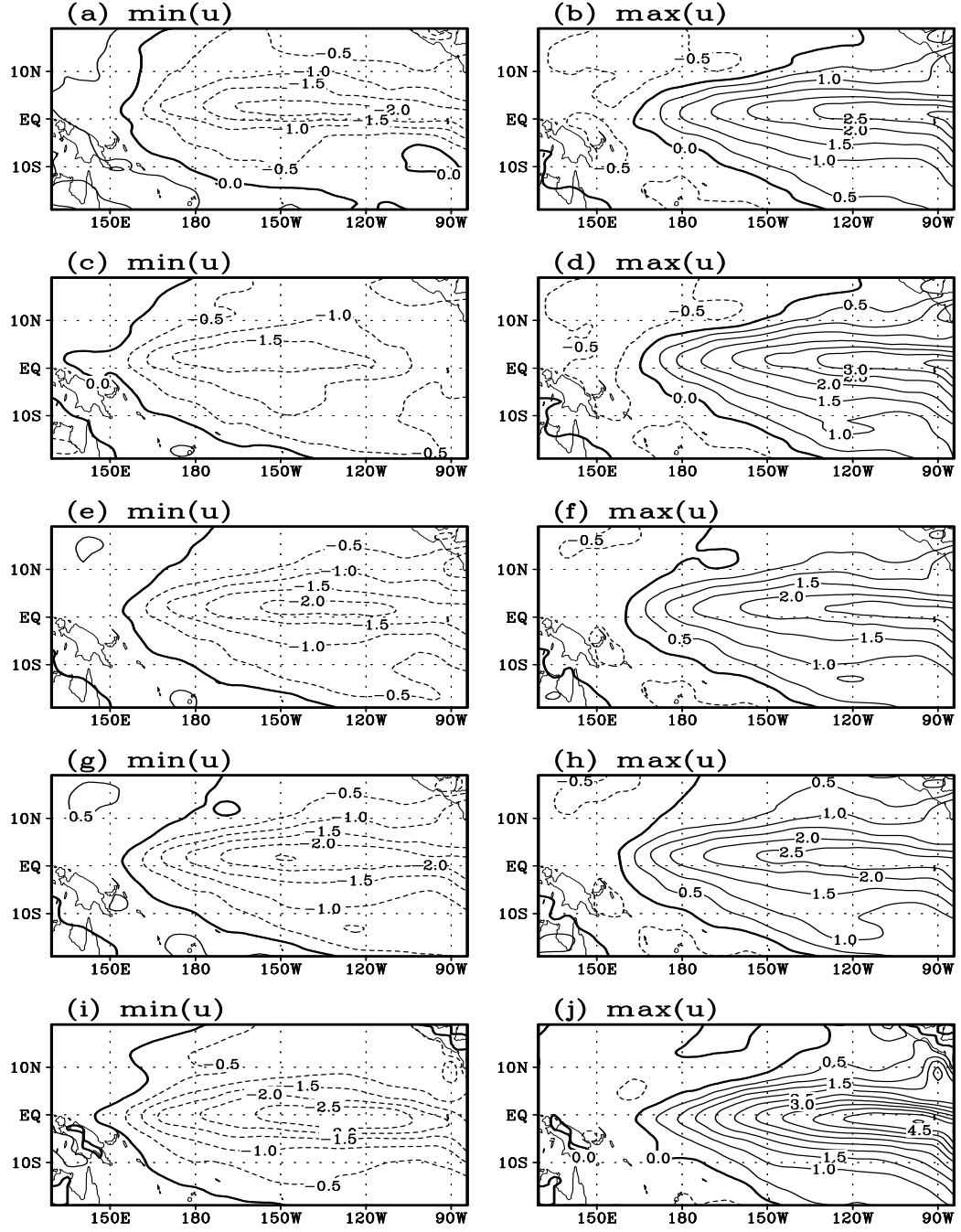


Figure 7: The SSTA patterns (in $^{\circ}\text{C}$) of the nonlinear principal component analysis (NLPCA) mode 1 when the NLPC u is minimum (corresponding to strong La Niña) (left column) or maximum (strong El Niño) (right column). The five rows from top to bottom present the SSTA from the NONA (a-b), NOLA (c-d), LONA (e-f), LOLA (g-h) models, and observations (i-j). The contour interval is 0.5°C .

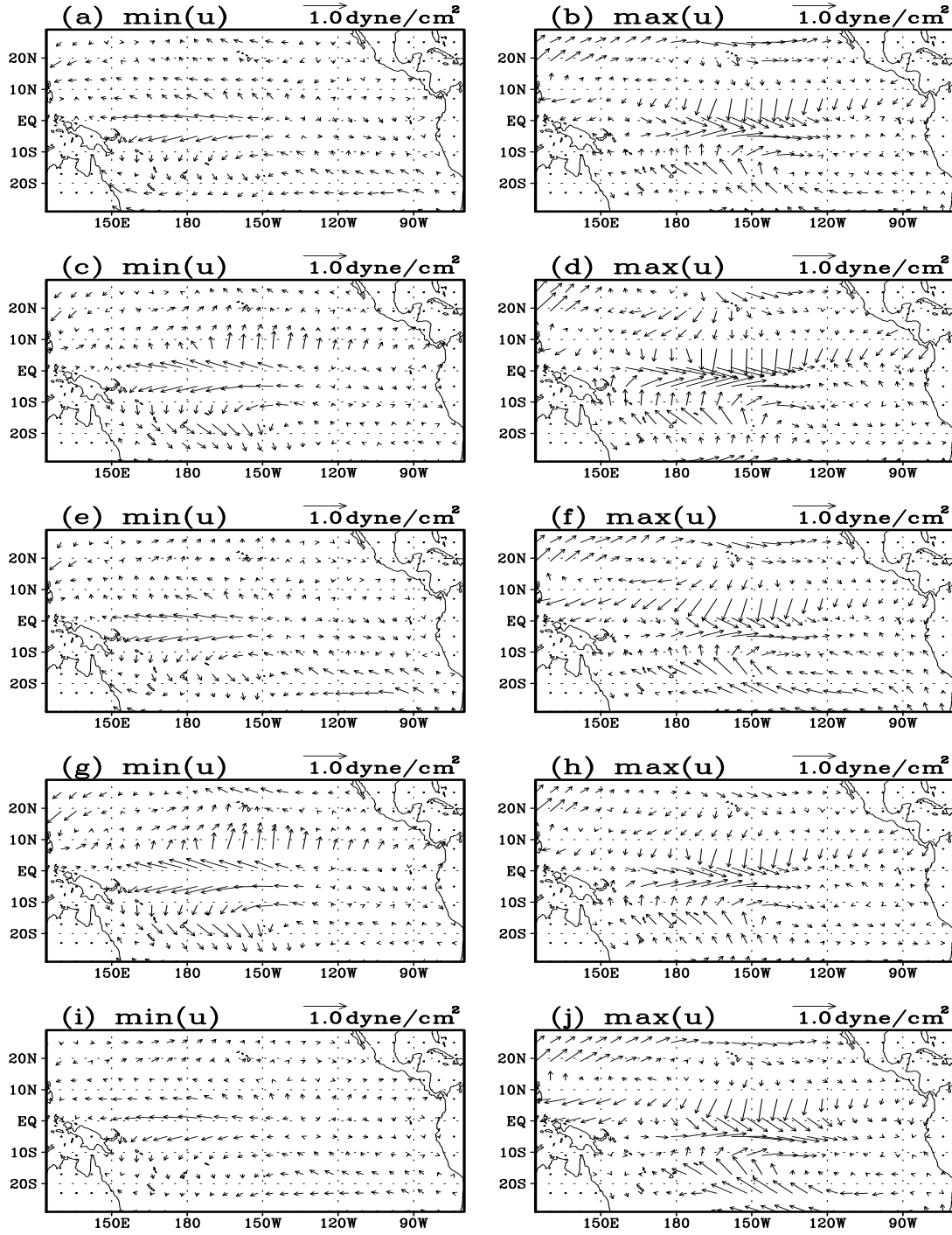


Figure 8: The WSA patterns of the nonlinear principal component analysis (NLPCA) mode 1 when the NLPC u is minimum (corresponding to strong La Niña) (left column) or maximum (strong El Niño) (right column). The five rows from top to bottom present the WSA from the NONA (a-b), NOLA (c-d), LONA (e-f), LOLA (g-h) models, and observations (i-j).