Multi-proxy reconstruction of the equatorial Pacific SST and zonal wind fields of the past 10,000 years using Mg/Ca and alkenone records

Emily C. Gill,[1,2] Balaji Rajagopalan,[1,2], Peter H. Molnar[2,3], and Thomas M. Marchitto[3,4]

Corresponding author: Emily C. Gill, Department of Civil Engineering, University of Colorado, Boulder, Colorado, USA. (emily.gill@colorado.edu)

[1] Department of Civil, Environmental and Architectural Engineering, University of Colorado at Boulder, Boulder, CO, USA.

[2] Cooperative Institute for Research in Environmental Science (CIRES), Boulder, CO, USA.

[3] Department of Geological Sciences, University of Colorado at Boulder, Boulder, CO, USA.

A multi-proxy, reduced-dimension methodology blends magnesium-calcium (Mg/Ca) and alkenone (U^k_{37}) paleo sea surface temperature (SST) records from the eastern and western equatorial Pacific, to recreate snapshots of full field SSTs and zonal winds from 10 to 2 ka BP in two-thousand year increments. In the multi-proxy reconstruction, the coldest SST anomalies of $\sim$ -1°C occur at 10 ka in the eastern equatorial Pacific with concurrent easterly maximum wind anomalies of $\sim$ 8 m s^{-1} throughout the central Pacific. The largest zonal temperature anomaly differences (average east Pacific SST minus average west Pacific SST) are -0.36°C and -0.33°C, which occur at 6 and 10 ka, respectively. The eastern equatorial Pacific warms and central Pacific easterly winds weaken gradually from 10 to 2 ka, but western Pacific SSTs fluctuate throughout the period. Although both Mg/Ca and U^k_{37} act as SST proxies, comparisons of the single proxy SST field reconstructions (Mg/Ca-only versus U^k_{37}-only) reveal differences in the timing and duration of maximum cooling across the equatorial Pacific. Due to this, the multi-proxy approach is beneficial in revealing the common patterns of variability between the two different SST proxies.

---

[4] Institute of Arctic and Alpine Research
(INSTARR), University of Colorado at
Boulder, Boulder, CO, USA.

1. Introduction

Paleoclimatic proxy records spanning the early- to mid-Holocene (∼ 10-6 thousand years ago, ka) suggest that the equatorial Pacific was in an "enhanced La Nia-like state" during this time, consistent with reduced El Nio - Southern Oscillation (ENSO) variability due to a cooler eastern [e.g. *Koutavas et al.*, 2006] and perhaps a slightly warmer western [e.g. *Stott et al.*, 2004] Pacific. Not all paleo records show reduced ENSO variability *Cobb et al.* [2013], and those that do tend to disagree on timing and duration. Clastic sediment records from lakes in the Andes [*Moy et al.*, 2002; *Rodbell et al.*, 1999] and marine records [*Koutavas et al.*, 2006; *Tudhope et al.*, 2001] show reduced ENSO variability around 6 ka, which agrees with GCM simulations [*Liu et al.*, 2000; *Otto-Bliesner et al.*, 2003; *Timmermann et al.*, 2007; *Zheng et al.*, 2008] and has been theorized to be tied to orbital forcing [*Clement et al.*, 1999; *Zheng et al.*, 2008; *Luan et al.*, 2012]. From pollen, *Shulmeister and Lees* [1995] inferred a drying of Australia at ∼ 4 ka, which they attributed to an onset of modern-like ENSO conditions. Other more recent marine records [*Carré et al.*, 2014; *McGregor et al.*, 2013] have challenged the timing of these results and suggest that ENSO variability was at its lowest around 4-5 ka.

Paleoclimatic proxy data are intrinsically point measurements, and therefore associating one paleoclimate time series with a process that reaches over a huge area takes a risk. Moreover, because of dissolution of carbonate tests and low sedimentation rates at great depth, useful cores for SSTs proxy records (Tables 1 and 2, Figures 1 and 2) are commonly limited to regions of relatively shallow depth, near coasts, from aseismic ridges, or over young oceanic crust. As a result, the absence of long-term records from the central Pacific makes it difficult to draw conclusions about large-scale spatial patterns of ENSO-relevant SST fields over paleoclimatic timescales. The records that exist along the ocean margins are often irregularly sampled (spatially and temporally), and different proxies from the same core can sometimes yield contradictory inferences. With these limitations in mind, we use all available SST information at various points from the east and west Pacific in order to reconstruct the full field of tropical Pacific SSTs and zonal winds (10°S to 10°N and 100°E to 60°W) over the past 10 ka.

Previous studies have used paleo-proxy data with various statistical approaches to reconstruct full fields [e.g. *Cook et al.*, 1999; *Kaplan et al.*, 1998; *Luterbacher et al.*, 2004; *Mann et al.*, 1998, 2008; *Tingley and Huybers*, 2010] or time-series [e.g. *Kaufman et al.*, 2009; *Lee et al.*, 2008; *Li et al.*, 2010; *Moberg et al.*, 2005; *Mann et al.*, 2008] of climate variables. We use reduced-dimension reconstruction approaches, in which contemporary patterns of variability are related to a sparse sampling of points to resolve full fields or time-series of climate variables. Reduced-dimension approaches have been used to reconstruct surface temperature [e.g. *Luterbacher et al.*, 2004; *Mann et al.*, 1998; *Rutherford et al.*, 2005], pressure [e.g. *Luterbacher et al.*, 2002], and drought [e.g. *Cook et al.*, 1999; *Zhang et al.*, 2004]. *Evans et al.* [2002] provide an example of a reduced-dimension SST field reconstruction using $\delta^{18}O$ proxy records from coral.

2. Data

Contemporary (1854-2013) gridded (2° by 2°) monthly SSTs were obtained from the NOAA NCDC Extended Reconstruction Sea Surface Temperature (ERSST) version 3b dataset [*Smith et al.*, 2008]. Contemporary (1949-2013) gridded (2.5° by 2.5°) monthly zonal winds were obtained from NOAA NCEP-NCAR CDAS-1 Global Reanalysis [*Kalnay et al.*, 1996]. Monthly anomalies were calculated for each dataset using 1981-2010 climatology. Monthly anomalies for SSTs and winds were then converted into annual averages by averaging from May to the following April in order to capture the annual ENSO cycle.

Two common proxies for paleothermometry in the Pacific exploit magnesium-calcium (Mg/Ca) ratios from foraminifera shells and alkenones (U^k_{37}) from coccolithophores to infer SSTs. Numerous Mg/Ca and U^k_{37} proxy records have been reported for the east and west Pacific (see *Leduc et al.* [2010] for a review). At high temperatures, more Mg is incorporated in the shells of foraminifera, and a ratio of Mg/Ca can be used to infer SSTs back through time [e.g. *Lea et al.*, 1999; *Nürnberg et al.*, 1996]. Alkenones are a trans-fat altered by cocclithophores, or algae, under various temperatures: lower temperatures lead to an increasein the degree of unsaturation. The ratio of di- to tri-unsaturated alkenones is used to create an Unsaturation Index named U^k_{37} [e.g. *Brassell et al.*, 1986; *Herbert*, 2003]. The remains of foraminifera and coccolithophores sink to the bottom of the ocean and are preserved in the sediment record. Mg/Ca (Table 1 and Figure 1) and U^k_{37} (Table 2 and Figure 2) SST records (Tables 1 and 2) were obtained from the archives of NCDC (http://www.ncdc.noaa.gov/data-access/paleoclimatology-data/datasets) and Pangaea (http://www.pangaea.de/).

Most proxy records listed in Tables 1 and 2 reach back to 10 ka, but some records stop prior to 0 ka. All the records were generated at various resolutions of at least one SST value per 1,000 years. In order to reconstruct 2, 4, 6, 8, and 10 ka snapshots of the equatorial pacific SST and wind fields, each paleoceanographic record was smoothed

(Figures 1) and 2) using a local polynomial method with a second order polynomial and a local neighborhood consisting of 70% of the nearest data points [e.g. *Loader et al.*, 1996]. These records were then converted to SST anomalies using the climatological mean temperature from ERSST at the grid-cell closest to the location of the proxy record.

3. Methodology

Our work is rooted in similar methodologies and assumptions to the work of [*Mann et al.*, 1998] who used a multi-proxy reduced-dimension approach to combine instrumental records with proxies from tree rings, ice cores, and corals to reconstruct global annual temperature patterns over the past six centuries. We use a principal component analysis (PCA) approach (Figure 3) to reconstruct SSTs and a canonical correlation analysis (CCA) approach to reconstruct winds.

3.1. Principal Component Analysis (PCA) Reconstruction of SSTs

In PCA, multivariate space-time data is decomposed into orthogonal space-time components via eigen-decomposition of the covariance matrix. These components, also referred to as patterns or modes, are ordered based on the percentage of total variance that each resolves. For most climatological datasets, the first few modes capture the majority of the variance of the original data thus, reducing the dimension of the data. Using bold-faced uppercase variables to denote matrices, with brackets used only when specifying the dimensions of the matrix, and lower-case bold-faced variables as vectors, the formulation for a PCA is as follows:

$$[\boldsymbol{X}]_{NxG} = [\boldsymbol{Y}]_{NxG}[\boldsymbol{U}]_{GxG} \tag{1}$$

$$[\boldsymbol{Y}]_{NxG} = [\boldsymbol{X}]_{NxG}[\boldsymbol{U}]^T_{GxG} \tag{2}$$

where $\boldsymbol{X}$ is the matrix of the original data, $\boldsymbol{Y}$ is a matrix of principal components (PCs), and $\boldsymbol{U}$ is a matrix of eigenvectors, N is the length of data at each location (i.e. the number of times sampled), and G is the number of grid-points. $\boldsymbol{U}$ is considered a transformation matrix, as it transforms the correlated data $\boldsymbol{X}$ into the orthogonal Principal Component (PC) space of $\boldsymbol{Y}$, or transforms the PCs back into the data-space. In theory, the eigenvectors are obtained sequentially by solving a constrained optimization problem. The first eigenvector, $\boldsymbol{u}_1$, is found such that $\epsilon_1 = E[\boldsymbol{X} - \boldsymbol{X}\boldsymbol{u}_1]^2$ is minimized and subjected to the orthonormality constraint of $\boldsymbol{u}_1\boldsymbol{u}_1^T = 1$. The subsequent eigenvectors are obtained to minimize the residual mean squared errors. Thus the variance of all the PCs sum to the variance of the original data. *Von Storch and Zwiers* [2001] give details of PCA and other multivariate analysis techniques. In practice, eigenvectors are obtained simultaneously by decomposing the covariance matrix into left-singular vectors or eigenvalues ($\boldsymbol{U}$), right singular vectors ($\boldsymbol{V}$) and a diagonal matrix of singular values (λ). The application steps of this method to our reconstruction is described below and shown as a flowchart (Figure 3).

3.1.1. Step (i)

Let $\boldsymbol{T}$ be an NxG matrix of contemporary average annual (May-Apr) SST anomalies composed of 160 years of data (1854-2014) at each of 973 grid-points spanning the equatorial Pacific from 10°S to 10°N and 100°E to 60°W (300°E). A PCA is performed on $\boldsymbol{T}$, such that:

$$[\boldsymbol{T}]_{NxG} = [\boldsymbol{Y}]_{NxG}[\boldsymbol{U}]_{GxG} \tag{3}$$

$$\boldsymbol{T} = \sum_{i=1}^{G} \lambda_i \boldsymbol{y}_i \boldsymbol{u}_i \tag{4}$$

whereby $\boldsymbol{T}$ is decomposed into G orthogonal eigenvectors, $\boldsymbol{U}$. For each i^{th} eigenvector or mode, we find an G-vector empirical orthogonal function (EOF) represented by $\boldsymbol{u}_i$ and an N-vector principal component (PC) represented by $\boldsymbol{y}_i$, which respectively describe the spatial and temporal variability of that particular eigenvector. In addition, λ_i is a scalar that provides the fraction of total variance of the original data that is resolved by the i^{th} mode. The first four EOFs (Figure 4a) and PCs (Figure 4b) explain almost 90% of the total variance of the full field SST data ($\lambda_1 = 0.688$, $\lambda_2 = 0.111$, $\lambda_3 = 0.071$, $\lambda_4 = 0.025$).

3.1.2. Step (ii)

Let $\boldsymbol{T}^{\dagger}$ be a matrix of contemporary SST anomalies at the locations for which we have proxy records (referred to as the "limited field" hereafter), which makes it order NxP, where P is 26. A PCA is performed on $\boldsymbol{T}^{\dagger}$, which similarly results in P orthogonal eigenvectors $\boldsymbol{U}^{\dagger}$, each with an P-vector $\boldsymbol{u}_i^{\dagger}$, N-vector $\boldsymbol{y}_i^{\dagger}$, and a scalar $\lambda_i^{\dagger}$.

$$[\boldsymbol{T}^{\dagger}]_{NxM} = [\boldsymbol{Y}^{\dagger}]_{NxM}[\boldsymbol{U}^{\dagger}]_{MxM} \tag{5}$$

$$\boldsymbol{T}^{\dagger} = \sum_{i=1}^{P} \lambda_i^{\dagger} \boldsymbol{y}_i^{\dagger} \boldsymbol{u}_i^{\dagger} \tag{6}$$

The first four modes (Figure 5) explain almost 95% of the total variance of the limited field SST data ($\lambda_1^{\dagger} = 0.782$, $\lambda_2^{\dagger} = 0.126$, $\lambda_3^{\dagger} = 0.022$, $\lambda_4^{\dagger} = 0.017$)

3.1.3. Step (iii)

We performed the following steps for each of six reconstruction periods: 2, 4, 6, 8, and 10 ka. It should be noted that although there are 29 independent proxy records, there are three locations at which the sediment core was analyzed for both Mg/Ca and U_{37}^{k}. For those three cases in the multi-proxy reconstruction, the Mg/Ca records were kept, and the U_{37}^{k} records were ignored to maintain equal representation from each proxy in each part of the Pacific. When removing the Mg/Ca records and keeping the U_{37}^{k}, however, differences in the results were negligible. Using the smoothed proxy records, all 26 values for, say, 10 ka are contained in matrix $\boldsymbol{R}$ of order 1 by P, where $P = 26$. $\boldsymbol{R}$ undergoes eigenvalue decomposition by multiplying it with the eigenvalue transformation matrix $\boldsymbol{U}^{\dagger}$ obtained from the limited field PCA in Step (ii):

$$[\boldsymbol{Y}^{R}]_{1xP} = [\boldsymbol{R}]_{1xP}[\boldsymbol{U}^{\dagger}]_{PxP}^{T} \tag{7}$$

where $\boldsymbol{Y}^{R}$ represents the PCs of the proxy SST values for 10 ka.

3.1.4. Step (iv)

For the reconstruction model, we use a few of the PCs of the limited field as predictors to model each of the first few full-field PCs. The decision to keep a certain number of PCs and predictors is somewhat arbitrary, but depends on the distribution of variance resolved across the modes. Although keeping more modes may increase the total resolved variance, inclusion of too many PCs can introduce noise into the calculated fields. Given the eigenvalue spectrum (EVS) of the full-field and limited-field λ-values (Figure 6), we were faced with a few reasonable options in deciding how many limited-field predictor PCs to use and how many full-field PCSs to reconstruct. First, we could limit the reconstruction to only the dominant mode, or first PC (represented by the green points on Fig. 6). In this case, we would model the first PC of the full field as a function of the first PC of the limited field. As a second option, we could choose some percentage of resolved variance as a threshold. If we choose PCs that resolve 10% or more of the variance, we would model the first two PCs of the full field (green dot and orange dot in Fig. 6a) each as functions of the first two PCs of the limited field (green dot and purple dot in Fig. 6b). If we chose 5% as a threshold, the number of predictors would remain the same, but we would now model the first three PCs of the full field (green, orange, and purple dots in Fig. 6a). A final option would be to use the "knee" in the eigenvalue spectra as thresholds. The "knee" (purple points in Figure 6) of each EVS is located at the PC just before each noise floor (grey points in Figure 6). Coincidentally, it turns out that using either 5% or the "knee" as a threshold achieves the same combination of PCs (first three) and predictor PCs (first two). Reconstructions of SSTs for each of these three scenarios were considered. Using only the first PC of each did not reconstruct modern day SSTs well enough because it omitted other important non-noise modes. Although the second (10% threshold) and third (5% threshold, or "knee") scenarios reconstructed both modern day and Holocene SSTs similarly, we chose to use the latter to include the most possible PCs. Under this combination, the first three full-field PCs capture 87% of the full-field variance, and the first two limited-field PCs capture 91% of the limited-field variance.

With this reasoning in mind, we model the first three PCs of the full field each as a linear function of the first two PCs of the limited field:

$$\boldsymbol{y}_1 = f(\boldsymbol{y}_1^\dagger, \boldsymbol{y}_2^\dagger) \qquad \boldsymbol{y}_2 = f(\boldsymbol{y}_1^\dagger, \boldsymbol{y}_2^\dagger) \qquad \boldsymbol{y}_3 = f(\boldsymbol{y}_1^\dagger, \boldsymbol{y}_2^\dagger) \tag{8}$$

The best linear regression model for each is identified via step-wise model selection based on Bayesian Information Criterion.

3.1.5. Step (v)

We use the PCs of the proxy data $\boldsymbol{Y}^R$ to reconstruct the first three PCs representative of 10 ka ($\hat{\boldsymbol{y}}_1$, $\hat{\boldsymbol{y}}_2$, $\hat{\boldsymbol{y}}_3$), and similarly for other times. The remaining G minus 3 reconstructed PCs ($\hat{\boldsymbol{y}}_{4 \dots G}$) are the means of the PCs from the full field ($\boldsymbol{y}_{4 \dots G}$).

3.1.6. Step (vi)

Finally, we transform the reconstructed PCs into a field of reconstructed SST anomalies via eigenvalue expansion, by multiplying by the original eigenvalue matrix from Step (i):

$$[\hat{\boldsymbol{T}}]_{1xG} = [\hat{\boldsymbol{Y}}]_{1xG}[\boldsymbol{U}]_{GxG} \tag{9}$$

Since $\hat{\boldsymbol{T}}$ is estimated using only the first few PCs, the reconstructed SSTs represent only the dominant signal. Steps (ii-vi) are repeated for 8, 6, 4, and 2 ka.

Standard errors from the regressions in Step (iv) are used to generate 500 ensembles of the first three PCs. 500 ensembles for each of the remaining PCs were created by bootstrapping values from the original PCs at each grid point. The standard deviation of these ensembles provides an estimate of uncertainty. As these errors were similar across the five reconstruction periods, we present the standard errors only from 2 ka in Figure 7. Standard errors are generally between 0.1-0.2°C throughout the central Pacific, with some errors reaching 0.3°C immediately along the coast of South America. Present-day variance is concentrated in the cold tongue of the eastern and east-central Pacific. In most places, the standard errors are only a small fraction, <10%, of the present-day variance (Figure 7).

3.2. Canonical Correlation Analysis (CCA) Reconstruction of Zonal Winds

The CCA-reconstruction method is quite similar to the PCA-based method depicted in Figure 3 and described in Section 3.1. CCA, however, is preferred over standard PCA in a scenario such as this where the patterns of one variable, e.g. SSTs, are being used as predictors of patterns for a related variable, e.g. winds. This method is widely used for multi-site forecasting of precipitation [e.g. *Barnston*, 1994; *Barnston and Smith*, 1996], temperature [e.g. *Barnett and Preisendorfer*, 1987; *Barnston*, 1994; *Barnston and Smith*, 1996; *Mo*, 2003], streamflow [e.g. *Salas et al.*, 2010], and SSTs [e.g. *Barnston and Ropelewski*, 1992]. See *Bretherton et al.* [1992], *Von Storch and Zwiers* [2001], and *Cherry* [1996] for details on CCA. Here, we use SST patterns to reconstruct zonal winds across the equatorial Pacific. We choose to reconstruct zonal winds because in this region zonal winds are much larger than meridional winds.

For the most part, the CCA method follows the PCA method described in Figure 3 through Steps (i-iii), and then differs from Step (iv) onward. For steps (i) and (ii), $\boldsymbol{T}$ becomes a matrix of 1949-2013 zonal wind anomalies, which we will call $\boldsymbol{Z}$, and $\boldsymbol{T}^\dagger$

remains a matrix of SST anomalies for the limited field. Since we only have 65 years of zonal wind data, N is now 65. As before, a PCA on both of these fields results in:

$$[\boldsymbol{Z}]_{NxG} = [\boldsymbol{Y}]_{NxG}[\boldsymbol{U}]_{GxG} \tag{10}$$

$$[\boldsymbol{T}^{\dagger}]_{NxP} = [\boldsymbol{Y}^{\dagger}]_{NxP}[\boldsymbol{U}^{\dagger}]_{PxP} \tag{11}$$

where $\boldsymbol{Y}$ and $\boldsymbol{U}$ are the full field PCs and eigenvectors of the zonal winds (Figures 6c and 8), and $\boldsymbol{Y}^{\dagger}$ and $\boldsymbol{U}^{\dagger}$ are still the limited field PCs and eigenvectors (Figures 5 and 6b). Step (iii) is the same: the proxy data $\boldsymbol{R}$ are decomposed into PCs, $\boldsymbol{Y}^{R}$, by multiplying by the limited field eigenvalues, $\boldsymbol{U}^{\dagger}$.

At this point, N_{PC} from $\boldsymbol{Y}$ and $\boldsymbol{Y}^{\dagger}$ are retained. Note that the numbers of full-field and of limited-field PCs retained must be equal. It is also worth noting that since a canonical correlation resolves the joint correlation between two PCs, it is acceptable to retain more PCs than one typically does during a standard PCA reconstruction. In fact, for zonal winds, keeping the first six PCs ($N_{PC} = 6$), which involves keeping PCs below the “knee” in Figure 6c, strengthened the model statistics (Section 3.3) and allowed us to account for 85% of the full-field variance. A canonical correlation is performed between the two to resolve the joint correlation between the two sets of PCs:

$$[\boldsymbol{Y}]_{NxN_{PC}} = [\boldsymbol{S}]_{NxN_{PC}}[\boldsymbol{A}]_{N_{PC}xN_{PC}} \tag{12}$$

$$[\boldsymbol{Y}^{\dagger}]_{NxN_{PC}} = [\boldsymbol{S}^{\dagger}]_{NxN_{PC}}[\boldsymbol{B}]_{N_{PC}xN_{PC}} \tag{13}$$

where $\boldsymbol{A}$ and $\boldsymbol{B}$ are now the canonical transformation functions, $\boldsymbol{S}$ and $\boldsymbol{S}^{\dagger}$ are N_{PC} pairs of canonical components (CCs), each of which explains more of the joint variance than the next, and all of which are uncorrelated to each other. The CC of the full-field are regressed against the corresponding CC of the limited field, such that:

$$s_j = \hat{a}_j + \beta_j s_j^{\dagger} \tag{14}$$

where j is 1 through N_{PC}, and $\hat{\boldsymbol{\beta}}$ is the solution to the least squares optimization. The first N_{PC} reconstructed PCs of the full wind field $\hat{\boldsymbol{Y}}$ are obtained through expansion by multiplication of first N_{PC} proxy PCs with the full-field canonical transformation matrix $\boldsymbol{A}$ and the $\hat{\boldsymbol{\beta}}$ solution to the regression:

$$[\hat{\boldsymbol{Y}}]_{1xN_{PC}} = [\boldsymbol{Y}^{R}]_{1xN_{PC}}[\boldsymbol{A}]_{N_{PC}xN_{PC}}[\hat{\boldsymbol{\beta}}]_{N_{PC}xN_{PC}} \tag{15}$$

We calculate the remaining $G - N_{PC}$ PCs ($\hat{y}_{7...M}$) by taking the means of the PCs from the full field. Finally, we obtain the reconstructed zonal winds $\hat{\boldsymbol{Z}}$ by eigenvalue expansion:

$$[\hat{\boldsymbol{Z}}]_{1xG} = [\hat{\boldsymbol{Y}}]_{1xG}[\boldsymbol{U}]_{GxG} \tag{16}$$

We repeated this process for each of the six time periods (10, 8, 6, 4, and 2 ka).

Implicit in this approach are a few assumptions. First, we assume that the calibration equations used by the authors of each proxy record represent the “best” calibration for that particular record. There are a number of calibration equations that can be used to transform Mg/Ca or U_{37}^{k} to SST. The “best” equation is typically chosen by considering proxy type, location, dissolution effects (in the case of Mg/Ca), and sampling method. We do not convert all records using the same calibration equation because we assume that the authors of each proxy record have chosen the calibration equation that is most appropriate for their particular record. Second, we assume that spatial patterns of the proxy network are linearly related to spatial patterns in the contemporary period. Third, we assume that the main patterns of variability that exist in the contemporary dataset have not changed over the past 10 ka. *Molnar and Cane* [2002, 2007] show that ENSO teleconnection patterns of present-day seems to have existed as far back as the early Pliocene ($\sim$ 5 Ma), suggesting that the major patterns of variability have not changed significantly over the relatively short period of time considered in this analysis.

3.3. Calibration and Verification

To assess the reliability of the procedure above to reproduce contemporary SSTs and zonal winds from the limited field contemporary SST data, we use the ‘resolved variance’ statistic β, given by:

$$\beta = 1 - \frac{\sum (y - \hat{y})^2}{\sum y^2} \tag{17}$$

where y is the contemporary data and $\hat{y}$ is the reconstructed data for the full period of each dataset (1854-2014 for SST, 1949-2013 for winds). We compute this statistic at each grid-point over the reconstruction domain (Figure 9, top). For a perfect fit, one would expect β = 1, and for two random series, one would expect β = -1. Across most of the reconstruction region, β is greater than 0.6 for both SSTs and zonal winds. The lowest β-values are found over the central Pacific for SSTs and over the landmasses in the west and east Pacific for zonal winds.

We also correlated data and calculated values at each grid-point, and plot maps of the squared correlation coefficient, R^2 (Figure 9, bottom). For SSTs, the east and west Pacific show the highest correlations, which we should expect, as those areas provide the SST data for the reconstruction. The western “cold tongue” region of the east-central Pacific, where there are no data, shows correlations greater than or equal to 0.5. The lowest correlation region for SSTs is the west-central Pacific with values of 0.2-0.3. For zonal winds, the R^2 values are largest in the central Pacific, which is perhaps not surprising, as the central Pacific winds depend on SSTs at either end of the Pacific. As with the β-test, the worst performance for winds is seen over continental landmasses east and west of the Pacific.

For additional calibration, we compare the actual and reconstructed SST and zonal wind fields for the anomalous ENSO years of 1988-1989 (strong La Niña with standardized SST anomalies of -1.5°C in the NINO3 index region of 150°W - 90°W and 5°S - 5°N) and 1997-1998 (strong El Niño with standardized NINO3 index SST anomalies of +2.8°C). For 1988-1989, the model captures anomalies slightly cooler than -1°C, which are not as cold and do not extend as far westward as the actual anomalies seen during this particular La Niña year (Figure 10, top). The reconstructed zonal wind pattern for this year follows a similar shape to that of actual zonal winds, with strongly negative anomalies (strengthened easterlies) across the central Pacific and into the western Pacific north of the equator, with positive anomalies over the Indonesian archipelago (Figure 11, top). The reconstruction captures the two regions of highest negative anomalies (-3.5 m s^{-1} for the central Pacific and -2.5 m s^{-1} for the western Pacific), but switches the magnitudes in the reconstruction. The negative anomalies in winds over the far eastern Pacific off the coast of South America are not captured by the reconstruction, which is consistent with the relatively low β- and R^2- values seen in that region (Figure 9).

The magnitude and spatial pattern of anomalous SSTs is reconstructed better for the strong El Niño year of 1997-1998 (Figure 10, bottom) than for 1988-1989. Consistent with actual warming, the reconstructed "warm tongue" extends just west of the dateline, the greatest magnitude of warming, which is off the coast of Peru in the far eastern Pacific, reaches approximately +3.5°C, and the distribution of temperatures within the reconstructed "warm tongue" closely resemble those of the observed SSTs. It is likely that the failure of the limited-field model to capture La Niña and El Niño SST anomalies across the entire Pacific is due to the restricted locations of the proxy records. The cold La Niña anomaly of 1988-1989 is centered in the east-central Pacific, where there are no proxy records. The maximum warm anomaly of the 1997-1998 El Niño, however, is located in the far eastern Pacific, which the paleoceanographic data sample well. Reconstructed winds for the El Niño year are arguably better than those for the La Niña year as well. The spatial patterns resemble one another, but the magnitude of the maximum reconstructed easterlies is underestimated by about 1 m s^{-1}.

For model verification, we train the PCA and CCA models using only the most recent data (1980-2013) and withhold the data from earlier periods (1854-1979 for SST, 1949-1979 for winds). Then, we reconstruct the fields from the earlier epoch using each respective model. As expected, the distribution of β and R^2 statistics (Figure 12) show smaller values than the calibration statistics for the complete data sets (Figure 9), but the regions that were reconstructed best during the calibration experiments (Figure 9) match those reconstructed best in the verification test (Figure 12). For SSTs, the training model reconstructs the earlier period everywhere except the west-central Pacific and areas near 10°S and 10°N in the central Pacific. For zonal winds, the central Pacific is the only region that is reconstructed using the training model. These patterns are consistent across both the β and R^2 statistics.

4. Multi-proxy Reconstruction of SSTs and Zonal Winds

We reconstruct full-field SSTs and zonal winds for 10, 8, 6, 4, and 2 ka using the multi-proxy approach described above (Figure 13). We choose not to reconstruct 0 ka since many of the proxy records do not extend to 0 ka and extrapolation of late Holocene SST trends introduces bias, especially in the east. Instead, we provide the mean present-day state (1854-2013 for SSTs and 1949-2013 for zonal winds) as reconstructed using the limited-field data (Figure 13). The multi-proxy SST reconstructions show the far-eastern Pacific to have been a maximum of -1°C cooler than today at 10 ka, along with a cooler (-0.8°C) "cold tongue" extending to about 140°E at 10 ka. The western Pacific was -0.6°C cooler for the same time period. From 10 to 6 ka, conditions changed little, both in magnitude and location, with the exception of a slight warming in the west and west-central Pacific. By 4 to 2 ka, the majority of the region had warmed by 0.4-0.6°C since 10 ka. The present-day reconstructions show small cool anomalies of 0.4C throughout the Pacific. Recall that anomalies in Figure 13 are relative to the 1981-2010 climatology of each variable.

The multi-proxy reconstruction of winds (Figure 13) reveals maximum easterlies (large negative zonal winds) at 10 ka, which is also the period of lowest SSTs in the eastern Pacific, consistent with increased easterlies seen during present-day La Niña events. Consistently throughout the reconstructions, the zone of strongest easterlies is centered around 140°W. Easterlies are strongest at 10 ka, as much as 8 m s^{-1} stronger than present. Between 8 and 2 ka, the difference between past and present winds decreases and the location of the maximum difference moves eastward. Present-day reconstructions of zonal winds show the weakest easterlies of the Holocene (only -1 m s^{-1}) concentrated in the east-central Pacific.

The zonal difference in SSTs from west to east is also of interest, as some [e.g. *Koutavas et al.*, 2006] have argued that an "enhanced zonal gradient" characterized the early- to mid-Holocene Pacific. To quantify such gradients, Table 3 provides various indices throughout the Pacific (WPAC, NINO4, NINO3.4, NINO3, and NINO1+2) as well as difference indices (TNI and WTNI). The domain used in calculating each index is depicted in Figure 14 and provided in the notes of Table 3. The TNI index quantifies the zonal SST difference by subtracting NINO4 from NINO1+2. NINO4, however, is farther east than all of the western proxy records. For this reason, we also report a difference index called the WTNI, which is calculated by subtracting the WPAC (120°E - 180°E and 4°S to 4°N) from NINO1+2. Although the period of maximum eastern Pacific cooling was 10 ka, both 6 and 10 ka appear to have had the largest zonal difference in SSTs with -0.36 ± 0.29°C and -0.33 ± 0.30°C, respectively (Table 3a). The next largest zonal difference occurs at 8 ka (-0.25°C ± 0.29°C). From 4 to 2 ka, the zonal difference was only -15°C ± 0.28°C.

A comparison of reconstructed SST values with proxy SST anomalies from the western (Figure 15) and eastern (Figure 16) Pacific reveals better matches at some locations than others, which we quantify with the residual sum of squares (RSS) in the top right of each scatterplot. Recall that the first two modes of the limited field were used as predictors of SSTs. Eigenvalues from the first and second modes under each scatterplot help explain why some proxy records are replicated better than others. These eigenvalues correspond to those plotted in the first two maps of Figure 5. Although not always the case, high RSS values, and hence poor fits, can often be attributed to low EOF values assigned to those locations, which suggests that these data contributed little to the reconstructed SST fields.

For the first mode (Figure 5, EOF no. 1), the signal is dominated by the east Pacific, with eigenvalues of much greater magnitude than those from the west Pacific (note the low EOF1 values on all of the west Pacific scatterplots of Figure 15). Of all the locations in the east Pacific, the three most northern records (proxy no. 11, 22, and 23) have the lowest EOF1 values, and of these, records 11 and 23 have high RSS values. Proxy no. 22, despite a low EOF1 value, is reconstructed quite well (RSS = 1.49°C^2). The two records with the lowest RSS values (RSS = 0.9°C^2), proxy no. 27 and 14, have two of the highest EOF1 values.

The second mode EOF no. 2 (Figure 5) is dominated by records from the western Pacific (note the low EOF2 values on the east Pacific scatterplots of Figure 16). It is important to remember, however, that the second mode explains only 12.6% of the limited-field variance (as compared to 78.2% by the first mode). Nevertheless, only 3 out of 15 reconstructions for the west Pacific show RSS values greater than 10°C^2, which suggests that reconstructions match proxy records better in the west than in the east, for which RSS >10°C^2 for 5 of 11 reconstructions.

5. Sensitivity to Proxy Type

Although Mg/Ca and U_{37}^{k} both provide reconstructions of SSTs, they differ in methodology and calibration. Furthermore they are produced by two different groups of organisms that may each reflect biases related to season of production, depth habitat, water column stratification, and preservation [*Leduc et al.*, 2010]. It is clear from Figures 1 and 2 that for a given location, records based on different proxies are not always consistent, but also within proxy type. In a case for which there are many records, a reduced-dimension reconstruction approach is particularly beneficial, assuming that biases across the various proxy reconstructions cancel out. By combining all records, this method extracts common features across the proxies and allows for a robust reconstruction of the most dominant patterns of variability. Additionally, including all the proxy records together allows for wider coverage in both space and time, than if one were to reconstruct using only one proxy type. For these reasons, we recommend the multi-proxy approach. We recognize, however, that comparing a reconstruction based only on Mg/Ca with one based only on U_{37}^{k} provides a test of the sensitivity of proxy type to the multi-proxy reconstruction. All 15 records listed in Table 1 were used in the Mg/Ca-only reconstruction (Figure 17, top), and all 14 records listed in Table 2 were used in the U^k_{37}-only reconstruction (Figure 17, bottom).

SST reconstructions based on either Mg/Ca or U^k_{37} show negative anomalies throughout 10-2 ka, and weak negative anomalies for the present-day 1854-2013 period. They differ, however, in magnitude and timing of cold anomalies. First, for Mg/Ca, the coldest eastern equatorial SST anomalies occur in a well-defined "cold tongue" pattern at 6 ka, which is also the period of largest zonal difference in SSTs of -0.77°C ± 0.30°C (Table 3b). For U^k_{37}, however, maximum eastern equatorial cooling occurs at 10 ka. The zonal difference during this time is slightly positive, though, since larger negative anomalies are seen in the western Pacific (Table 3c).

Second, the evolution of SSTs in the west Pacific is strikingly different: In the Mg/Ca reconstruction, the western Pacific cold anomalies are always smaller than those of the eastern Pacific, but in the U^k_{37} reconstruction, west Pacific cold anomalies is almost always as large as and sometimes larger than those of the eastern Pacific. This characteristic in the U^k_{37} reconstruction is inconsistent with a typical La Niña-like SST configuration, where the western Pacific is warmer than the eastern Pacific, and is also inconsistent with model results for the early- to mid-Holocene [*Liu et al.*, 2000; *Otto-Bliesner et al.*, 2003; *Zheng et al.*, 2008].

Lastly, only the U^k_{37} reconstruction shows warmer eastern equatorial SSTs during 2 ka than present-day. This would be consistent with marine records *Koutavas et al.* [2006]; *Tudhope et al.* [2001] and sediment records *Moy et al.* [2002]; *Rodbell et al.* [1999] that have suggested an abrupt return of ENSO variability (i.e. more El Niños and thus, warmer eastern equatorial SSTs) at around 2 ka.

6. Discussion

We apply PCA-based and CCA-based reduced-dimension reconstructions of the full field of SSTs and ENSO variability during the past 10 ka. Using a limited spatial field of 26 locations provides enough data to resolve the major patterns of variability for the entire Pacific SSTs and most of the central Pacific zonal winds.

For the multi-proxy reconstruction, the period with the coolest eastern equatorial Pacific (-1°C) and strongest easterly winds (-8 m s^{-1}) across the central Pacific occurred at 10 ka, with similar patterns, but gradually decreasing magnitudes persisting until 6 ka. By 4 to 2 ka, the Pacific remained in a state cooler than present, but with SST anomalies around 6 to -4 °C.

Assuming decreased ENSO variability is consistent with an enhanced La Niña-like state, the timing of maximum cooling in these findings is consistent with the timing reconstructed by both [*Rodbell et al.*, 1999] and [*Moy et al.*, 2002], who used records of sedimentation in lakes in Ecuador to infer ENSO variability over the past 15 and 12 ka, respectively, and who showed an increasing occurrence of ENSO following 7 ka along with a low frequency of ENSO events prior to that. Additionally, in a comparison of east and west Pacific proxy SST reconstructions, *Koutavas et al.* [2006] claimed that the period of an enhanced zonal gradient across the Pacific, and thus, an enhanced La Niña-like state, existed from 11 ka until 5 ka. [*Clement et al.*, 2000], who used an orbitally forced model to reconstruct ENSO variability during the past 12 ka, found that the amplitude of

ENSO events increased from 10 ka to present, but that both the amplitude and frequency increased during the latter half of that period. Our reconstructions are consistent with this if one assumes that increased ENSO amplitude and frequency would result in warmer anomalies during the mid- to late-Holocene.

Koutavas et al. [2006] emphasized the role of the zonal difference between the east and west Pacific that helped create persisting La Niña conditions, and they inferred that although the enhanced zonal SST difference lasted from 11-5 ka, it was strongest at 6 ka. This calls attentino to the importance of considering not only periods of greatest cooling, but also the periods when the SST difference between eastern and western Pacific was greatest. As noted in Table 3a, the zonal difference in the multi-proxy reconstructions was largest at 6 and 10 ka. This timing coincides well with the timing suggested in the multi-proxy reconstructions: (1) the largest zonal difference, and presumably lowest ENSO variability, occurs at 10 and 6 ka, (2) between 10 and 6 ka, the zonal difference decreases slightly, but remains "La Niña-like," and (3) the late Holocene (4 to 2 ka) show only small negative zonal differences.

The PCA and CCA based methods offer several advantages: they are conceptually simple, and have a well-developed theoretical foundation and rich history of application to climate field reconstructions. They are, however, limiting in other ways: the reconstructions are based on only linear projections, quantification of uncertainty in the reconstructions can be unwieldy and, it is not straightforward to incorporate uncertainty information of the proxy data. Emerging Bayesian Hierarchical Modeling framework offers attractive alternatives to addressing these limitations, especially in estimating uncertainties in a robust manner by including data and model uncertainties.

Acknowledgments. The authors thank M. Cane and Y. Kushnir for valuable discussions of this work. This research was supported in part by the National Science Foundation under grants EAR-1211378 and OCE-1103482. All of the contemporary data used in this manuscript can be found on the IRI/LDEO Data Library (http://iridl.ldeo.columbia.edu/). All paleo-SST proxy records can be found in either the archives of NCDC (http://www.ncdc.noaa.gov/data-access/paleoclimatology-data/datasets) or Pangaea (http://www.pangaea.de/).

References

Anand, P., H. Elderfield, and M. H. Conte (2003), Calibration of mg/ca thermometry in planktonic foraminifera from a sediment trap time series, *Paleoceanography*, *18*(2).

Barnett, T., and R. Preisendorfer (1987), Origins and levels of monthly and seasonal forecast skill for united states surface air temperatures determined by canonical correlation analysis, *Monthly Weather Review*, *115*(9), 1825–1850.

Barnston, A. G. (1994), Linear statistical short-term climate predictive skill in the northern hemisphere, *Journal of Climate*, *7*(10), 1513–1564.

Barnston, A. G., and C. F. Ropelewski (1992), Prediction of enso episodes using canonical correlation analysis, *Journal of climate*, *5*(11), 1316–1345.

Barnston, A. G., and T. M. Smith (1996), Specification and prediction of global surface temperature and precipitation from global sst using cca, *Journal of Climate*, *9*(11), 2660–2697.

Benway, H. M., A. C. Mix, B. A. Haley, and G. P. Klinkhammer (2006), Eastern pacific warm pool paleosalinity and climate variability: 0–30 kyr, *Paleoceanography*, *21*(3).

Bolliet, T., A. Holbourn, W. Kuhnt, C. Laj, C. Kissel, L. Beaufort, M. Kienast, N. Andersen, and D. Garbe-Schönberg (2011), Mindanao dome variability over the last 160 kyr: Episodic glacial cooling of the west pacific warm pool, *Paleoceanography*, *26*(1).

Brassell, S., G. Eglinton, I. Marlowe, U. Pflaumann, and M. Sarnthein (1986), Molecular stratigraphy: a new tool for climatic assessment, *Nature*, *320*, 129–133.

Bretherton, C. S., C. Smith, and J. M. Wallace (1992), An intercomparison of methods for finding coupled patterns in climate data, *Journal of climate*, *5*(6), 541–560.

Carré, M., J. P. Sachs, S. Purca, A. J. Schauer, P. Braconnot, R. A. Falcón, M. Julien, and D. Lavallée (2014), Holocene history of enso variance and asymmetry in the eastern tropical pacific, *Science*, *345*(6200), 1045–1048.

Cherry, S. (1996), Singular value decomposition analysis and canonical correlation analysis, *Journal of Climate*, *9*(9), 2003–2009.

Clement, A., R. Seager, and M. Cane (1999), Orbital controls on the el niño/southern oscillation and the tropical climate, *Paleoceanography*, *14*(4), 441–456.

Clement, A. C., R. Seager, and M. A. Cane (2000), Suppression of el niño during the mid-holocene by changes in the earth's orbit, *Paleoceanography*, *15*(6), 731–737.

Cobb, K. M., N. Westphal, H. R. Sayani, J. T. Watson, E. Di Lorenzo, H. Cheng, R. Edwards, and C. D. Charles (2013), Highly variable el niño–southern oscillation throughout the holocene, *Science*, *339*(6115), 67–70.

Conte, M. H., M.-A. Sicre, C. Rühlemann, J. C. Weber, S. Schulte, D. Schulz-Bull, and T. Blanz (2006), Global temperature calibration of the alkenone unsaturation index (uk 37) in surface waters and comparison with surface sediments, *Geochemistry, Geophysics, Geosystems*, *7*(2).

Cook, E. R., D. M. Meko, D. W. Stahle, and M. K. Cleaveland (1999), Drought reconstructions for the continental united states*, *Journal of Climate*, *12*(4), 1145–1162.

de Garidel-Thoron, T., Y. Rosenthal, F. Bassinot, and L. Beaufort (2005), Stable sea surface temperatures in the western pacific warm pool over the past 1.75 million years, *Nature*, *433*(7023), 294–298.

de Garidel-Thoron, T., Y. Rosenthal, L. Beaufort, E. Bard, C. Sonzogni, and A. C. Mix (2007), A multiproxy assessment of the western equatorial pacific hydrography during the last 30 kyr, *Paleoceanography*, *22*(3).

Dekens, P. S., D. W. Lea, D. K. Pak, and H. J. Spero (2002), Core top calibration of mg/ca in tropical foraminifera: Refining paleotemperature estimation, *Geochemistry, Geophysics, Geosystems*, *3*(4), 1–29.

Evans, M. N., A. Kaplan, and M. A. Cane (2002), Pacific sea surface temperature field reconstruction from coral δ18o data using reduced space objective analysis, *Paleoceanography*, *17*(1), 7–1.

Hastings, D., M. Kienast, S. Steinke, and A. Whitko (2001), A comparison of three independent paleotemperature estimates from a high resolution record of deglacial sst records in the tropical south china sea, in *AGU Fall Meeting Abstracts*, vol. 1, p. 10.

Herbert, T. (2003), Alkenone paleotemperature determinations, Elsevier.

Kalnay, E., M. Kanamitsu, R. Kistler, W. Collins, D. Deaven, L. Gandin, M. Iredell, S. Saha, G. White, J. Woollen, et al. (1996), The ncep/ncar 40-year reanalysis project, *Bulletin of the American meteorological Society*, *77*(3), 437–471.

Kaplan, A., M. A. Cane, Y. Kushnir, A. C. Clement, M. B. Blumenthal, and B. Rajagopalan (1998), Analyses of global sea surface temperature 1856–1991, *Journal of Geophysical Research: Oceans (1978–2012)*, *103*(C9), 18,567–18,589.

Kaufman, D. S., D. P. Schneider, N. P. McKay, C. M. Ammann, R. S. Bradley, K. R. Briffa, G. H. Miller, B. L. Otto-Bliesner, J. T. Overpeck, B. M. Vinther, et al. (2009), Recent warming reverses long-term arctic cooling, *Science*, *325*(5945), 1236–1239.

Kienast, M., S. Steinke, K. Stattegger, and S. Calvert (2001), Synchronous tropical south china sea sst change and greenland warming during deglaciation, *Science*, *291*(5511), 2132–2134.

Kienast, M., S. S. Kienast, S. E. Calvert, T. I. Eglinton, G. Mollenhauer, R. François, and A. C. Mix (2006), Eastern pacific cooling and atlantic overturning circulation during the last deglaciation, *Nature*, *443*(7113), 846–849.

Koutavas, A., and J. P. Sachs (2008), Northern timing of deglaciation in the eastern equatorial pacific from alkenone paleothermometry, *Paleoceanography*, *23*(4).

Koutavas, A., G. C. Olive, J. Lynch-Stieglitz, et al. (2006), Mid-holocene el niño–southern oscillation (enso) attenuation revealed by individual foraminifera in eastern tropical pacific sediments, *Geology*, *34*(12), 993–996.

Lea, D. W., and P. A. Martin (1996), A rapid mass spectrometric method for the simultaneous analysis of barium, cadmium, and strontium in foraminifera shells, *Geochimica* *et Cosmochimica Acta*, *60*(16), 3143–3149.

Lea, D. W., T. A. Mashiotta, and H. J. Spero (1999), Controls on magnesium and strontium uptake in planktonic foraminifera determined by live culturing, *Geochimica et Cosmochimica Acta*, *63*(16), 2369–2379.

Lea, D. W., D. K. Pak, and H. J. Spero (2000), Climate impact of late quaternary equatorial pacific sea surface temperature variations, *Science*, *289*(5485), 1719–1724.

Lea, D. W., D. K. Pak, C. L. Belanger, H. J. Spero, M. A. Hall, and N. J. Shackleton (2006), Paleoclimate history of galapagos surface waters over the last 135,000 yr, *Quaternary Science Reviews*, *25*(11), 1152–1167.

Leduc, G., L. Vidal, K. Tachikawa, F. Rostek, C. Sonzogni, L. Beaufort, and E. Bard (2007), Moisture transport across central america as a positive feedback on abrupt climatic changes, *Nature*, *445*(7130), 908–911.

Leduc, G., R. Schneider, J.-H. Kim, and G. Lohmann (2010), Holocene and eemian sea surface temperature trends as revealed by alkenone and mg/ca paleothermometry, *Quaternary Science Reviews*, *29*(7), 989–1004.

Lee, T. C., F. W. Zwiers, and M. Tsao (2008), Evaluation of proxy-based millennial reconstruction methods, *Climate Dynamics*, *31*(2-3), 263–281.

Levi, C., L. Labeyrie, F. Bassinot, F. Guichard, E. Cortijo, C. Waelbroeck, N. Caillon, J. Duprat, T. de Garidel-Thoron, and H. Elderfield (2007), Low-latitude hydrological cycle and rapid climate changes during the last deglaciation, *Geochemistry, Geophysics, Geosystems*, *8*(5).

Li, B., D. W. Nychka, and C. M. Ammann (2010), The value of multiproxy reconstruction of past climate, *Journal of the American Statistical Association*, *105*(491), 883–895.

Liu, Z., J. Kutzbach, and L. Wu (2000), Modeling climate shift of el nino variability in the holocene, *Geophysical Research Letters*, *27*(15), 2265–2268.

Loader, C. R., et al. (1996), Local likelihood density estimation, *The Annals of Statistics*, *24*(4), 1602–1618.

Luan, Y., P. Braconnot, Y. Yu, W. Zheng, and O. Marti (2012), Early and mid-holocene climate in the tropical pacific: seasonal cycle and interannual variability induced by insolation changes, *Climate of the Past*, *8*(3), 1093–1108.

Lückge, A., M. Mohtadi, C. Rühlemann, G. Scheeder, A. Vink, L. Reinhardt, and M. Wiedicke (2009), Monsoon versus ocean circulation controls on paleoenvironmental conditions off southern sumatra during the past 300,000 years, *Paleoceanography*, *24*(1).

Luterbacher, J., E. Xoplaki, D. Dietrich, R. Rickli, J. Jacobeit, C. Beck, D. Gyalistras, C. Schmutz, and H. Wanner (2002), Reconstruction of sea level pressure fields over the eastern north atlantic and europe back to 1500, *Climate Dynamics*, *18*(7), 545–561.

Luterbacher, J., D. Dietrich, E. Xoplaki, M. Grosjean, and H. Wanner (2004), European seasonal and annual temperature variability, trends, and extremes since 1500, *Science*, *303*(5663), 1499–1503.

Mann, M. E., R. S. Bradley, and M. K. Hughes (1998), Global-scale temperature patterns and climate forcing over the past six centuries, *Nature*, *392*(6678), 779–787.

Mann, M. E., Z. Zhang, M. K. Hughes, R. S. Bradley, S. K. Miller, S. Rutherford, and F. Ni (2008), Proxy-based reconstructions of hemispheric and global surface temperature variations over the past two millennia, *Proceedings of the National Academy of Sciences*, *105*(36), 13,252–13,257.

McGregor, H., M. Fischer, M. Gagan, D. Fink, S. Phipps, H. Wong, and C. Woodroffe (2013), A weak el niño/southern oscillation with delayed seasonal growth around 4,300 years ago, *Nature Geoscience*, *6*(11), 949–953.

Mo, K. C. (2003), Ensemble canonical correlation prediction of surface temperature over the united states, *Journal of climate*, *16*(11), 1665–1683.

Moberg, A., D. M. Sonechkin, K. Holmgren, N. M. Datsenko, and W. Karlén (2005), Highly variable northern hemisphere temperatures reconstructed from low-and high-resolution proxy data, *Nature*, *433*(7026), 613–617.

Molnar, P., and M. A. Cane (2002), El niño's tropical climate and teleconnections as a blueprint for pre-ice age climates, *Paleoceanography*, *17*(2), 11–1.

Molnar, P., and M. A. Cane (2007), Early pliocene (pre–ice age) el niño–like global climate: Which el niño?, *Geosphere*, *3*(5), 337–365.

Moy, C. M., G. O. Seltzer, D. T. Rodbell, and D. M. Anderson (2002), Variability of el niño/southern oscillation activity at millennial timescales during the holocene epoch, *Nature*, *420*(6912), 162–165.

Müller, P. J., G. Kirst, G. Ruhland, I. von Storch, and A. Rosell-Melé (1998), Calibration of the alkenone paleotemperature index u 37 k based on core-tops from the eastern south atlantic and the global ocean (60 n-60 s), *Geochimica et Cosmochimica Acta*, *62*(10), 1757–1772.

Nürnberg, D., J. Bijma, and C. Hemleben (1996), Assessing the reliability of magnesium in foraminiferal calcite as a proxy for water mass temperatures, *Geochimica et Cosmochimica Acta*, *60*(5), 803–814.

Otto-Bliesner, B. L., E. C. Brady, S.-I. Shin, Z. Liu, and C. Shields (2003), Modeling el niño and its tropical teleconnections during the last glacial-interglacial cycle, *Geophysical Research Letters*, *30*(23).

Pahnke, K., J. P. Sachs, L. Keigwin, A. Timmermann, and S.-P. Xie (2007), Eastern tropical pacific hydrologic changes during the past 27,000 years from d/h ratios in alkenones, *Paleoceanography*, *22*(4).

Pelejero, C., and J. O. Grimalt (1997), The correlation between the uk'37 index and sea surface temperatures in the warm boundary: The south china sea, *Geochimica et Cosmochimica Acta*, *61*(22), 4789–4797.

Pelejero, C., J. O. Grimalt, S. Heilig, M. Kienast, L. Wang, et al. (1999), High-resolution uk37 temperature reconstructions in the south china sea over the past 220 kyr, *Paleoceanography*, *14*(2), 224–231.

Pena, L., I. Cacho, P. Ferretti, and M. Hall (2008), El niño–southern oscillation–like variability during glacial terminations and interlatitudinal teleconnections, *Paleoceanography*, *23*(3).

Prahl, F. G., L. A. Muehlhausen, and D. L. Zahnle (1988), Further evaluation of long-chain alkenones as indicators of paleoceanographic conditions, *Geochimica et Cosmochimica Acta*, *52*(9), 2303–2310.

Rodbell, D. T., G. O. Seltzer, D. M. Anderson, M. B. Abbott, D. B. Enfield, and J. H. Newman (1999), An˜ 15,000-year record of el niño-driven alluviation in southwestern ecuador, *Science*, *283*(5401), 516–520.

Rosenthal, Y., and G. P. Lohmann (2002), Accurate estimation of sea surface temperatures using dissolution-corrected calibrations for mg/ca paleothermometry, *Paleoceanography*, *17*(3), 16–1.

Rosenthal, Y., D. W. Oppo, and B. K. Linsley (2003), The amplitude and phasing of climate change during the last deglaciation in the sulu sea, western equatorial pacific, *Geophysical Research Letters*, *30*(8).

Rutherford, S., M. Mann, T. Osborn, K. Briffa, P. D. Jones, R. Bradley, and M. Hughes (2005), Proxy-based northern hemisphere surface temperature reconstructions: sensitivity to method, predictor network, target season, and target domain, *Journal of Climate*, *18*(13), 2308–2329.

Salas, J. D., C. Fu, and B. Rajagopalan (2010), Long-range forecasting of colorado streamflows based on hydrologic, atmospheric, and oceanic data, *Journal of Hydrologic Engineering*.

Shulmeister, J., and B. G. Lees (1995), Pollen evidence from tropical australia for the onset of an enso-dominated climate at c. 4000 bp, *The Holocene*, *5*(1), 10–18.

Smith, T. M., R. W. Reynolds, T. C. Peterson, and J. Lawrimore (2008), Improvements to noaa's historical merged land-ocean surface temperature analysis (1880-2006), *Journal of Climate*, *21*(10), 2283–2296.

Sonzogni, C., E. Bard, F. Rostek, R. Lafont, A. Rosell-Mele, and G. Eglinton (1997), Core-top calibration of the alkenone index vs sea surface temperature in the indian ocean, *Deep Sea Research Part II: Topical Studies in Oceanography*, *44*(6), 1445–1460.

Steinke, S., M. Kienast, J. Groeneveld, L.-C. Lin, M.-T. Chen, and R. Rendle-Bühring (2008), Proxy dependence of the temporal pattern of deglacial warming in the tropical south china sea: toward resolving seasonality, *Quaternary Science Reviews*, *27*(7), 688–700.

Stott, L., K. Cannariato, R. Thunell, G. H. Haug, A. Koutavas, and S. Lund (2004), Decline of surface temperature and salinity in the western tropical pacific ocean in the holocene epoch, *Nature*, *431*(7004), 56–59.

Timmermann, A., Y. Okumura, S.-I. An, A. Clement, B. Dong, E. Guilyardi, A. Hu, J. Jungclaus, M. Renold, T. Stocker, et al. (2007), The influence of a weakening of the atlantic meridional overturning circulation on enso, *Journal of Climate*, *20*(19), 4899–4919.

Tingley, M. P., and P. Huybers (2010), A bayesian algorithm for reconstructing climate anomalies in space and time. part i: Development and applications to paleoclimate reconstruction problems, *Journal of Climate*, *23*(10), 2759–2781.

Tudhope, A. W., C. P. Chilcott, M. T. McCulloch, E. R. Cook, J. Chappell, R. M. Ellam, D. W. Lea, J. M. Lough, and G. B. Shimmield (2001), Variability in the el niño-southern oscillation through a glacial-interglacial cycle, *Science*, *291*(5508), 1511–1517.

Visser, K., R. Thunell, and L. Stott (2003), Magnitude and timing of temperature change in the indo-pacific warm pool during deglaciation, *Nature*, *421*(6919), 152–155.

Von Storch, H., and F. W. Zwiers (2001), *Statistical analysis in climate research*, Cambridge university press.

Zhang, Z., M. E. Mann, and E. R. Cook (2004), Alternative methods of proxy-based climate field reconstruction: application to summer drought over the conterminous united states back to ad 1700 from tree-ring data, *The Holocene*, *14*(4), 502–516.

Zhao, M., C.-Y. Huang, C.-C. Wang, and G. Wei (2006), A millennial-scale u 37 k sea-surface temperature record from the south china sea (8 n) over the last 150 kyr: Monsoon and sea-level influence, *Palaeogeography, Palaeoclimatology, Palaeoecology*, *236*(1), 39–55.

Zheng, W., P. Braconnot, E. Guilyardi, U. Merkel, and Y. Yu (2008), Enso at 6ka and 21ka from ocean–atmosphere coupled model simulations, *Climate Dynamics*, *30*(7-8), 745–762.

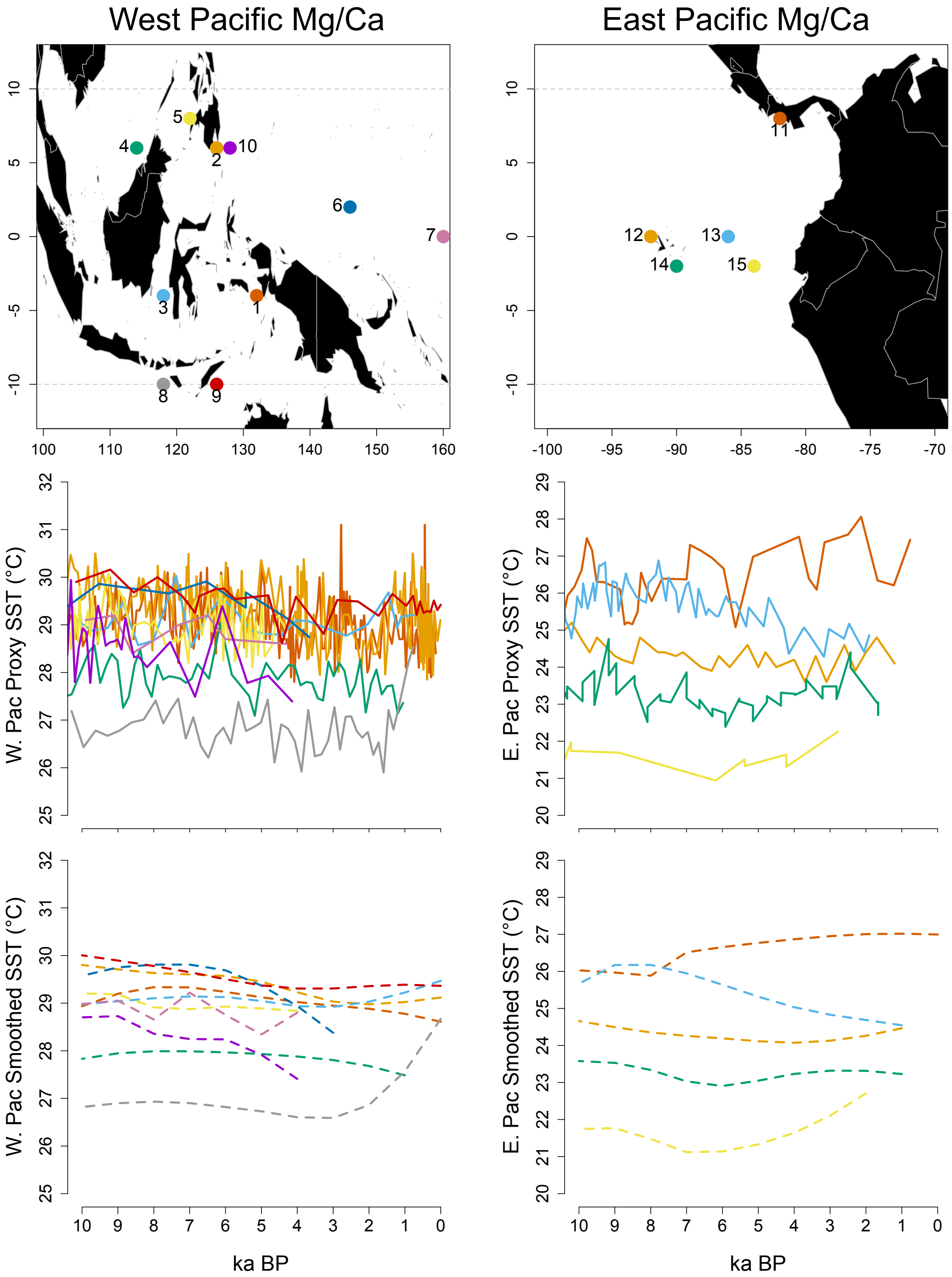


Figure 1. Maps of western and eastern Mg/Ca based SST records. The second row provides raw SST time series as calibrated by the authors listed in Table 1. The third row provides the records smoothes by a second order local polynomial method with a local neighborhood of 70% nearest data points.

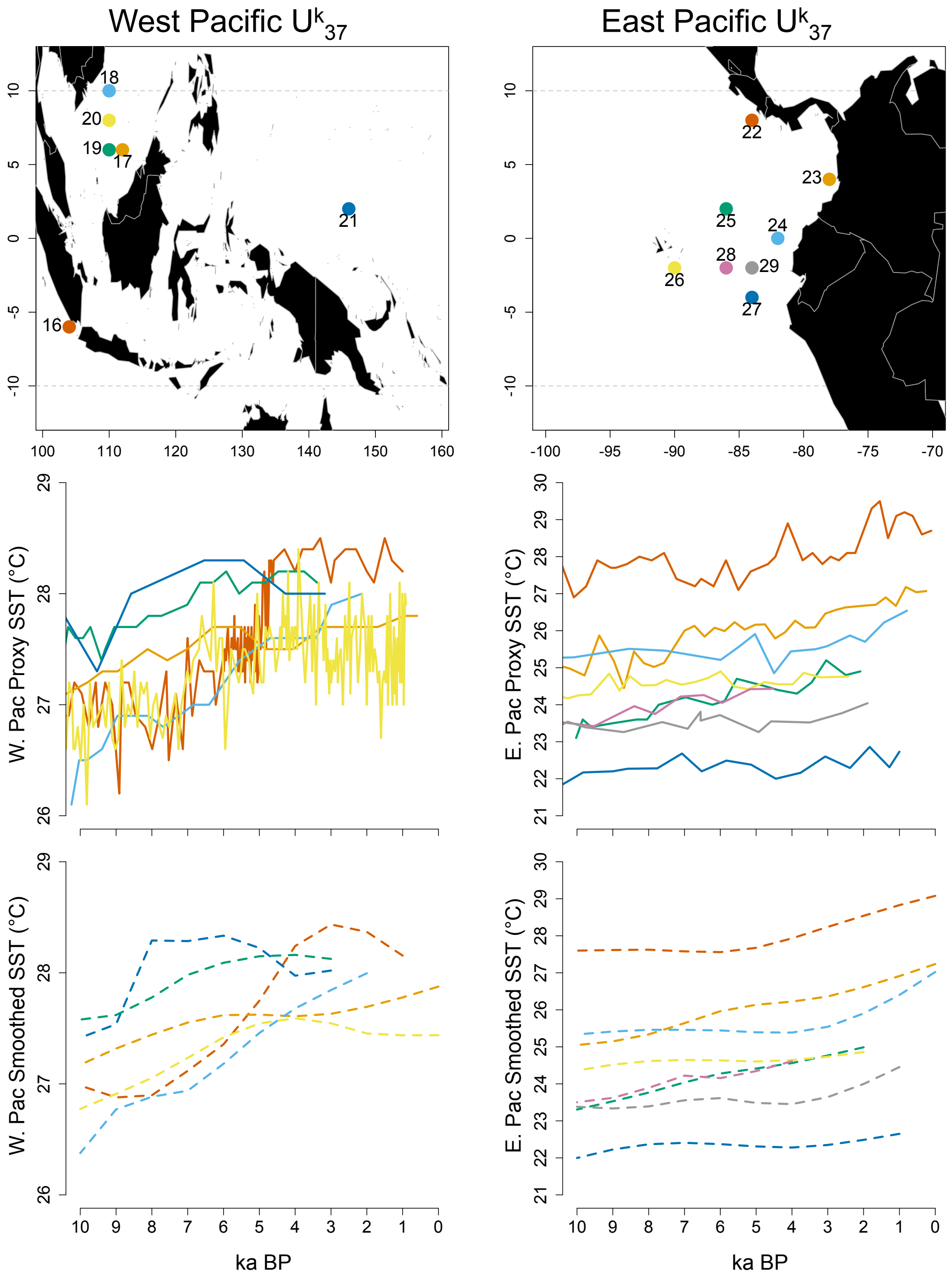


Figure 2. Maps of western and eastern U_{37}^{k} based SST records. The second row provides raw SST time series as calibrated by the authors listed in Table 1. The third row provides the records smoothes by a second order local polynomial method with a local neighborhood of 70% nearest data points.

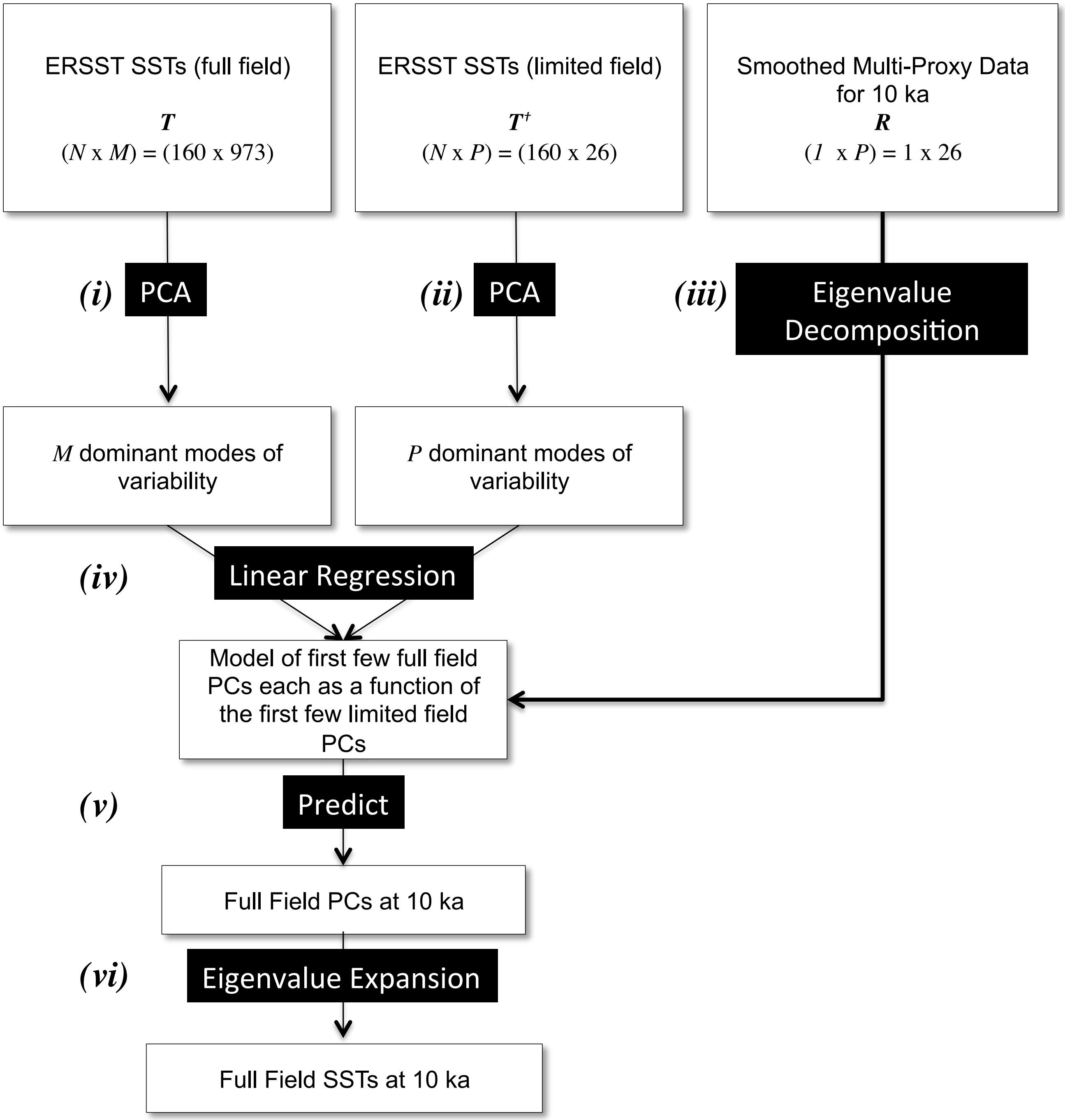


Figure 3. Flow chart of principal component analysis approach used for the multi-proxy SST reconstruction. The steps are given for producing SSTs representative of 10 ka. Steps $(iii - vi)$ are repeated for 8, 6, 4, and 2 ka.

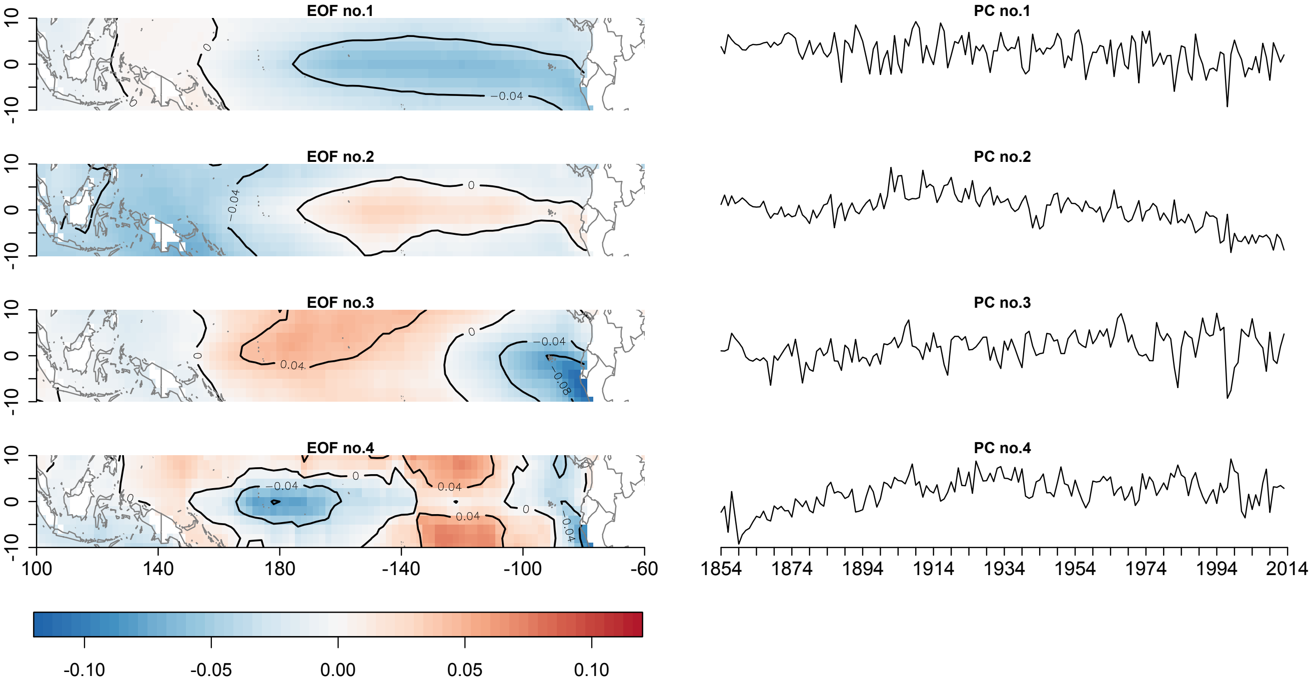


Figure 4. This figure provides the EOFs and PCs of the four leading modes of the PCA performed on the full field of equatorial Pacific annually averaged (May to April) SSTs obtained from the 1854-2013 2° x 2° ERSST dataset [*Smith et al.*, 2008]. The first four eigenvectors account for 68.8%, 11.1%, 7.1%, and 2.5%, respectively, of the total variability in the full field SST data.

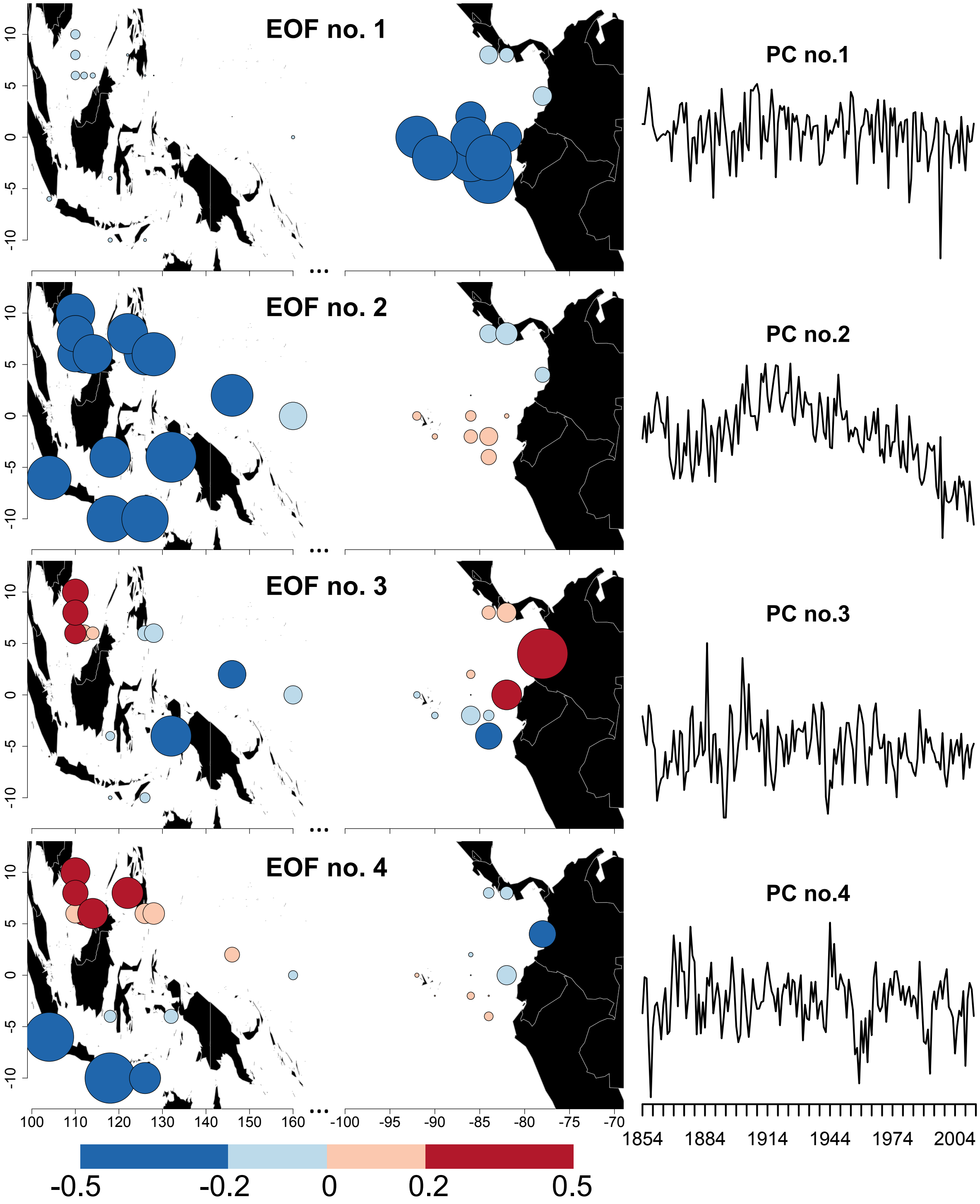


Figure 5. This figure is the same as Figure 4, but for the limited field of equatorial Pacific annually averaged (May to April) SSTs obtained from the 1854-2013 2° x 2° ERSST dataset [*Smith et al.*, 2008]. The first four eigenvectors account for 78.2%, 12.6%, 2.2%, and 1.7%, respectively, of the total variability in the limited field SST data. The radius of each circle is scaled to the magnitude of the eigenvalue.

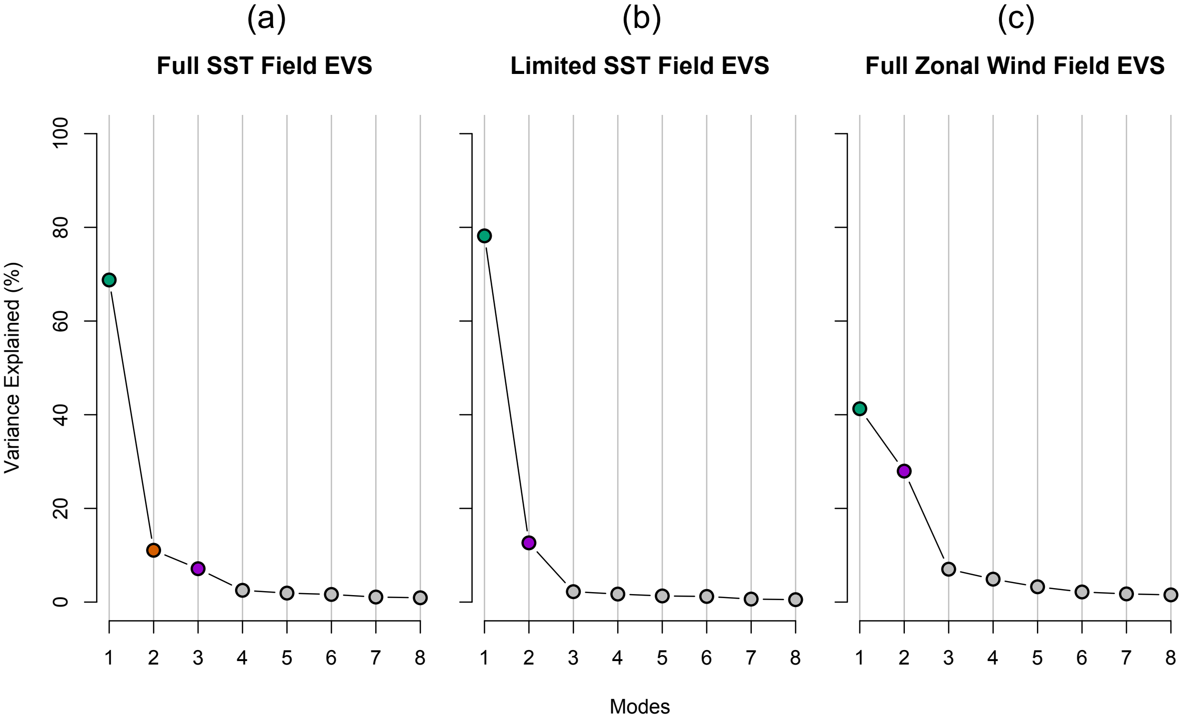


Figure 6. Eigenvalue spectrums (EVS) for the (a) full SST field (1854-2013), (b) limited SST field (1854-2013), and (c) full zonal wind field (1949-2013). The limited SST field consists of contemporary SSTs only at locations for which there exists either Mg/Ca or alkenone SST records. The green points represent the variance explained by the first mode of each field. The purple dots represent the variance explained at the knee, which is the point just before the noise floor (represented by the grey dots). The orange dot on the full field EVS represents the mode that would be included in the model if the threshold for variance resolved was at least 10%. Description of how these are used to determine the number of retained PCs and predictors used in the model is provided in the text under Section 3.1.

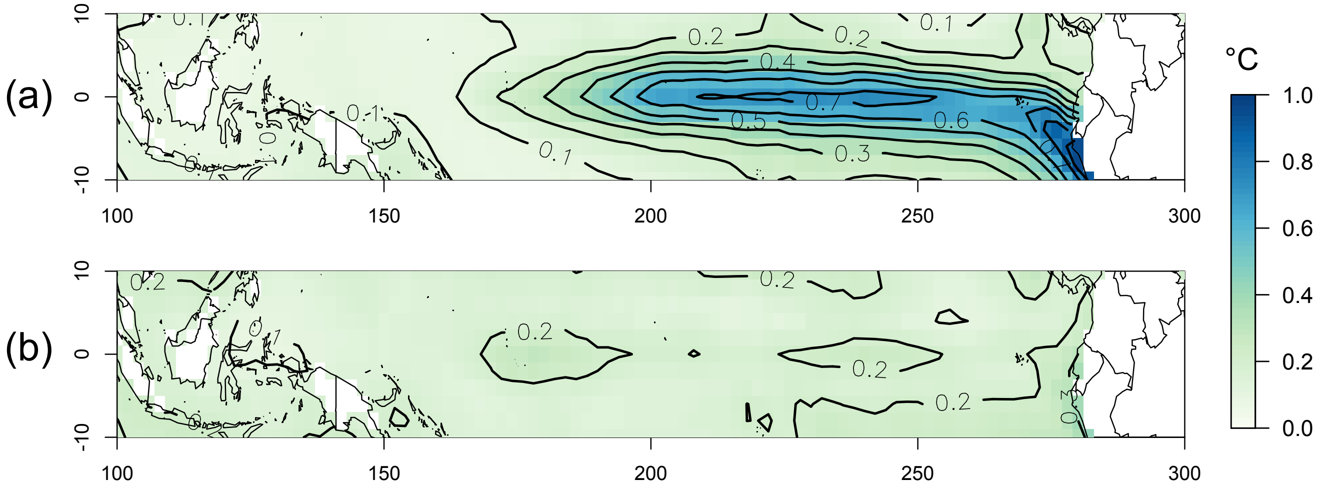


Figure 7. (a) Variance of 1854-2013 observed SSTs and (b) standard error SST map for 2 ka. The standard error map was produced by using the standard errors from the model regression and generating 500 ensembles of the first three PCs. The remaining PCs were bootstrapped from the original PCs to complete ensembles for each grid-point. This map is similar to those for other time periods reconstructed.

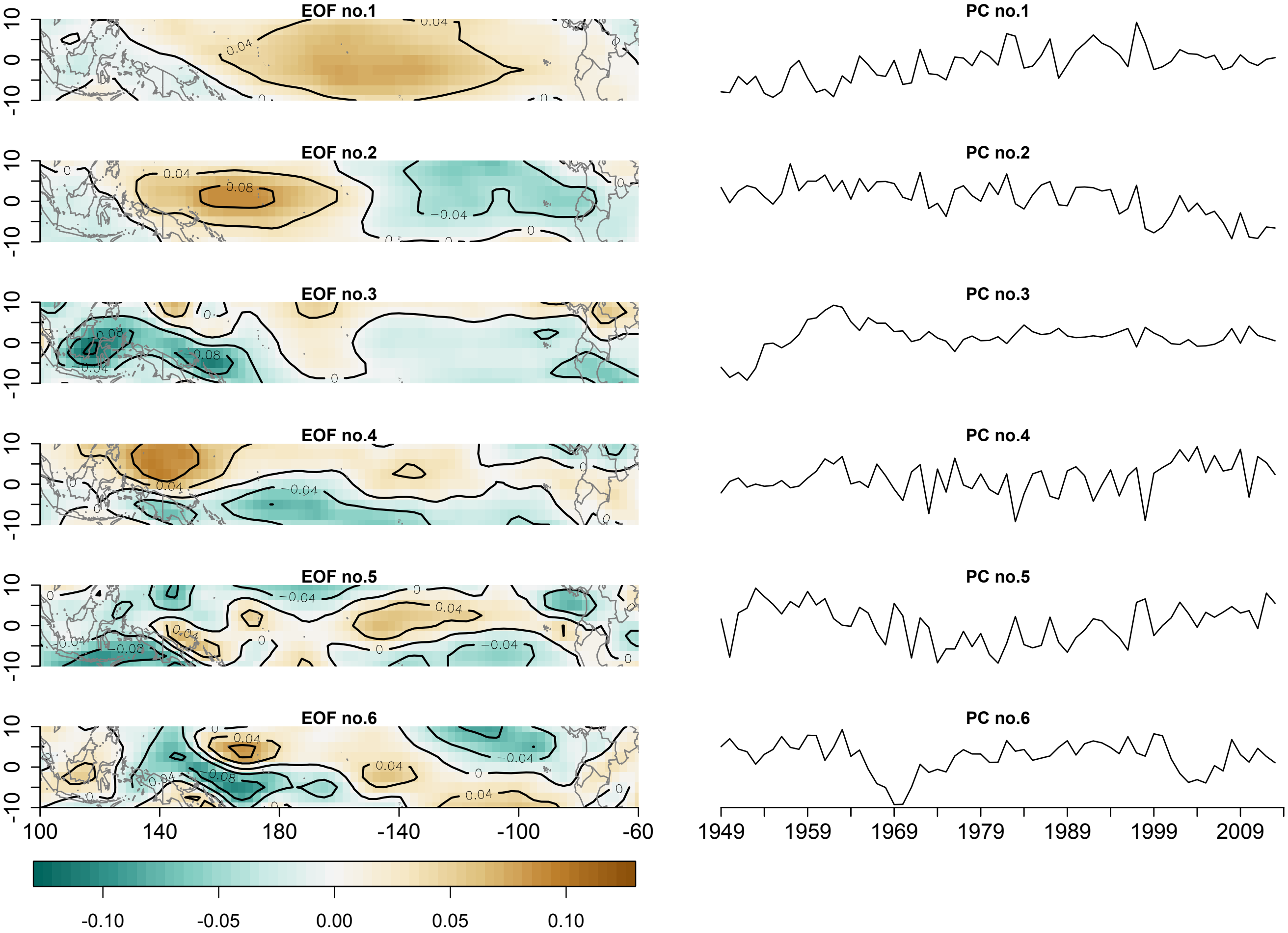


Figure 8. EOFs and PCs of the six leading modes of the PCA performed on the full field of equatorial Pacific annually averaged (May to April) zonal winds obtained from the 1949-2013 2.5° x 2.5° NCEP-NCAR Reanalysis [*Kalnay et al.*, 1996]. The first four eigenvectors account for 41.3%, 28.0%, 7.0%, and 4.9%, respectively, of the total variability in the full field zonal wind data.

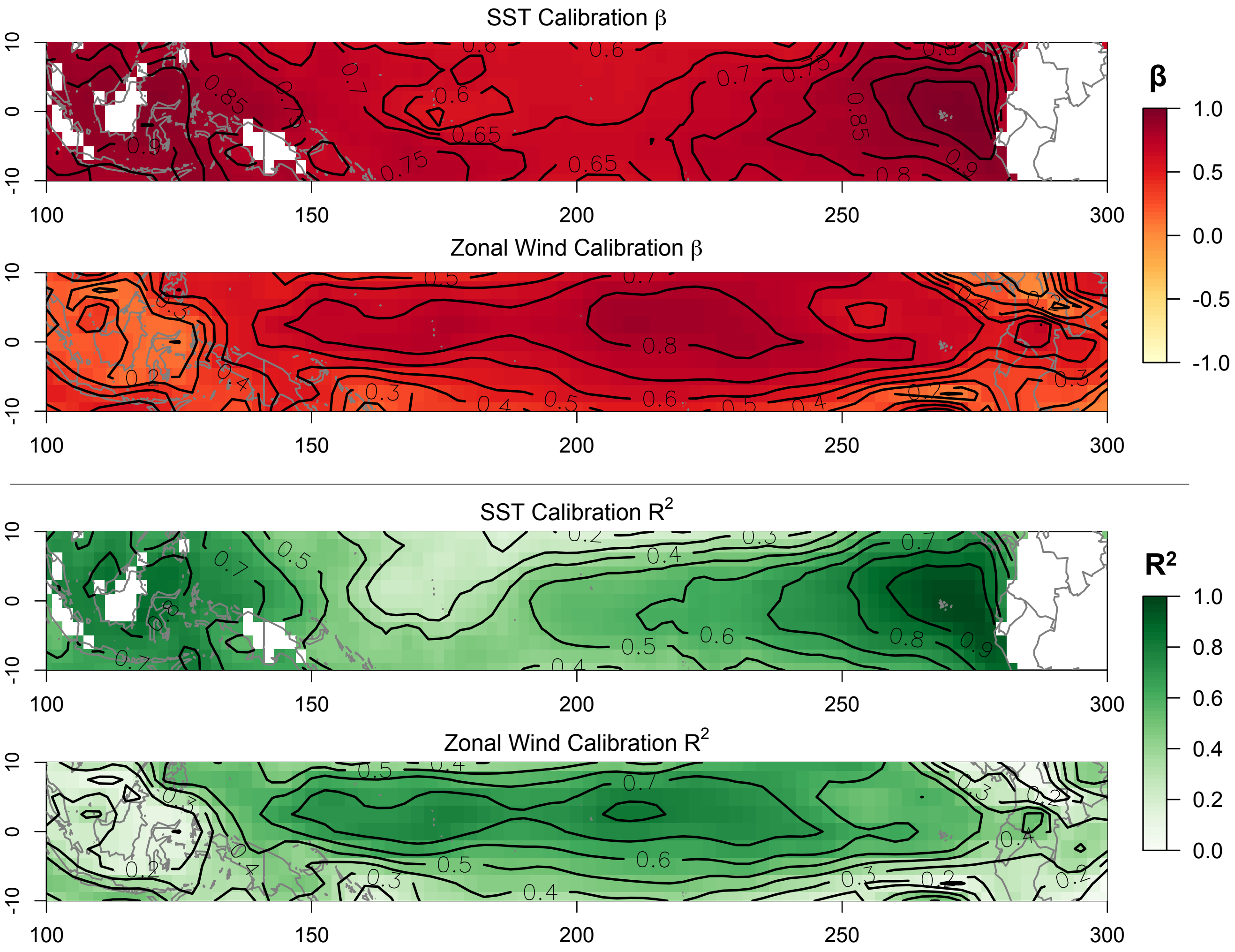
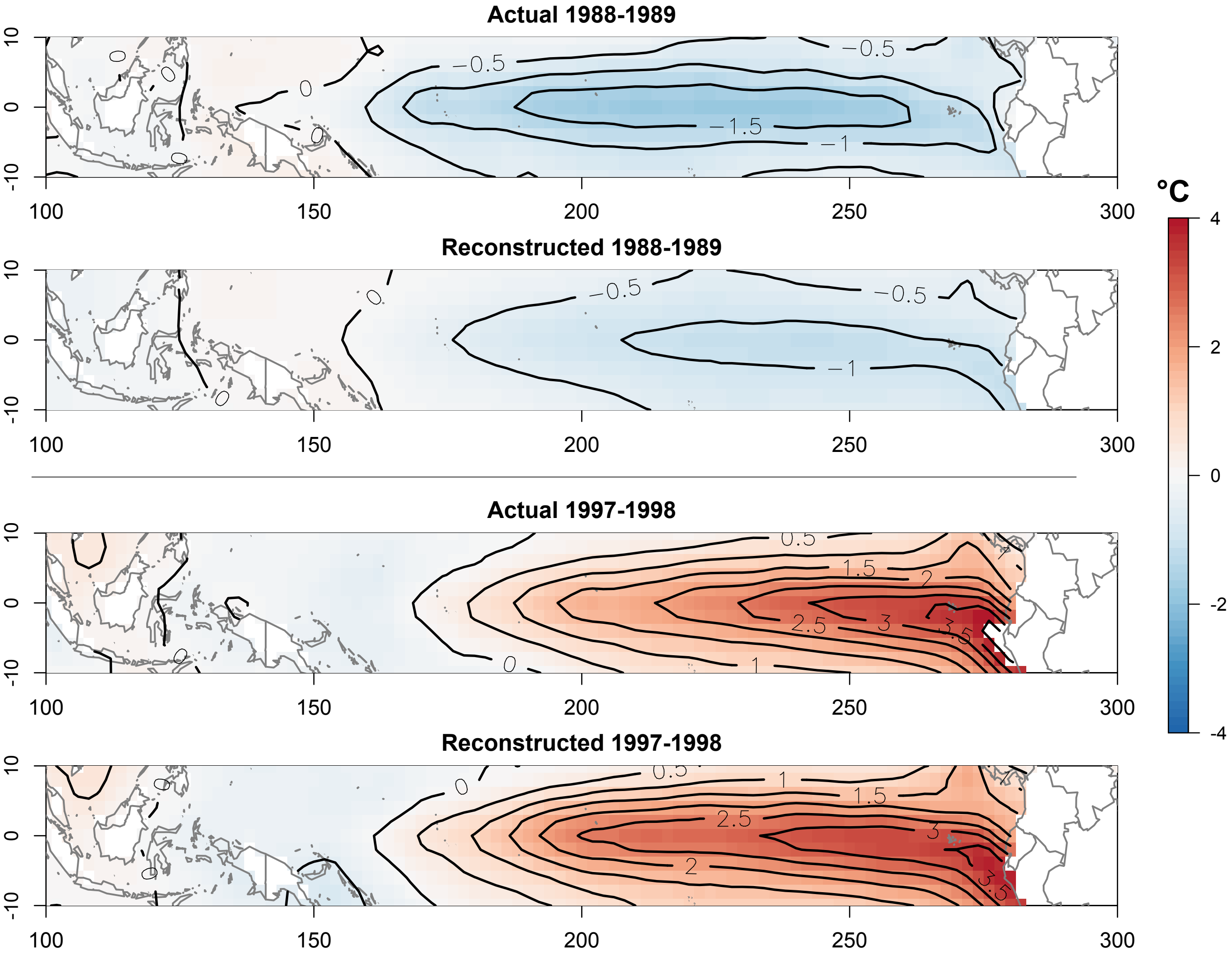


Figure 9. Model calibration statistics for the PCA-based SST model and the CCA-based zonal wind model showing the skill of the models in reconstructing each contemporary dataset (1854-2013 for SSTs and 1949-2013 for zonal winds). The β-statistic represents the resolved variance captured by the reconstructed contemporary data. The R^2 provides the correlation between the observed contemporary data and the reconstructed contemporary data.

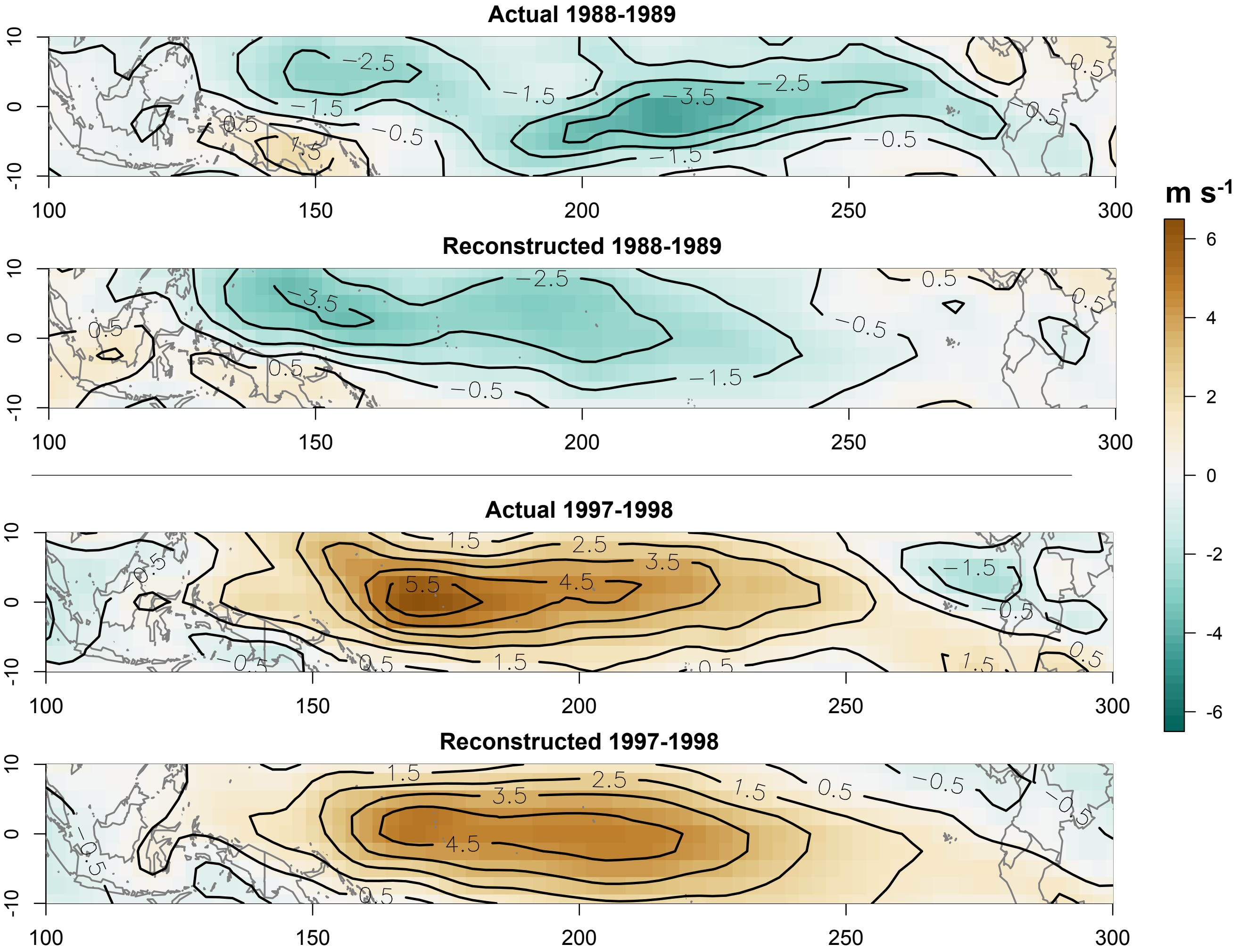

Actual 1988-1989

Reconstructed 1988-1989

Actual 1997-1998

Reconstructed 1997-1998

Figure 10. Actual and reconstructed SST anomalies (°C) for 1988-1989 (a strong La Niña year) and 1997-1998 (a strong El Niño year).



Actual 1988-1989

Reconstructed 1988-1989

Actual 1997-1998

Reconstructed 1997-1998

Figure 11. Actual and reconstructed zonal wind anomalies (m s^{-1}) for 1988-1989 (a strong La Niña year) and 1997-1998 (a strong El Niño year).

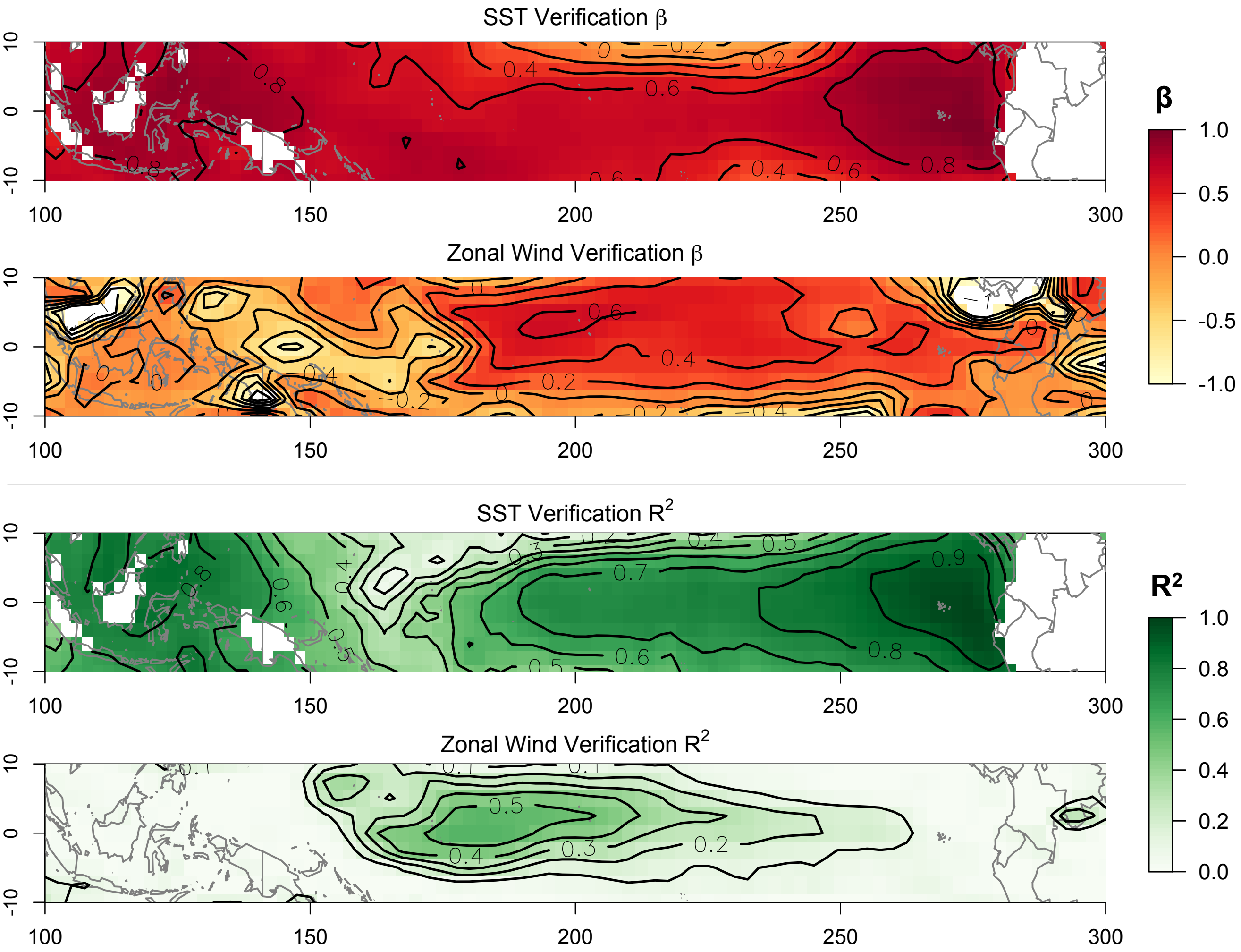


Figure 12. Model verification performed by training the model on 1980-2013 data, and using that model to validate SSTs and zonal winds from the period prior (1854-1979 for SSTs and 1949-1979 for zonal winds). The β- and R^2-statistics are used once again to quantify model skill.

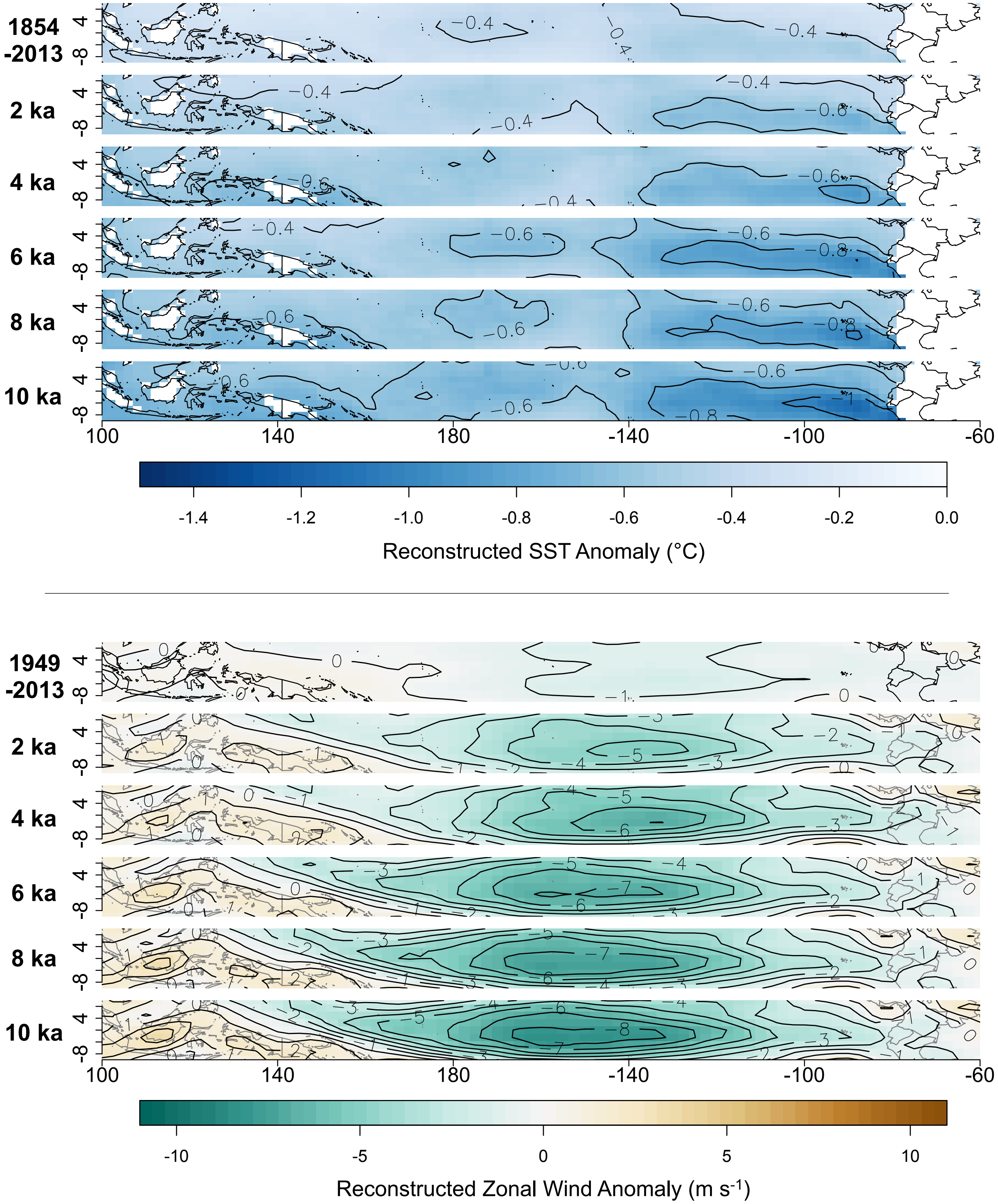


Figure 13. Multi-proxy reconstructed SSTs and zonal winds for 10, 8, 6, 4, and 2 ka. The top panel of each provides a reconstruction of present-day SSTs (1854-2013) and zonal winds (1949-2013) using the limited-field data.

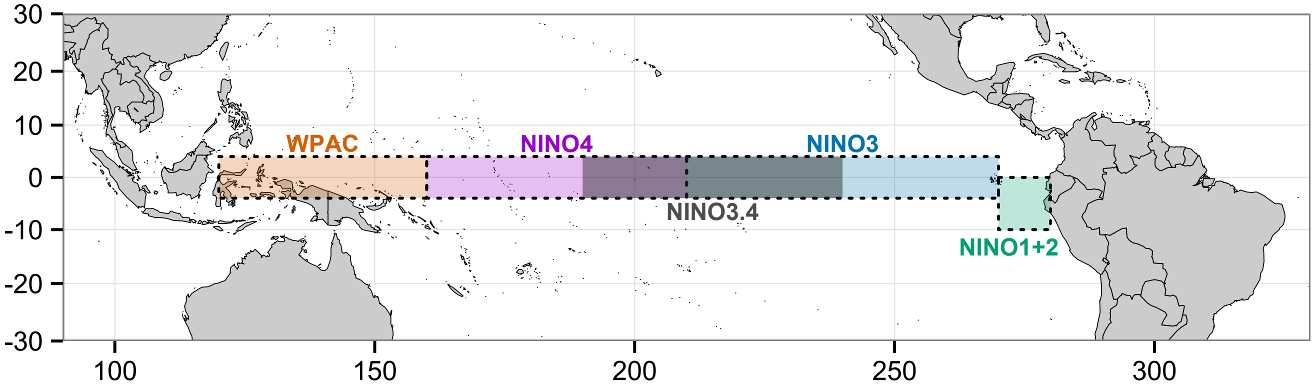


Figure 14. Regions used for the various ENSO indices reported in Table 3.

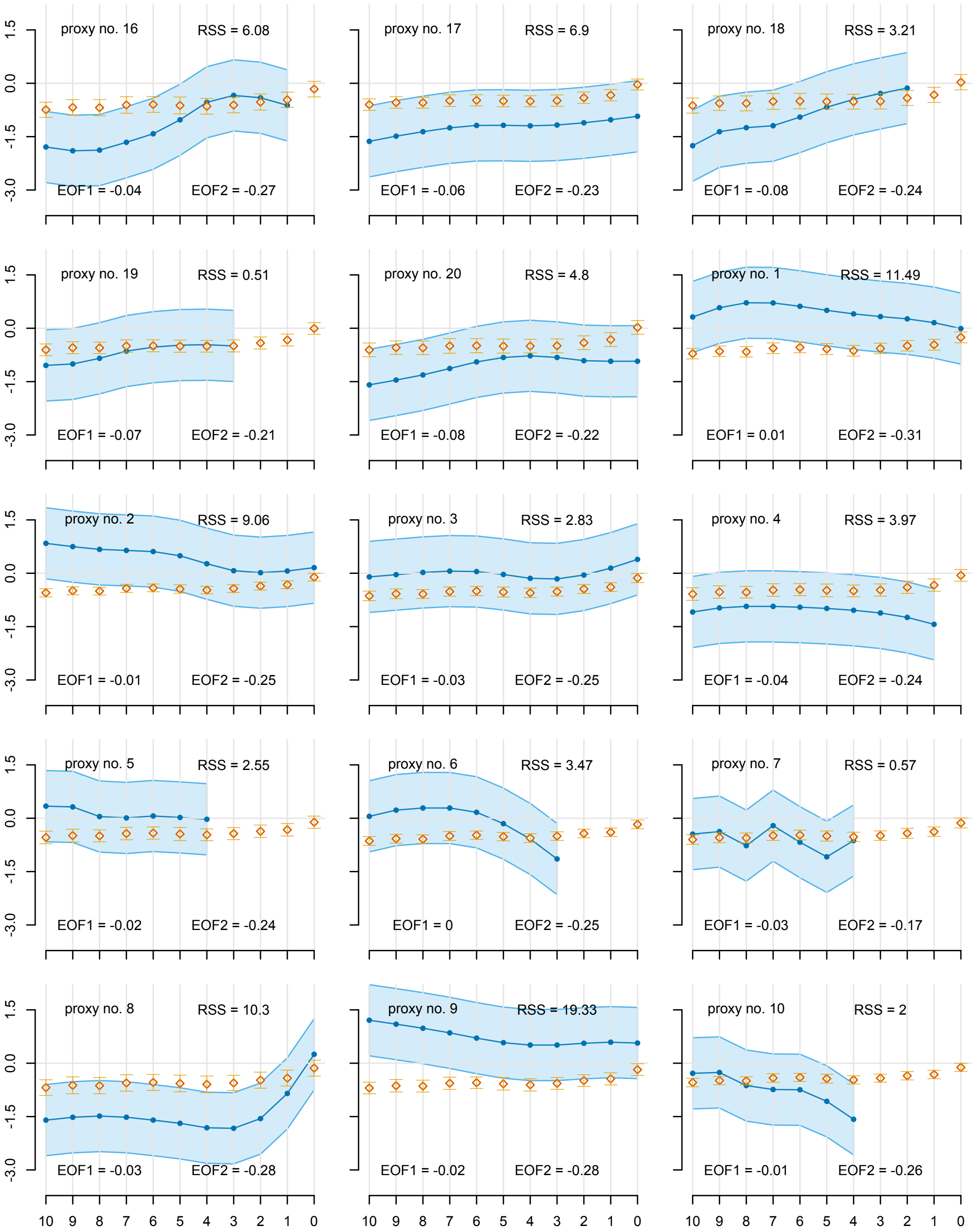


Figure 15. Scatterplots the each original proxy record from the western Pacific (blue) along with the reconstructed SST values for the grid-point nearest each record (orange). A nominal error value of ±1°C is shaded in light blue around each proxy record. Standard errors from the reconstruction model are plotted as orange whiskers. Proxy numbers in the upper left correspond to the numbers in Tables 1 and 2 and Figures 1 and 2. The residual sum of squares is provided to quantify how closely the reconstructed SSTs match the proxy SSTs. Recall that the first two PCs of the limited field were used as predictors for the PCA-based reconstruction model. The first two eigenvalues are written on the lower part of each scatterplot.

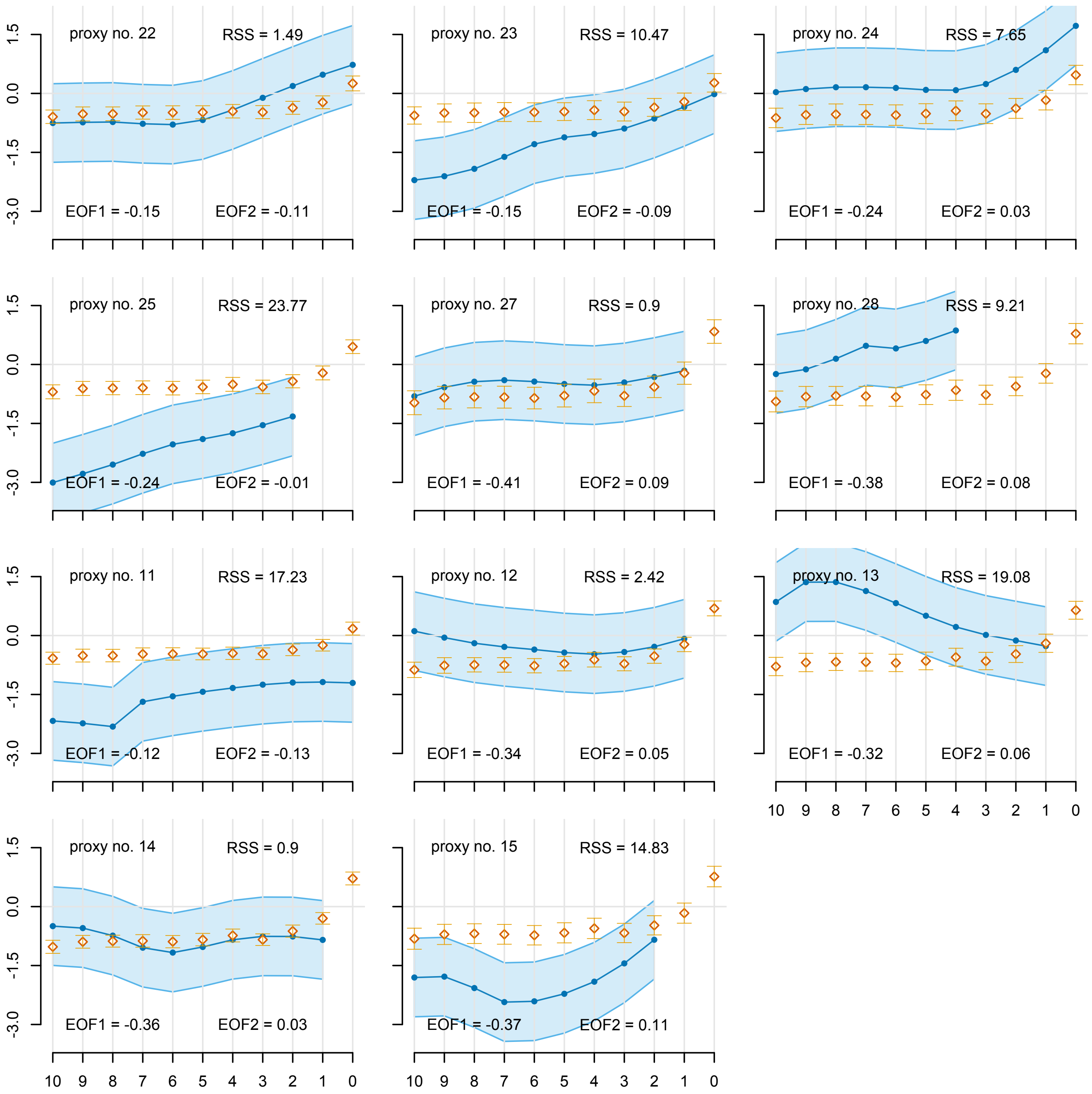


Figure 16. Same as Figure 15 but for eastern Pacific records.

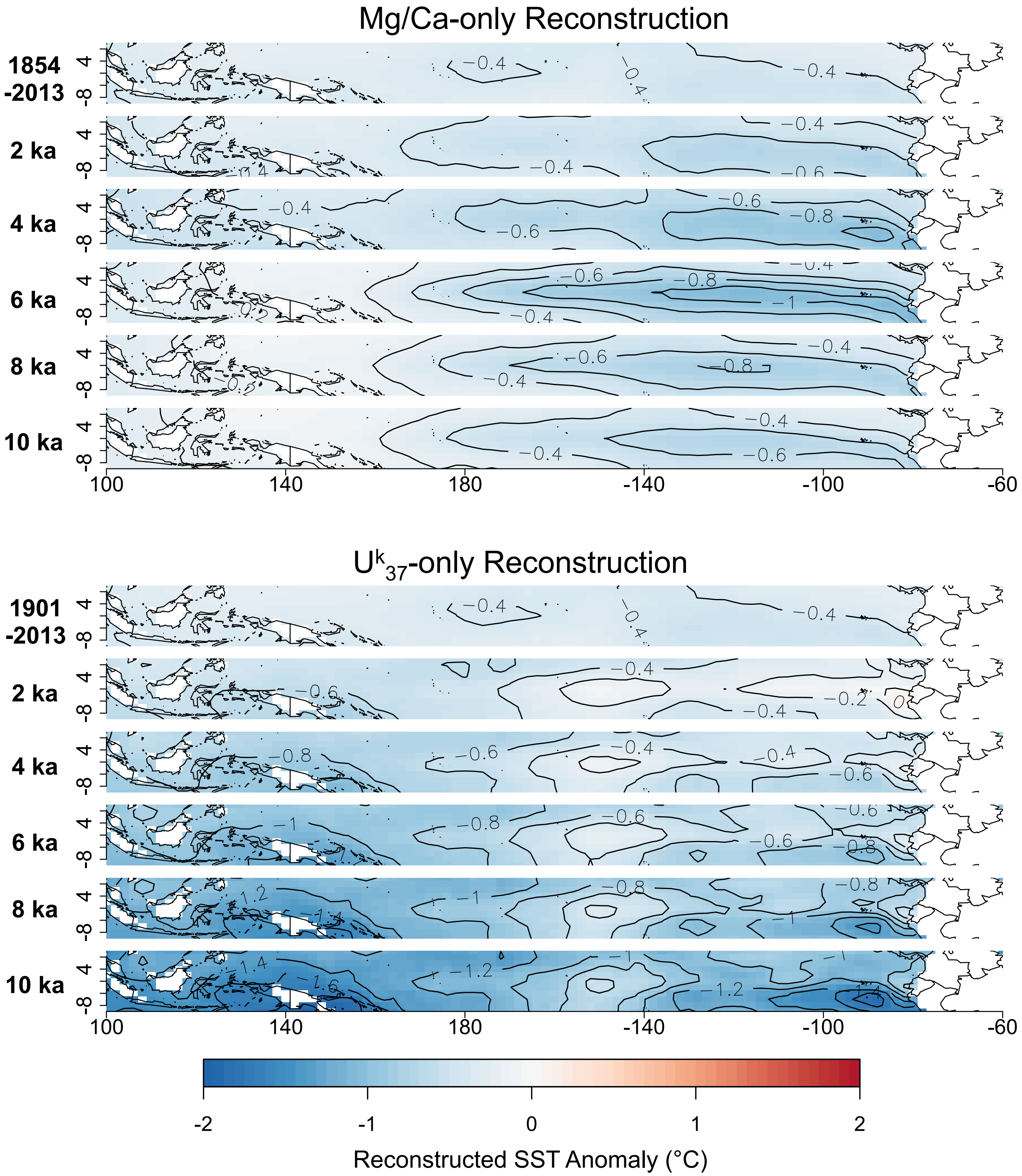


Figure 17. Single proxy reconstructions of 10, 8, 6, 4, and 2 ka SSTs using only Mg/Ca proxies (top) and U^k_{37} proxies (bottom). The top panel of each provides a reconstruction of present-day (1854-2013) SSTs using the limited-field data.

Table 1. A list of the Mg/Ca-based proxy records from the western and eastern equatorial Pacific used in the SST and zonal wind reconstruction.

Record #	**Core**	**Lat**	**Lon**	**Calibration**	**Reference**
Western Equatorial Pacific (Figure 1, left)					
1	MD98 - 2176	-5.00	133.45	*Nürnberg et al.* [1996]	*Stott et al.* [2004]
2	MD98 - 2181	6.30	125.83	*Nürnberg et al.* [1996]	*Stott et al.* [2004]
3	MD98 - 2162	-4.69	117.90	*Hastings et al.* [2001]	*Visser et al.* [2003]
4	MD01 - 2390	6.64	113.41	*Dekens et al.* [2002]	*Steinke et al.* [2008]
5	MD97 - 2141	8.80	121.30	*Rosenthal and Lohmann* [2002]	*Rosenthal et al.* [2003]
6	MD97 - 2138 *	1.25	146.14	*de Garidel-Thoron et al.* [2005]	*de Garidel-Thoron et al.* [2007]
7	ODP - 806b	0.32	159.35	*Lea and Martin* [1996]	*Lea et al.* [2000]
8	MD98 - 2165	-9.65	118.34	*Dekens et al.* [2002]	*Levi et al.* [2007]
9	MD98 - 2170	-10.59	125.39	*Nürnberg et al.* [1996]	*Stott et al.* [2004]
10	MD06 - 3067	6.51	126.50	*Anand et al.* [2003]	*Bolliet et al.* [2011]
Eastern Equatorial Pacific (Figure 1, right)					
11	ODP - 1242	7.86	-83.61	*Anand et al.* [2003]	*Benway et al.* [2006]
12	TR163 - 22	0.52	-92.40	*Dekens et al.* [2002]	*Lea et al.* [2006]
13	ODP - 1240	0.02	-86.45	*Dekens et al.* [2002]	*Pena et al.* [2008]
14	V21 - 30 *	-1.22	-89.68	*Dekens et al.* [2002]	*Koutavas et al.* [2006]
15	V19 - 28 *	-2.51	-84.65	*Dekens et al.* [2002]	*Koutavas et al.* [2006]

Table 2. A list of the U^{k}_{37}-based proxy records from the western and eastern equatorial Pacific used in the SST and zonal wind reconstruction.

Record #	**Core**	**Lat**	**Lon**	**Calibration**	**Reference**
Western Equatorial Pacific (Figure 2, left)					
16	SO139 - 74KL	-6.54	103.83	*Conte et al.* [2006]	*Lückge et al.* [2009]
17	GIK 17964	6.16	112.21	*Müller et al.* [1998]	*Pelejero et al.* [1999]
18	GIK 18252-3	9.23	109.38	*Pelejero and Grimalt* [1997]	*Kienast et al.* [2001]
19	GIK 18287-3	5.65	110.65	*Pelejero and Grimalt* [1997]	*Kienast et al.* [2001]
20	MD97 - 2151	8.72	109.87	*Pelejero and Grimalt* [1997]	*Zhao et al.* [2006]
21	MD97 - 2138 *	1.25	146.14	*Prahl et al.* [1988]	*de Garidel-Thoron et al.* [2007]
Eastern Equatorial Pacific (Figure 2, right)					
22	MD02 - 2529	8.21	-84.12	*Sonzogni et al.* [1997]	*Leduc et al.* [2007]
23	KNR176 - JPC32	4.85	-77.96	*Prahl et al.* [1988]	*Pahnke et al.* [2007]
24	V19 - 27	-0.47	-82.67	*Prahl et al.* [1988]	*Koutavas and Sachs* [2008]
25	ME0005A - 24JC	1.50	-89.68	*Prahl et al.* [1988]	*Kienast et al.* [2006]
26	V21 - 30 *	-1.22	-89.68	*Prahl et al.* [1988]	*Koutavas and Sachs* [2008]
27	V19 - 30	-3.38	-83.52	*Prahl et al.* [1988]	*Koutavas and Sachs* [2008]
28	RC11 - 238	-1.52	-85.82	*Prahl et al.* [1988]	*Koutavas and Sachs* [2008]
29	V19 - 28 *	-2.51	-84.65	*Prahl et al.* [1988]	*Koutavas and Sachs* [2008]

Table 3. Reconstructed SST anomalies for various NINO indices.

ka BP	Single Indices					Trans-Pacific Indices	
	WPAC	**NINO4**	**NINO3.4**	**NINO3**	**NINO 1+2**	**TNI**	**WTNI**
a. Multi-proxy Reconstruction							
10	-0.64 ± 0.13	-0.65 ± 0.18	-0.72 ± 0.17	-0.81 ± 0.17	-0.97 ± 0.27	-0.33 ± 0.33	-0.33 ± 0.30
8	-0.59 ± 0.13	-0.58 ± 0.18	-0.63 ± 0.18	-0.71 ± 0.17	-0.84 ± 0.26	-0.26 ± 0.32	-0.25 ± 0.29
6	-0.48 ± 0.13	-0.57 ± 0.18	-0.67 ± 0.18	-0.74 ± 0.17	-0.84 ± 0.26	-0.27 ± 0.32	-0.36 ± 0.29
4	-0.56 ± 0.12	-0.52 ± 0.18	-0.54 ± 0.17	-0.60 ± 0.17	-0.71 ± 0.26	-0.18 ± 0.31	-0.15 ± 0.28
2	-0.44 ± 0.13	-0.46 ± 0.18	-0.50 ± 0.18	-0.54 ± 0.17	-0.60 ± 0.27	-0.14 ± 0.33	-0.16 ± 0.30
b. Mg/Ca-only Reconstruction							
10	-0.12 ± 0.12	-0.41 ± 0.18	-0.58 ± 0.17	-0.61 ± 0.17	-0.57 ± 0.26	-0.16 ± 0.32	-0.44 ± 0.29
8	-0.14 ± 0.12	-0.45 ± 0.17	-0.65 ± 0.17	-0.69 ± 0.17	-0.67 ± 0.26	-0.22 ± 0.31	-0.52 ± 0.29
6	-0.16 ± 0.13	-0.57 ± 0.18	-0.84 ± 0.17	-0.91 ± 0.17	-0.93 ± 0.27	-0.37 ± 0.32	-0.77 ± 0.30
4	-0.46 ± 0.12	-0.59 ± 0.18	-0.73 ± 0.17	-0.81 ± 0.17	-0.91 ± 0.25	-0.31 ± 0.31	-0.45 ± 0.28
2	-0.34 ± 0.13	-0.47 ± 0.19	-0.57 ± 0.18	-0.62 ± 0.18	-0.65 ± 0.27	-0.18 ± 0.33	-0.31 ± 0.30
b. U^{k}_{37}*-only Reconstruction*							
10	-1.52 ± 0.15	-1.07 ± 0.22	-0.91 ± 0.24	-1.05 ± 0.23	-1.42 ± 0.31	-0.36 ± 0.38	+0.09 ± 0.35
8	-1.28 ± 0.14	-0.85 ± 0.20	-0.67 ± 0.22	-0.77 ± 0.22	-1.04 ± 0.29	-0.19 ± 0.35	+0.24 ± 0.32
6	-1.04 ± 0.14	-0.67 ± 0.20	-0.48 ± 0.20	-0.55 ± 0.20	-0.73 ± 0.28	-0.07 ± 0.33	+0.32 ± 0.29
4	-0.82 ± 0.13	-0.52 ± 0.19	-0.36 ± 0.18	-0.39 ± 0.18	-0.51 ± 0.26	+0.02 ± 0.33	+0.32 ± 0.29
2	-0.58 ± 0.13	-0.41 ± 0.19	-0.24 ± 0.18	-0.20 ± 0.18	-0.14 ± 0.27	+0.27 ± 0.33	+0.45 ± 0.30

[a] Single indices are calculated by averaging the mean SST anomaly over the following regions: WPAC (120°E to 160°E, 4°S to 4°N), NINO4 (160°E to 190°E, 4°S to 4°N), NINO3.4 (190°E to 240°E, 4°S to 4°N), NINO3 (210°E to 270°E, 4°S to 4°N), NINO1+2 (270°E to 280°E, 10°S to 0°). Note that typically, NINO4, NINO3.4, and NINO3 are averaged over 5°S to 5°N, but the ERSST dataset used in this study is gridded only for every 2°, and we therefore average over 4°S to 4°N.TNI is calculated as NINO1+2 minus NINO4. WTNI is calculated as NINO1+2 minus WPAC.