AGU PUBLICATIONS

Geophysical Research Letters

RESEARCH LETTER
10.1002/2014GL062331

Key Points:
- Constraints on ENSO amplitude derived from proxy variance are uncertain
- Estimates of these uncertainties may be made from a simple statistical framework
- Defines proxy variance changes needed to robustly infer ENSO amplitude changes

Supporting Information:
- Readme
- Text S1
- Figure S1

Correspondence to:
T. Russon,
tom.russon@ed.ac.uk

Citation:
Russon, T., A. W. Tudhope, M. Collins, and G. C. Hegerl (2015), Inferring changes in ENSO amplitude from the variance of proxy records, *Geophys. Res. Lett.*, *42*, 1197–1204, doi:10.1002/2014GL062331.

Received 24 OCT 2014
Accepted 7 JAN 2015
Accepted article online 10 JAN 2015
Published online 19 FEB 2015

Inferring changes in ENSO amplitude from the variance of proxy records

T. Russon[1], A. W. Tudhope[1], M. Collins[2], and G. C. Hegerl[1]

[1]School of GeoSciences, University of Edinburgh, Edinburgh, UK, [2]College of Engineering, Mathematics and Physical Sciences, University of Exeter, Exeter, UK

Abstract One common approach to investigating past changes in El Niño–Southern Oscillation (ENSO) amplitude is through quantifying the variance of ENSO-influenced proxy records. However, a component of the variance of all such proxies will reflect influences that are unrelated to the instrumental climatic indices from which modern ENSO amplitudes are defined. The unrelated component of proxy variance introduces a fundamental source of uncertainty to all such constraints on past ENSO amplitudes. Based on a simple parametric approach to modeling this uncertainty, we present guidelines for the magnitudes of proxy variance change required to robustly infer the following: (i) any change at all in ENSO amplitude and (ii) a change in ENSO amplitude that exceeds the plausible range of unforced variability. It is noted that more extreme changes in proxy variance are required to robustly infer decreases, as opposed to increases, in past ENSO amplitude from modern levels.

1. Introduction

A common problem of interest to paleoclimatic studies is whether the amplitudes of phenomena observed in the modern system, such as the El Niño–Southern Oscillation (ENSO), have changed through time and if so, whether these changes exceed the range of unforced variability. Such studies have the potential to provide empirical constraints on the relationship between ENSO amplitude and the climatic mean state [*Guilyardi*, 2006; *Choi et al.*, 2011; *Cobb et al.*, 2003; *Collins et al.*, 2010; *Adams et al.*, 2003]. One approach to defining the modern amplitude of a climatic phenomena is through the variance of an instrumental "index" time series. For example, the variance of the interannual anomalies of eastern equatorial Pacific sea surface temperature (NINO3 SSTA) is widely used as a metric of ENSO amplitude [*Bellenger et al.*, 2014; *Watanabe et al.*, 2012; *Guilyardi et al.*, 2009]. For preinstrumental periods, variance-based constraints on ENSO amplitudes in the real (as opposed to modeled) climate system can only be obtained from paleoclimatic proxy records that are sensitive to one or more aspects of the ENSO system. However, all proxy systems respond to multiple climatic influences and contain at least some degree of climatic/nonclimatic noise. Consequently, any ENSO-relevant proxy will necessarily contain some component of its variance that is unrelated to NINO3 SSTA. The unrelated component of proxy variance is certainly not the only source of uncertainty in proxy variance-derived constraints on past ENSO amplitudes [*Thompson et al.*, 2011; *Stevenson et al.*, 2013; *Comboul et al.*, 2014], but it is a fundamental one, in that it is intrinsic to the very definition of a proxy system. The present study uses a simple parametric approach to exploring this uncertainty, with the aim of developing practical guidance on the magnitudes of proxy variance change required to make robust inferences about past changes in ENSO amplitude. The results and their implications are illustrated using the practical examples provided by one set of ENSO-relevant proxy records [*Cobb et al.*, 2003]. However, insofar as the statistical requirements of the simple parametric model used to generate these results are met, they may also be applied to the relationship between any climatic index and any other time series, be that a single-proxy record, a multiproxy synthesis, or indeed an instrumental record from another location.

2. Methods

A simple model for the relationship between a proxy variable (**Y**) and an instrumental index variable (**X**), both expressed as deviations from their respective means, is a linear model with slope β, such as equation (1). The individual values of the "noise" component (ϵ_i) of proxy variability are unknown but

are required to be independent of (i.e., unrelated to) the index and normally distributed with zero mean and variance σ_ϵ^2. It is recognized that these requirements will not be fully met in real-world applications where the climatic phenomena of interest is nonstationary and/or the proxy-index relationship is nonlinear. Nonetheless, simple linear models are widely used in paleoclimate science and provide an accessible framework in which to explore the relative contributions to proxy variance arising from variability that is, and is not, related to that of the index.

$$y_i = \beta x_i + \epsilon_i \tag{1}$$

For a time interval comprising n independent values of the index variable, the sample variance of the modeled proxy variable (s_y^2) is given by equation (2). In real-world time series applications it must be remembered that n will generally be less than the number of data points, to an extent that will depend on both the physical processes involved and the time series processing methods applied to the data. To avoid such dependency on data type and processing method, interval lengths in the present study are discussed in terms of degrees of freedom.

$$s_y^2 = \frac{1}{n-1}\sum_{i=1}^{n}(\beta x_i + \epsilon_i)^2 \tag{2}$$

The expected value of the modeled proxy sample variance is $\beta^2 s_x^2 + \sigma_\epsilon^2$, where s_x^2 is the index sample variance. However, the unknown values of the n noise terms also lead to a sampling distribution around this value. In other words, any given value of the index sample variance could be associated with a range of proxy sample variances, due to the additional contribution from climatic/nonclimatic factors that are independent of the index itself. These "proxy variance sampling distributions" follow a noncentral χ^2 form, with $n-1$ degrees of freedom and noncentrality parameter λ, which depends in turn on n, s_x^2, β, and σ_ϵ^2. The slope and noise variance model parameters may be rewritten in terms of the square of Pearson's moment correlation coefficient (r^2) and the population variances of the index (σ_x^2) and proxy (σ_y^2) variables [*von Storch and Zwiers*, 2003]. Such a substitution allows for the proxy variance sampling distribution to be expressed in terms of dimensionless variance ratios and a measure of the proxy-index calibration strength, as shown in equation (3). The derivation of this expression, along with that of a central χ^2 approximation [*Patnaik*, 1949] for which the distributional properties are more widely available, is given in Appendix A.

$$\frac{s_y^2}{\sigma_y^2} \sim \frac{1-r^2}{n-1}\chi_\lambda^2(n-1) \qquad \lambda = \frac{s_x^2}{\sigma_x^2}\frac{r^2(n-1)}{1-r^2} \tag{3}$$

In real-world applications the calibration strength (r^2) is not known exactly, although an unbiased estimate can be obtained from Ordinary Least Squares regression of the proxy and index data over the available period of their overlap [*von Storch and Zwiers*, 2003]. Associated estimates of the long-term variances of the proxy and index variables are then supplied by their sample variances over the calibration interval. The present study neglects any effect that sampling uncertainties in these estimates, arising from the finite duration of the available calibration interval, may then have on the resultant proxy variance sampling distributions. These additional uncertainties can always be minimized through the use of a calibration interval that spans the entirety of the available instrumental period. However, the extent of intercentennial variability in unforced ENSO amplitude seen within climate model simulations [*Gallant et al.*, 2013; *Stevenson et al.*, 2011; *Wittenberg*, 2009; *Russon et al.*, 2014] suggests that substantial uncertainties may persist even with a full century of calibration data. Under the present approach of assuming a perfectly known calibration relationship (i.e., perfectly known β and r^2), the variance ratio terms in equation (3) refer to comparisons between the interval of interest and the population variance, but where the latter is constrained to be equal to the sample variance over the calibration interval. In the following section, we will make use of such variance ratios to explore ENSO amplitude comparisons between a preinstrumental interval and a modern calibration interval, which often comprises the paleoclimatic problem of interest. Variance comparisons between two noncalibration intervals can also be considered within the present framework and the relevant sampling distributions for these applications are described in Appendix B.

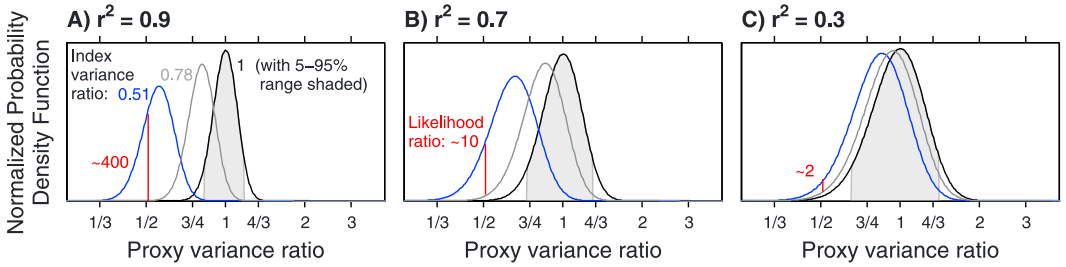

Figure 1. (a–c) Selected proxy variance ratio sampling distributions for the comparison of a preinstrumental record with a modern one that spans the proxy-index calibration period. The preinstrumental interval length is based on the estimate of degrees of freedom present within the 71 year duration 2–7 year band-pass-filtered twelfth C Palmyra coral record. The three panels refer to different proxy-index calibration strengths, with the middle one (b) being that derived from the band-pass-filtered modern Palmyra coral and HadISST NINO3 SSTA records. On each panel, the three sampling distributions then represent different scenarios for the change in NINO3 SSTA variance. The black lines represent no change and are relevant to the null hypothesis approach discussed in section 3.1. The under-curve shading in these cases illustrates the 5–95% confidence intervals derived from the associated cumulative distribution functions. The blue lines represent a change equal to that reconstructed from the actual proxy data comparison (a variance ratio of 0.51). The grey lines represent a change that is considered to remain within the plausible range of unforced variability (in this case a variance ratio of 0.78). Comparisons between the blue and grey distributions forms the basis of the likelihood ratio approach discussed in section 3.2. The vertical red lines illustrate the comparison of the likelihoods of obtaining the reconstructed value under these two scenarios, with the likelihood ratios annotated in red (given to 1 Significant Figure (s.f.)). The proxy variance ratio axis is logarithmic.

3. Results and Discussion

3.1. Proxy Variance Changes Needed to Infer Any Change in ENSO Amplitude

Given a reconstructed change in proxy variance between a preinstrumental interval and a modern calibration interval, let us consider what can be robustly inferred regarding the associated change in NINO3 SSTA variance, and hence ENSO amplitude. The simplest approach is to consider whether we can infer any change at all in ENSO amplitude between the two intervals. A sampling distribution for the proxy variance ratios to be expected under a null hypothesis of no change may be obtained by setting the index variance ratio to be unity ($\frac{s_x^2}{\sigma_x^2} = 1$) in equation (3). Such a distribution may then be used to define the critical magnitudes of proxy variance change required to robustly infer the occurrence of some (unspecified) change in ENSO amplitude.

To illustrate this concept, we make use of the real-world examples provided by the suite of monthly resolution last millennium Palmyra coral $\delta^{18}O$ records [*Cobb et al.*, 2003]. The 2–7 year band-pass-filtered modern coral record provides an r^2 value of 0.7 when calibrated to the similarly filtered Hadley Centre Sea Ice and Sea Surface Temperature (HadISST) NINO3 SSTA record. Band-pass filtering is applied as the coral $\delta^{18}O$ records contain pronounced decadal/longer timescale variability that may arise from controls that are not sea surface temperature [*Cobb et al.*, 2003]. A rule-of-thumb estimate for the effective number of degrees of freedom present in the filtered records is that 2 years of such data provide 1 degree of freedom [*Russon et al.*, 2014]. For the particular example of comparing the twelfth C fossil coral record with the modern, the form of the proxy variance ratio sampling distributions under the null hypothesis of no change in ENSO amplitude is shown by the black line in Figure 1b. The reconstructed variance ratio of 0.51 is seen to lie well outside of the central 90% of the cumulative distribution function for such values (shading in Figure 1b), such that a change in ENSO amplitude can indeed be robustly inferred. The other panels in Figure 1 illustrate that the critical magnitudes of proxy variance change increase as the calibration strength falls, due to the influence of r^2 on the width of the proxy variance sampling distributions.

The critical proxy variance ratios such that the null hypothesis of no change in ENSO amplitude may be rejected at the $p = 0.10$ level are plotted, as a function of calibration strength and preinstrumental interval length, in Figures 2a and 2b. The upper-tail critical variance ratios (Figure 2a), which relate to situations where the preinstrumental variance is greater than the modern, are not simply the multiplicative inverse of the lower-tail ones (Figure 2b), which relate to situations where the preinstrumental variance is less than the modern. This "tail asymmetry" arises from the tendency of the random noise contribution to modeled proxy variance to dampen whatever variability is present in the index. The blue markers locate the examples provided by the four fossil Palmyra coral records, with their record lengths (as given in parentheses) having

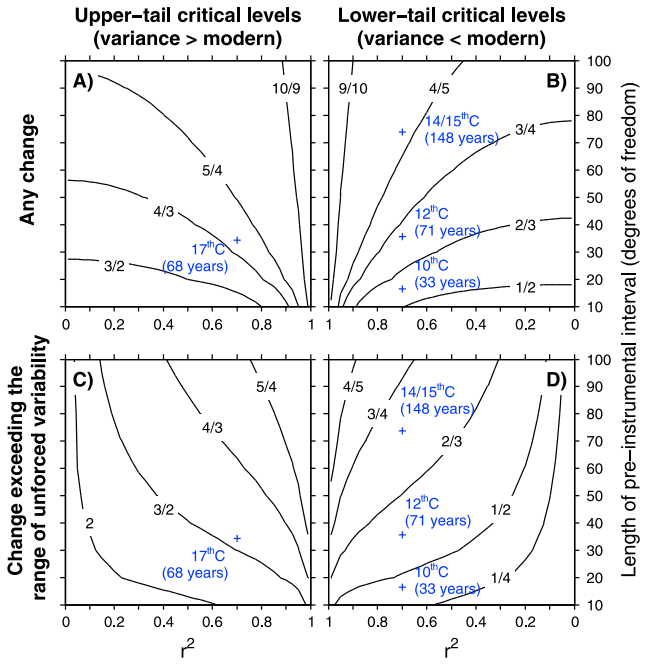

Figure 2. The critical proxy variance ratios required to make certain robust inferences regarding changes in ENSO amplitude between a preinstrumental interval and a modern calibration interval, plotted as a function of proxy-index calibration strength and preinstrumental interval length (in degrees of freedom). (a and b) The critical levels required to robustly infer some (unspecified) change in NINO3 SSTA variance, and hence ENSO amplitude (as described in section 3.1). (c and d) The critical levels required to robustly infer a change in NINO3 SSTA variance, and hence ENSO amplitude, that lies outside the plausible range of unforced variability (as described in section 3.2). The upper-tail panels in Figures 2a and 2c (lower-tail panels in Figures 2b and 2d) show the critical variance ratios when the preinstrumental variance is greater (less) than over the calibration interval. The calibration strength axes for the lower-tail panels are reversed, so that weaker calibration strengths always lie at the edges of the plot. The blue crosses approximately locate the examples provided by comparing each of the four fossil Palmyra coral records to the (112 year duration) modern coral [*Cobb et al.*, 2003]. As these comparisons are based on the use of 2–7 year band-pass-filtered records, the durations of the fossil corals have been converted to estimated degrees of freedom on the *y* axes. The four fossil Palmyra coral records are the following: (i) a tenth C record (33 years duration, band-pass-filtered variance ratio between fossil and modern coral of 0.46), (ii) a twelfth C record (71 years duration, variance ratio of 0.51), (iii) a fourteenth/fifteenth C record (148 years duration, variance ratio of 0.39), and (iv) a seventeenth C record (68 years duration, variance ratio of 1.05).

been converted to estimated degrees of freedom on the *y* axes. Figure 2a shows that a proxy variance ratio of more than 1.30 would be needed to robustly infer a change in ENSO amplitude for the comparison of the seventeenth C fossil coral to the modern. In this example, the reconstructed variance ratio of 1.05 is insufficient to allow for such an inference to be made. However, the reconstructed variance changes associated with all three of the fossil coral records with variance lower than the modern (Figure 2b) do allow for such inferences to be made.

The results in Figures 2a and 2b illustrate the potential difficulty of robustly inferring anything at all regarding changes in ENSO amplitude, when based on modest changes in the variance of proxies with modest calibration strengths to NINO3 SSTA. Furthermore, given that many ENSO-relevant proxy records tend to show decreases in preinstrumental variance relative to the modern [*Cobb et al.*, 2003; *Li et al.*, 2011; *Tudhope et al.*, 2001; *Fowler et al.*, 2012; *McGregor et al.*, 2010, 2013], the tail asymmetry present in these results should be borne in mind when such changes are interpreted in terms of ENSO amplitude. However, even when the occurrence of a change in ENSO amplitude can be robustly inferred under this approach, such a result does not contain any information at all regarding the magnitude of this change. As such, a more sophisticated method is required if we seek to address the principle question of interest, which is whether a change in ENSO amplitude of sufficient magnitude so as to lie outside of the plausible range of unforced variability has occurred.

3.2. Proxy Variance Changes Needed to Infer a Change in ENSO Amplitude Exceeding the Range of Unforced Variability

The present statistical framework may also be used to compare the relative likelihoods of obtaining a reconstructed proxy variance ratio under two different scenarios for the change in NINO3 SSTA variance, and hence ENSO amplitude. If the likelihood of obtaining the reconstructed value under one scenario is 3 (or more) times greater than that of obtaining it under the other, then there is considered to be "substantial" relative evidence for the first scenario [*Jeffreys*, 1961]. Paleoclimatic studies often interpret changes in ENSO-relevant proxy variance directly in terms of ENSO amplitude, and this assumption provides a first scenario to explore, namely that of a NINO3 SSTA variance change equal to that reconstructed from the proxy data. This scenario may then be compared to one in which the change in NINO3 SSTA variance would not be considered to be of climatological interest. The simplest such scenario would be that of no change at all and results derived on this basis are presented as supplementary information. However, a scenario with more direct relevance to the principle question of interest is that of a change in NINO3 SSTA variance which is sufficiently small as to remain consistent with unforced climate variability. The use of such a scenario requires the definition of a range of NINO3 SSTA sample variances that could plausibly occur within an unforced climate system. A conservative constraint on this range is provided by the 20–80% confidence interval associated with the central χ^2 sampling distribution for the sample variance of a given finite interval length. In other words, NINO3 SSTA sample variances within this range would not be considered statistically significant against random time domain sampling within a linear and stationary model of an unforced ENSO system. If the likelihood of obtaining the proxy data under a change in NINO3 SSTA variance equal to that reconstructed is substantially greater than that under this scenario, then this provides a useful step toward robustly inferring a change in ENSO amplitude exceeding the range of unforced variability. It is emphasized, however, that such a result does not imply that the true, or even most likely, magnitude of ENSO amplitude change is given by the reconstructed proxy variance ratio.

The likelihood ratio approach can be visualized, for the example comparison of the twelfth C Palmyra coral record with the modern, by comparing the blue and grey proxy variance ratio sampling distributions in Figure 1b. The blue distribution represents the scenario of a NINO3 SSTA variance ratio equal to that reconstructed and the grey one that of a NINO3 SSTA variance change of the same sense, but remaining within the plausible range of unforced variability. In this example, the likelihood of obtaining the reconstructed variance change under the first scenario is seen to be around 10 times greater than under the second scenario, such that a change in ENSO amplitude exceeding the range of unforced variability can be robustly inferred. The other panels of Figure 1 illustrate that if the proxy-index calibration were to have yielded an r^2 value of 0.3 (a value not atypical of many interannual proxy systems), then the same change in proxy variance would not have allowed for the same inference to be made.

The critical proxy variance ratios at which a change in ENSO amplitude exceeding the range of unforced variability can be robustly inferred are plotted as a function of calibration strength and preinstrumental interval length in Figures 2c and 2d. It should be noted that larger magnitude critical variance ratios would result from any definition for the plausible range of unforced variability in NINO3 SSTA variance that is more liberal than the parametric one outlined above. Among the fossil Palmyra corals, only the comparisons to the modern associated with the fourteenth/fifteenth C (variance ratio of 0.39) and twelfth C (variance ratio of 0.51) records allow for the robust inference of a decrease in ENSO amplitude that exceeds the plausible range of unforced variability. The comparison of the tenth C coral record to the modern (variance ratio of 0.46) provides an example where some (unspecified) change in ENSO amplitude can be robustly inferred (Figure 2b), but a change exceeding the plausible range of unforced variability cannot be (Figure 2d).

The critical proxy variance ratios associated with the likelihood ratio approach (Figures 2c and 2d) show the same sense of tail asymmetry as was seen for the hypothesis test approach (Figures 2a and 2b). However, the critical variance ratios for the likelihood ratio approach are seen to depend more strongly on calibration strength. While the two sets of values are not directly comparable to one another, due to the differing concepts of significance level associated with the two approaches, this basic difference may be understood from the different questions being asked by the two approaches. For the hypothesis test approach, the essential question is the amount of uncertainty introduced to modeled proxy variance by the noise terms, as seen in the width of the black distributions in Figure 1. This is bounded at low calibration strengths (e.g., Figure 1c) by the situation where the noise component constitutes the entirety of modeled proxy variance, as represented by a central χ^2 distribution. For the likelihood ratio approach, the essential question is

the extent to which the reconstructed proxy variance ratio can be considered more likely to have arisen under different potential NINO3 SSTA variance ratio scenarios. At the limit of low calibration strengths, the sampling distributions associated with all possible index variance ratio scenarios become increasingly hard to distinguish from one another, as they converge toward the central χ^2 form. In other words, as the calibration strength falls to zero, the variability of the proxy contains increasingly little information regarding the variability of the index, and this is reflected in the critical variance ratios for the likelihood ratio becoming extremely large in magnitude.

3.3. Comparing Two Preinstrumental Intervals

If the ENSO amplitude comparison of interest is between two preinstrumental intervals, then larger magnitude changes in proxy variance will be needed to robustly infer the same results regarding ENSO amplitude than those that have been described for comparisons between a preinstrumental interval and a modern calibration interval. This greater difficulty arises because we must now consider the influence of the component of proxy variability that is unrelated to NINO3 SSTA in both of the compared intervals, rather than just one of them. Appendix B describes the sampling distributions necessary to determine the critical proxy variance ratios associated with such comparisons. The bounding properties of these distributions mean that these critical levels must be greater in magnitude than those shown in the panels of Figure 2, evaluated at the duration of the shorter of the two intervals under consideration. For the example of a comparison between the twelfth and seventeenth C Palmyra fossil coral records, the reconstructed variance ratio of 2.08 exceeds the critical proxy variance ratio of 1.85 required to robustly infer a change in ENSO amplitude exceeding the plausible range of unforced variability. In contrast, the critical variance ratio required to make the same inference for a comparison between the seventeenth C record and a modern calibration interval would only have been 1.46.

4. Conclusions

Unforced (internal) variability within the climate system introduces uncertainty to all attempts to detect forced changes in ENSO amplitude. However, constraints on past changes in ENSO amplitude derived from proxy variance also contain a further fundamental source of uncertainty, arising from the component of proxy variance that is unrelated to the instrumental indices (such as NINO3 SSTA) from which modern ENSO amplitudes are defined. We present two sets of critical levels for the magnitudes of proxy variance change required to robustly infer certain results regarding past changes in ENSO amplitude in the face of this uncertainty. Given that we are generally interested in the magnitudes of past changes in ENSO amplitude, rather than simply the existence of any (unspecified) change, those derived from the "likelihood ratio approach" (Figures 2c and 2d) should provide the more useful criteria against which to evaluate proxy data. By neglecting other sources of uncertainty and making certain bounding assumptions, these critical levels are likely to represent best-case scenarios for the true difficulties associated with making such inferences. The examples used to illustrate the present study demonstrate that robust inferences regarding past changes in ENSO amplitude are indeed possible on the basis of existing high-resolution proxy records, with relatively long record lengths and high calibration strengths with NINO3 SSTA. However, our results also demonstrate the substantial difficulties associated with making such inferences on the basis of proxies with short record lengths and/or modest calibration strengths. This simple-to-use approach provides a tool for paleoclimate researchers to evaluate the climatic significance of changes in proxy variance through time.

Appendix A: Derivation of Sampling Distribution

As the noise terms (ϵ_i) are assumed to be normally distributed, it follows directly from equation (2) that s_y^2 is distributed in the noncentral χ^2 form shown by equation (A1).

$$s_y^2 \sim \frac{\sigma_\epsilon^2}{n-1}\chi_\lambda^2(n-1) \qquad \lambda = \sum_{i=1}^{n}\frac{(\beta x_i)^2}{\sigma_\epsilon^2} = \frac{\beta^2 s_x^2(n-1)}{\sigma_\epsilon^2} \tag{A1}$$

The relationships required to rewrite β and σ_ϵ^2 in terms of r^2, σ_x^2, and σ_y^2 are the following: $\beta^2\sigma_x^2 = r^2\sigma_y^2$ and $\sigma_\epsilon^2 = (1-r^2)\sigma_y^2$. Such substitutions transform equation (A1) into equation (3). For readers without ready access to the properties of the noncentral distribution, a central χ^2 approximated based on the approach

of *Patnaik* [1949], as shown in equation (A2), may provide a more accessible alternative. The centrally approximated form leads to only modest loss of precision in the hypothesis test approach but potentially introduces greater uncertainty to the likelihood ratio approach.

$$\frac{s_y^2}{\sigma_y^2} \sim (r^2\frac{s_x^2}{\sigma_x^2} + 1 - r^2)\left(\frac{k}{n-1}\right)\chi^2\left(\frac{n-1}{k}\right)$$
$$k = \left(\frac{(2r^2\frac{s_x^2}{\sigma_x^2} + 1 - r^2)(1-r^2)}{(r^2\frac{s_x^2}{\sigma_x^2} + 1 - r^2)^2}\right) \quad \text{(A2)}$$

If the index variance ratio in equation (A2) is set to be unity, as used by the hypothesis test approach, then the approximated form simplifies to equation (A3).

$$\frac{s_y^2}{\sigma_y^2} \sim \left(\frac{1-r^4}{n-1}\right)\chi^2\left(\frac{n-1}{1-r^4}\right) \quad \text{(A3)}$$

Appendix B: Comparing Two Noncalibration Intervals

If the proxy variance sampling distribution for the comparison of one noncalibration interval to the calibration interval follows a noncentral χ^2 distribution, then the comparison of two (independent) noncalibration intervals (denoted by subscripts 1 and 2) can be achieved through the ratio of two such distributions, which follows a doubly noncentral *F* distribution. These proxy variance sampling distributions contain two extra dimensions relative to equation (3), due to the presence of the two preinstrumental interval lengths (n_1 and n_2) and the two associated index sample variances (s_{x1}^2 and s_{x2}^2). In order for such distributions to be used in comparably simple applications to those described in the text, it is necessary to make an assumption regarding the relationship between the two index sample variances, such that they can be expressed in terms of the variance ratio between the two intervals under consideration ($\frac{s_{x1}^2}{s_{x2}^2}$). One approach is to assume that $\frac{s_{x1}^2}{\sigma_x^2} = \frac{\sigma_x^2}{s_{x2}^2} = \sqrt{\frac{s_{x1}^2}{s_{x2}^2}}$, meaning that the two preinstrumental intervals are taken as being evenly spaced in relative variance terms around the long-term value. The use of this (or any other) assumption entails greater structural uncertainty in such applications than for the comparison to calibration interval applications described in the main text. Nonetheless, under the present assumption the proxy variance sampling distribution associated with comparing two noncalibration intervals may be written as equation (B1). For users without access to the properties of the doubly noncentral *F* distribution, a centrally approximated form (not shown) may be obtained from the ratio of two distributions of the form of equation (A2).

$$\frac{s_{y1}^2}{s_{y2}^2} \sim F_{\lambda_1,\lambda_2}(n_1-1, n_2-1) \quad \lambda_1 = \sqrt{\frac{s_{x1}^2}{s_{x2}^2}}\frac{r^2(n_1-1)}{1-r^2} \quad \lambda_2 = \sqrt{\frac{s_{x2}^2}{s_{x1}^2}}\frac{r^2(n_2-1)}{1-r^2} \quad \text{(B1)}$$

For the hypothesis test approach, under which the index variance ratio is taken as unity, the centrally approximated form simplifies to equation (B2). Comparison of equations (A3) and (B2) makes clear that the critical levels for the comparison of two noncalibration intervals with intervals of length n_1 and n_2 (where $n_1 < n_2$) must always be greater in magnitude than those for the comparison of one noncalibration interval of length n_1 to the calibration interval.

$$\frac{s_{y1}^2}{s_{y2}^2} \sim F\left(\frac{n_1-1}{1-r^4}, \frac{n_2-1}{1-r^4}\right) \quad \text{(B2)}$$

For the likelihood ratio approach, the range of index variance ratios (between two noncalibration intervals) that are considered to remain consistent with unforced climate variability is defined as the 20–80%

confidence interval associated with the random sampling of two such intervals within a linear and stationary process. This amounts to using the critical levels of the central *F* distribution, with degrees of freedom $n_1 - 1$ and $n_2 - 1$, rather than those of the central χ^2 distribution with degrees of freedom $n_1 - 1$. The relationship between the critical proxy variance ratios associated with the comparison to calibration and comparison of two noncalibration interval cases is less intuitively clear than was the case for the hypothesis test approach, but the same bounding relationship between the two sets of critical levels is again seen to exist. Calculation of the critical levels associated with the comparison of two noncalibration intervals under the likelihood ratio approach requires the use of either equation (B1) or its centrally approximated form (not shown). It should be noted that the comparison of two noncalibration intervals case will always apply when comparing two preinstrumental intervals but may also be relevant when making comparisons of the preinstrumental to the modern and should the available calibration fail to span the entirety of the instrumental period.

Acknowledgments

This work was funded through NERC grant NE/H009957/1. The University of Exeter Statistical Science Group are thanked for constructive discussions of this work. The MATLAB code used to perform the analysis is available upon request.

The Editor thanks two anonymous reviewers for their assistance in evaluating this paper.

References

Adams, J. B., M. E. Mann, and C. M. Ammann (2003), Proxy evidence for an El Niño-like response to volcanic forcing, *Nature*, *426*(6964), 274–278.

Bellenger, H., E. Guilyardi, J. Leloup, M. Lengaigne, and J. Vialard (2014), ENSO representation in climate models: From CMIP3 to CMIP5, *Clim. Dyn.*, *42*(7–8), 1999–2018, doi:10.1007/s00382-013-1783-z.

Choi, J., S.-I. An, and S.-W. Yeh (2011), Decadal amplitude modulation of two types of ENSO and its relationship with the mean state, *Clim. Dyn.*, *38*(11–12), 2631–2644, doi:10.1007/s00382-011-1186-y.

Cobb, K. M., C. D. Charles, H. Cheng, and R. L. Edwards (2003), El Niño/Southern Oscillation and tropical Pacific climate during the last millennium, *Nature*, *424*(6946), 271–276.

Collins, M., et al. (2010), The impact of global warming on the tropical Pacific ocean and El Niño, *Nat. Geosci.*, *3*(6), 391–397.

Comboul, M., J. Emile-Geay, M. N. Evans, N. Mirnateghi, K. M. Cobb, and D. M. Thompson (2014), A probabilistic model of chronological errors in layer-counted climate proxies: Applications to annually banded coral archives, *Clim. Past*, *10*, 825–841, doi:10.5194/cp-10-825-2014.

Fowler, A. M., G. Boswijk, A. M. Lorrey, J. Gergis, M. Pirie, S. P. J. McCloskey, J. G. Palmer, and J. Wunder (2012), Multi-centennial tree-ring record of ENSO-related activity in New Zealand, *Nat. Clim. Change*, *2*(3), 172–176.

Gallant, A. J. E., S. J. Phipps, D. J. Karoly, A. B. Mullan, and A. M. Lorrey (2013), Nonstationary Australasian teleconnections and implications for paleoclimate reconstructions, *J. Clim.*, *26*, 8827–8849.

Guilyardi, E. (2006), El Niño–mean state–seasonal cycle interactions in a multi-model ensemble, *Clim. Dyn.*, *26*, 329–348.

Guilyardi, E., A. Wittenberg, A. Fedorov, M. Collins, C. Wang, A. Capotondi, G. van Oldenborgh, and T. Stockdale (2009), Understanding El Niño in ocean-atmosphere general circulation models: Progress and challenges, *Bull. Am. Meteorol. Soc.*, *90*, 325–340.

Jeffreys, H. (1961), *The Theory of Probability* 3rd ed., Clarendon, Oxford, U. K.

Li, J., S.-P. Xie, E. R. Cook, G. Huang, R. D'Arrigo, F. Liu, J. Ma, and X.-T. Zheng (2011), Interdecadal modulation of El Niño amplitude during the past millennium, *Nat. Clim. Change*, *1*(2), 114–118.

McGregor, S., A. Timmermann, and O. Timm (2010), A unified proxy for ENSO and PDO variability since, *Clim. Past*, *6*(1), 1–17, doi:10.5194/cp-6-1-2010.

McGregor, S., A. Timmermann, M. H. England, O. Elison Timm, and A. T. Wittenberg (2013), Inferred changes in El Niño–Southern Oscillation variance over the past six centuries, *Clim. Past*, *9*, 2269–2284.

Patnaik, P. B. (1949), The non-central χ^2 and F-distributions and their approximations, *Biometrika*, *26*, 202–232.

Russon, T., A. W. Tudhope, G. C. Hegerl, M. Collins, and A. Schurer (2014), Assessing the significance of changes in ENSO amplitude using variance metrics, *J. Clim.*, *27*(13), 4911–4922, doi:10.1175/JCLI-D-13-00077.1.

Stevenson, S., B. Fox-Kemper, M. Jochum, R. Neale, C. Deser, and G. Meehl (2011), Will there be a significant change to El Niño in the twenty-first century?, *J. Clim.*, *25*(6), 2129–2145, doi:10.1175/JCLI-D-11-00252.1.

Stevenson, S., H. V. McGregor, S. J. Phipps, and B. Fox-Kemper (2013), Quantifying errors in coral-based ENSO estimates: Toward improved forward modeling of $\delta^{18}O$, *Paleoceanography*, *28*, 633–649, doi:10.1002/palo.20059.

Thompson, D. M., T. R. Ault, M. N. Evans, J. E. Cole, and J. Emile-Geay (2011), Comparison of observed and simulated tropical climate trends using a forward model of coral $\delta^{18}O$, *Geophys. Res. Lett.*, *38*, L14706, doi:10.1029/2011GL048224.

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

von Storch, H., and F. W. Zwiers (2003), *Statistical Analysis in Climate Research*, Cambridge Univ. Press, Cambridge, U. K.

Watanabe, M., J.-S. Kug, F.-F. Jin, M. Collins, M. Ohba, and A. T. Wittenberg (2012), Uncertainty in the ENSO amplitude change from the past to the future, *Geophys. Res. Lett.*, *39*, L20703, doi:10.1029/2012GL053305.

Wittenberg, A. T. (2009), Are historical records sufficient to constrain ENSO simulations?, *Geophys. Res. Lett.*, *36*, L12702, doi:10.1029/2009GL038710.