Two-hundred-fifty years of reconstructed and modeled tropical temperatures

Rob Wilson, Alexander Tudhope, Philip Brohan, Keith Briffa, Timothy Osborn, and Simon Tett

Received 28 July 2005; revised 26 April 2006; accepted 26 June 2006; published 14 October 2006.

[1] Recent large-scale palaeoclimate reconstructions of past temperature have been essentially biased to the extratropics owing to a paucity of proxy data in tropical regions. Herein we describe the first coral-based reconstruction of sea surface temperatures (SSTs) for the whole of the tropics (30°N–30°S). It was developed from 14 disparate coral records located in the Indian and Pacific oceans. Over the most replicated period, the reconstruction explains 57% of the tropical SST variance. However, the strength of this signal weakens markedly as the number of coral records decreases. The reconstruction is robust between 1850 and 1993, but some fidelity is indicated back as far as the mid 18th century. These results suggest that ambiguities in the low frequency domain of δ18O measurements can be partially overcome by pooling together multiple time series from different locations around the tropics. Agreement with simulations from two general circulation models indicates that the late 20th century is likely the warmest period in the tropics for the last 250 years, and that this recent warming can only be explained by anthropogenic forcing. The high frequency variability is dominated by the El Niño–Southern Oscillation. The reconstruction, owing to the small number of coral records, is unfortunately restricted both in time and space. Therefore we hope that this study will spur the palaeoclimate community to develop new and longer proxy series to improve the current meager data-base of temperature sensitive series in the tropics.


1. Introduction

[2] In recent years, much attention has been focused on the reconstruction of large-scale hemispheric mean temperatures [Jones et al., 1998; Mann et al., 1999; Briffa, 2000; Esper et al., 2002; Moberg et al., 2005; D’Arrigo et al., 2006a] using mostly extratropical proxy data sets (e.g., high-latitude tree ring chronologies). In the tropics, palaeoclimate studies, using annually resolved proxy records, have focused mainly on the reconstruction of local sea surface temperatures (SSTs) [Dunbar et al., 1994; Quinn et al., 1998; Evans et al., 1999; Kuhnert et al., 1999, 2000; Cole et al., 2000; Linsley et al., 2000a; D’Arrigo et al., 2006b], the El Niño–Southern Oscillation (ENSO) [Stahle et al., 1998; Mann et al., 2000; Evans et al., 2002; Cobb et al., 2003; D’Arrigo et al., 2005a] and decadal variability [Felis et al., 2000; Cobb et al., 2001; Evans et al., 2001; Linsley et al., 2004; D’Arrigo et al., 2005b]. However, except for the spatial SST field reconstructions of Evans et al. [2002] in the Pacific Basin, no reconstruction of past large-scale tropical SSTs has been attempted, despite the known importance of low-latitude oceanic sites for large-scale temperature reconstruction [Bradley, 1996]. This situation not only reflects the lack of long, annually resolved climate proxy records from these regions, but also ambiguities in the climate response of the dominant proxy type, corals, particularly in the low-frequency domain, owing to competing influences of temperature and salinity [Crowley et al., 1999; Lough, 2004]. Although it has been shown that warming in the world’s oceans over the past 40 years is probably “human induced” [Barnett et al., 2005], this observation has not yet been placed into a longer-term context.

[3] In this study, we present a first attempt at reconstructing large-scale mean tropical (30°N–30°S) SSTs, at annual to centennial timescales, using a modest network of annually resolved coral proxy records. The resulting reconstruction is compared with general circulation model (GCM) simulated temperatures for the equivalent region and time period and inferences about tropical climate variability and radiative forcing over recent centuries are made. Comparative analysis is also made with an independent tree ring based reconstruction of NINO3 SSTs to quantify how well the coral reconstruction and GCM time series capture the
higher-frequency ENSO signal. We conclude by emphasizing the importance of the tropics in palaeoclimate studies, as it is the main source of energy for the global atmospheric/ocean system, and highlight the need to extend the currently meager proxy database for the lower latitudes.

2. Coral Data and Screening for a Temperature Signal

[4] From the NOAA Palæoclimatology Program (http://www.ncdc.noaa.gov/paleo/corals.html) and unpublished sources, there are more than 60 annual-to-monthly resolved coral records that start in the 19th century or earlier. These records represent a variety of measurement types (e.g., calcification, $\delta^{18}$O, Sr/Ca, U/Ca, Ba/Ca and fluorescence data) and portray varying environmental signals. To realize the main aim of this study, the coral records (annualized) were screened for a temperature signal by comparison with local $1 \times 1$ degree gridded annual SSTs (HadISST) [Rayner et al., 2003]. Correlation analysis was made using both unfiltered and first differenced versions of the data. Only coral records that correlated at the 95% confidence limit (C.L.) using both the original and transformed data were considered for further analysis. This approach guards against the incorrect attribution of a significant temperature response in series that in reality merely display coincidental similarity in their long-term trend.

[5] A total of 16 coral records, located in the Pacific and Indian oceans, passed this screening process (Figure 1, Table 1, and auxiliary material Figure S1). Of these, 1

---

1Auxiliary material is available at ftp://ftp.agu.org/apend/jc/2005jc003188. Other auxiliary material files are in the HTML.
15 represent skeletal $\delta^{18}O$ composition (a proxy that records the combined effects of temperature and salinity changes during the growth of the coral; see below). The other coral time series (HVI) is a record of skeletal fluorescence, which is documented as a proxy of rainfall and river run-off, although some of the variance is also explained by maximum temperatures [Isdale et al., 1998]. It should be noted that Hendy et al. [2003] identified minor dating errors in the original Isdale et al. [1998] series. However, this record was still used in this study as it significantly correlates with local and tropical temperatures. There are also probably unknown dating errors in the early portions of the other coral records which are not quantifiable.

[6] Instrumental local SSTs from many of these coral locations correlate significantly with mean tropical (30°N–30°S) SSTs and therefore it is not surprising that many of the coral records also correlate significantly with large-scale tropical temperatures (Figure 1a). By utilizing the first differenced transforms, the correlation analysis also shows that many of the coral records cohere with large-scale temperatures even at high frequencies (Figure 1b). This observation presumably reflects the importance of the ENSO signal in both the tropical average temperatures and many of the proxy records. The characteristic ENSO pattern is clearly observed in the spatial correlations shown in Figure 1. Only RUS, NGR and HAI show first differenced correlations with mean tropical temperatures that are not significant at the 95% C.L., although they do all show the same sign.

[7] We note that for the LAI and RAR $\delta^{18}O$ coral records, however, the sign of the correlations against local temperatures are opposite to those against large-scale tropical mean temperatures (Figure 1). These two coral records are located in a broad region from New Guinea to the south central Pacific where SSTs are inversely correlated at high frequencies with large-scale mean tropical temperatures (Figure 1b), implying that these two coral records are behaving as would be expected, at least at high frequencies. In this study, however, we did not consider utilizing these two records for the coral-based reconstruction, which is essentially developed from simple averaging, as they would have to be band-pass filtered [Osborn and Briffa, 2000] to optimize their signal in the frequency domain.

[8] Coral skeletal $\delta^{18}O$ records generally, but not always, represent a mixed signal of local sea surface temperature and salinity (SSS) conditions and so it is not always possible to interpret such records strictly as temperature proxies [Lough, 2004]. To highlight this problem, Table 2a shows ordinary least squares (OLS) regression results between each coral record and its respective local HadISST temperature grid. In every case there are problems in the model residuals (either significant linear trends or autocorrelation) suggesting that a significant amount of the variability (at decadal or lower frequencies) in the coral time series cannot be explained by local temperatures. To further highlight local-scale calibration problems, Table 2b presents correlations between each coral series and its local SSS grid for the independent periods, 1897–1924, 1925–1952 and 1953–1981. Significant correlations (95% C.L.) are only found in all periods for the RUS, PAL, MAI and TAR records, although by comparing each subperiod correlation value with the mean of the other two, only six of the coral series show a significant temporal instability in their relationship with local SSS. These analyses (Tables 2a and 2b) indicate that, at the local scale, it is not possible to interpret these coral series as pure nonbiased proxies of past temperature. These results likely reflect a combination of (1) unknown influences of salinity changes at the local site, (2) poor quality of the early instrumental SST record, (3) possible time instabilities in the coral/SSS relationship and (4) potential artifacts in skeletal $\delta^{18}O$ related to marine diagenesis [Muller et al., 2001]. However, at least in the South West Pacific, the instrumental SST record is well supported by comparison with independent instrumental air-temperature data [Folland et al., 2003], so limitations of the instrumental SSTs are unlikely to be a significant obstacle. Also, we note that for many of the coral sites, SST and SSS are positively correlated on at least seasonal and interannual (ENSO) timescales, due largely to the impact of SST on convective rainfall. In these cases, changes in salinity may have a less detrimental effect on the fidelity of the empirically derived temperature reconstructions, although the degree to which this is so depends upon the stability of the SST:SSS relationship as a function of timescale. For this study we therefore hypothesize that by pooling these coral

---

**Table 1. Coral Records Used in This Paper**

<table>
<thead>
<tr>
<th>Site Name</th>
<th>Site Code</th>
<th>Record Length</th>
<th>Longitude</th>
<th>Latitude</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Urvina Bay</td>
<td>URV</td>
<td>1607–1981</td>
<td>91.14°W</td>
<td>0.24°S</td>
<td>Dunbar et al. [1994], Shen et al. [1992]</td>
</tr>
<tr>
<td>Havannah Island</td>
<td>HVI</td>
<td>1844–1986</td>
<td>146.33°E</td>
<td>18.51°S</td>
<td>Isdale et al. [1998]</td>
</tr>
<tr>
<td>Ras Umm Sidd</td>
<td>RUS</td>
<td>1751–1995</td>
<td>34.18°E</td>
<td>27.50°N</td>
<td>Felis et al. [2000]</td>
</tr>
<tr>
<td>Houtman Abrolhos Islands</td>
<td>HAI</td>
<td>1794–1994</td>
<td>113.46°E</td>
<td>28.28°S</td>
<td>Kuhnert et al. [1999]</td>
</tr>
<tr>
<td>Malindi</td>
<td>MAL</td>
<td>1801–1992</td>
<td>40.00°E</td>
<td>3.00°N</td>
<td>Cole et al. [2000]</td>
</tr>
<tr>
<td>Maiana</td>
<td>MAI</td>
<td>1840–1994</td>
<td>173.00°E</td>
<td>1.00°N</td>
<td>Urban et al. [2000]</td>
</tr>
<tr>
<td>Mahe, Seychelles</td>
<td>MAH</td>
<td>1846–1995</td>
<td>55.00°E</td>
<td>4.37°S</td>
<td>Charles et al. [1997]</td>
</tr>
<tr>
<td>Jarvis</td>
<td>JAR</td>
<td>1850–1998</td>
<td>159.59°W</td>
<td>0.22°S</td>
<td>A. Tudhope et al. (manuscript in preparation, 2006)</td>
</tr>
<tr>
<td>Suvarrow Atoll</td>
<td>SUW</td>
<td>1881–1998</td>
<td>163.06°W</td>
<td>13.15°S</td>
<td>A. Tudhope et al. (manuscript in preparation, 2006)</td>
</tr>
<tr>
<td>Palmyra Island</td>
<td>PAL</td>
<td>1886–1996</td>
<td>162.08°W</td>
<td>5.52°N</td>
<td>Cobb et al. [2001]</td>
</tr>
<tr>
<td>Tarawa</td>
<td>TAR</td>
<td>1894–1988</td>
<td>172.00°E</td>
<td>1.00°N</td>
<td>Cole and Fairbanks [1990], Cole et al. [1993]</td>
</tr>
<tr>
<td>Nauru Island</td>
<td>NAI</td>
<td>1897–1994</td>
<td>166.00E</td>
<td>0.30S</td>
<td>Guilderson and Schrag [1999]</td>
</tr>
<tr>
<td>Ratonga</td>
<td>RAR</td>
<td>1726–1996</td>
<td>159.00W</td>
<td>21.00S</td>
<td>Linsley et al. [2004]</td>
</tr>
<tr>
<td>Laing Island</td>
<td>LAI</td>
<td>1885–1992</td>
<td>144.53E</td>
<td>4.09S</td>
<td>Tudhope et al. [2001]</td>
</tr>
</tbody>
</table>

Note that URV denotes composite of Urvina Bay (1607–1953/1962–1981 [Dunbar et al., 1994)] and Punta Pitt (1936–1981 [Shen et al., 1992]). Both corals are located closely together (r = 0.49) on the Galapagos Islands. Both data sets have missing values in the 20th century. To fill these gaps, the two time series were normalized over the period (values) of overlap and averaged, so creating an almost continuous time series from 1607–1981. Three values (1907–1909) which were still missing, were linearly interpolated using the five annual values on either side of the gap.
Table 2a. OLS Regression Results (1897–1981) Between Each Coral Record and Its Respective Local Annual (January–December) HadISST Temperature Grid*

<table>
<thead>
<tr>
<th></th>
<th>URV</th>
<th>HVI</th>
<th>RUS</th>
<th>HAI</th>
<th>MAL</th>
<th>MAH</th>
<th>JAR</th>
<th>NGR</th>
<th>SUW</th>
<th>PAL</th>
<th>MAI</th>
<th>TAR</th>
<th>CLP</th>
<th>NAI</th>
</tr>
</thead>
<tbody>
<tr>
<td>r: 1897–1981</td>
<td>-0.40</td>
<td>-0.20</td>
<td>-0.32</td>
<td>-0.35</td>
<td>-0.29</td>
<td>-0.49</td>
<td>-0.43</td>
<td>-0.28</td>
<td>-0.11</td>
<td>0.09</td>
<td>-0.61</td>
<td>-0.55</td>
<td>-0.37</td>
<td>-0.33</td>
</tr>
<tr>
<td>DW</td>
<td>1.83</td>
<td>1.44</td>
<td>0.73</td>
<td>0.95</td>
<td>0.78</td>
<td>1.23</td>
<td>1.06</td>
<td>1.40</td>
<td>1.27</td>
<td>1.10</td>
<td>1.08</td>
<td>1.39</td>
<td>1.65</td>
<td>1.14</td>
</tr>
<tr>
<td>LT</td>
<td>-0.35</td>
<td>0.39</td>
<td>0.05</td>
<td>0.20</td>
<td>0.74</td>
<td>0.36</td>
<td>-0.39</td>
<td>0.00</td>
<td>-0.22</td>
<td>-0.24</td>
<td>-0.28</td>
<td>0.06</td>
<td>-0.26</td>
<td>-0.08</td>
</tr>
</tbody>
</table>

*Abbreviations: r, Pearson’s correlation of linear relationship; DW, Durbin-Watson test statistic for residual autocorrelation (see Appendix A); LT, linear trend as calculated from the correlation between the residual values and time. Italicized entries for either DW or LT denote autocorrelation or a linear trend (95% C.L.) in the regression model residuals.

Tables together, we maximize the common temperature signal while minimizing the unquantifiable effects of uncorrelated salinity changes, and other sources of uncorrelated error.

3. Coral-Based Reconstruction of Mean Tropical Temperatures

[9] Although only a relatively small number of corals, records, from the Pacific and Indian oceans, were found to portray a significant SST signal, Evans et al. [1998] showed that robust reconstructions of SST could be developed from such a limited data set. Our reconstruction was developed using the following four steps.

3.1. Combining the MAI and TAR Records

[10] The MAI and TAR records come from almost the same location (Figure 1). Therefore making a simple average of all the coral records would result in a bias to this location as it is effectively sampled twice. To remove this bias we combined the two records into a single composite called MTA. This is the mean of MAI and TAR for 1894–1981. Prior to 1894, the MAI series was used on its own.

3.2. Averaging the Proxy Data and Calibration

[11] The shortest coral series (NAI) starts in 1897, and the last year for which there is data for all the coral series is 1981, so any reconstruction using all the coral series is restricted to the period 1897–1981. A mean series was made over this common 1897–1981 period by first scaling each of the coral series so they had a mean of zero and a variance of one (i.e., standardized to z scores), and then averaging the normalized series. The mean coral series was then calibrated against the instrumental (HadISST) mean tropical temperatures (which are available since 1870) over the 1897–1981 period using OLS regression. The 2 sigma error bars were derived using the standard deviation of the regression residuals; the uncertainty in the regression coefficient is small in comparison and has been neglected. We should emphasize, however, that these calculated uncertainties are optimistic estimates as they were calculated using the calibration residuals and do not incorporate the additional uncertainty of the regression coefficients and the possibility of diminished dating precision in the early portions of the coral records.

3.3. Extending the Reconstruction Backward

[12] Utilizing only the common period of all the proxy data would result in a reconstruction covering only the 1897–1981 period. Therefore, to extend the reconstruction as far back in time as possible using coral series of different length (the number of coral sites included in the proxy average will reduce as we extend the reconstruction back in time), we employed a variant of the so-called “nesting” procedure [Meko, 1997; Cook et al., 2002]. As the regression model, developed using the average of all the series (back to 1897), will not be appropriate for the average of smaller sets of coral series extending farther back in time, this nesting approach allows the development of multiple regression models as each coral series leaves the data matrix.

[13] First, using the common period of all the coral series, a reconstruction is developed back to 1897 (see above); this is the first nest reconstruction. Next, we discard the shortest coral series, NAI, leaving those series that extend back to 1894. In the same way as for the first nest, a mean series is developed for the 1894–1981 period and calibrated (1897–1981) against the instrumental temperature as before. This gives us a new second nest reconstruction, with error ranges, for 1894–1981. The values for the 1894–1896 period from this second nest need to be spliced onto the first nest reconstruction. To blend the two nest reconstructions, the mean and variance of the second reconstructed time series was scaled (i.e., transformed to have the same mean and variance) to that of the first (i.e., over the 1897–1981 period), and the values from the second series reconstruction for 1894–1896 were spliced on to the first reconstruction (1897–1981). The rescaling of the data removes artificial changes in the variance of the final time series, owing to the decrease in the number of coral series and resultant weakening in explained variance, while retaining potential real changes in variance that may represent a response to actual changes in climatic variability. The


<table>
<thead>
<tr>
<th></th>
<th>URV</th>
<th>HVI</th>
<th>RUS</th>
<th>HAI</th>
<th>MAL</th>
<th>MAH</th>
<th>JAR</th>
<th>NGR</th>
<th>SUW</th>
<th>PAL</th>
<th>MAI</th>
<th>TAR</th>
<th>CLP</th>
<th>NAI</th>
</tr>
</thead>
<tbody>
<tr>
<td>r: 1897–1924</td>
<td>-0.04</td>
<td>-0.01</td>
<td>-0.54</td>
<td>-0.18</td>
<td>-0.14</td>
<td>-0.19</td>
<td>-0.37</td>
<td>-0.26</td>
<td>-0.08</td>
<td>-0.70</td>
<td>-0.63</td>
<td>-0.55</td>
<td>-0.72</td>
<td>-0.14</td>
</tr>
<tr>
<td>r: 1925–1952</td>
<td>-0.55</td>
<td>-0.37</td>
<td>-0.39</td>
<td>-0.15</td>
<td>-0.37</td>
<td>-0.43</td>
<td>-0.71</td>
<td>-0.23</td>
<td>-0.71</td>
<td>-0.76</td>
<td>-0.56</td>
<td>-0.64</td>
<td>-0.16</td>
<td>-0.54</td>
</tr>
<tr>
<td>r: 1953–1981</td>
<td>-0.86</td>
<td>-0.17</td>
<td>-0.40</td>
<td>-0.62</td>
<td>-0.59</td>
<td>0.01</td>
<td>-0.76</td>
<td>-0.16</td>
<td>-0.57</td>
<td>-0.79</td>
<td>-0.68</td>
<td>-0.55</td>
<td>-0.34</td>
<td>-0.66</td>
</tr>
</tbody>
</table>

*Italicized entries denote correlations that are not significantly (95% C.L.) different from zero. Boldface values indicate correlations that are significantly (95% C.L.) different from the mean of the other two period correlations, implying that for these coral records a significant temporal instability in the coral:SST relationship exists.
2 sigma error bars were also adjusted (inflated), using the same scaling function as used for each nested series, to account for the decrease in explained variance in each nest. The error bars and verification statistics (see below) for the first nest apply to the spliced series for 1897–1981, and those for the second nest apply to the spliced series for 1894–1896. This nesting procedure is repeated several times, each time removing one more coral series from the data matrix, and extending the spliced reconstruction further back in time. The most replicated nest, containing all the coral series, extends back to 1897 and the least replicated nest, containing only the longest series (URV), extends back to 1607.

3.4. Verification

As the coherence of each coral record to local SSTs is not time stable (Table 1), a strict verification procedure is essential to quantify the robustness of the large-scale mean reconstruction [Crowley et al., 1999]. For this purpose, we used the Pearson’s correlation coefficient (r), reduction of error (RE), coefficient of efficiency (CE), and sign test (ST) statistics, commonly used in dendroclimatology [Cook and Kairiukstis, 1990; Cook et al., 1994]. These statistics are calculated for the period of overlap of the coral data and the instrumental series, and are described in detail in Appendix A. Each statistic is calculated separately for each nest.

As well as full period calibration (1897–1981), split period calibration and verification was made over the periods 1897–1939 and 1940–1981 (i.e., calibrated on the early period with verification on the late period and vice versa). For the mean nest series extending prior to 1870, an extra verification period (1870–1896) was also used as it represents a stringent verification using data that were independent of the

![Figure 2](image-url). Coral reconstruction. (a) Actual and reconstructed January–December mean tropical (30°N–30°S) temperatures with 2 sigma error bars. The shaded histogram denotes the number of coral series utilized through time. (b) Full period (1897–1981) calibration and verification (1870–1896). (c) Early period (1897–1939) calibration and late period (1940–1980) verification. (d) As in Figure 2c but periods are reversed. Here $r^2$ is square of the multiple correlation coefficient, RE is reduction of error, and CE is coefficient of efficiency. See Appendix A for an explanation of these statistics and see auxiliary Table S1 for all other verification results.
initial screening process (Figure 1). However, it should also be emphasized that these early data are based on much sparser observational data and the quality of this instrumental SST record for this period may be poor.

[17] To test the robustness of the long-term signal in the reconstructed nested series, assessment of the regression model residuals was also made by calculating the linear trend of the regression residual time series as well as testing for autocorrelation using the Durbin-Watson (DW) statistic [Shaw and Wheeler, 1985; Draper and Smith, 1998]. The DW statistic is detailed in Appendix A.

[18] Figure 2 presents the full mean coral reconstruction along with its 2 sigma error bars and associated calibration and verification statistics (Table S1 in the auxiliary material details the full complement of calibration and verification statistics). Over the most replicated nest (1897–1981) the reconstruction explains 57% of the tropical SST variance. However, the calibration and verification statistics weaken markedly as the number of coral records used in the reconstruction decreases. Verification is very robust for those nests between 1850 and 1993, with positive CE values and no linear trend in the model residuals. Prior to 1850, although CE values become negative, RE values are still positive for the nested models that extend back to 1644, suggesting some reconstructive confidence over this extended period. However, we advise caution. First, prior to 1840, the explained variance of the reconstruction is quite low and drops below 30%. Second, the 1703–1750 period is weakly verified over the earlier 1870–1896 verification period and all nest models prior to 1846 show a significant linear trend in their residuals suggesting that not all low frequency information is being captured robustly (see auxiliary Table S1).

[19] The results for the 1635–1702 period appear improved (i.e., narrower error bars and stronger verification statistics). For this period, we included a “fossil” section (1635–1702) of coral data for the PAL record. The calibration/verification statistics for this period were generated using the living PAL data. However, this use of living data essentially does not test the robustness of the fossil coral data. The actual dating accuracy of the PAL fossil data is approximately ±5–10 years [Cobb et al., 2003]. Therefore the narrower error bars over this period are optimistic. In light of these results, we suggest that the coral reconstruction is robust between 1850 and 1993, and portrays some fidelity back to 1751, a period where the reconstruction is calculated from at least three coral records.

[20] To test if we would expect a valid reconstruction from using so few data points throughout the tropics, a tropical mean temperature reconstruction using the gridded SST data from the coral proxy locations was generated using the same averaging procedure. This instrumental-based reconstruction explains 84% of the temperature variance (1897–1981) and passes split period (1897–1939 and 1940–1981) calibration/verification (results not shown) as well as extra verification over the 1870–1896 and 1982–2003 independent periods (Figure 3). Despite these strong validation results, the model fails residual analysis with significant autocorrelation (DW = 1.13) and a significant linear increasing trend in the regression residuals of 0.1°C/100 years (compared to a non significant decrease of 0.05°C/100 years in the most replicated nest of the coral reconstruction). These results, therefore, show that a relatively robust large-scale reconstruction can be developed from this sparse network of coral records (provided that they show a significant relationship with local SSTs), although trends in the model residuals imply a tendency, by using this sparse network, to not capture all the low-frequency information in the mean tropical temperature time series. This observation is empirically derived, however, and is qualitatively not obvious from Figure 3. Interestingly, between 1850 and 1993, no significant linear trend is observed in the residuals from the coral-based calibrations, although all the coral nested reconstructions have significant autocorrelation in their model residuals (see auxiliary Table S1). Evans et al. [1998] showed that robust reconstructions of tropical SSTs could be developed from a limited network of coral proxy records, specially located in

Figure 3. Calibration and verification results of the instrumental-based reconstruction of tropical SSTs. See Appendix A for a description for RE, CE, ST, and DW. LT is linear trend as calculated from the relationship (using the Pearson’s correlation coefficient) between the residual values and time.
the central/eastern Pacific and Indian oceans. Our results support these conclusions, although stringent residual analysis suggests that there are biases in the lower frequency domain. This is perhaps not an unexpected result given the known ambiguities of $\delta^{18}O$ at lower frequencies. The fact that between 1850 and 1993, in the coral reconstruction, no linear trends are observed from the regression modeling, may simply reflect the partial canceling out of two conflicting biases, i.e., too much trend in the $\delta^{18}O$ series and not enough trend noted in the SSTs from the sparse proxy network used.

4. Comparison With GCM Model Output

[21] To explore forcings that may explain the trend of tropical temperatures over recent centuries, we compared the coral reconstruction with model output from the coupled climate models HadCM3 [Gordon et al., 2000; Pope et al., 2000; Tett et al., 2006] and ECHO-G [Legutke and Voss, 1999; González-Rouco et al., 2003; von Storch et al., 2004].

[22] The coral reconstruction (Figure 4) indicates that between 1750 and 1899 mean tropical SSTs were likely about 0.3°C cooler than the reference 1961–1990 period, with coolest conditions found near the beginning of the 19th century. Temperatures increased, with some decadal fluctuations, from this cool phase until the 1990s which is the warmest reconstructed period over the last 250 years. Correlations (1756–1989) between the reconstruction and the model outputs (Table 3) indicate the strongest relationship with the ALL forcing runs of the ECHO-G ($r = 0.74$) and HadCM3 ($r = 0.57$) modeled mean tropical SSTs. These correlations, generally driven by the low frequency trends in the series, reflect the close similarity between the series (Figure 4b) in that they show cool conditions in the early 19th century followed by a warming phase until present.
Prior to 1870, however, the coral reconstruction appears to track the HadCM3 model better as the ECHO-G ALL model simulates cooler conditions compared to the other time series. The natural forcing runs (Figure 4a) do not model 20th century warming.

[23] The amplitude (Figure 4b) of the coral reconstruction (0.52°C) is lower than for the ALL forcing runs of ECHO-G (0.97°C) and HadCM3 (0.79°C). This discrepancy may result from the bias from using OLS regression for calibrating linear relationships, which results in predicted data with lower variance than the original instrumental data, leading to an underestimation of the amplitude [von Storch et al., 2004; Esper et al., 2005], but may also imply that the models are too sensitive to the forcing parameters. The ALL forcing models incorporate the radiative forcing of solar irradiance, aerosols from volcanic eruptions, anthropogenic aerosols (not included in ECHO-G) and greenhouse gas concentrations [von Storch et al., 2004; Tett et al., 2006]. The variability in the models prior to ~1850 is largely determined by volcanic and solar forcing (i.e., natural) indicating that the cool conditions in the early 19th century are a result of the combined influence of a period of known volcanic activity and low solar irradiance (the Dalton Minimum).

[24] It is interesting to note that the reconstruction does not show a response to the 1815 Tambora eruption (Figure 4b), although there are clear negative relative deviations in 1809/10 (unknown) and 1883 (Krakatoa) that coincide with the other tropical volcanic events [Chenoweth, 2001]. The fact that the coral reconstruction does not show cooling in 1816, a year in which ship-based instrumental data indicate significantly cooler marine air temperature conditions [Chenoweth, 2001], may reflect the spatial nature of the temperature response to the Tambora event, the lack of correspondence between locations of the ship-based observations and the coral records, and the effect of the ~1817 ENSO warm event [Ortlieb, 2000].

5. Comparison With NINO3 SSTS

[25] The previous section, which compared the coral reconstruction with model output from two coupled climate models, concluded that natural forcings (mainly solar and volcanic) explain much of the past low frequency SST tropical variability until the mid 20th century. However, the natural forcing runs could not track recent warming observed in tropical SSTs. Only by including anthropogenic forcings (mainly greenhouse gas concentrations) into the models do they agree with the coral reconstruction in the recent period. However, such analysis only explores the long term controls upon tropical SST, and does not reflect the situation at higher frequencies. Figures 5a–5c present Multitaper method [Mann and Lees, 1996] spectral analysis results for the coral reconstruction and ALL forcings output for HADCM3 and ECHO-G. At the 99% confidence limit, the spectra for the coral reconstruction and HADCM3 are quite similar, with common spectral peaks at >90, 5.5, ~3.7–4.3 and 3.0–3.6 years. The former spectral peak presumably reflects low-frequency trends within the reconstruction and models related to external and anthropogenic forcing. The latter peaks are situated within the classical ENSO bandwidth.

[26] To further explore the ENSO signal, the coral reconstruction and models were compared to an independent tree ring based reconstruction of NINO3 (1407–1978 [D’Arrigo et al., 2005a]). To remove non ENSO related secular scale variability, the time series were high-pass filtered with an 8-year Gaussian filter. Although the NINO3 reconstruction portrays a different season (December–February), the correlation between tropical annual SSTs and the NINO3 instrumental data is 0.76 (after high-pass filtering) over the 1870–1978 period. The correlation between the respective reconstructions over the same period is 0.47. As expected from the spectral analysis (Figure 5a), therefore, the coral reconstruction does cohere reasonably with NINO3 SSTs within the classical ENSO bandwidth. Running 31-year correlations made between the coral reconstruction and NINO3 reconstruction (Figure 5d), however, identifies a temporally changeable coherence between both time series, specifically with nonsignificant or negative correlations over the periods ~1625–1675, ~1700–1775 and ~1825–1875. These periods of noncoherence may reflect (1) quality issues within the proxy data used (e.g., dating errors in the early portions of the coral proxies and general weakening in reconstruction confidence back in time); (2) the fact that the tree ring NINO3 reconstruction was calibrated during the recent period using teleconnected relationships that may not be time stable; (3) that the coral data were not specifically screened for an ENSO signal, so some records may not be good proxies for this phenomenon; (4) the differing influence of volcanic events on proxies from differing locations; (5) a real pattern of less coherency in the Pacific atmosphere-ocean system over these periods prior to the 20th century. Further examination of these noncoherent periods is beyond the scope of this paper.

[27] Min et al. [2005] and Collins et al. [2001] showed that both the ECHO-G and HADCM3 models simulate the ENSO spatial pattern reasonably well. Collins et al. [2001] also noted ENSO variability with an irregular 3–4 year cycle, results similar to ours (Figure 5b). However, the HADCM3 ALL run high-pass filtered time series does not correlate with actual or reconstructed NINO3 SSTs (Figure 5d). This result is not unexpected, but highlights

### Table 3. Correlations Between the Coral Reconstruction and NATURAL and ALL Forcing Runs From the HadCM3 and ECHO-G Models

<table>
<thead>
<tr>
<th>Period</th>
<th>Had SST nat (adf = 81, p = 0.03)</th>
<th>Had SST all (adf = 98, p = 0.00)</th>
<th>Echo SST nat (adf = 62, p = 0.00)</th>
<th>Echo SST all (adf = 79, p = 0.00)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1760–1898</td>
<td>0.57 (p = 0.00)</td>
<td>0.39 (p = 0.00)</td>
<td>0.74 (p = 0.00)</td>
<td></td>
</tr>
<tr>
<td>1800–1898</td>
<td>0.57 (p = 0.00)</td>
<td>0.47 (p = 0.00)</td>
<td>0.75 (p = 0.00)</td>
<td></td>
</tr>
<tr>
<td>1850–1898</td>
<td>0.46 (p = 0.00)</td>
<td>0.29 (p = 0.18)</td>
<td>0.68 (p = 0.00)</td>
<td></td>
</tr>
<tr>
<td>1760–1850</td>
<td>0.33 (p = 0.06)</td>
<td>0.38 (p = 0.06)</td>
<td>0.44 (p = 0.00)</td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: nat, NATURAL; all, ALL; Had, HadCM3; Echo, ECHO-G. The degrees of freedom have been adjusted (adf) to account for the first-order autocorrelation in the time series [Dawdy and Matalas, 1964].

8 of 13
Figure 5. Multitaper method spectral analysis [Mann and Lees, 1996] for (a) the coral reconstruction, (b) HADCM3-ALL, and (c) and ECHO-G-ALL calculated over the 1760–1989 period. Dashed significance lines are 95% and 99% C.L. Significant peaks are labeled for the 99% C.L. (d) Running 31-year correlations between the coral reconstruction, HADCM3-ALL and ECHO-G-ALL series with the D’Arrigo et al. [2005a] NINO3 reconstruction. The series have been high-pass filtered prior to analysis using an 8-year Gaussian filter.
that although the model represents ENSO variability well in model space, even for the forced runs, the timing of El Niño and La Niña years will not be the same in the model output as the real world. The ECHO-G ALL forcing run shows weaker ENSO variability with peaks at 4.6–4.7, 5.1–5.3 and 6–7 years (significant at the 95% C.L.; Figure 5c), and, like the HADCM3 series, no strong correspondence with the NINO3 reconstruction is observed (Figure 5d).

6. Discussion and Conclusions

[28] We have detailed the development of the first coral-based, large-scale temperature reconstruction of mean annual SSTs for the whole tropics. Over the most replicated period, the reconstruction explains 57% of the instrumental temperature variance and robustly verifies when compared to independent data. Although the fidelity of the reconstruction quickly weakens back in time due to the shortness of the coral records, the strong calibration and verification results since 1850 (a period utilizing eight or more records) not only indicate the importance of coral records for palaeoclimatic reconstruction, but that ambiguities in the low-frequency domain of δ18O measurements can be partially overcome by pooling together multiple time series from disparate locations around the tropics. However, we must stress that trends in the residuals from the regression modeling suggest that this reconstruction of large-scale tropical SSTs is not problem free and that future attempts need to try and further address biases within the frequency domain.

[29] Low-frequency agreement between the reconstruction and two global coupled-climate models indicates that the 1990s were the warmest period in the last 250 years in the tropics and that there has been a ~0.5°–1.0°C increase in annual tropical SSTs from the coolest period in the early 19th century to the present (Figure 4b). Comparative analysis suggests that prior to the 20th century, natural forcing (the combined influence of solar and volcanic forcing) explains much of the reconstructed variability in tropical climate. However, the natural forcing used in these simulations cannot explain 20th century warming, which is only reproduced by simulations that include anthropogenic forcing (essentially greenhouse gas concentrations), an observation supporting the results of Barnett et al. [2005] and other studies [e.g., Crowley, 2000; Stott et al., 2000; Bauer et al., 2003]. However, we also stress caution with respect to the stronger coherence, compared to HADCM3, between the coral reconstruction and the ECHO-G model, which, unrealistically, does not include aerosol-induced cooling as a forcing factor [Osborn et al., 2006]. The higher correlation may simply reflect more trend within the ECHO-G modeled data, and the greater amplitude (Figure 4b) noted for ECHO-G compared to the coral reconstruction and HADCM3 may reflect this exclusion of aerosol forcing in the ALL-forcing run.

[30] As expected, comparison with an independent tree ring based reconstruction of NINO3 SSTs and related instrumental data shows that high-frequency variability in tropical mean temperatures is dominated by ENSO. This coherence is relatively stable (r = ~0.5) since the late 19th century, but is nonsignificant or negative over the periods ~1625–1675, ~1700–1775 and ~1825–1875, likely reflecting changes in spatial teleconnections or the quality of the reconstructions through time. MTM spectral analysis shows that the HADCM3 model expresses ENSO-like variability within the classical band width (similar to the coral reconstruction), but this variability is not so strong in ECHO-G.

[31] The tropics, because of its relatively homogenous climate compared to the high latitudes, is a perfect test bed from which large-scale temperature reconstructions can be developed from sparse proxy networks [Evans et al., 1998], and where models may perform more robustly, as the signal-to-noise ratio of forced to internal variability is strongest [Tett et al., 2006]. In this study, utilizing a slightly different coral proxy data set (~25% common data overlap), we have added to earlier work [Evans et al., 1998, 2000, 2002] and evaluated the potential for reconstructing large-scale tropical temperatures by utilizing a modest network of coral proxy series. Although a significant new contribution to the understanding of tropical temperature variability, the current reconstruction is restricted both in space (i.e., there are no data that correlate significantly with local SSTs in the Atlantic, and no data have been included from climatically important locations, e.g., the Indonesia Warm Pool region [D’Arrigo et al., 2006b]) and length, highlighting the need for further effort in sampling new long coral series and also possibly including fossil coral records [Tudhope et al., 2001; Cobb et al., 2003] so long as dating control is robust and precise.

Appendix A: Statistical Tests Used to Assess the Coral Reconstruction

A1. Reduction of Error (RE)

[32] The RE statistic tests whether a reconstruction provides a better estimate of climatic variability than climatology (i.e., the mean of the meteorological data in the calibration period [Cook and Kairiukstis, 1990; Cook et al., 1994]). It is calculated as

\[
RE = 1.0 - \frac{\sum_{i=1}^{n} (x_i - \bar{x}_i)^2}{\sum_{i=1}^{n} (x_i - \bar{x})^2},
\]

where \(x_i\) and \(\bar{x}_i\) are the actual and estimated data in year \(i\) of the verification period, and \(\bar{x}_i\) is the mean of the actual data in the calibration period. The value of RE can range from negative infinity to a maximum value of 1.0 which indicates perfect estimation. If the total difference between the estimates and the actual data is less than the total difference between the calibration mean and the actual data, the RE statistic will be positive. It should be noted that if there is trend within the instrumental data (i.e., the mean of the calibration and verification periods are different), RE can be greater than the square of the Pearson’s correlation coefficient. This behavior suggests that RE should be interpreted cautiously when the instrumental data portrays a strong trend and for this reason we also utilized the Coefficient of Efficiency.

A2. Coefficient of Efficiency (CE)

[33] The CE statistic is similar to RE except that its benchmark for determining model skill is the verification
period and not the calibration period (i.e., the difference between RE and CE is in the denominator term). It can be described as an expression of the true $r^*$ of a regression equation when applied to new data [Cook and Kairiukstis, 1990; Cook et al., 1994]. The equation used to calculate CE is expressed as

$$CE = 1.0 - \frac{\sum (x_i - \bar{x})^2}{\sum (x_i - \bar{x}_r)^2},$$

where $\bar{x}$ is the mean of the actual data in the verification period. Like RE, CE can range from negative infinity to a maximum value of 1.0 which indicates perfect estimation. Again a positive value signifies that the regression model has some skill. When $\bar{x} = \bar{x}_r$, $CE = RE$. When $\bar{x} \neq \bar{x}_r$, RE will be greater than CE by a factor related to the difference in means. In comparison to RE, CE is more difficult to pass, and is therefore a more rigorous verification statistic.

A3. Sign Test (ST)

[34] The ST is a nonparametric method based upon the number of agreements and disagreements in sign of departure from the mean in the observed and reconstructed series [Cook et al., 1994]. As a general rule, if the number of agreements exceeds the number of disagreements by greater than expected by chance alone, the reconstruction passes. When the number (N) of observations in the verification period is $\geq 50$, the number of agreements required for the test to indicate significance at the 95% confidence limit is

$$\left( N - 1 - 1.96\sqrt{N} \right)/2$$

For $N > 50$, the normal distribution is used to derive significance. When $N < 50$, the cumulative distribution tables for the binomial distribution can be used instead to test significance. A significance table is detailed by Fritts [1976].

A4. Durbin-Watson (DW) Statistic

[35] The DW statistic tests for the presence of first-order autocorrelation in the residuals of a regression equation [Shaw and Wheeler, 1985; Draper and Smith, 1998]. The test compares the residual for time period $t$ with the residual from time period $t-1$ and measures the significance of the correlation between these successive comparisons. The statistic is formulated thus:

$$DW = \frac{\sum_{i=2}^{n} (e_i - e_{i-1})^2}{\sum_{i=1}^{n} (e_i)^2},$$

where $e$ is the residual value ($Y_i - \hat{Y}_i$) and $t$ is the time period. The statistic, ranging from 0 to 4, can test for the presence of both positive and negative autocorrelation in the residuals. A value of 2 denotes no autocorrelation while values less (greater) than 2 denoting positive (negative) autocorrelation. Significance tables are detailed by Shaw and Wheeler [1985] and Draper and Smith [1998].

[36] Acknowledgments. This research was funded by the European Union (grant EVK2-CT-2002-00160, SO&P); the two Hadley Centre authors were also funded by the Public Meteorological Service Research and Development Contract. The ECHO-G NATURAL run was provided by Irina Fast of the Institute for Meteorology, Freie Universitat Berlin, and was performed at the German Climate Computing Centre (DKRZ), Hamburg. The ECHO-G ALL run was provided along with comments and advice from Eduardo Zorita and Julie Jones of the GKSS Forschungszentrum, Hamburg. The final version of this paper was greatly improved by comments and reviews from Mike Evans, Rosanne D’Arrigo, Dave Frank, two anonymous reviewers, and the Associate Editor.

References


Dawdy, D. R., and N. C. Matalas (1964), Statistical and probability analysis of hydrologic data, part III: Analysis of variance, covariance and time-

K. Briffa and T. Osborn, Climatic Research Unit, School of Environmental Sciences, University of East Anglia, Norwich NR4 7TJ, UK. (k.briffa@uea.ac.uk; t.osborn@uea.ac.uk)

P. Brohan, Hadley Centre for Climate Prediction and Research, Met Office, Exeter EX1 3PB, UK. (philip.brohan@metoffice.gov.uk)

S. Tett, Hadley Centre (Reading Unit), Met Office, University of Reading, Meteorology Building, Reading RG6 6BB, UK. (simon.tett@metoffice.gov.uk)

A. Tudhope and R. Wilson, School of GeoSciences, Grant Institute, University of Edinburgh, West Mains Road, Edinburgh EH9 3JW, UK (sandy.tudhope@ed.ac.uk; rob.wilson@ed.ac.uk)