Australian Government
Bureau of Meteorology

The Centre for Australian Weather and Climate Research
A partnership between CSIRO and the Bureau of Meteorology

CSIRO

A new basin-wide multi-proxy index of the El-Nino Southern Oscillation, A.D. 1525-1982

CAWCR Technical Report No. 005

Karl Braganza, Joëlle L. Gergis, Scott B. Power, James S. Risbey and Anthony M. Fowler

December 2008

www.cawcr.gov.au

A new basin-wide multi-proxy index of the El-Nino Southern Oscillation, A.D. 1525-1982

CAWCR Technical Report No. 005

Karl Braganza[1], Joëlle L. Gergis[2], Scott B. Power[1], James S. Risbey[3] and Anthony M. Fowler[4]

December 2008

[1] Centre for Australian Weather and Climate Research, Bureau of Meteorology, Melbourne GPO Box 1289, Melbourne Victoria 3001, Australia.

[2] School of Earth Sciences, University of Melbourne, Victoria Melbourne VIC 3010, Australia.

[3] Centre for Australian Weather and Climate Research, CSIRO, GPO Box 1538 Hobart Tasmania 7001, Australia

[4] School of Geography and Environmental Science, University of Auckland, Private Bag 92019 Auckland, New Zealand.

Enquiries should be addressed to:

Dr Karl Braganza
National Climate Centre
Bureau of Meteorology
GPO Box 1289K, Melbourne
Victoria 3001, Australia
k.braganza@bom.gov.au

Phone: 61 3 9669 4344

ISSN: 1836-019X

ISBN: 978-1-921424-80-9

Series: Technical report (Centre for Australian Weather and Climate Research.) ; no. 5.
Also issued in print.

Copyright and Disclaimer

Contents

List of Tables

List of Figures

ABSTRACT

An understanding of past variability in the El Niño Southern Oscillation (ENSO) would be useful in assessing the significance of recent observed changes to ENSO and in determining the realism of coupled climate model simulations. While the instrumental data cover a period of 100 years or more, this record is insufficient for the assessment of changes in the magnitude, frequency and duration of ENSO variability on inter-decadal and longer timescales.

Using tree-ring, coral and ice-core data, we reconstruct a proxy-based ENSO index between A.D. 1525 and 1982. Unlike most previous studies, which have drawn climate proxies from limited geographic regions, our network is Pacific basin-wide, using ENSO sensitive proxies from the western equatorial Pacific, New Zealand, the central Pacific and sub-tropical North America. This network provides a more robust proxy ENSO signal.

The common signal recorded in the multi-proxy network has a high correlation with the Southern Oscillation Index (SOI), Niño 3.4 Sea Surface Temperature (Niño3.4 SST) and a combined ocean-atmosphere ENSO index (CEI). The proportion of instrumental variance explained is 48% for the SOI, 48% for Niño3.4 SST and 53% for the CEI. The proxy-ENSO index also displays skill in reproducing warm and cold extremes of the SOI.

Spectral analysis of the proxy-ENSO index over the last four hundred and fifty years shows considerable amplitude and frequency modulation in the three to ten year band on multi-decadal timescales. There is a relative reduction in the amplitude of high frequency variability during the sixteenth, early-seventeenth and mid-eighteenth centuries. In contrast, high frequency ENSO variability has increased over the last two hundred years. Variability during the first half of the twentieth century is similar to that evident in the nineteenth century.

1. INTRODUCTION

For many people, the most important impacts of climate change are related to changes in regional temperature and precipitation. However the contribution of anthropogenic climate change to regional climate change is clouded by the presence of naturally occurring variability. For rainfall and temperature, a major source of this variability on 2-7 year timescales is the El Niño-Southern Oscillation (ENSO) (Folland et al, 2001). A major difficulty in the task of separating the anthropogenic signal from the background of natural, ENSO related variability is the relative brevity of the instrumental record, which covers a period of one hundred years or less. Extended, multi-century reconstructions of ENSO can help us to assess how unusual 20th century variability in ENSO has been in a longer term context.

In comparison to reconstructions of global and hemispheric surface temperature, reconstructions of ENSO variability have received less attention (Trenberth and Hoar, 1996; Trenberth and Hoar, 1997; Crowley, 2000; Folland et al, 2001; Mann, 2003). Palaeo-reconstructions have generally relied on calibration techniques that relate ENSO sensitive variability in proxy indicators to an instrumental ENSO index, typically the Southern Oscillation Index (SOI) or Niño region sea surface temperature (SST). Stahle et al (1998) produced a reconstruction of the December to February SOI for the period A.D. 1706-1977 using subtropical North American tree-ring data. D'Arrigo et al, (2005) reconstructed December to February Niño 3 SST for A.D. 1525-1982 using an expanded and updated version of the Stahle et al (1998) data. Mann et al (2000) reconstructed October to March Niño3 SSTs for the period A.D. 1650-1980 using a restricted tropical subset of the Mann et al (1998) global multi-proxy data. A review of recent approaches to reconstructing ENSO from proxy records is provided by Gergis et al (2006).

Unlike surface temperature, which is a directly measurable quantity, ENSO is a coupled atmosphere-ocean climate phenomenon. How best to define and monitor ENSO is a question that has been posed for some time (Rasmusson and Carpenter, 1982; Allan et al, 1996, Trenberth, 1997). This question also presents difficulties for the reconstruction of ENSO from climate proxies. When employing a set of multiple, diverse proxy indicators, part of the difficulty in calibration arises from understanding what aspect of ENSO variability is being registered by the proxies. For example coral proxies will register climate variability as a function of sea surface temperature, while tree-ring segments may reflect variability in local rainfall, temperature or both. Of relevance to ENSO reconstructions calibrated to single (uncoupled) indices is the fact that ocean and atmosphere components of the system need not be synchronous in their timing or amplitude (for example see Fairbanks et al, 1997; Lyon and Barnston, 2005; Gergis and Fowler 2005; and elsewhere). Events that register strongly in some instrumental indices and proxies may be absent in others, with obvious implications for proxy calibration. For example, Gergis and Fowler (2005) demonstrated that the identification of ENSO events was dependent upon the index employed. The presence of noise in the instrumental indices, such as the Madden Julian Oscillation (MJO) in the SOI (Trenberth 1997), is also a confounding factor in the calibration of proxies.

A second major problem in ENSO reconstruction arises because there is no typical climate response to ENSO forcing (Allan et al, 1996, Power et al, 1998, van Oldenborgh, 2005). The strength of ENSO teleconnections is likely to vary in both strength and persistence, and from region to region, for each event. In addition, the nature of the atmospheric and oceanic response to ENSO has been shown to be sign (warm or cold phase) dependent in some locations (Hoerling et al, 1997, Power et al, 2006). Hence, changes in ENSO climate teleconnections from event to event are also likely to confound reconstruction attempts, particularly if proxy data are restricted to isolated (single) regions. To date, the coverage of proxy data has been limited to predominately eastern Pacific teleconnection regions with limited representation of sites influenced by western Pacific teleconnections (Stahle et al, 1998; D' Arrigo et al, 2005). A number of studies have suggested that a reconstruction of ENSO derived from a number of widely-spaced regional proxies, with sufficient east and west Pacific representation, is more likely to be representative of ENSO than more restricted networks (Baumgartner et al, 1989; D'Arrigo et al, 1994; Diaz and Pulwarty, 1994; Whetton and Rutherford, 1994; Gedalof and Mantua, 2002).

In this paper we use a set of multi-proxy indicators from locations that span a broader area of the Pacific basin than has been attempted previously. In this manner, we aim to represent climate signals from a range of different regions (tropics versus extra tropics, terrestrial versus marine environments) that have proven responses to ENSO forcing. Specifically, the incorporation of new western Pacific proxies are introduced, including two high quality tree-ring records from New Zealand (Fowler et al, 2000; Fenwick, 2003), and a well replicated Australian coral record from the Great Barrier Reef (Hendy et al, 2002; Hendy et al, 2003). These records provide an important counterpart to the chronologies developed from Eastern Pacific locations.

We use the Pacific wide network of ENSO sensitive proxies to identify a common, annually resolved ENSO signal, which we employ as a multi-century index of ENSO. Whereas previous studies have used calibration techniques to reconstruct the SOI or Niño region SST anomalies, our ENSO index is determined solely by ENSO related co-variability (the ENSO mode) in the network. By not calibrating this index against 20th century observations, we expect greater homogeneity over the length of the record, since there is less likelihood that the index is tuned to the signal of 20th century variability. We compare this proxy-based index to the SOI, Niño3.4 SST and the coupled ocean-atmosphere ENSO index (CEI) of Gergis and Fowler (2005).

The various proxy records used in our network are described in section 2. In section 3, we outline the use of empirical orthogonal function (EOF) analysis to identify a proxy-based ENSO index. In detailing our method, we attempt to address two questions that arise in relation to previous attempts at ENSO index reconstruction from proxies. The first question relates to the validity of our approach compared with direct calibration methods. We use synthetic data to test the ability of our EOF method to extract known climate signals embedded in a set of proxies that include signal plus noise. The second question relates to the filtering of proxy data. Since a number of filtering methods exist in the literature, we test the sensitivity of our method, and the spectra of our ENSO index, to various different filtering methods that have been attempted previously. In section 4, we compare the proxy-based ENSO index to instrumental indices over the 20th century. We also investigate the ability of our index to capture low-frequency (decadal and longer) variability through comparison with the Interdecadal Pacific

Oscillation (IPO) and the low frequency component of the SOI. Finally, in section 5, we provide a brief discussion of ENSO variability since A.D. 1525.

2. DATA

In this section, we describe the data records selected for the Pacific basin-wide, multi-proxy network that we use to develop our historic ENSO index. We also briefly describe the three instrumental indices (the SOI, Niño3.4 SST and the CEI) to which we compare this index over the 20th century.

2.1 Proxy data and their preparation

A number of criteria are used for the selection of proxy records used in this study:

1. Each record is a seasonally or annually resolved published record from areas that have been identified by previous researchers as containing an ENSO signal (e.g. Stahle et al, 1998, Fowler et al, 2003; Hendy et al, 2003). Coral records were limited to those identified by Lough (2004) as showing significant correlation with ENSO over the 20th century.

2. Proxies are selected so that we maintain a network that spans the Pacific Basin (including eastern, western and central Pacific sites) and to ensure that individual regions are not over-represented.

3. Each proxy in the network must have a continuous record that is free of documented inhomogeneities.

By maximising criteria 2 and 3 we define two multi-proxy networks. The largest network contains eight proxy records (Figure 1) from A.D.1727 to 1982. The other network is a subset of this network that contains five continuous records from A.D. 1525 to 1982.

The five longest records (A.D. 1525 to 1982) are the Kauri and Pink Pine tree-ring chronologies from New Zealand (western Pacific, Southern Hemisphere), the Douglas Fir and Pinyon Pine (Stahle et al, 1998; Cleaveland et al, 2003) tree-ring sequences from sub-tropical North America (eastern Pacific, Northern Hemisphere) and the Quelccaya ice-core (Thompson et al, 1984; Thompson et al, 1985; Thompson et al, 2000) from Peru (Eastern Pacific, Equator).

The larger network (A.D. 1727 to 1982) contains the five records described above and three additional coral records. These records are the Great Barrier Reef and New Caledonia (Quinn et al, 1998) coral records (western Pacific) and the Rarotonga (Linsley et al, 2000) coral record (central Pacific).

All tree ring sequences used in this study are regional ‘master chronologies’ that represent the mean of multiple core samples from multiple trees. For each species, these were produced using the method of Fowler and Boswijk (2003). Tree-ring sequences are usually high-pass filtered to remove the low frequency signals of biological origin (ie. not related to climate variability). Since such filtering is somewhat arbitrary, we seek to employ tree-ring sequences with little or no filtering, with the aim of retaining as much low-frequency climate information as possible. In practice (ensuring all records were subjected to the same data treatment), only tree-ring sequences filtered using a high-pass, 200 year cubic spline filter are available for this investigation. The preservation of variability on timescales less than 200 years from records with maximum length of 400 years should be sufficient for capturing any low frequency ENSO variability. All coral and ice-core records were filtered using a three-year low-pass Gaussian filter, which attenuates year-to-year noise in the individual timeseries. This approach is consistent with previous studies (Diaz et al, 2000, Goodwin et al, 2004, Linsley et al, 2004; Lough, 2004) that use similar records to infer inter-annual and longer-scale climate variability.

With the exception of the Kauri, Pink Pine and the Great Barrier Reef records, all data has been obtained from the National Oceanic and Atmospheric Administration’s (NOAA) World Data

Center-A for Paleoclimatology (details in Table 1). The Kauri, Pink Pine and Great Barrier Reef chronologies are sourced via individual authors (Fenwick, 2003; Hendy et al, 2003; Fowler et al, 2004; Gergis et al, 2005b; 2005a).

The limitations and/or potential biases that are specific to each type of proxy are well understood (Jones and Mann, 2004). The temporal resolving power (ie. seasonal, annual, decadal etc) is a key consideration in the reconstruction of ENSO variability. Proxy records commonly resolve an annual signal that is registered during one particular growth season. On interannual timescales, ENSO proxies rarely capture more than 50% of instrumental variance. Furthermore, they are unable to register variance equally well across a number of frequency domains (Bradley, 1996). For instance, the ability of tree-ring climate proxies to faithfully capture low frequency climate variability is potentially restricted due to physical limitations related to segment length. This issue is extensively discussed elsewhere (Esper et al, 2002; Hughes, 2003; Cook et al, 2004). To address this, we compare our proxy-based ENSO index to interannual instrumental ENSO, as well as the low frequency component of ENSO, in section 4.

2.2 Instrumental ENSO indices

To verify the usefulness of our proxy-based ENSO index, we use the SOI and Niño3.4 SST and the Gergis and Fowler (2005) Coupled ENSO Index (CEI) index to represent ENSO over the instrumental period.

The SOI time series (1871-2003) from the Australian Bureau of Meteorology is based on the Troup method (Troup 1965). This index is the standardised anomaly of the mean sea level pressure difference between Tahiti and Darwin, presented in standard deviation units around a mean of zero, with significant positive/negative departures representing La Niña/El Niño conditions.

Raw data for the Niño 3.4 SST region (5°N-5°S, 120°W-170°W), for the period 1950-2003, has been obtained from the NOAA's Climate Prediction Center. Pre-1950, we use the Trenberth and Stepaniak interpolation-based reconstruction of Niño3.4 SSTs, which is derived from the 1-degree gridded HadISST (Hadley Centre Sea Ice and Sea Surface Temperature) dataset. It should be noted that spatial variability of SSTs prior to 1950 suffers from data quality issues in all existing maritime observational data sets (for example, see Power and Colman 2006).

The CEI is a composite index based on the oceanic (Niño3.4 SST) and atmospheric (SOI) indices described above. Anomalies expressed in either N3.4 SST or the SOI alone (and therefore perhaps indicative of decoupled or out of phase changes) are maintained in the CEI, while fully coupled ocean-atmospheric anomalies result in an amplification of the index.

3. METHOD

In this section we describe a simple method for extracting a common ENSO signal from our multi-proxy network using Empirical Orthogonal Function (EOF) analysis. This approach is compared with various regression based calibration approaches using synthetic data. We demonstrate that the EOF method of signal extraction provides a simple means of removing non-ENSO related noise from proxy records.

3.1 Extracting the ENSO signal from a multi-proxy network

A large proportion of the total variability, on any given timescale, in individual palaeo climate records is due to non climate factors. For example, biological processes account for a significant proportion of lower-order autocorrelation in tree ring sequences. Extracting the climate signal from this background variability (noise) in a single palaeo-climate record is a difficult task. Previous studies have taken a similar conceptual approach to developing ‘transfer functions’ that relate proxy records to ENSO. Almost all have sought to ‘calibrate’ individual proxy records, or multi-proxy networks, using multivariate regression against instrumental observations.

For instance Stahle et al (1998) use a so-called ‘PC-based regression’, whereby leading modes (EOFs) from a multi-proxy data set are regressed against the instrumental SOI over the 20th century. This regression model is then used to hindcast SOI variability prior to the 20th century. Mann et al (1998, 2000), take the so-called ‘inverse’ approach to calibrating their proxy data. They first decompose 20th century tropical Pacific SSTs into multiple spatial EOFs. A calibration model is then defined by regressing individual proxy records from a multi-proxy data set onto the timeseries of the leading observed modes. Whether one is regressing the timeseries of leading multi-proxy modes of variability onto 20th century climate indices or one is training the proxy data to spatially determined modes of 20th century observed variability (inverse method), an underlying assumption is made that fundamental relationships between the proxy indicators and the real climate during the 20th century is similar to climate of the past.

In this study, we take a relatively simple complementary approach to defining an ENSO index from a multi-proxy network. We reason that an appropriately chosen network, following the spatial distribution requirements we outlined earlier and with each record having been previously identified as carrying a strong ENSO signal, should be robust enough to resolve a common mode of covariability that is directly comparable to one or more ENSO indices. This approach makes use of the fact that background noise has very short spatial scales and is therefore not present as variability that is synchronous or coherent across the whole network. This fact is further exploited by selecting a multi-proxy network where noise due to internal biological processes would be expected to have no relationship across the diverse group of indicators.

Hence, we use EOF analysis to decompose the basin-wide multi-proxy network into leading modes of covariability. Any common signal in the network (most likely to be climate-related) will be represented by the lower-order eigenvectors, while uncorrelated noise is separated into higher-order eigenvectors. The orthogonality requirement of EOF analysis means that the appropriate ENSO signal should be isolated by a single leading eigenvector (or EOF), the timeseries of which represents our 'proxy-ENSO' index. In effect, the entire 'calibration' is performed by simply applying the EOF analysis and dependent upon the selection of an appropriate network. We show that this simple approach leads to skill in simulating ENSO variability in independent data at comparable levels to previous studies.

3.2 Proxy-ENSO Index

As outlined above, the index of ENSO that we 'reconstruct' here is simply the ENSO related EOF from our multi-proxy network. Using the correlation matrix of anomalies from the proxy network described in section *2.1.*,

$$R = (N-1)^{-1} Z'^{T} Z' \quad , \qquad \text{Eqn 1}$$

Where $\boldsymbol{Z}$ represents the proxy matrix $\boldsymbol{X}$ linearly transformed into standardized (unit variance) deviations, we define a set of orthonormal basis vectors ($\boldsymbol{E}_j$, $j=1,9$) or eigenvectors;

$$R\tilde{E}_j = [\lambda_j]\tilde{E}_j \text{..,} \qquad \text{Eqn 2}$$

where λ_j (eigenvalue) is the *jth* diagonal element of the transpose of the correlation matrix $\boldsymbol{R}$.

Each eigenvector (EOF) represents an orthogonal mode of covariability in the proxy data set $\boldsymbol{X}$, with the lower order eigenvectors explaining the largest fraction of the total variance. Timeseries of each EOF are generated by projecting each eigenvector onto $\boldsymbol{X}$ at each timestep, representing the amplitude of each mode through time. The relevant mode relating to ENSO is determined through comparison with the SOI over the 20th century. Hence, the timeseries of this EOF may be considered as a ‘proxy-ENSO’ index that is analogous to instrumental ENSO indices. In the actual proxy data, ENSO variability is represented by the leading EOF and no other modes were found to have a clear association with instrumental ENSO indices.

For each proxy, the signal of ENSO is registered annually during the ‘growth’ season. Growth seasons typically last three to six months and occur at different times of the year, depending on the type of proxy and where it is located. This is an issue for Southern Hemisphere tree-rings, where growth seasons can be spread over two calendar years, corresponding to the Southern Hemisphere summer. Dating convention determines which year the growth season is recorded in, typically the previous year if the growth season covers two calendar years. Hence, prior to calculating the correlation matrix for the proxy network, each timeseries was adjusted such that the appropriate growth seasons where annually synchronous. In practice, only the Kauri record from New Zealand requires adjustment by a 1 year lag to ensure it was synchronous with all other records.

Two proxy-ENSO indices are defined. One using the long (A.D. 1525-1982) network of five proxies (hereafter referred to as R5) and the other from (A.D. 1727-1982) using the extended network of eight proxies (hereafter referred to as R8). Note that R5 is a spatial subset of R8 and hence is not independent.

3.3 Synthetic data testing of the EOF signal extraction method

In this section we attempt to show that simply extracting the relevant mode of covariability from a set of proxies using EOF analysis provides a robust method of signal estimation. We conduct a series of synthetic data experiments to evaluate the simple decomposition method described above against more direct calibration methods. Decomposition of a correlation matrix into a set of orthonormal basis vectors is analogous to total least squares regression. However, whereas regression or calibration methods make use of some sub-sample of the timeseries to predict the signal elsewhere, EOF decomposition allows the signal to be defined from the total variability. Using synthetic data, we test the power of the EOF method to extract the appropriate mode of covariability (signal) against a multiple linear regression model of signal against proxies from some sub-sample of the total timeseries.

We also attempt to test whether including multiple modes of *proxy* variability in a calibration type regression model does anything to improve the resultant reconstruction. For example Stahle et al (1998) regress multiple EOFs (EOFS 1 to 4) from their multi-proxy network against the SOI to build a hindcast model. Similarly, Mann et al (2000) use multiple spatial modes of observed tropical SSTs to calibrate their proxy timeseries. In the specific case of ENSO, it is unclear whether fitting additional modes of variability (particularly fitting multiple modes of *proxy* variability) adds any skill to the reconstruction of an index such as the SOI. While fitting multiple modes may improve the model performance over the calibration period, it may in fact degrade the model outside of this period. We use simple synthetic data testing to evaluate whether the fitting of additional proxy modes improves the overall skill of the regression model, or whether this approach simply tunes the model to the calibration period.

We use a Monte Carlo type approach to test each of these questions. We define a standardised timeseries (***S_t***, *t=1,500)* as our 500 year synthetic ENSO index. This timeseries is constructed using a Gaussian random number generator and has unit variance. A first order autocorrelation (AR1) structure, equivalent to that of observed interannual SOI (AR1 = 0.05), is added to this timeseries according to;

$$S_t = aS_{t-1} + \sqrt{1-a^2}N_t \ , \qquad \text{Eqn 3}$$

where *N(t)* represents Gaussian random noise and *a* is the AR1 first order autocorrelation. An ensemble of one thousand synthetic timeseries are created.

We also define a second synthetic ENSO timeseries in which we allow *a* to vary on inter-decadal timescales. While, there is no evidence that this type of structure occurs in the actual ENSO signal, we contrive this timeseries to explicitly introduce periodic, low frequency variability. Since we have no estimate of ENSO prior to the instrumental record, it is unknown whether variability during the last 100 years (the period used to calibrate proxy records) is significantly different to previous periods. Using a periodic function, we set the autocorrelation *a* in (3) to,

$$a_t = A\sin(2\pi t/\lambda) + A \ , \qquad \text{Eqn 4}$$

Where the amplitude *A* is set to *0.25* and the wavelength λ is set to 100 years. The autocorrelation of the embedded signal therefore varies between 0 and 0.5 sinusoidally over a century. The amplitude and wavelength values are somewhat arbitrary. Here, we use values that, in practice, we judge to introduce a discernible and realistic decadal signal to the data. We use a random number between *0* and λ to initialize *t* in order to phase-shift the sequence for different iterations within the ensemble of synthetic sequences.

A set of eight (matching the number of proxies in our actual proxy network) synthetic proxy timeseries (***X**$_{jt}$*, *j=1,8, t=1,500*) are also constructed. The real proxy data contains both a climate signal, which is invariably red as in (3), as well as red noise that is not associated with climatic influences. Hence we embed the synthetic 'observed' signal *S* in synthetically generated red noise, following;

$$\mathbf{X}_{jt} = \alpha S_t + \sqrt{1-\alpha^2}\,\mathbf{RN}_{jt} \,, \qquad \text{Eqn 5}$$

where ***RN**$_j$(t)* is a synthetic red noise timeseries generated as in (3) but with the AR1 coefficient set to 0.4 to simulate persistence that approximates the autocorrelation of the actual proxy records previously described. The signal-to-noise ratio within the synthetic proxy data is set by the value of *α*, with *α=1* determining all signal and α=0 determining no signal. In practice, *α* values in the range of 0.2 to 0.4 produce synthetic data that are statistically similar to the actual proxies. As with the synthetic signal, an ensemble of one thousand, five hundred-year timeseries of synthetic proxies are constructed.

Two different sets of synthetic proxies are constructed, one with a constant signal-to-noise ratio for the 500 year period and another where we allow the strength of the signal-to-noise ratio, *α*, to vary on inter-decadal timescales analogously to (4). This second set of proxies is constructed to test a contrived case where the strength of the signal in the proxies is attenuated on multi-decadal timescales.

Firstly, we test the EOF derived signal against the multiple linear (least squares) regression model. We derive a set of eigenvectors and their associated timeseries from our synthetic proxy matrix, ***X***, following (1) and (2) in section 3.2. Table 2 shows the percentage of variance explained by each eigenvector (EOFs 1 -5) for the set of synthetic proxies ***X*** and for a range of signal-to-noise ratios (alpha (α) weightings). Values shown in Table 2 represent the mean of

1000 analyses. A clear signal (whereby most of the variability is explained in a single EOF) is obtained using an alpha (α) weighting of 0.6 or higher in the synthetic proxies. Alpha (α) weightings of 0.2 to 0.4 are statistically more reflective of the actual proxies. For example, an alpha weighting of 0.4 produces a similar partitioning of variance in the EOFs to the actual proxy network (see Table 6). Furthermore, the relationship between the leading synthetic EOF and the synthetic observations (*S*) is similar to the leading proxy EOF and the SOI in the real data (see Table 7).

The correlation and RMS error between the synthetic signal (S_t, $t=500$) and the leading EOF (EOF1) from the synthetic proxies (X_{jt}, $j=8$, $t=500$), for different signal weightings, is shown in Table 3. We compare EOF1 to a multiple regression of ***X*** onto *S* using the full 500 years of data, following;

$$S_t = \sum_{j=1}^{n} \beta_j \mathbf{X}_{jt} + \xi_t \ , \qquad \text{Eqn 6}$$

Also shown in Table 3 is the correlation between the synthetic signal, *S*, and the predictand, $\hat{\boldsymbol{S}}$**,** from (6). Correlations and RMS errors represent the mean of 1000 analyses with the standard deviation derived from resampling. For almost all realistic signal-to-noise weightings (ie. $\alpha = 0.2\text{-}0.4$), both forms of signal analysis produce similar results. In other words, a linear decomposition of the total variance in the proxy data set produces a very similar signal to that predicted by a multiple regression model.

The above experiment is also conducted using a set of synthetic proxies constructed using periodic weighting of the signal-to-noise ratio (4). Neither the EOF method nor the multiple linear regression method is able to extract the embedded signal with any fidelity (not shown). Importantly, this result indicates that a reasonably consistent signal-to-noise ratio must be present in actual proxies for them to be of use, and the successful extraction of any signal would indicate that this condition is reasonably true.

Next we test the case of a regression model built upon a small sub-sample of the total timeseries. This is analogous to a realistic situation where only 100 years of observations exist against a much longer proxy data set. Using synthetic data, we are able to gauge how well a multiple regression model using a sample of the data predicts the actual signal for the rest of the timeseries. Further, we test how well this prediction compares to the leading EOF decomposed using the *full* 500 year timeseries.

Using the last 100 years of synthetic data ($S_{t,}$ $t=401\text{-}500$) for our regression model, we predict a signal, $\hat{\boldsymbol{S}}$, for the first 400 years and compare this to the synthetic signal, ($S_{t,}$ $t=1\text{-}400$). We also compare $\hat{\boldsymbol{S}}$ to EOF1 from the covariance matrix of the synthetic proxies. As before, we use 1000 separate realisations of the synthetic data to conduct our test. Mean correlations and RMS errors for EOF1 and $\hat{\boldsymbol{S}}$ and the synthetic signal S are shown in Table 4. Results again show that the two methods are similar. The leading EOF provides very similar, or slightly higher, correlation with the synthetic signal for the first 400 years of data compared to the regression model prediction. RMS errors are slightly larger for EOF1 than $\hat{\boldsymbol{S}}$**.** This experiment was repeated using a synthetic signal with an explicitly set, time-varying autocorrelation structure (3) with very similar results.

As most studies have noted, a caveat of multivariate regression calibration using spatially distributed proxies, is that they must assume a degree of stationarity in the climate modes that force the covariability in the proxy data. It is unknown whether large changes have occurred in the climate modes that the proxies represent, such that the relationship of covariability during the calibration period is not representative of periods outside of that window. Given that the leading eigenvector is determined from the entire timeseries, the EOF method is perhaps the most conservative in terms of reducing possible sources of error (due to non-stationarities), compared with a regression model built upon some fraction of that timeseries. In other words, while both methods rely on projecting the strength of some stationary linear combination of proxies onto each time-step to extract the climate signal, the EOF method should provide a better estimate of relative changes in the context of the *entire period*. Non-linear methods using a similar approach may be expected to yield better results.

The third experiment we perform is intended to test the validity of so called 'principal component regression' (for example see Stahle et al, 1998) for the purpose of calibrating multi-proxy records. As described previously, this approach regresses *multiple* modes of covariability (multiple EOFS) from multi-proxy networks onto 20th century observations to build a prediction model for past ENSO. We show that for ENSO, which is essentially a single mode of climate variability, it is unnecessary to fit additional modes of proxy variability to the observations. Furthermore, since the additional modes are not shown to be ENSO-related, they may act to tune the regression model to the calibration period.

Using the synthetic data, we construct a multiple regression model analogous to principal component regression. We use the last 100 years of EOF1 (the signal mode) from the synthetic proxies and *four* additional white noise timeseries in a regression model, following;

$$S_t = \sum_{l=1}^{n} \beta_l \mathbf{N}_{lt} + \beta_{n+1} EOF1_t + \xi_t \ , \qquad \text{Eqn 7}$$

where S_t ($t=401,500$) is the synthetic observations and $\boldsymbol{N}_{lt}$ ($l=1,4,\ t=401,500$) is synthetic white noise. Fitting additional white noise, $\boldsymbol{N}$, is analogous to fitting additional modes of variability (such as additional EOFs) that have no physical relationship to the signal we are trying to reconstruct. We do this to demonstrate that such noise fitting may improve the model over the calibration period, but is not the optimal approach for reducing uncertainties during the hindcast period. This model is used to predict the synthetic ENSO signal, $\hat{\boldsymbol{S}}$, for the first 400 years of data. A comparison is then made between S and $\hat{\boldsymbol{S}}$ during the first 400 years and during the calibration period.

Table 5 shows correlation and RMS errors between the synthetic observations and the predictand, $\hat{\boldsymbol{S}}$, for a range of realistic (see Table 2 and above) signal weightings. During the calibration period, fitting additional white noise to the model marginally improves the correspondence between $\hat{S}$ and the synthetic observations, as will adding *any* additional timeseries. Since this is the period in which verification skill of the reconstruction is tested, fitting additional modes of proxy variability in a 'principle component' regression (even though those modes may not be related to the ENSO signal) maybe seen as advantageous. However, comparison between S and $\hat{S}$ over the first 400 years (hindcast period) shows that correspondence marginally decreases as more non-signal related noise is fitted to the data. Hence this approach is less optimal (in reducing potential errors outside the period of calibration) than simply taking the leading EOF, particularly if the additional proxy modes contain lower order autocorrelation.

In summary, the synthetic tests show that a common signal may be confidently extracted from a set of proxy indicators (each of which include signal plus noise) using EOF analysis. This is predicated by two modest data requirements in the proxy indicators:

(1) that the signal is distinct from the noise (α values ≥ 0.2) in each of the proxies and

(2) that the non-climate noise is unrelated from proxy to proxy.

For ENSO, the later is ensured by choosing an appropriately widespread network, while the modest signal-to-noise requirement is easily met by proxies with an established ENSO signal.

Calibrating proxy indicators to observations is less important for the reconstruction of ENSO indices than it is for surface temperature, since the (absolute) scale of variability is not meaningful. As well as being straightforward, the EOF method may provide more certain temporal continuity for assessing relative changes in the signal over time.

3.4 Is pre-filtering of the proxy data necessary?

In the preceding sections (*3.1* and *3.2*) our methodology for extracting the ENSO signal from the multi-proxy network is described. This methodology requires no pre-filtering of the individual proxy timeseries, since noise is successfully separated from the signal in the EOF analysis. In fact, the use of raw, unfiltered data provides the best chance that low frequency components of the signal are retained in the final analysis. For single site calibrations, where proxies from a small geographical area are used to estimate ENSO variability, it is necessary to pre-process the data to ensure that the predictors (proxies) are likely to contain the majority of their association with the predictand in the relevant signal that we wish to reconstruct (ie. ENSO). Typically this involves high-pass filtering or detrending of the proxy timeseries to remove low frequency variability that may be associated with other climate processes or noise. For example, for tree-ring records, variability on scales longer than 20 years has been identified with biological processes (Fowler et al, 2004) and is therefore often removed using smoothing splines or similar techniques. Following single site calibration studies, pre-filtering techniques have also been employed variously in compiling multi-proxy networks. Hence we consider it useful to briefly describe the sensitivity of our index to several common pre-filtering methods.

Three forms of data processing are common in the literature for the preparation of proxy data prior to the construction of transfer functions. These are pre-whitening, detrending and high-pass filtering (see for example, Cook and Peters, 1981; Stahle et al, 1998; Fowler et al, 2004; D'Arrigo et al, 2005). Pre-whitening (removing lower order auto-correlation from the timeseries) and high-pass spline filtering are typically used to remove low frequency signals (a function of biological growth) from tree-ring sequences. Linear detrending is most commonly used to remove long term drift in the timeseries that is not associated with the relevant signal, for example the removal of a global warming signature when trying to isolate ENSO variability.

In this study we construct our proxy network (matrix) using one, all or various combinations of pre-processing described above, to investigate whether these improved the comparison of EOF1 (our proxy-ENSO index) with the SOI, Niño3.4 SST and the CEI. We find (not shown) that none of these techniques significantly improve the comparison between EOF1 and any of the instrumental ENSO indices. Special mention is made here of the spline filtering techniques, since these are commonly used in widely available statistical toolboxes for the preparation of tree-ring sequences. We compare the resultant EOF1 from proxy matrices using 200-year spline and 20-year spline filtering of the proxy tree-ring sequences described previously. While 20-year spline filtering improves the signal-to-noise ratio of EOF1, the ability to represent decadal and longer variability is severely attenuated. In practice, the improved signal-to-noise ratio achieved using 20-year splines marginally improved the ability of EOF1 to faithfully reconstruct extreme events (not shown). On balance, we considered that preserving low frequency characteristics of ENSO is of greater importance and hence 200-year splines are employed for this study. The ability to extract an ENSO signal that retains lower frequency variability is a significant advantage of the multi-proxy network based approach over single site approaches. The issue of low frequency ENSO is explored in greater detail in section 4.4.

4. RESULTS

We compare the timeseries of our proxy-ENSO index (described in section 3.2) to instrumental ENSO indices over the 20th century. To allow for seasonal differences, we compare the annually resolved multi-proxy data against seasonal DJF, MAM, JJA and SON values of the SOI, Niño3.4 SST and the CEI. In the following section we present a comparison with observations based on temporal correlation, correspondence in the spectral domain, and skill in the ability of the proxy-ENSO index to simulate threshold based El Niño and La Niña events.

4.1 Correlation between 'proxy-ENSO' index and observations

Here we present the correlation (R) and the proportion of common variance explained (R^2) between the proxy-ENSO index and observed indices of ENSO. Close association between our index and instrumental indices over the 20th century infers that the index is a suitable proxy for ENSO variability. Table 6 shows variance explained by the leading four EOFs from both of the multi-proxy networks, R5 (1525 to 1982) and R8 (1727 to 1982). While the total variance of the proxy records is evenly spread across the four leading modes of variability, only the first mode is correlated with any/all of the ENSO indices (Table 8). The uniform distribution of the variance in multiple directions highlights the degree of noise on inter-annual to multi-decadal timescales that exists within the palaeo data. Based upon the distribution of variance explained in the synthetic data sets presented previously, the signal-to-noise ratio is of the order 1:3.

The geographic distribution of proxies is obviously an important component in accurately capturing ENSO related variability. Table 7 shows the relative loading (eigenvectors) of each the proxy chronologies from the 1727-1982 network and their lead/lag correlation with the associated R8 (EOF1) timeseries. The loadings represent a measure of coherent variability associated with the direction of the leading EOF and hence indicate which sites contribute most to the overall variance within this mode. The variance associated with R8 is dominated by covariability in all of the tree-ring sequences, the Quelccaya ice-core and the Rarotonga coral record. Importantly, the resolved signal is not dominated by a particular hemisphere and seems to capture a coherent, basin-wide signal.

Table 8 shows the correlation between the leading four annually resolved EOFS from the A.D. 1727-1982 network and the SOI, CEI and Niño3.4 SST. R8 (EOF1) is highly correlated with each of the ENSO indices from September of the previous year, through to May of the concurrent year. This relationship peaks in DJF for the SOI and CEI and in MAM season for Niño3.4 SST. This result is consistent with previous studies which show that proxy association with ENSO indices is greatest in December through to May. The strength of the associations weakens for the remainder of the concurrent year in the SOI and CEI, but remains strong through JJA for Niño3.4 SST.

Table 9 shows correlations and common variance explained for R5 and R8 and each of the ENSO indices in all seasons. The proxy-ENSO index resolves a high proportion of variability in all three of the ENSO indices, the closest association occurring with the CEI coupled index. Comparison between R5 and R8 indicates that the inclusion of more records in the network improves the subsequent reconstruction in terms of coherence with 20th century ENSO. The proportion of instrumental variance resolved by R5 and R8 compares well with previous efforts at ENSO index reconstruction. Using 3 dominant eigenvectors, Mann et al (2000) were able to resolve 34-42% of October to March instrumental Niño 3 SST. Stahle et al (1998) capture a maximum of 53% for December to February instrumental SOI. Most recently, D'Arrigo et al (2005) resolve between 43-52% of Niño 3 DJF SST. Here, the R8 index resolves 48% of DJF SOI variability, 53% of DJF CEI variability and 49% of MAM Niño3.4 SST. R5 resolves 42% of 20th century CEI variability. Figure 2 shows 20th century R8 scaled to instrumental DJF SOI, DJF CEI and MAM Niño3.4 SST. Residuals from the reconstructions (Figure 3) are normally distributed and free of significant auto-correlation.

4.2 Correspondence in the Spectral Domain

Next we compare the spectral characteristics of the instrumental SOI and the proxy-ENSO index. Power spectra are calculated from anomaly data and smoothed using multiple passes of a centrally weighted, moving three-point window (1-2-1 weighting) across the frequency domain. The background red noise spectrum is estimated from the theoretical AR1 or Markov process noise (Gilman, 1963; Torrence and Compo, 1998). The significance of spectral peaks relative to background noise is estimated from the AR1 noise spectrum and the χ-squared/degrees of freedom distribution (Bath, 1974), with the degrees of freedom dependent on the size of the effective bandwidth after smoothing. Results in this study were compared with other methods of power spectra and background noise estimation, most notably the multi-taper method (Mann and Lees, 1996), and found to be similar.

Figure 4 shows spectral power for the instrumental SOI from 1871-1982. Power spectra for Niño3.4 SST and the CEI (not shown) are very similar to the SOI. From Figure 4, peaks in spectral power for the SOI occur at ~2.9, ~3.5, ~4.0, ~6.2 and ~9.3 year frequencies. Spectral analysis of R8 for the period A.D. 1727-1982 is shown in Figure 5. While no direct comparison can be made between Figures 4 and 5 (since spectral results are sensitive to the time period), R8 over the last 300 years contains many of the spectral signals seen in the 20th century SOI, as well greater low frequency variability. Figures 6 and 7 compare the spectra over the A.D. 1871-1982 period only. In general, the reconstructions reproduce the 20th century spectral signal well, with cross-spectral analysis indicating significant coherence at observed ENSO frequencies.

4.3 ENSO Event Reconstruction

Previous researchers have used palaeo-climate and historical evidence to document past El Niño and La Niña events (for example Qiunn and Neal, 1992; Whetton and Rutherfurd, 1994; Ortlieb, 2000). Similarly, studies that have attempted to reconstruct continuous (annual) indices of ENSO have also sought to assess how well their records reproduce individual events. Stahle et al (1998) (SOI) and Mann et al (2000) (Tropical SSTs) use a single (one) standard deviation threshold to define and compare El Niño and La Niña events in their records and the observations.

A number of factors affect the ability of proxy-based ENSO indices to capture signature ENSO events. One of the biggest limitations is that proxy reconstructions tend to underestimate the amplitude of ENSO events due to the loss of variance associated with regression-based prediction. This is because linear regression implicitly disregards statistical outliers. This same limitation applies to defining orthonormal basis vectors from a covariance or correlation matrix. In terms of reproducing past ENSO, this loss of large amplitude variance means that not all of the observed extreme events will be captured by the proxy index. In Figure 8, we compare the 10-year moving mean standard deviation for R8 and observed DJF SOI. Both Niño3.4 SST and the CEI exhibit similar changes in variance to the SOI in Figure 8. In comparison with the instrumental record, R8 variance is smaller and displays less pronounced decadal variability. This is most likely a result of the proxy-index failing to capture the magnitude of very large ENSO events, which will tend to influence 10-year mean historic variance.

The type of instrumental index used to identify warm and cold events can also be expected to affect comparisons between proxy-ENSO and observed ENSO. Here we choose to use the SOI to identify observed events, principally because it provides a higher quality record over the earlier part of the 20th century. Choosing an appropriate temporal average for the SOI is an important consideration. While the proxy-ENSO index has its highest association with the December-February SOI, this is not the most suitable period for identifying ENSO events from a circulation index, due to the influence of noise from factors such as the MJO. For this reason we use the June-November monthly-mean SOI to identify El Niño and La Niña events in the observations. This period is traditionally associated with the peak maturity of El Niño events in the tropical circulation. Based on Table 8, the proxy-ENSO index has its highest association with June-November SOI in the *previous* calendar year. R5 has a correlation of 0.50 with June-November SOI of the previous year, while R8 has a correlation of 0.58.

To assess the accuracy of discrete event capture in the proxy-ENSO index, we follow the approach of previous studies (Stahle et al, 1998; Mann et al 2000) and use a single standard deviation threshold to define El Niño (negative SOI anomaly) and La Niña (positive SOI anomaly) events from both R5 and R8, as well the June-November SOI. For the list of instrumental events, we also indicate a number of marginal or weak events, that fail to meet the single standard deviation threshold but have been documented elsewhere (for example the Australian Bureau of Meteorology's list of El Niño and La Niña events that is available on the world wide web). These results are summarised in Figure 9 for the period A.D. 1876 to 1982 (the period with the highest quality observations).

Firstly we concentrate on El Nino events identified since 1876. Twenty one events (including five weak events) are recorded in the SOI over this time period. Five of these occur prior to the 20^{th} century, 1877, 1881, 1885, 1888 and 1896. The first and last of these events, in 1877 and 1896, are captured by both R5 and R8. The events of 1855 and 1888 are registered in the preceding year by the proxies, perhaps reflecting the strongly negative SOI prior to June of that year. The event of 1881 is not present in the reconstructions.

Sixteen events are identified in the period 1900 to 1982. Six of these events are registered in the proxy-ENSO indices, those of 1905, 1911, 1919, 1925, 1940 and 1965. The event of 1977 is registered in the following year in both R5 and R8 reflecting a persistently negative SOI into December through February of 1977-1978. Two periods of continuously negative SOI, recorded from 1940-1943 and 1963-1966, are captured by both R5 and R8. R5 registers the weak El Niño events of 1902, 1913, 1951 and 1957, which are not identified by the single standard deviation threshold in the SOI. Conversely, the moderate, short-lived El Nino events of 1923 and 1946 are recorded by the single standard deviation threshold in the SOI but are not registered by the proxy-indices. The event of 1972 marginally fails to reach the threshold value in the proxy-indices. False El Niño events (ones not recorded in the SOI or elsewhere) are reproduced by R5 in 1930 and 1980 and by both R5 and R8 in 1967.

Fifteen La Niña events (including one weak event) are identified from the SOI from 1876 to 1982. Prior to the 20^{th} century R5 and R8 correctly identify the event of 1878. The event of 1889 is present in the proxies, but marginally fails to reach the threshold value. Both proxies fail to register the recorded event of 1886. A possible weak event in 1892-1893 is also identified by both indices. During the 20^{th} century, strong La Niña events were correctly identified in 1909, 1916, 1950 (R8), 1955 (R8) and 1973. A short-lived positive anomaly in the SOI during 1903-1904 is registered as an event in both R5 and R8. The event of 1971 is registered in the preceding year in the proxy index, reflecting concurrent positive (but below the threshold) SOI anomalies from that period. Events in 1906, 1924, 1938, 1975 and a weak La Niña during 1964 are all missed by the proxies. False positive identification occurs in 1928 (R5), 1933 and 1952 (R5). There is also a tendency for the proxies to register the year preceding a La Niña as above the standard deviation threshold, even when the observed SOI does not display sustained positive anomalies. This occurs for the years 1908, 1915, 1949 (R8) and 1969-70.

In general, one may expect differences in the ability of palaeo-climate indicators to accurately register El Niño and La Niña events. Non-linearity in ENSO teleconnections and sign-dependent strengths in the regional response to ENSO in different proxies and regions has been noted by previous researchers (Hoerling et al, 1997, Power et al, 2006). Power et al (2006), for instance, show that the magnitude of La Niña SST anomalies are closely correlated with the magnitude of the response in area-averaged Australian rainfall changes, whereas the magnitude of El Niño SST anomalies are not. While the linear relationship between individual proxy chronologies and the observed ENSO indices is varied, Figure 10 shows that the relationship between the proxy-ENSO index and observed SOI and CEI over the 20th century is linear for both phases of ENSO. This perhaps reflects the fact that a range of teleconnection regions are represented in the R8 proxy data.

In summary, out of sixteen El Niño events identified from 1900-1982, thirteen (including 1978) are correctly identified by R5 and R8 and three events are missed. Only three clear false positive El Niño events (not associated with a negative SOI period in the observations) are recorded in 1930, 1967 and 1980 (R5). Out of eleven identified La Niña events from 1900-1982, six (including 1971) were correctly identified by the proxy-ENSO indices. Four proxy-based La Niña events were falsely reconstructed, including one (1903) that is ambiguous. The proxy-ENSO index also identifies the year preceding an observed La Niña as a threshold event on six occasions. The reconstructions produced here are consistent with previous studies in that they under-represent the observed number of events. The reconstructions show higher skill in reproducing warm phase ENSO events than cold phase events.

4.4 Low frequency ENSO variability and the IPO

ENSO reconstructions should, ideally, faithfully represent lower frequency (decadal or greater) changes in ENSO. Both ENSO and ENSO teleconnectivity have been shown to display low frequency variability (Allan et al, 1996, Power et al, 1999). Power et al, (1999) show that the relationship between ENSO and Australian rainfall on interannual timescales varies on interdecadal timescales, in correspondence to changing phases of the Interdecadal Pacific Oscillation (IPO). The IPO is defined as a decadal mode of Pacific SST variability that bears pattern similarity to ENSO SST variability (Folland, 1999, Power et al, 1999) and closely resembles low frequency ENSO in the interdecadal component of the Pacific Decadal Oscillation (PDO) (Mantua et al, 1997; Folland et al, 2002; Power et al, 2006). Hence, the ability to characterise low frequency ENSO would greatly improve the usefulness of the proxy-ENSO index as a diagnostic tool for past climate.

Several studies (Hughes, 2002; Cook et al, 2004) have made note of potential limitations in the ability of climate proxies to register low frequency climate variability. However Esper et al (2002) demonstrate that through the careful selection and treatment of tree-ring chronologies, it is possible to preserve temperature variability on multi-centennial timescales. Here we investigate the low-frequency variability of R8 and R5 in comparison to low frequency ENSO. Figure 11 shows the index of the IPO over the 20th century compared to the low frequency component of observed SOI and low frequency R8 and R5 over the 20th century. Low frequency SOI and proxy-ENSO indices have been estimated using a low pass filter using the spectral method described by Power et al (1999).

The proxy-indices capture 39% of the decadal-scale variability shown in the instrumental SOI, with notable differences in amplitude, particularly during the first 40 years of the 20th century. Comparison between R8, the SOI and the IPO shows that the reconstruction provides a reasonable measure of IPO-like variability. The largest differences between low frequency ENSO and the proxy indices occur during 1900-1940. The R5 reconstruction estimates greater decadal variability during this period.

5. ENSO OVER THE LAST FOUR CENTURIES

Here we describe various changes in the proxy-ENSO indices over the last 450 years. Figure 12 shows the extended timeseries for R5 (1525-1982) and R8 (1727-1982). Figure 13 shows the low frequency component (as defined in section 4.4) of R5 and R8 for the same periods. Temporal changes in the frequency domain for R5 are shown in Figure 14, which uses a 50-year moving window to calculate the spectral power of the proxy-ENSO index since 1525. Figure 15 shows the 50-year moving variance (spectral power) for R5 in three separate frequency bands (2-4 years, 4-8 years and 8-20 years), as well as the total variance, expressed as anomalies from the long-term 1525-1982 mean. Finally, Figure 16 presents the fraction of total variance captured by each of these bands. Taken collectively, Figures 12 through 16 suggest considerable amplitude and frequency modulation in ENSO over the past four centuries. Results for the spectral analysis are insensitive to window lengths between 30 and 80 years. In interpreting Figures 15 and 16, it is important to note that the frequency bands do not represent discrete modes of variability. These bands were chosen based on the apparent structure of changes in the power spectrum shown in Figure 14. Results using R8 were similar to those of R5.

There are a number of notable changes in ENSO variability during the reconstruction period. In particular, there are extended periods of relative quiescence in high-amplitude, interannual (~3-4 year) variability during the sixteenth and early seventeenth centuries, and again during the eighteenth century. ENSO variability during these periods is dominated by interdecadal or longer timescales. The signal of higher frequency variability, particularly as a fraction of total variability, increases from the beginning of the nineteenth century until the mid-twentieth century, the end of our record. Results here reflect the more general conclusion of previous palaeo-ENSO studies (most notably Mann et al. 2000) in that suggested changes in the amplitude of ENSO variability are perhaps more notable than frequency modulation. Very few threshold-based El Niño events (events being defined using a single standard deviation threshold, and using a moving 50-year window to determine the standard deviation) are recorded during the sixteenth century and during the last fifty years of the eighteenth century.

The timing of changes suggested by the R5 index is not consistent with previous ENSO reconstructions. While Mann et al (2000) also noted pronounced breakdowns in internannual variability of tropical SST, these occurred during the early to mid nineteenth century of their record. Documentary records of ENSO events, for example Quinn et al, (1987), show a reasonably consistent occurrence of interannual El Niño events from 1525 to near-present. Several reasons may exist for these differences. Most obviously, the proxy-network employed here differs to previous ENSO index reconstructions through the inclusion of several south-western Pacific sites.

Documentary records of El Niño events such as that of Quinn et al (1987) rely on subjective classification of events, which suffer from the shortcomings associated with quantifying anecdotal information, and often have a poor correspondence with instrumental indices. For example Rasmusson et al (1995) conclude that recurrence statistics derived from the Quinn et al compilation of El Niño events cannot be considered a reliable index of basin-scale ENSO variability. Similarly however, the designation of events using a simple threshold in the proxy-ENSO indices used here may be too simplistic to correctly identify past El Niño and La Niña occurrences. This is especially true given that the long record may be subject to low frequency variability on multi-decadal timescales that is an artefact of the analysis (as a result, for example, of by chance association between background noise in the proxy network).

Dynamical studies have suggested that solar, volcanic and anthropogenic radiative forcing changes have influenced past ENSO variability, particularly a tendency toward El Niño-like conditions during periods of radiative cooling (Clement et al, 1996; Cane et al, 1997; Mann et al, 2005). However given the complexity of the atmosphere-ocean feedbacks involved in ENSO and the inconsistency in current ENSO modelling efforts (Collins, 2005) causal relationships based upon the simplistic correlation of apparent changes must be viewed with circumspection. We find it difficult to isolate any synchronicity between past radiative forcing changes and changes in ENSO variability in our proxy-index. Pronounced changes in ENSO in the late 19th/early 20th century that have been noted by previous researchers (Stahle et al, 1998; Allan, 2000; Trenberth and Caron, 2000) are also not found in this study, although we do find evidence of greater high frequency variability over the last two-hundred years.

6. SUMMARY

In this study we reconstruct a proxy record of ENSO-related variability over the last five centuries, using a network of palaeo-climate indicators spanning a broad geographic region of the Pacific. An eight-member multi-proxy data set (R8), comprising of two sub-tropical North American tree-rings, two New Zealand tree-rings, three western/central Pacific coral records and a single South American ice-core, is used to reconstruct ENSO variability from A.D. 1727 to 1982. The tree ring and ice core records are also used in a five-member network (R5) to reconstruct a longer record of ENSO variability from A.D. 1525 to 1982. These two proxy networks differ from previous reconstructions in containing proxy sites with a distinct ENSO signal from the south-west Pacific (New Zealand tree rings and Australian corals).

Our index of ENSO variability is defined simply as the timeseries of the leading mode of covariability from the proxy network. This approach differs to most previous studies, which have chosen to 'calibrate' proxy ENSO variability by regressing their palaeo-indicators against observed 20th century indices. For ENSO, where the absolute scale of the index is meaningless, we find that such direct calibration of the proxy indicators is unnecessary. Further, using synthetic data testing and reasoning, we conclude that direct calibration of the index to 20th century observations does not significantly improve the fidelity of the proxy-index and may introduce errors to the reconstruction outside of the calibration period.

The time-amplitude of the leading EOF from each of the proxy data sets is well correlated with the SOI, Niño3.4 SST and a combined ocean-atmosphere index (CEI) over the 20th century. Both indices R5 and R8 compare well with previously published reconstructions in terms of the proportion of common variance explained and the ability to register threshold-defined El Niño events. In common with previous efforts, there is less skill in capturing La Niña events. Both palaeo-ENSO indices also reproduce the main spectral characteristics of the instrumental ENSO indices. Of the three instrumental indices investigated here (SOI, CEI and Niño3.4 SST), highest correspondence/skill is found with the composite land-ocean CEI index.

The R8 ENSO index (which includes western Pacific coral records) explains a higher proportion of instrumental SOI (52%) than the R5 index (43%). We conclude, more generally, that the inclusion of proxy sites across a broader region of the Pacific provides a more robust reconstruction of ENSO. This is primarily a reflection of the fact that the strength of ENSO regional teleconnections is non-stationary and that a larger network of proxies provides more degrees of freedom, thereby reducing potential errors due to site-specific biases.

In terms of the reality of the extended ENSO reconstructions, a number of factors need to be considered. The loss of variance associated with reconstruction methods is perhaps the largest single cause of uncertainty in terms of estimating the magnitude of past ENSO variability and individual warm/cold events. Hence, reconstructions may be expected to be more robust in the frequency domain, since this component is less sensitive to the regression-based calibration than amplitude. This means that attempts to characterise past changes in ENSO should not rely solely on ‘event capture’ statistics. Changes in ENSO are perhaps more accurately gauged from the amplitude and frequency modulation of continuous, high temporal-resolution indices.

The apparent amplitude and frequency modulations of the R5 proxy-ENSO index is suggestive of some interesting changes in ENSO over the last four centuries. The amplitude of high frequency (~2-4 year) ENSO variability is damped during the sixteenth, early-seventeenth and mid-eighteenth centuries. There are also extended periods of the record where interdecadal (~8-20 year) variability is dominant. High frequency variability has increased over the last two-hundred years. In the context of the entire record, we find no pronounced signal of twentieth century climate change in ENSO variability. It must be noted however, that our proxy-ENSO index does not extend to the period after 1982, which coincides with a period of rapid warming in global temperature.

ACKNOWLEDGMENTS

Many thanks to Erica Hendy (Lamont Doherty Earth Observatory) for Great Barrier Reef coral luminescence record, Pavla Fenwick (Lincoln University, New Zealand) for Pink Pine tree-ring data, Bruce Bauer and Connie Woodhouse (World Data Center for Paleoclimatology/National Climatic Data Center Paleoclimatology Branch) for data acquisition and assistance. The authors thank Harry Hendon, Matthew Wheeler (Bureau of Meteorology, Centre for Australian Weather and Climate Research) and Grant Beard (Bureau of Meteorology, National Climate Centre) for useful discussions. We also thank internal reviewers Doerte Jakob and Brad Murphy (Bureau of Meteorology). For part of this study, JLG was supported at UNSW by an Australian Postgraduate Award. AMF was funded by the Royal Society of New Zealand Marsden Fund (Contract UOA108). JSR has been part supported by the CSIRO Wealth from Oceans Flagship program.

REFERENCES

Allan, R., Lindsay, J. and Parker, D. (1996). *El Niño Southern Oscillation and climate variability.* Melbourne, Australia, CSIRO.

Allan, R. (2000). ENSO and climatic variability in the past 150 years. *El Niño and the Southern Oscillation: Multiscale Variability and Global and Regional Impacts.* H. Diaz and V. Markgraf. Cambridge, Cambridge University Press**:** 3-35.

Bath, M. (1974) Spectral Analysis in Geophysics, Elsevier Scientific publishing, pp 563.

Baumgartner, T., Michaelsen, J., Thompson, L., Shen, G., Soutar, A. and Casey, R. (1989). The recording of interannual climatic change by high-resolution natural systems: tree-rings, coral bands, glacial ice layers and marine varves. *Aspects of Climate Variability in the Pacific and Western Americas. Geophysical Monograph.* D. Peterson. Washington, American Geophysical Union. **55:** 1-14.

Bradley, R. (1996). Are there optimum sites for global palaeotemperature reconstruction? *Climate Variations and Forcing Mechanisms of the Last 2000 years.* P. Jones, R. Bradley and J. Jouzel. Berlin, Springer-Verlag**:** 603-624.

Cane, M., Clement, A., Kaplan, A., Kushnir, Y., Pozdnyakov, D., Seager, R., Zebiak, S. and Murtugudde, R. (1997). Twentieth-century sea surface temperature trends. *Science,* **275,** 957–960.

Cleaveland, M., Stahle, D., Therrell, M., Villanueva-Diaz, J. and Burns, B. (2003). Tree-ring Reconstructed Winter Precipitation and Tropical Teleconnections in Durango, Mexico. *Climatic Change* **59**: 369–388.

Clement, A. C., R. Seager, M. A. Cane, and S. E. Zebiak, (1996) An ocean dynamical thermostat. *J. Climate,* **9,** 2190–2196.

Collins, M. (2005). El Niño or La Niña-like climate change? *Climate Dynamics* **24**: 89-104.

Cook E. R., and K. Peters, (1981) The smoothing spline: A new approach to standardizing forest interior tree-ring width series for dendroclimatic studies. *Tree-Ring Bull*, **42**, 45–53.

Cook, E.R., Woodhouse, C.A., Eakin, M.C., Meko, D.M., Stahle, D.W., (2004) Long-term Aridity Changes in the Western United States. *Science*, 306, 1015-1018.

Crowley, T. (2000). Causes of Climate Change over the Past 1,000 years. *Science* **289**: 270-277.

D' Arrigo, R., Jacoby, G. and Krusic, P. (1994). Progress in Dendroclimatic Studies in Indonesia. *Terrestrial, Atmospheric and Oceanographic Sciences* **5**: 349-363.

D' Arrigo, R., Cook, E., Wilson, R., Allan, R. and Mann, M. (2005). On the variability of ENSO over the past six centuries. *Geophysical Research Letters* **32** (L03711): 1-4.

Diaz, H. and Pulwarty, R. (1994). An analysis of the time scales of variability in centuries-long ENSO-sensitive records in the last 1000 years. *Climatic Change* **26**: 317–342.

Diaz, H. and Markgraf, V. (2000). *El Niño and the Southern Oscillation; Multiscale Variability and Global and Regional Impacts.* Cambridge, Cambridge University Press.

Dunbar, R., Wellington, G., Colgan, M. and Glynn, P. (1994). Eastern Pacific sea surface temperature since 1600 A.D.: the d18O record of climate variability in Galapagos corals. *Palaeoceanography* **9**: 291–316.

Esper, J., Cook, E.R., Schweingruber, F.H., (2002) Low frequency Signals in Long Tree-Ring Chronologies for Reconstructing Past Temperature Variability. *Science*, 295, 2250-2253.

Fairbanks, R., Evans, M., Rubenstone, J., Mortlock, R., Broad, K., Moore, M. and Charles, C. (1997). Evaluation of climate indices and their geochemical proxies measured in corals. *Coral Reefs* **16** (Supplement): S93-S100.

Fenwick, P. (2003). Reconstruction of past climates using pink pine (Halocarpus biformus) tree-ring chronologies. Soil Plant and Ecological Sciences. Christchurch, New Zealand, Linclon University.

Folland, C.., Parker, D., Colman, A. and Washington, R. (1999). Large scale modes of ocean surface temperature since the late nineteenth century. Chapter 4, pp73-102 of *Beyond El Nino: Decadal and Interdecadal Climate Variability*. Ed: A. Navarra. Springer-Verlag, Berlin, pp 374.

Folland, C., Karl, T., Christy, J., Clarke, R., Gruza, G., Jouzel, J., Mann, M., Oerlemans, J., Salinger, M. and Wang, S. (2001). Observed Climate Variability and Change. *Climate Change 2001: The Scientific Basis. Contribution of Working Group 1 to the Third Assessment Report of the Intergovernmental Panel on Climate Change.* J. Houghton, Y. Ding, D. Griggset al. United Kingdom and New York, Cambridge University Press.

Folland, C., Renwick, J., Salinger, M.J., and Mullan, A.B. (2002). Relative influences of the Interdecadal Pacific Oscillation and ENSO on the South Pacific Convergence Zone. Geophysical Research Letters, 29 (13): 10.1029/2001GL014201. Pages 21-1 - 21-4

Fowler, A.M., Palmer, J., Salinger, J. and Ogden, J. (2000). Dendroclimatic interpretation of tree-rings in Agathis australis (Kauri) 2; Evidence of a significant relationship with ENSO. *Journal of Royal Society of New Zealand* **30** (3): 277-292.

Fowler, A.M. and Boswijk, G. (2003). Chronology stripping as a tool for enhancing the statistical quality of tree-ring chronologies. *Tree-ring Research* **59** (2): 53-62.

Fowler, A.M., Boswijk, G. and Ogden, J. (2004). Tree-ring studies on Agathis australis (Kauri): a synthesis of development work on Late Holocene chronologies. *Tree-ring Research* **60** (1): 15-29.

Fowler, A.M., Boswijk, G., Gergis, J. and Lorrey, A. (2008). ENSO history recorded in *Agathis australis* (Kauri) tree-rings Part A: Kauri's potential as an ENSO proxy. *International Journal of Climatology* 28(1):1–20.

Fowler, A.M. (2008). ENSO history recorded in *Agathis australis* (kauri) tree-rings Part B: 422 years of ENSO robustness. *International Journal of Climatology*: **28**(1):21–35.

Gedalof, Z. and Mantua, N. (2002). A multi-century perspective of variability in the Pacific Decadal Oscillation: new insights from tree-rings and coral. *Geophysical Research Letters* **29** (24): 57/1-57/3.

Gergis, J. and Fowler, A. (2005). Classification of synchronous oceanic and atmospheric El Niño-Southern Oscillation (ENSO) events for palaeoclimate reconstruction. *International Journal of Climatology* 25: 1541–1565.

Gergis, J., Boswijk, G. and Fowler, A. (2005a). An update of modern Northland Kauri (*Agathis australis*) tree-ring chronologies 1: Puketi State Forest. New Zealand tree-ring Site Report No.19, School of Geography and Environmental Science Working Paper 29, University of Auckland, New Zealand.

Gergis, J., Boswijk, G. and Fowler, A. (2005b). An update of modern Northland Kauri (*Agathis australis*) tree-ring chronologies 2: Trounson Kauri Park. New Zealand tree-ring Site Report No.20, School of Geography and Environmental Science Working Paper 30, University of Auckland, New Zealand.

Gergis, J., Braganza, K., Fowler, A., Risbey, J. and Mooney, S. (2006). Reconstructing El Niño-Southern Oscillation (ENSO) from high-resolution palaeoarchives. *Journal of Quaternary Science* **21** (7): 707-722.

Gergis, J and Fowler, A.M. (2008). A history of ENSO events since A.D. 1525: implications for future climate change. Climatic Change, *in press*.

Goodwin I., Van Ommen T., Curran M., Mayewski P. (2004). Mid latitude winter climate variability in the South Indian and Southwest Pacific regions since 1300 AD *Climate Dynamics*, **22** 783-794.

Gilman, D.L., Fuglister, F.J. and Mitchell, J.M. (1963) On the power spectrum of red noise. *Journal of Atmos Sci.,* **20**, 182-184.

Hendy, E., Gagan, M., Alibert, C., Mc Culloch, M., Lough, J. and Isdale, P. (2002). Abrupt Decrease in Tropical Pacific Sea Surface Salinity at End of Little Ice Age. *Science* **295** (5559): 1511-1514.

Hendy, E., Gagan, M. and Lough, J. (2003). Chronological control of coral records using luminescent lines and evidence for non-stationarity ENSO teleconnections in northeastern Australia. *The Holocene* **13** (2): 187-199.

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

Hughes, M. K., (2002) Dendrochronology in climatology- the state of the art. *Dendrochronologia*, 20/1-2, 95-116.

Jolliffe, I.T. and Stephenson, D.B. (2003) Forecast Verification. A Practitioner's Guides in Atmospheric Science. *John Wiley and Sons*, 240 pages.

Jones, P. and Mann, M. (2004). Climate over past millenia. *Review of Geophysics* **42**: 1-42.

Linsley, B., Wellington, G. and Schrag, D. (2000). Decadal Sea Surface Temperature Variability in the Subtropical South Pacific from 1726 to 1997 A.D.. *Science* **290**: 1145-1149.

Lough, J. (2004). A strategy to improve the contribution of coral data to high-resolution palaeoclimatology. *Palaeogeography, Palaeoclimatology, Palaeoecology* **204**: 115-143.

Lyon, B. and Barnston, A. (2005). The evolution of the weak El Niño of 2004-2005. *US CLIVAR Variations* **3** (2): 1-4.

Mann, M. and Lees, J. (1996). Robust estimation of background noise and signal detection in climatic time series. *Climatic Change* **33**: 409-445.

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

Mann, M., Bradley, R. and Hughes, M. (2000). Long-term variability in the El Niño/Southern Oscillation and associated teleconnections. *El Niño and the Southern Oscillation; Multiscale Variability and Global and Regional Impacts.* H. Diaz, and Markgraf, V. Cambridge, Cambridge University Press**:** 327-372.

Mann, M. (2003). On Past Temperatures and Anomalous Late-20th Century Warmth. *Eos* **84** (27): 1-3.

Mann, M., Cane, M., Zebiak, S. and Clement, A. (2005). Volcanic and Solar Forcing of the Tropical Pacific over the Past 1000 Years. *Journal of Climate* **18**: 447-456.

Mantua, N.J., S.R. Hare, Y. Zhang, J.M. Wallace, and R.C. Francis, (1997) A Pacific decadal climate oscillation with impacts on salmon. *Bulletin of the American Meteorological Society*, Vol. 78, pp 1069-1079.

Moberg, A., Sonechkin, D., Holmgren, K., Datsenko, N. and Karlen, W. (2005). Highly variable Northern Hemisphere temperature reconstructed from low and high resolution proxy data. *Nature* **433**: 613-617.

Oldenborgh, G.J. van and G. Burgers, (2005b) Searching for decadal variations in ENSO precipitation teleconnections. *Geophys. Res. Lett*. **32**, 15,

Ortlieb, L., 2000: The documented historical record of El Niño events in Peru: An update of the Quinn record (Sixteenth through Ninenteenth centuries). *El Niño and the Southern Oscillation, Multiscale Variability and Global and Regional Impacts,* H. F. Diaz and V. Markgraf, Eds., Cambridge Univ. Press, 207-295.

Osborn, T. and Briffa, K. (2004). The Real Color of Climate Change? *Science* **306**: 621-622.

Power, S., Tseitkin, F., Torok, S., Lavery, B., and B. McAvaney, (1998) Australian temperature, Australian rainfall, and the Southern Oscillation, 1910-1996: Coherent variability and recent changes. *Australian Meteorological Magazine* **47**, 85-101.

Power, S., Folland, C., Colman, A., and V. Mehta, (1999) Inter-decadal modulation of the impact of ENSO on Australia. *Climate Dynamics*, **15**, 319-324.

Power, S.B., and R. Colman, (2006) Multi-year predictability in a coupled general circulation model. *Climate Dynamics*, **26**, 247-272.

Power, S., M. Haylock, R. Colman, and X. Wang, (2006) The predictability of interdecadal changes in ENSO and ENSO teleconnections. *Journal of Climate*, **8**, 2161-2180.

Quinn, W. and Neal, V. (1992). The historical record of El Niño events. *Climate Since A.D. 1500*. Bradley, R. and Jones, P. London, Routledge.**:** 623–648.

Quinn, T., Crowley, T., Taylor, F., Henin, C., Joannot, P. and Join, Y. (1998). A multicentury stable isotope record from a New Caledonia coral: Interannual and decadal SST variability in the southwest Pacific since 1657. *Paleoceanography* **13** (4): 412-426.

Rasmusson, E.M. and Carpenter, T. (1982). Variations in tropical sea surface temperature and surface wind fields associated with the Southern Oscillation/El Niño. *Monthly Weather Review* **110**: 354-384.

Rasmusson, E.M., Wang, X. and Ropelewski, C.F. (1995) Secular variability of the ENSO cycle. In: Martinson, D.G. et al (eds) National Academy Press, Washington, D.C. USA, pp 458-469.

Stahle, D., D' Arrigo, R., Krusic, P., Cleaveland, M., Cook, E., Allan, R., Cole, J., Dunbar, R., Therrell, M., Gay, D., Moore, M., Stokes, M., Burns, B., Villanueva-Diaz, J. and Thompson, L. (1998). Experimental dendroclimatic reconstruction of the Southern Oscillation. *Bulletin of the American Meteorological Society* **79** (10): 2137-2152.

Thompson, L., Mosley-Thompson, E. and Morales Arnao, B. (1984). El Niño-Southern Oscillation events recorded in stratigraphy of the tropical Quelccaya Ice Cap. *Science* **276**: 50-52.

Thompson, L., Mosley-Thompson, E., Bolzan, J. and Koci, B. (1985). A 1500 year record of tropical precipitation in ice-cores from the Quleccaya ice cap, Peru. *Science* **229**: 971-973.

Thompson, L., Henderson, K., Mosley-Thompson, E. and Lin, P. (2000). The Tropical Ice-core Record of ENSO. *El Niño and the Southern Oscillation: Multiscale Variability and Global and Regional Impacts.* H. Diaz and V. Markgraf. Cambridge, Cambridge University Press.

Torrence C. and Compo, G. (1998). A practical guide to wavelet analysis, *Bulletin American Met Soc*, **79**, 61-78

Trenberth, K. and Hoar, T. (1996). The 1990-1995 El Niño Southern Oscillation event: Longest on Record. *Geophysical Research Letters* **23** (1): 57-60.

Trenberth, K. and Hoar, T. (1997). El Niño and climate change. *Geophysical Research Letters* **24** (23): 3057-3060.

Trenberth, K. (1997) The definition of El Niño. *Bulletin of American Met. Society* **78**, 2771-2777

Trenberth, K. and Caron, J. (2000). The Southern Oscillation revisited: sea level pressures, surface temperatures, and precipitation. *Journal of Climate* **13**: 4358-4365.

Trenberth, K. and Stepaniak, D. (2001). Indices of El Niño evolution. *Journal of Climate* **14**: 1697-1701.

Troup, A. (1965). The Southern Oscillation. *Quarterly Journal of the Royal Meteorology Society*. **91**: 490-506.

Von Storch, H., Zorita, E., Jones, J., Dimitriev, Y., Gonza´lez-Rouco, F. and Tett, S. (2004). Reconstructing Past Climate from Noisy Data. *Nature* **306**: 679-682.

Whetton, P. and Rutherfurd, I. (1994). Historical ENSO teleconnections in the Eastern Hemisphere. *Climatic Change* **28**: 221-253.

TABLES

Table 1. Proxy Data used in this study

(a) Fowler et al. (2000; 2004; 2008), Gergis et al. (2005, b), Fowler (2008).
(b) Fenwick (2003).
(c) Cleaveland et al.(2003). Stahle and Cleaveland, (2002), IGBP Pages/WDC-A for Paleoclimatology Contribution Series 2002-004. Greybill (1994) IGBP Pages/WDC-A for Paleoclimatology Contribution Series 1994-003. Grissino-Mayer and Swetnam (1992) IGBP Pages/WDC-A for Paleoclimatology Contribution Series 1992-012. Dean (1993) IGBP Pages/WDC-A for Paleoclimatology Contribution Series 1993-021. Grow (2000) IGBP Pages/WDC-A for Paleoclimatology Contribution Series 2003-094
(d) Stahle et al., (1998), IGBP Pages/WDC-A for Paleoclimatology Contribution Series 2002-004.
(e) D' Arrigo et al.(1994) IGBP Pages/WDC-A for Paleoclimatology Contribution Series 1999-063.
(f) Hendy et al. (2002), IGBP Pages/WDC-A for Paleoclimatology Contribution Series 2002-009.
(g) Quinn et al. (1998), IGBP Pages/WDC-A for Paleoclimatology Contribution Series 1999-003.
(h) Linsley et al. (2000), IGBP Pages/WDC-A for Paleoclimatology Contribution Series 2000-065.
(i) Thompson et al. (1992) IGBP Pages/WDC-A for Paleoclimatology Contribution Series Data Contribution Series 1992-008.

Proxy Record	Dates (A.D.)	Latitude/Longitude	Data Smoothing	ENSO Zone	Proxy Variable
Tree-rings					
New Zealand Kauri [a]	1525-2002	35°-37°S, 173°-175°E	200-year Spline	West Pacific	Total Ring Widths
New Zealand Pink Pine [b]	1525-1998	42°-47°S, 167°-174°E	200-year Spline	West Pacific	Total Ring Widths
Mexican Douglas Fir [c]	1525-1998	19°-30°S, 97°–108°W	200-year Spline	East Pacific	Total Ring Widths
SW USA Pinyon Pine [d]	1525-2000	33°-37°S, 106°-112°W	200-year Spline	East Pacific	Total Ring Widths
Indonesian Teak [e]	1841-1995	8°S, 113°E	200-year Spline	West Pacific	Total Ring Widths
Coral					
Great Barrier Reef [f]	1612-1985	20°S, 147°E	3-year Gaussian	West Pacific	Luminescence SSS
New Caledonia [g]	1658-1992	22°S, 166°E	3-year Gaussian	West Pacific	δO18 SST
Rarotonga [h]	1727-1997	21°S, 159°W	3-year Gaussian	Central Pacific	δO18 SST
Ice					
Quelccaya Ice-core [i]	1525-1984	14°S, 71°W	3-year Gaussian	East Pacific	δO18

Table 2. Proportion (fraction) of variance explained.

Fraction of variance explained in EOFS 1 to 5 from the synthetic proxy correlation matrix. Synthetic proxies were constructed using a synthetic signal defined in Equation 3 and embedded in red noise according to Equation 5. Results represent the mean of 1000 samples. Alpha weightings of 0.2 to 0.4 are statistically similar to the real proxy data.

Signal Weight	Noise Weight	Proportion of Variance Explained				
α	$\sqrt{(1-\alpha^2)}$	EOF1	EOF2	EOF3	EOF4	EOF5
0.0	1.00	0.16	0.15	0.14	0.13	0.12
0.1	0.99	0.16	0.15	0.14	0.13	0.12
0.2	0.98	0.17	0.15	0.13	0.13	0.12
0.3	0.95	0.21	0.14	0.13	0.12	0.11
0.4	0.92	0.27	0.13	0.12	0.11	0.10
0.5	0.87	0.35	0.11	0.11	0.10	0.09
0.6	0.80	0.44	0.10	0.09	0.08	0.08
0.7	0.71	0.55	0.08	0.07	0.07	0.06
0.8	0.60	0.68	0.06	0.05	0.05	0.04
0.9	0.44	0.83	0.03	0.03	0.03	0.02

Table 3. Correlation between Signal, EOF1 and $\hat{S}$ from multiple linear regression.

Comparison of correlation and RMS error for synthetic observations and the reconstructed signal $\hat{S}$ (multiple regression predictand) and EOF1 (from the synthetic proxy correlation matrix). The multiple regression model and EOF1 have been constructed using the full 500 years of synthetic data. Synthetic tests were carried out 1000 times, with R and RMS values representing mean values and uncertainties (standard deviation) calculated from resampling. Synthetic proxies were constructed using a synthetic signal defined in Equation 3 and embedded in red noise according to Equation 5. Results using periodic autocorrelation for the construction of the synthetic signal (Equation 4) are similar. Results are shown for a range of realistic signal weighting in the proxies (alpha).

Correlation with Signal (Years 1-500)				
Signal Weight	**Correlation (R)**		**RMS Error**	
α	**$\hat{S}$ (Multiple Regression)**	**EOF1**	**$\hat{S}$ (Multiple Regression)**	**EOF1**
0.2	**0.54 ± 0.03**	**0.43 ± 0.18**	**0.84 ± 0.03**	**1.16 ± 0.14**
0.3	**0.69 ± 0.03**	**0.68 ± 0.03**	**0.72 ± 0.03**	**0.98 ± 0.04**
0.4	**0.80 ± 0.02**	**0.79 ± 0.02**	**0.60 ± 0.03**	**0.95 ± 0.03**

Table 4. Correlation between Signal, EOF1 and $\hat{S}$ from multiple linear regression for the 'hindcast' period.

Comparison of correlation and RMS error for synthetic observations and the reconstructed signal $\hat{S}$ (multiple regression predictand) and EOF1 (from the synthetic proxy correlation matrix). $\hat{S}$ has been constructed using the last 100 years (years 401-500) of the synthetic proxies and observations in a multiple regression model to predict the first 400 years. EOF1 has been constructed using the normalised covariance matrix of the synthetic proxies using the full 500 years of synthetic data. Correlations are shown for years 1-400. Synthetic tests were carried out 1000 times, with R and RMS values representing mean values and uncertainties (standard deviation) calculated from resampling. Synthetic proxies were constructed using a synthetic signal defined in Equation 3 and embedded in red noise according to Equation 5. Results using periodic autocorrelation for the construction of the synthetic signal (Equation 4) are similar. Results are shown for a range of realistic signal weighting in the proxies (alpha).

Correlation with Signal (Years 1-400)							
Signal Weight		**Correlation (R)**			**RMS Error**		
α		**$\hat{S}$ (Multiple Regression)**	**EOF1**		**$\hat{S}$ (Multiple Regression)**	**EOF1**	
0.2		**0.46 ± 0.05**	**0.44 ± 0.18**		**0.81 ± 0.04**	**1.04 ± 0.13**	
0.3		**0.65 ± 0.04**	**0.68 ± 0.03**		**0.69 ± 0.03**	**0.88 ± 0.04**	
0.4		**0.77 ± 0.02**	**0.79 ± 0.02**		**0.58 ± 0.03**	**0.85 ± 0.03**	

Table 5. Synthetic data test for the 'principal component regression' approach to calibrating proxy records.

Correlations and RMS errors between EOF1 from the synthetic proxy network and the synthetic signal for realistic signal weightings (alpha coefficients) are shown. Also shown are correlations and RMS errors for a multiple regression model that fits EOF1 plus 1 to 4 additional white noise timeseries, following Equation 7. Thus, column N1 represents a multiple regression model using EOF1 and a single additional white noise timeseries, while N4 represents a multiple regression model using EOF1 and 4 additional white noise timeseries. Correspondence results are shown for the calibration period (the last 100 years of the 500 year timeseries) and the hindcast period (the first 400 years). Results represent the mean of 1000 iterations, with the uncertainty (standard deviation) calculated from resampling. The synthetic signal has been constructed using Equation 3 and the proxies using Equation 5. While the addition of additional white noise marginally improves the performance of the regression model during the calibration period (noise fitting) it does not improve (marginally degrades) the model for the hindcast period. Results using periodic autocorrelation for the construction of the synthetic signal (Equation 4) are similar.

Table 5a.

Calibration Period (Years 401-500) Multiple Regression Correspondence									
Correlation with Signal									
Signal Strength		Single Mode		Additional Noise Fitting					Uncert.
α		EOF1		N1	N2	N3	N4		SDEV
0.20		0.46		0.47	0.48	0.49	0.50		±0.10
0.30		0.68		0.68	0.69	0.69	0.69		±0.06
0.40		0.79		0.79	0.79	0.80	0.80		±0.04
RMS Error									
Signal Strength		Single Mode		Additional Noise Fitting					Uncert.
α		EOF1		N1	N2	N3	N4		SDEV
0.20		0.39		0.39	0.39	0.38	0.38		±0.03
0.30		0.32		0.32	0.32	0.32	0.32		±0.02
0.40		0.27		0.27	0.27	0.26	0.26		±0.02

Table 5b.

Hindcast Period (Years 1-400) Multiple Regression Correspondence									
Correlation with Signal									
Signal Strength		Single Mode		Additional Noise Fitting					Uncert.
α		EOF1		N1	N2	N3	N4		SDEV
0.20		0.46		0.45	0.44	0.44	0.43		±0.08
0.30		0.68		0.67	0.67	0.66	0.66		±0.03
0.40		0.79		0.79	0.79	0.78	0.78		±0.02
RMS Error									
Signal Strength		Single Mode		Additional Noise Fitting					Uncert.
α		EOF1		N1	N2	N3	N4		SDEV
0.20		0.79		0.80	0.80	0.81	0.81		±0.04
0.30		0.66		0.67	0.67	0.67	0.68		±0.03
0.40		0.55		0.55	0.56	0.56	0.56		±0.02

Table 6. Proportion of variance explained (fraction) in the first four principal components (EOFs) of multiproxy data.

Multiproxy Data Set	EOF1	EOF2	EOF3	EOF4
R8 (1727 – 1982)	0.23	0.15	0.13	0.12
R5 (1525 – 1982)	0.30	0.21	0.19	0.17

Table 7. Eigenvector (EOF1) loading values.

The annual correlation between individual proxy chronologies and the EOF1 timeseries (shown with lead and lag time of one year).

Individual Proxy records represented in the R8 Multiproxy Dataset	EOF Loading (eigenvector)	Correlation		
PC 1		t-1	t	t+1
New Zealand Kauri	**0.34**	**0.04**	**0.46**	**0.10**
New Zealand Pink Pine	**-0.39**	**-0.07**	**-0.53**	**-0.23**
Mexican Douglas Fir	**0.53**	**0.20**	**0.71**	**0.17**
SW USA Pinyon Pine	**0.40**	**0.18**	**0.54**	**0.04**
Quelccaya Ice-core	**0.33**	**0.33**	**0.45**	**0.27**
Great Barrier Reef	**-0.18**	**-0.11**	**-0.24**	**0.07**
Rarotonga	**0.34**	**0.22**	**0.46**	**0.29**
New Caledonia	**0.18**	**0.14**	**0.25**	**0.23**

Table 8. Correlations between the annual R8 EOFs and seasonal SOI, N3.4 and CEI indices over the 1871-1982 reference period.

Correlations are shown for synchronous years (t = 0),a as well as at lead (t = +1) and lag (t = -1) with observed indices. Annual lead/lag correlations are shown, with maximum correlations highlighted. Note that the DJF season represents the year in which the month of January falls.

	EOF1			EOF2			EOF3			EOF4		
JJA	**t-1**	**t=0**	**t+1**	**t-1**	**t=0**	**t+1**	**t-1**	**t=0**	**t+1**	**t-1**	**t=0**	**t+1**
CEI	**-0.04**	**-0.40**	**-0.54**	**-0.01**	**0.02**	**0.00**	**0.03**	**-0.17**	**-0.16**	**0.21**	**0.18**	**0.03**
SOI	**-0.04**	**-0.38**	**-0.51**	**0.04**	**0.06**	**-0.02**	**0.09**	**-0.11**	**-0.14**	**0.23**	**0.22**	**0.05**
N3.4	**0.14**	**0.59**	**0.12**	**0.14**	**0.13**	**0.16**	**0.05**	**0.16**	**0.12**	**-0.19**	**-0.03**	**0.01**
SON	**t-1**	**t=0**	**t+1**	**t-1**	**t=0**	**t+1**	**t-1**	**t=0**	**t+1**	**t-1**	**t=0**	**t+1**
CEI	**0.02**	**-0.22**	**-0.67**	**-0.01**	**0.06**	**0.05**	**0.06**	**-0.10**	**-0.20**	**0.17**	**0.25**	**0.06**
SOI	**0.00**	**-0.19**	**-0.62**	**-0.02**	**0.07**	**0.01**	**0.13**	**-0.05**	**-0.16**	**0.14**	**0.27**	**0.05**
N3.4	**0.05**	**0.37**	**0.49**	**0.06**	**0.02**	**-0.02**	**0.03**	**0.20**	**0.17**	**-0.17**	**-0.12**	**-0.01**
DJF	**t-1**	**t=0**	**t+1**	**t-1**	**t=0**	**t+1**	**t-1**	**t=0**	**t+1**	**t-1**	**t=0**	**t+1**
CEI	**-0.19**	**-0.72**	**0.00**	**0.01**	**0.00**	**-0.12**	**-0.03**	**-0.18**	**-0.11**	**0.31**	**0.08**	**0.06**
SOI	**-0.17**	**-0.69**	**-0.05**	**0.04**	**0.01**	**-0.13**	**0.00**	**-0.17**	**-0.09**	**0.32**	**0.11**	**0.04**
N3.4	**0.24**	**0.65**	**-0.11**	**-0.05**	**-0.06**	**0.04**	**0.11**	**0.20**	**0.14**	**-0.24**	**-0.06**	**-0.07**
MAM	**t-1**	**t=0**	**t+1**	**t-1**	**t=0**	**t+1**	**t-1**	**t=0**	**t+1**	**t-1**	**t=0**	**t+1**
CEI	**-0.13**	**-0.63**	**-0.23**	**-0.05**	**-0.05**	**-0.13**	**0.01**	**-0.16**	**-0.12**	**0.25**	**0.09**	**0.02**
SOI	**-0.10**	**-0.57**	**-0.32**	**0.05**	**0.05**	**-0.08**	**0.08**	**-0.14**	**-0.11**	**0.28**	**0.16**	**0.04**
N3.4	**0.20**	**0.70**	**-0.03**	**0.02**	**0.01**	**0.11**	**0.05**	**0.17**	**0.11**	**-0.27**	**-0.06**	**-0.07**

Table 9. Correlation (R) and proportion of common variance (R^2) of reconstructions with observed from 1871-1982 for SOI, Niño 3.4 SST and CEI for each season.

ENSO Index	ENSO Index	Correlation R8	Proportion Variance (%) Explained R8	Correlation R5	Proportion Variance (%) Explained R5
CEI	CEI				
JJA (year-1)	JJA	0.54	0.29	0.48	0.23
SON (year-1)	SON	0.67	0.44	0.60	0.36
DJF	DJF	0.72	0.52	0.66	0.43
MAM	MAM	0.63	0.40	0.61	0.38
SOI	SOI				
JJA (year-1)	JJA	0.51	0.26	0.44	0.19
SON (year-1)	SON	0.62	0.38	0.55	0.30
DJF	DJF	0.69	0.47	0.62	0.38
MAM	MAM	0.57	0.33	0.54	0.29
Niño3.4 SST	Niño3.4 SST				
SON (year-1)	SON	0.49	0.24	0.46	0.21
DJF	DJF	0.65	0.42	0.59	0.34
MAM	MAM	0.70	0.48	0.64	0.40
JJA	JJA	0.59	0.35	0.59	0.34

FIGURES

Figure 1. Location of proxy records used in this study shown with regard to El Niño teleconnection characteristics.

Rainfall anomalies are represented light grey shading (dry) and dark shading (wet). Temperature anomalies indicated by ‘c’ (cool) and ‘w’ (warm) annotation. 'T' denotes tree-ring chronologies (squares), 'C' coral sequences (circles) and 'I' ice-core data (triangle). Note that R5 (1525) proxies are unshaded and the R8 (1727) proxies are grey hatched. Details of each record are provided in Table 1. Note that the El Niño teleconnection base-map is adapted from Allan *et al.* (1996).

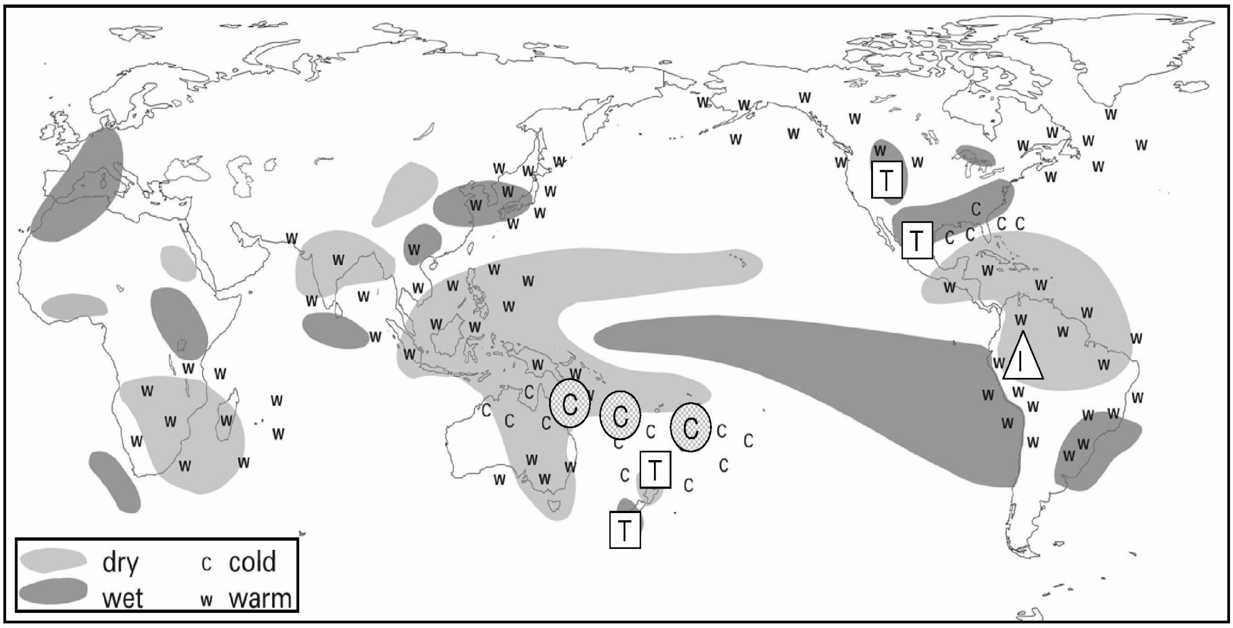


Figure 2. Timeseries of instrumental DJF CEI, DJF SOI and MAM N3.4 SST scaled to R8 Proxy-ENSO index over the period 1871-1982.

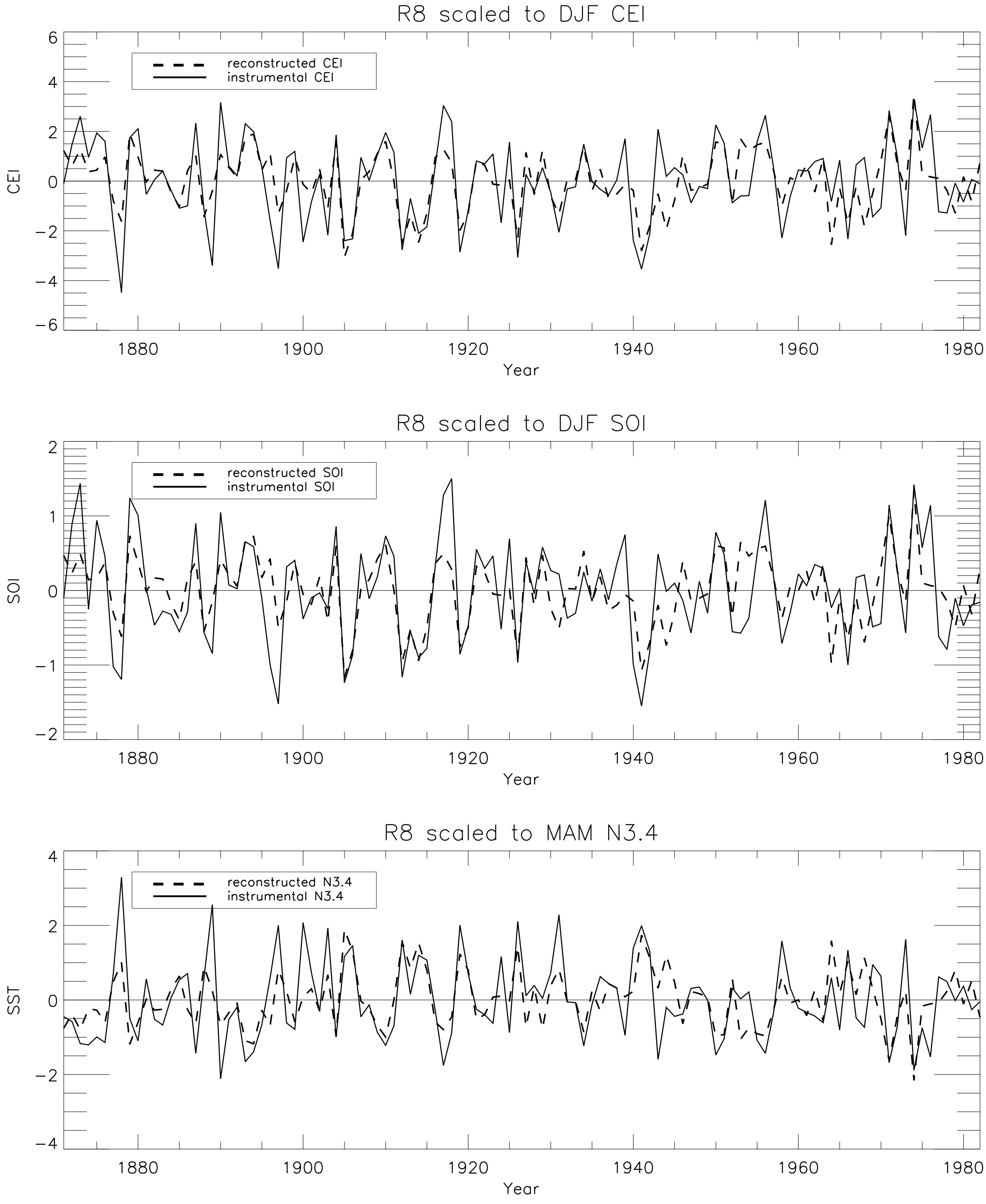


Figure 3. Frequency distribution of R8 and R5 residual errors for the period 1871-1982 for seasonal SOI reconstruction versus instrumental observations.

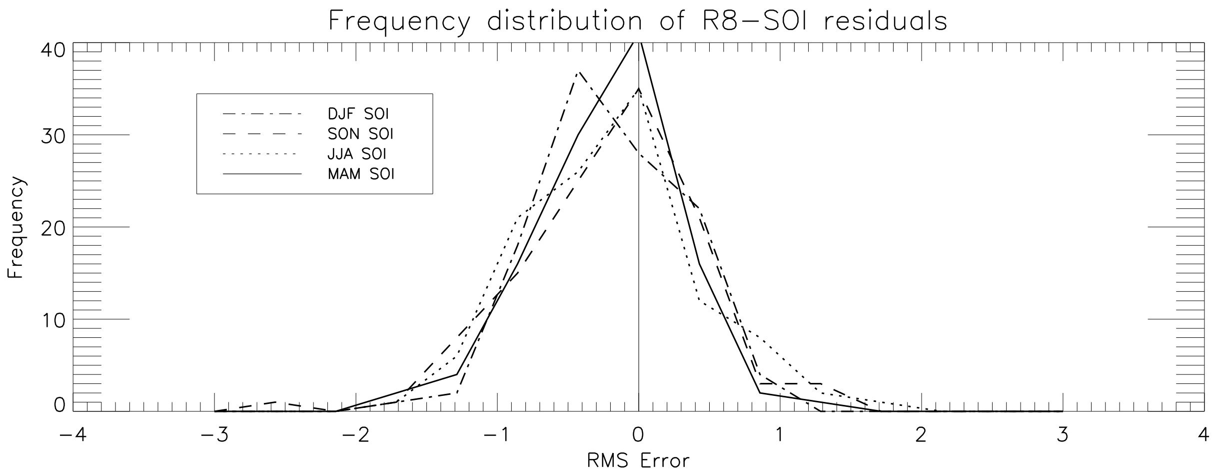


Figure 4. Power spectrum for normalised instrumental DJF SOI from 1871-1982. Background AR1 red noise (dashed line) and the 95% significance level relative (dotted line) are also shown. Effective bandwidth after smoothing = 2.66/N cycles/year.

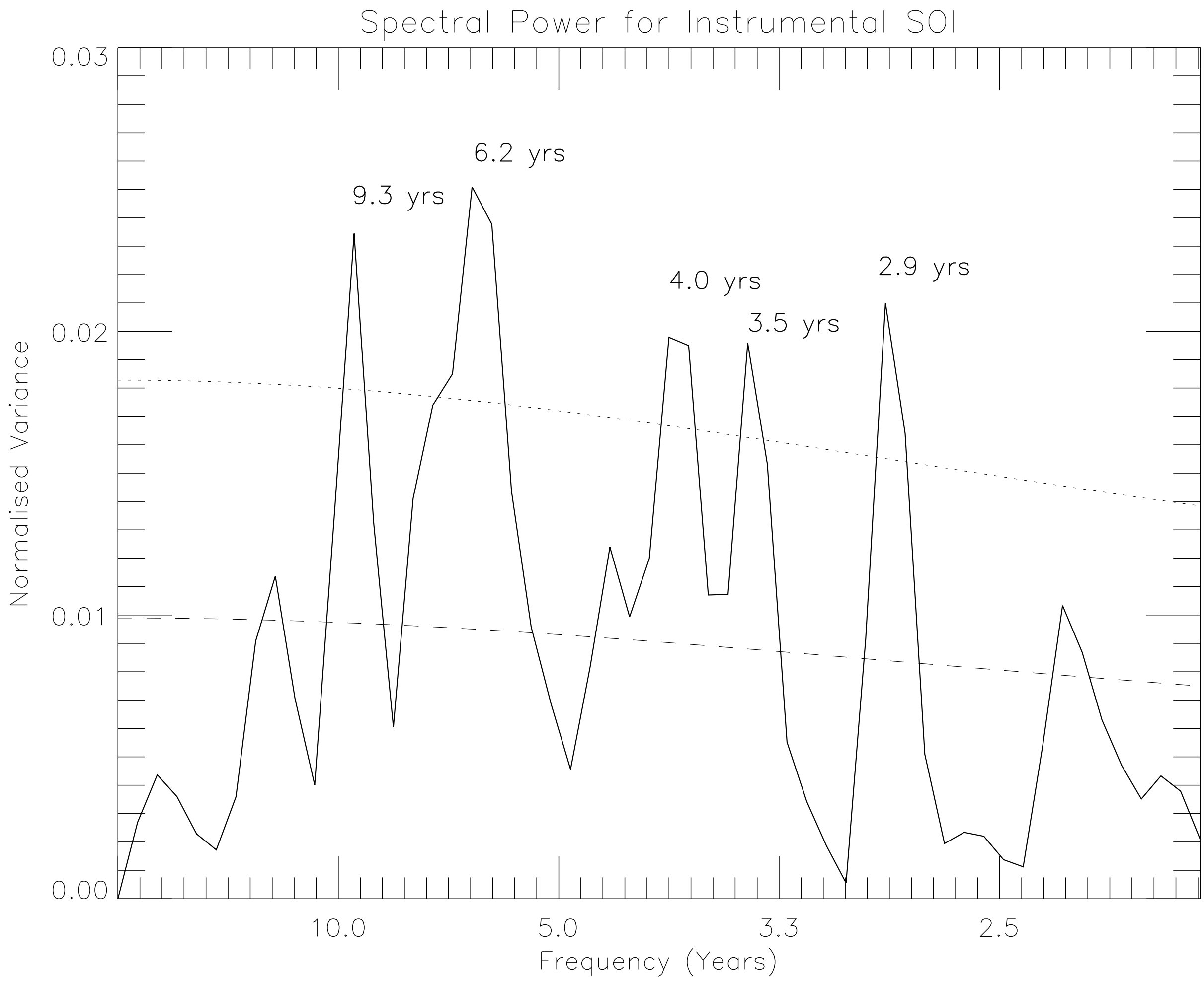


Figure 5. Power spectrum (un-normalised variance) for R8 Proxy-ENSO index (EOF1) for the period 1727-1982.

Significance at the 90% and 95 % (dotted lines) level is indicated relative to estimated background AR1 noise (solid line). Effective Bandwidth after smoothing = 3.2/N cycles/year.

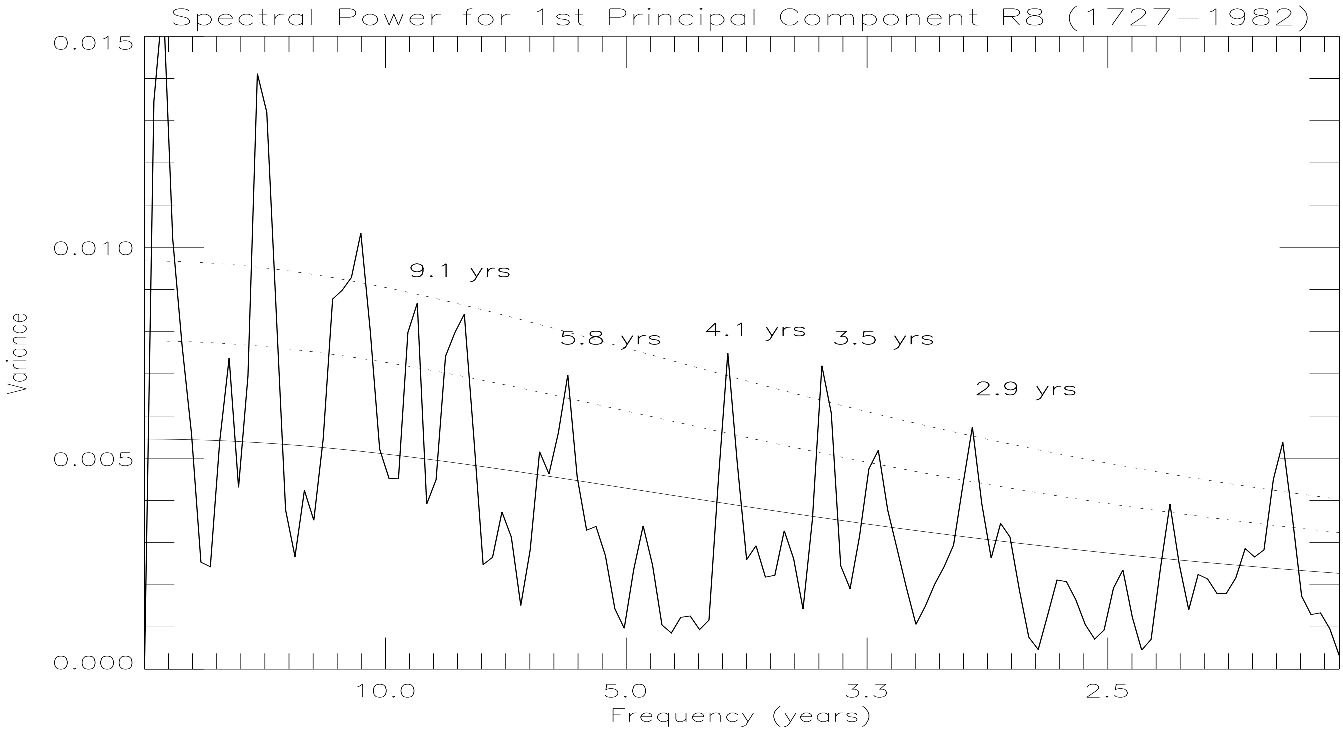


Figure 6. Power spectrum for instrumental DJF SOI and proxy-ENSO indices R5 and R8 for the period 1871-1982.

Significance at the 90% and 95 % (dotted lines) level is indicated relative to estimated background AR1 noise for the instrumental SOI. Effective Bandwidth after smoothing = 2.66/N cycles/year.

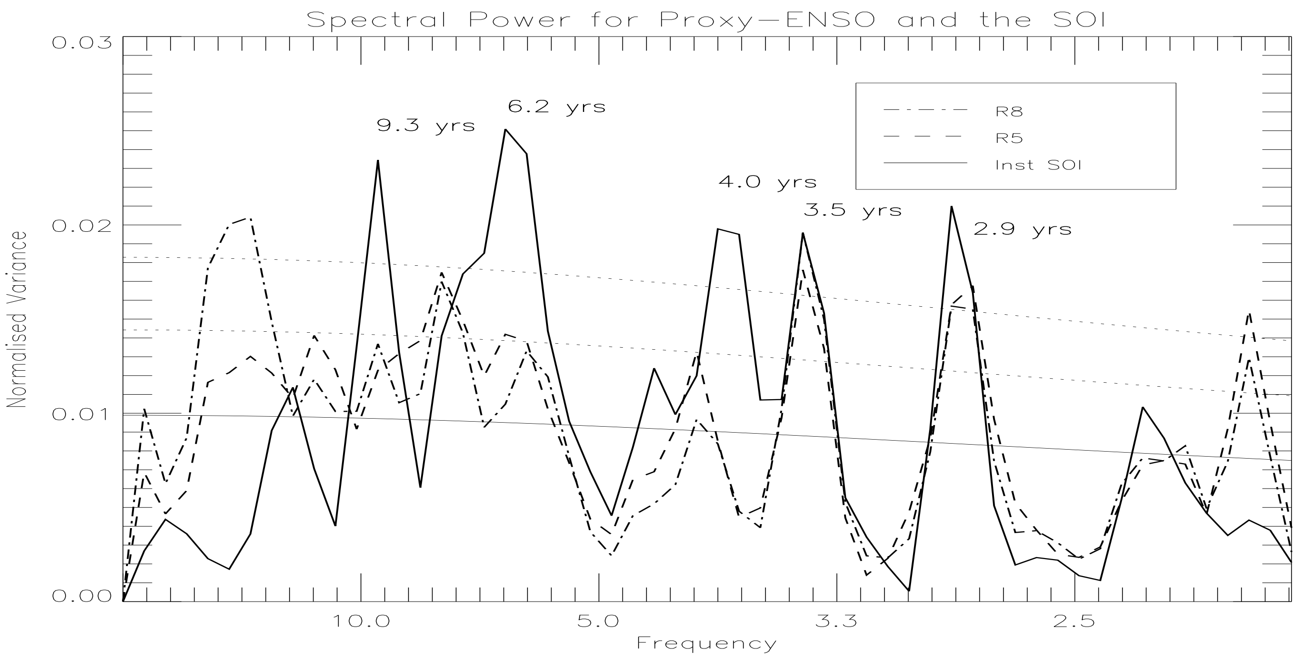


Figure 7. Coherency Squared for proxy-ENSO indices R5 and R8 and instrumental SOI for the period 1871-1982. 95% significance is indicated (dotted line).

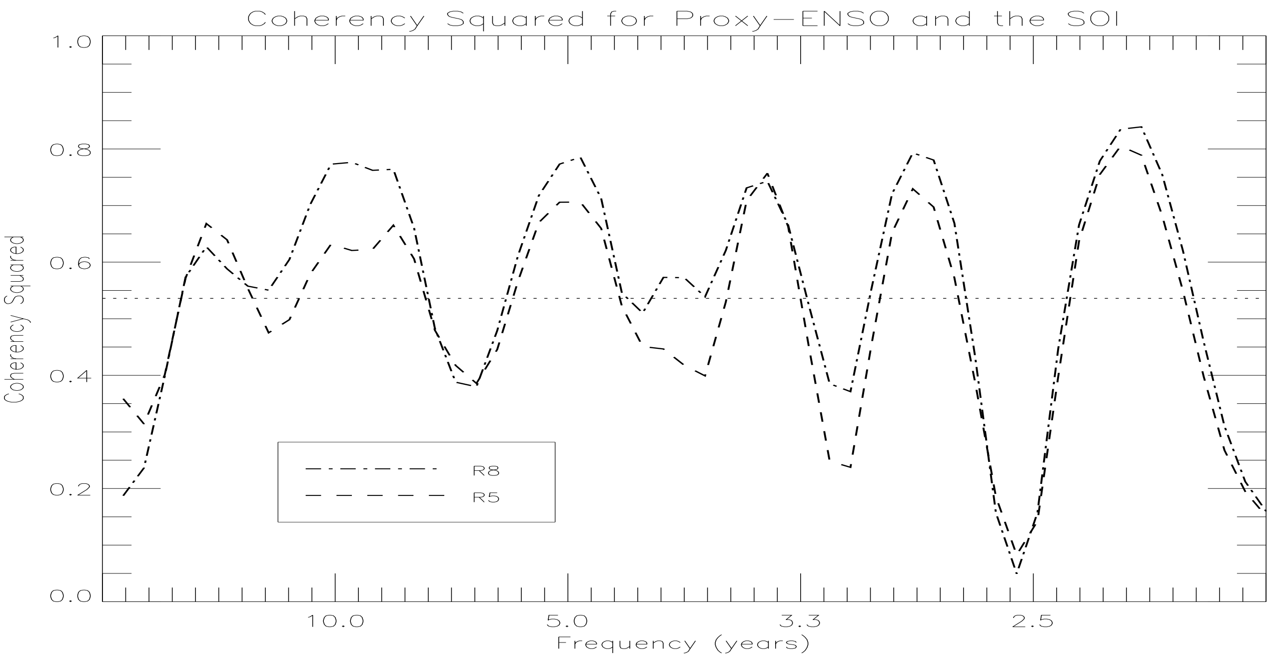


Figure 8. 10-year moving variability (standard deviation) for instrumental DJF SOI and Proxy-ENSO indices R5 and R8 for the period 1871-1982.

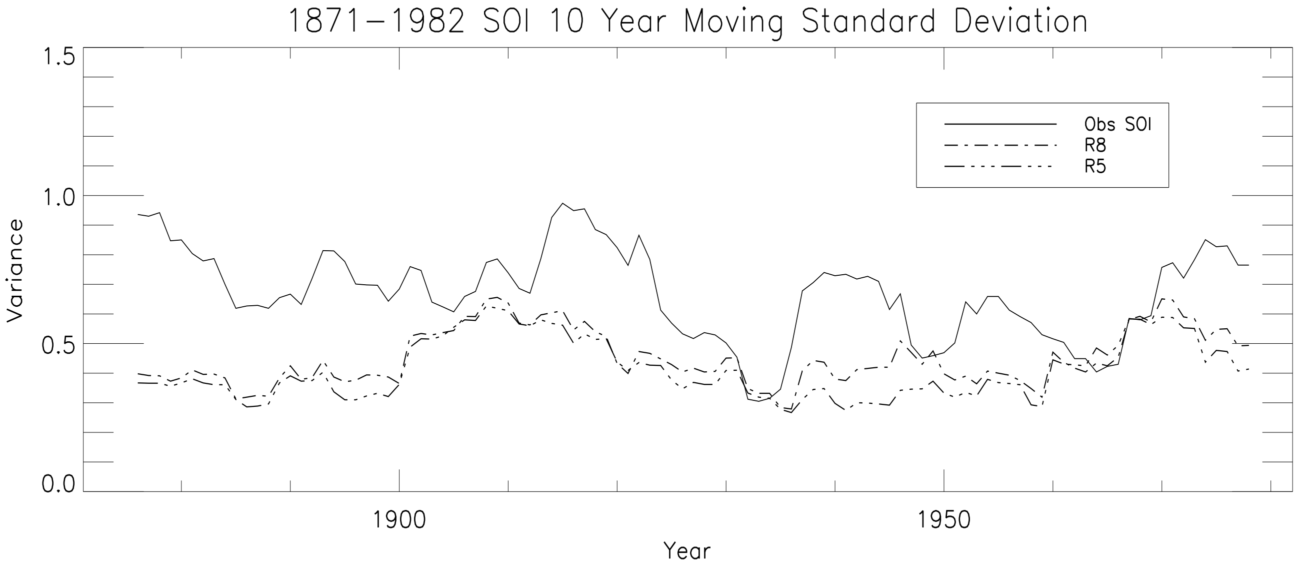


Figure 9. El Niño (top panel) and La Niña (bottom panel) events recorded since 1875 in the SOI, R5 and R8 proxy-ENSO indices.

Events are represented as solid dark grey squares. Events are defined as episodes (years) which are less than a single standard deviation (El Niño) or greater than a single standard deviation (La Niña) relative to variability from 1875-1982 (SOI), 1525-1982 (R5) and 1727-1982 (R8). Light grey squares represent weak events that did not reach the threshold in the SOI. Smaller black squares represent 'false positive' events in the R5 and R8 proxy-ENSO indices.

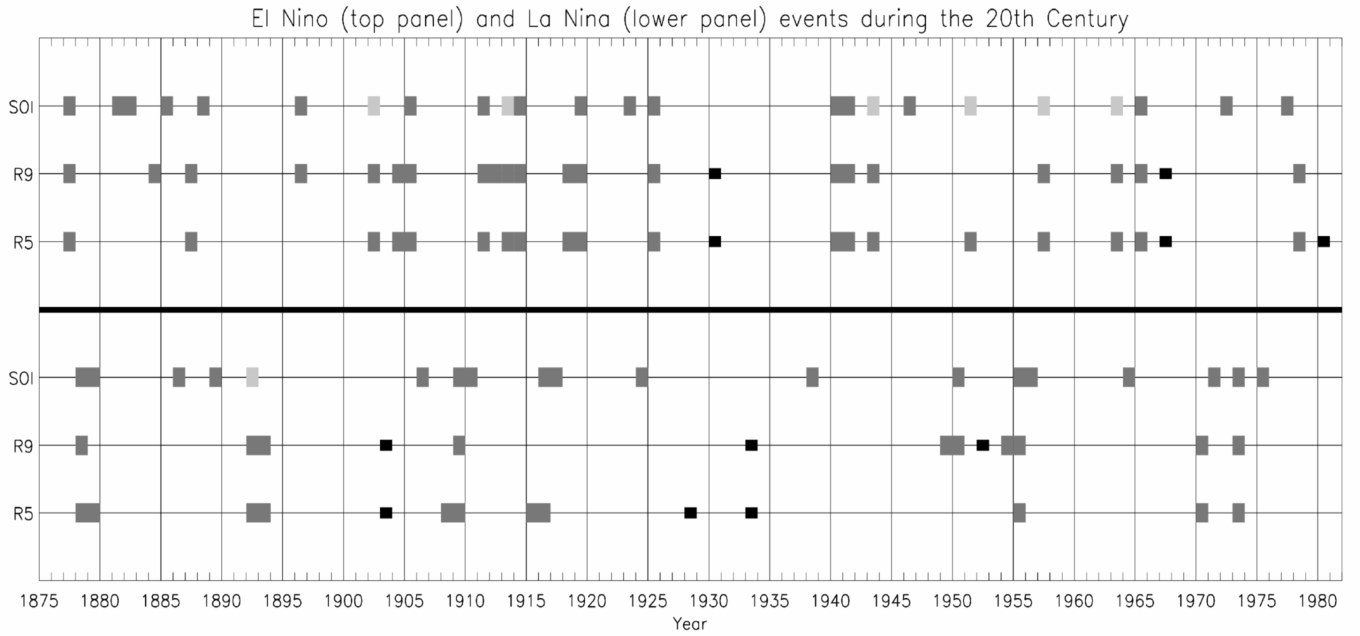
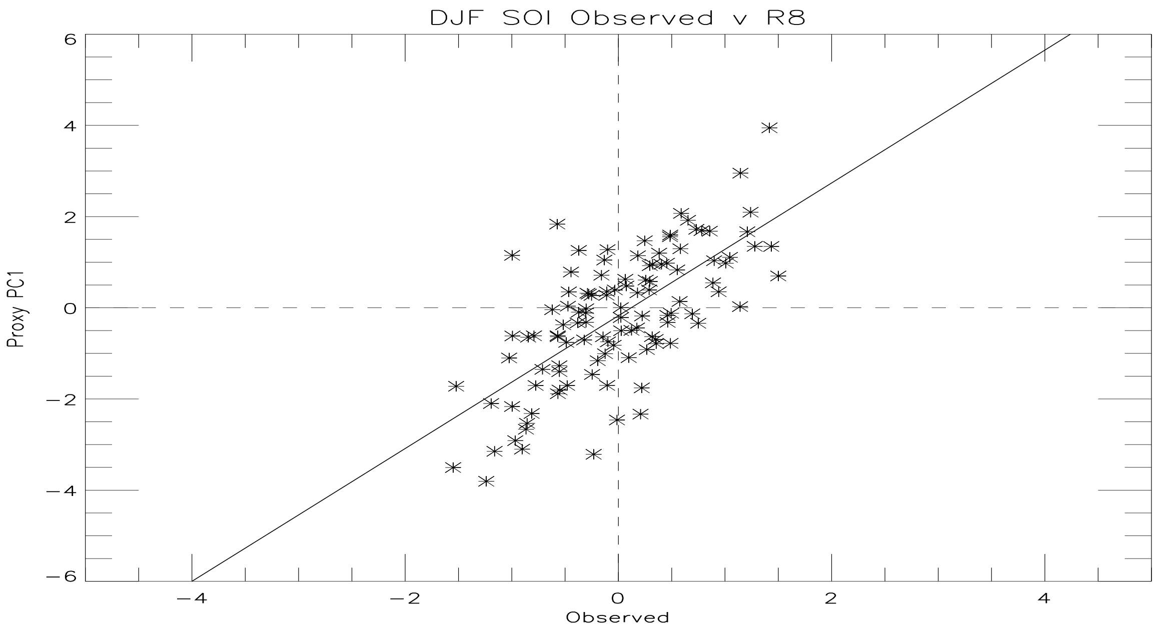


Figure 10. Scatter plot of Proxy-ENSO index (EOF1) R8 with observed DJF SOI and CEI for 1871-1982.



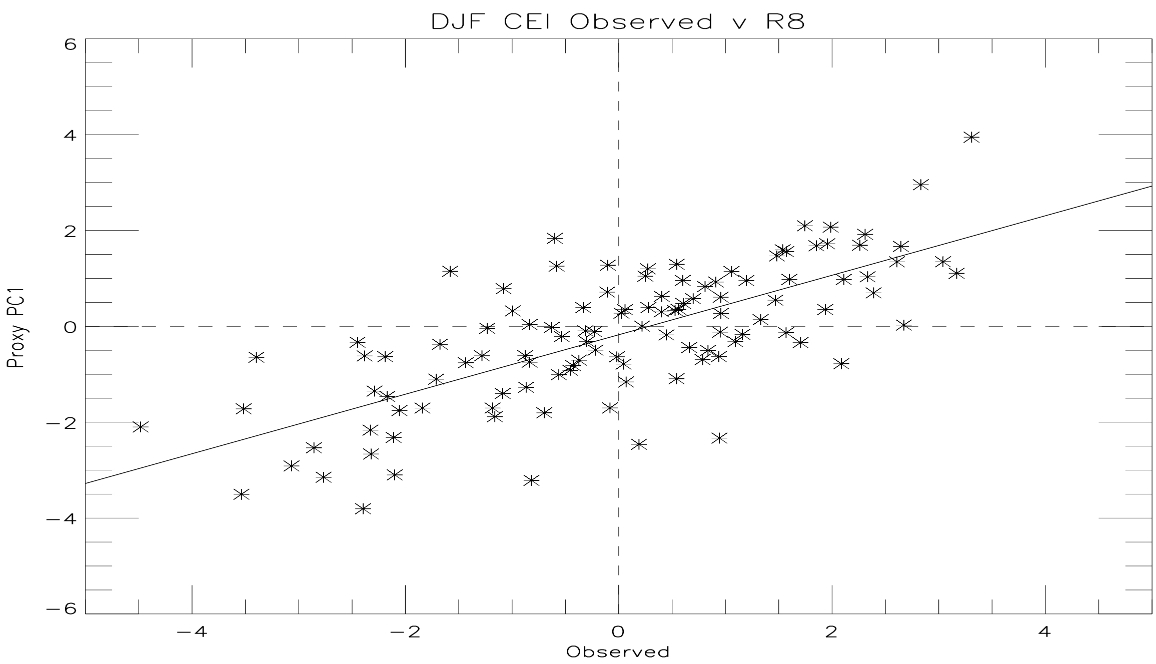


Figure 11. Low frequency ENSO from 1876-1982.

Proxy-ENSO indices R5 and R8 are shown together with the low frequency component of the SOI [each timeseries is filtered using a low pass filter using the spectral method described by Power et al (1999)] and the negative phase of the IPO.

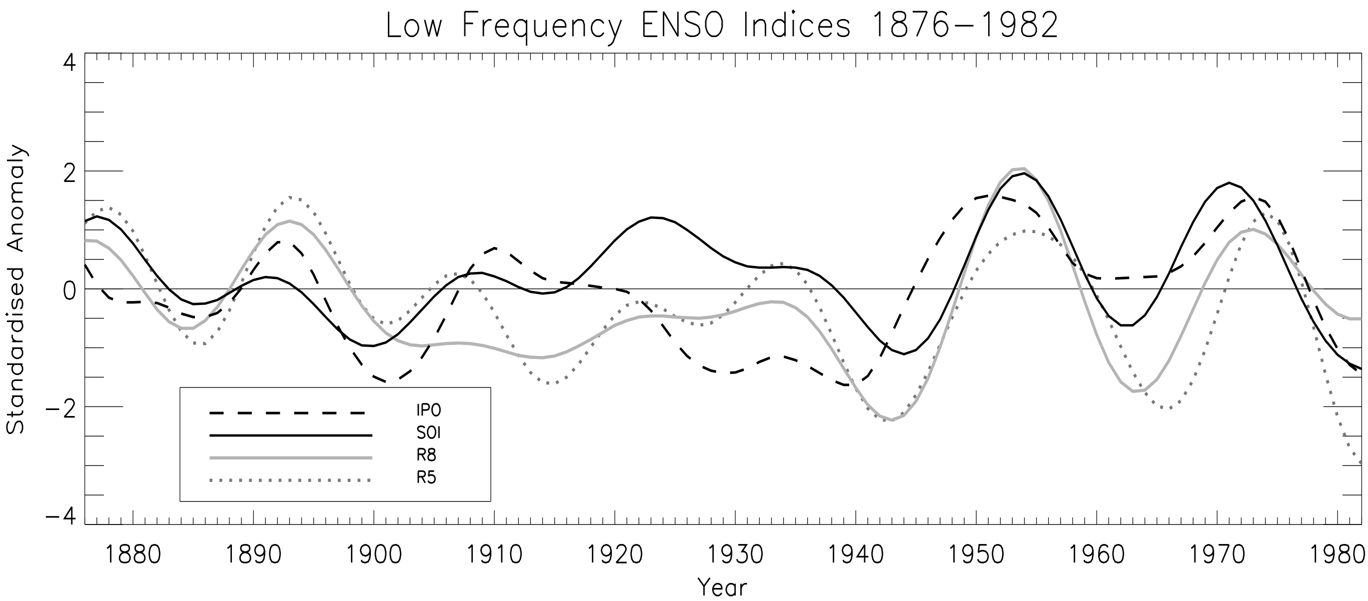


Figure 12. Proxy-ENSO indices R5 (dotted) and R8 (solid) scaled to DJF SOI for the period 1525-1982.

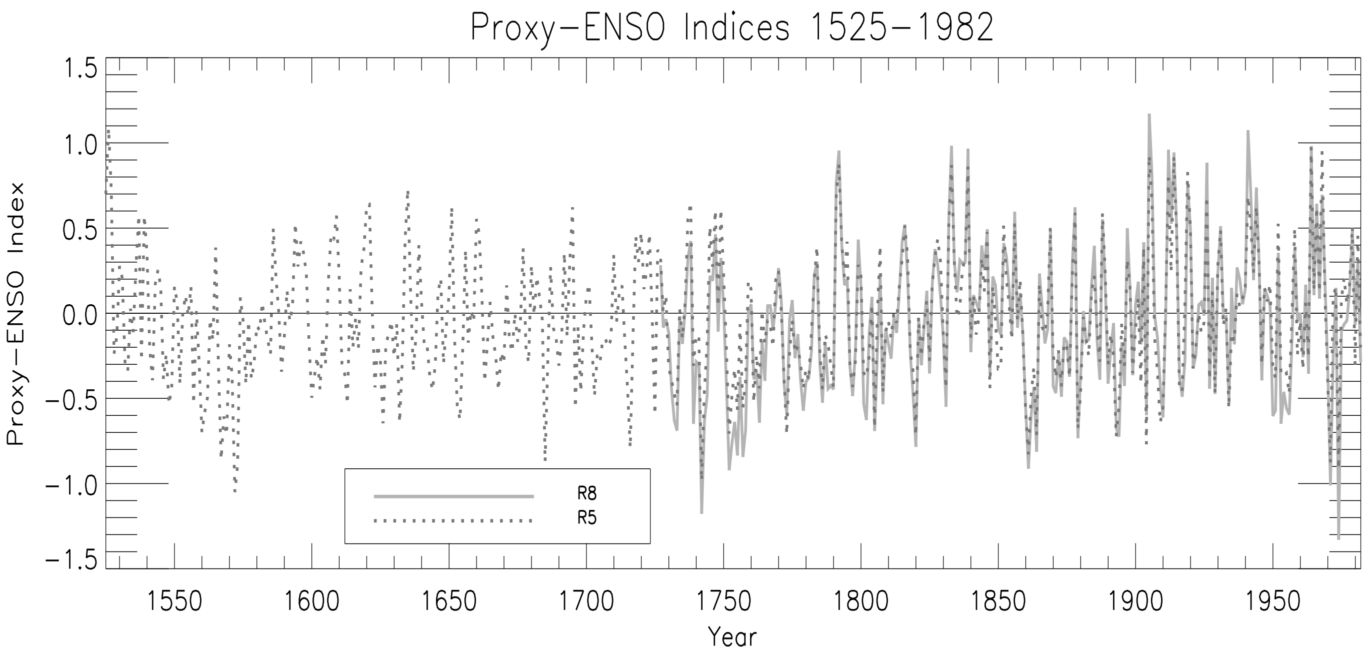


Figure 13. Low Frequency Proxy-ENSO indices R5 and R8 from 1525 to 1982.

Each timeseries is low-pass filtered using the spectral method described by Power et al (1999)

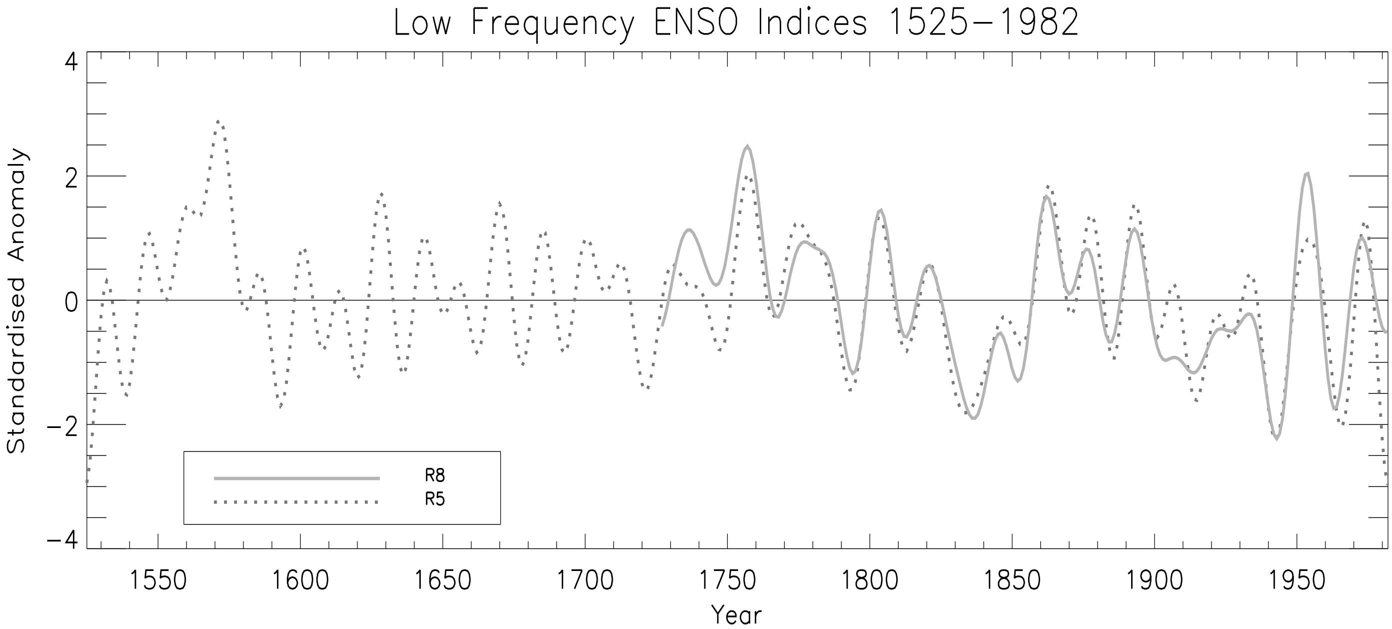


Figure 14. Hovmoller of the 50 year moving (window) power spectrum for Proxy-ENSO for 1525-1982.

Effective bandwidth after smoothing = 2.66/N cycles/year.

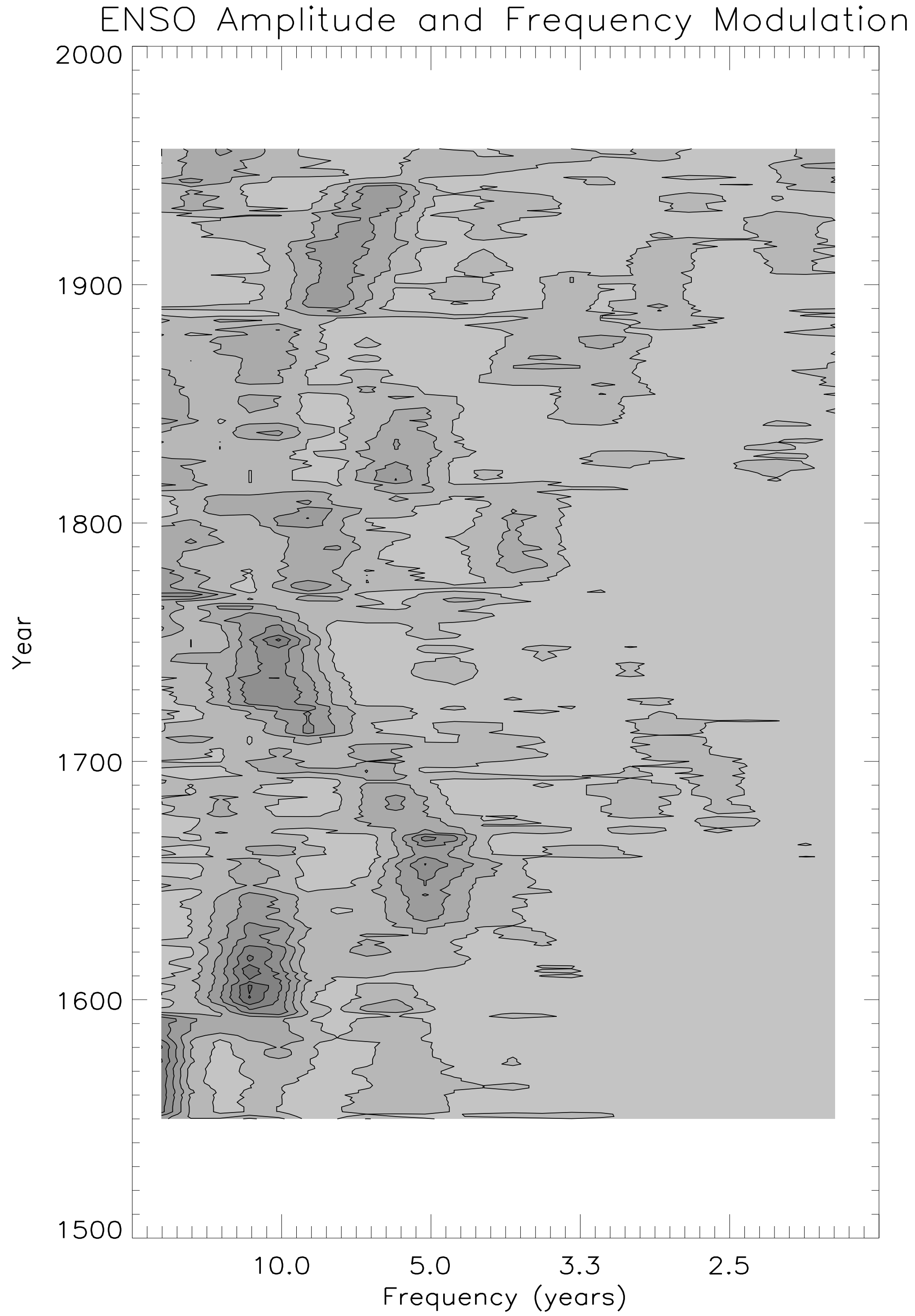


Figure 15. Variance (anomaly relative to long term mean) associated with frequency bands 2-4, 4-8 and 8-20 cycles/year, as well as total variance for Proxy-ENSO.

Spectral power was estimated in 50 year moving windows from 1525-1982.

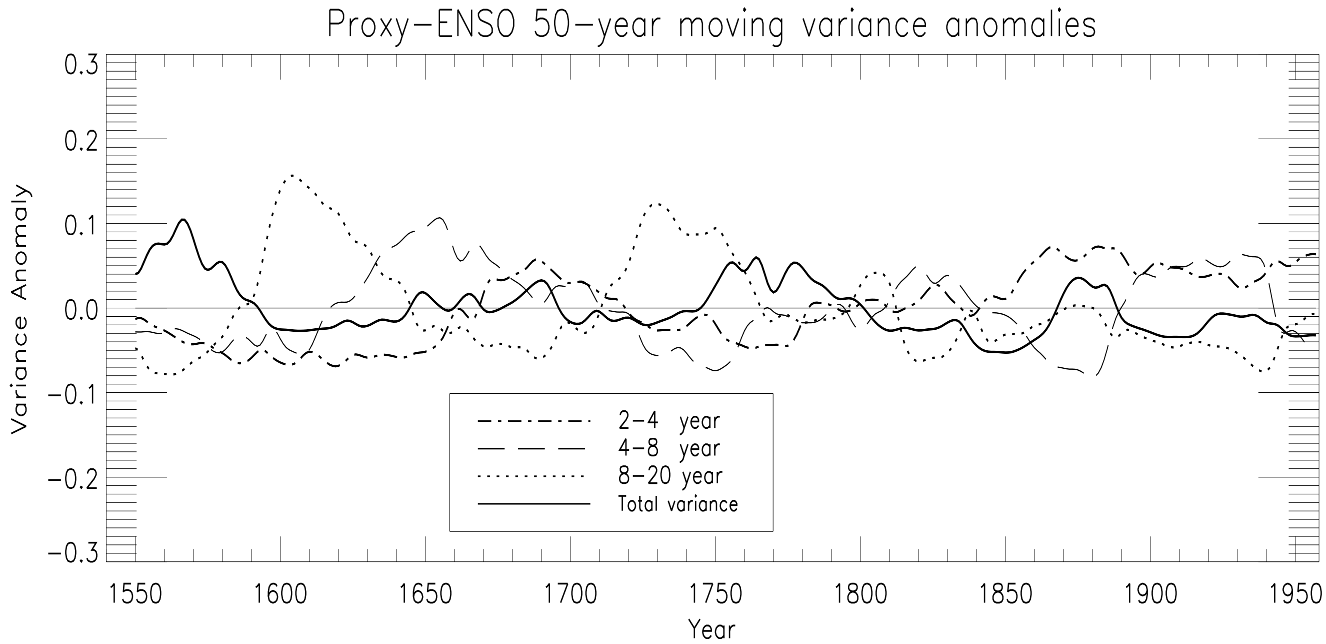
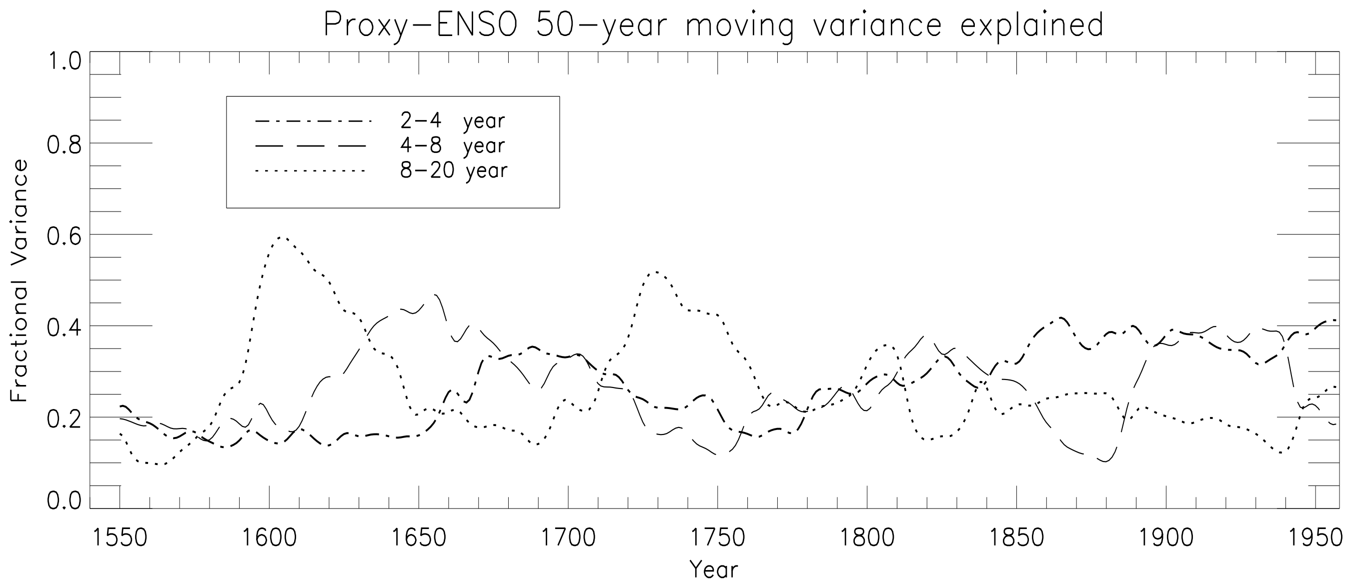


Figure 16. Fraction of variance in Proxy-ENSO associated with frequency bands 2-4, 4-8 and 8-20 cycles/year relative to total variance.

Spectral power was estimated in 50 year moving windows from 1525-1982.



The Centre for Australian Weather and Climate Research is a partnership between CSIRO and the Bureau of Meteorology.