Nonstationary Australasian Teleconnections and Implications for Paleoclimate Reconstructions

AILIE J. E. GALLANT
School of Earth Sciences, University of Melbourne, Parkville, and School of Geography and Environmental Science, Monash University, Clayton, Victoria, Australia

STEVEN J. PHIPPS
Climate Change Research Centre, and ARC Centre of Excellence for Climate System Science, University of New South Wales, Sydney, New South Wales, Australia

DAVID J. KAROLY
School of Earth Sciences, University of Melbourne, Parkville, Victoria, Australia

A. BRETT MULLAN
National Institute of Water and Atmospheric Research Ltd., Wellington, New Zealand

ANDREW M. LORREY
National Institute of Water and Atmospheric Research Ltd., Auckland, New Zealand

(Manuscript received 12 June 2012, in final form 18 April 2013)

ABSTRACT

The stationarity of relationships between local and remote climates is a necessary, yet implicit, assumption underlying many paleoclimate reconstructions. However, the assumption is tenuous for many seasonal relationships between interannual variations in the El Niño–Southern Oscillation (ENSO) and the southern annular mode (SAM) and Australasian precipitation and mean temperatures. Nonstationary statistical relationships between local and remote climates on the 31–71-yr time scale, defined as a change in their strength and/or phase outside that expected from local climate noise, are detected on near-centennial time scales from instrumental data, climate model simulations, and paleoclimate proxies.

The relationships between ENSO and SAM and Australasian precipitation were nonstationary at 21%–37% of Australasian stations from 1900 to 2009 and strongly covaried, suggesting common modulation. Control simulations from three coupled climate models produce ENSO-like and SAM-like patterns of variability, but differ in detail to the observed patterns in Australasia. However, the model teleconnections also display nonstationarity, in some cases for over 50% of the domain. Therefore, nonstationary local–remote climatic relationships are inherent in environments regulated by internal variability. The assessments using paleoclimate reconstructions are not robust because of extraneous noise associated with the paleoclimate proxies.

Instrumental records provide the only means of calibrating and evaluating regional paleoclimate reconstructions. However, the length of Australasian instrumental observations may be too short to capture the near-centennial-scale variations in local–remote climatic relationships, potentially compromising these reconstructions. The uncertainty surrounding nonstationary teleconnections must be acknowledged and quantified. This should include interpreting nonstationarities in paleoclimate reconstructions using physically based frameworks.

1. Introduction

Climatic teleconnections describe dynamical links between the states of two remote atmosphere–ocean systems...
The processes defining teleconnections are numerous but include the propagation of atmospheric and/or oceanic wave trains, such as Kelvin or Rossby waves (Gill 1980; Hoskins and Karoly 1981) and/or the mass movement of transient eddies such as weather systems (Seager et al. 2005; Li and Wettstein 2012). This propagation induces remote effects, resulting in hemispheric or even global responses in climate.

Teleconnections are responsible for a significant proportion of Australasian climate variability, defined as the region of the southwest Pacific that includes Australia and New Zealand (Fig. 1). Annual and subannual fluctuations in the Australasian atmosphere and ocean state have been linked to the El Niño–Southern Oscillation (ENSO) (McBride and Nicholls 1983; Gordon 1986; Karoly 1989; Drosdowsky 1993; Mullan 1995; Kidson and Renwick 2002; Kidson et al. 2002), the southern annular mode (SAM) (Karoly 1990; Hendon et al. 2007; Kidston et al. 2009), modes of variability in the Indian Ocean (Nicholls 1989; Saji et al. 1999; Ansell and Reason 2000; Verdon and Franks 2005; Ashok et al. 2007; Ummenhofer et al. 2009), and the Madden–Julian oscillation (Wheeler and Hendon 2004; Wheeler et al. 2009).

The relationships between local and remote climates may be altered because of the effects of local small-scale transients (e.g., synoptic-scale weather systems), influences from other dependent and independent dynamical systems (Meyers et al. 2007; Fogt et al. 2011), and changes in the large-scale state of the coupled atmosphere–ocean system. Therefore, the strengths of Australasian teleconnections (i.e., the relative influence of the remote dynamical mechanism on regional and local climates) are not perfectly stationary in time (Mullan 1995; Nicholls et al. 1996; Verdon and Franks 2006; Risbey et al. 2009).

Reconstructions of the preinstrumental climate using paleoclimate proxies have utilized teleconnection patterns to estimate remote (local) climate variations from a local (remote) source. For example, past variations in ENSO have been inferred from tree-ring widths from New Zealand and Australia (Cook et al. 2006; Fowler 2008; Fowler et al. 2012). Teleconnections to the Pacific Ocean (Verdon and Franks 2006; McGowan et al. 2009), Antarctica (van Ommen and Morgan 2010), and multiple remote source regions (Gallant and Gergis 2011; Gergis et al. 2012) have been used to infer past variations in Australian hydroclimates.

These and other paleoclimate studies are based on the paradigm of uniformitarianism on sufficiently long time scales (Nairn 1965), which assumes that present-day climatic processes are key to understanding past climate variations. Based on this principle, the spatial and temporal coherence in the climate system is presumed to be large (i.e., stable). The assumption is made because mechanistic studies of the present-day climate have shown partial regulation of local climates by large-scale dynamical mechanisms (e.g., ENSO) that have remote and unique effects, for example, through altering local circulation patterns (Drosdowsky 1993; Kidson 2000).

Most paleoclimate research acknowledges the potential problems surrounding the assumption of stationarity but does not, or cannot, perform any rigorous testing. In practice, those past paleoclimate studies previously described have assumed that local–remote climate relationships are static and have characterized these relationships using instrumental data for calibration. However, several studies have described large variations in the behavior of Australasian local–remote climate relationships on interdecadal and longer time scales in the instrumental record (Mullan 1995; Nicholls et al. 1996; Power et al. 1999; Pezza et al. 2007; Ummenhofer et al. 2011). If these variations represent changes to Australasian teleconnection patterns that are outside the influence of a change induced by local climate noise, then a local–remote climate relationship...
should be considered dynamic, not static. Thus, the assumptions of stationarity applied to paleoclimate research that reconstructs remote climate drivers from local information, or vice versa, become tenuous.

This study tests the validity of the assumption that Australasian teleconnections are stationary in time. We define these teleconnections as the statistical relationship between interannual variations in seasonal Australasian temperature and precipitation and two remote dynamical drivers, ENSO and SAM. Stationarity, in other words, a static relationship, is defined when the multidecadal (31–71 year) variations in these statistical local–remote climatic relationships are consistent with stochastic variation associated with local climate noise.

Our study is presented as follows. First, the data and methods are introduced in sections 2 and 3. We then provide an overview of the relationships between the local Australasian climate and ENSO and SAM from the instrumental data in section 4. Stationarity in these local–remote climate relationships is tested using instrumental, climate model, and paleoclimate proxy data in sections 5 and 6. The associated issues for paleoclimate research are presented in section 7, and a discussion of our findings in section 8.

2. Data

Instrumental stations, paleoclimate proxies, and climate model simulations were employed. Australasia is defined as the area encompassing the Australian and New Zealand landmasses approximately covered by the domain 55°S–0°, 110°E–175°W (Fig. 1).

a. Instrumental station data

Monthly mean temperature and precipitation data from stations in Australia and New Zealand were used. Australian high-quality station data were obtained from the Australian Bureau of Meteorology (http://www.bom.gov.au/climate/change/datasets/datasets.shtml). These represent the highest quality data available and are free from artificial discontinuities and trends. Further information on these precipitation and temperature data and a full list of stations are in Lavery et al. (1997) and Trewin (2001), respectively. Data for New Zealand were provided by the National Institute of Water and Atmospheric Research (NIWA) and represent the best and longest station records available in the country (Table 1). The data have been extensively checked and, although some precipitation records represent a composite of multiple stations, artificial inhomogeneities were not detected (G. Griffiths, National Institute of Water and Atmospheric Research, 2011, personal communication).

A total of 325 precipitation and 109 temperature stations were available for Australia and New Zealand (Fig. 1). There were 307 and 18 precipitation records and 103 and 6 temperature records from Australia and New Zealand, respectively. While many precipitation stations began recording in the late nineteenth and early twentieth centuries, most temperature records did not begin until the late 1950s and early 1960s.

The links between Australasian climate variables and the two largest sources of interannual variability in the Southern Hemisphere, ENSO (Diaz and Markgraf 2000) and the SAM (Karoly 1990; Thompson and Wallace 2000), were examined. Indices were used that represent variations in the centers of action of these modes, away from Australasia. ENSO was represented by the Niño-3.4 index (Trenberth 1997), defined as the mean sea surface temperature anomaly in the central-equatorial Pacific Ocean (5°S–5°N, 120°–170°W) and is available from the U.S. National Center for Atmospheric Research

Table 1. Data from the New Zealand monthly-mean temperature and precipitation stations used in this study. Station records denoted as a composite include at least two records associated with a slight change in site location and/or observing practice (e.g., a change from manual recording to an automatic weather station; G. Griffiths, National Institute of Water and Atmospheric Research, 2011, personal communication).

<table>
<thead>
<tr>
<th>Station</th>
<th>Lon</th>
<th>Lat</th>
<th>Period</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean temperature</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Milford Sound</td>
<td>167.92°E</td>
<td>44.67°S</td>
<td>1934–2011</td>
</tr>
<tr>
<td>Invercargill Airport</td>
<td>168.33°E</td>
<td>46.42°S</td>
<td>1948–2011</td>
</tr>
<tr>
<td>New Plymouth Composite</td>
<td>174.18°E</td>
<td>39.01°S</td>
<td>1944–2011</td>
</tr>
<tr>
<td>Paraparaumu Airport</td>
<td>174.98°E</td>
<td>40.91°S</td>
<td>1953–2011</td>
</tr>
<tr>
<td>Tauranga Composite</td>
<td>176.16°E</td>
<td>37.69°S</td>
<td>1913–2011</td>
</tr>
<tr>
<td>Napier–Nelson Park</td>
<td>176.91°E</td>
<td>39.50°S</td>
<td>1905–2011</td>
</tr>
<tr>
<td>Precipitation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Milford Sound</td>
<td>167.92°E</td>
<td>44.67°S</td>
<td>1929–2010</td>
</tr>
<tr>
<td>Invercargill Airport</td>
<td>168.33°E</td>
<td>46.42°S</td>
<td>1939–2010</td>
</tr>
<tr>
<td>Queenstown</td>
<td>168.66°E</td>
<td>45.04°S</td>
<td>1930–2010</td>
</tr>
<tr>
<td>Campbell Island</td>
<td>169.15°E</td>
<td>52.55°S</td>
<td>1941–2010</td>
</tr>
<tr>
<td>Dunedin–Musselburgh Composite</td>
<td>170.51°E</td>
<td>45.90°S</td>
<td>1918–2010</td>
</tr>
<tr>
<td>Hokitika Airport</td>
<td>170.99°E</td>
<td>42.72°S</td>
<td>1964–2010</td>
</tr>
<tr>
<td>Christchurch Airport</td>
<td>172.54°E</td>
<td>43.49°S</td>
<td>1943–2010</td>
</tr>
<tr>
<td>Nelson Airport</td>
<td>173.23°E</td>
<td>41.30°S</td>
<td>1941–2010</td>
</tr>
<tr>
<td>New Plymouth Composite</td>
<td>174.18°E</td>
<td>39.01°S</td>
<td>1944–2010</td>
</tr>
<tr>
<td>Wellington–Kelburn</td>
<td>174.77°E</td>
<td>41.29°S</td>
<td>1928–2010</td>
</tr>
<tr>
<td>Paraparaumu Airport</td>
<td>174.98°E</td>
<td>40.91°S</td>
<td>1945–2010</td>
</tr>
<tr>
<td>Wanganui–Spiggens Park</td>
<td>175.04°E</td>
<td>39.94°S</td>
<td>1937–2010</td>
</tr>
<tr>
<td>Hamilton–Ruakura</td>
<td>175.31°E</td>
<td>37.78°S</td>
<td>1907–2010</td>
</tr>
<tr>
<td>Tauranga Airport</td>
<td>176.20°E</td>
<td>37.67°S</td>
<td>1941–2010</td>
</tr>
<tr>
<td>Napier–Nelson Park</td>
<td>176.91°E</td>
<td>39.50°S</td>
<td>1964–2010</td>
</tr>
<tr>
<td>Gisborne Airport</td>
<td>177.99°E</td>
<td>38.66°S</td>
<td>1937–2010</td>
</tr>
<tr>
<td>Raoul Island</td>
<td>177.92°W</td>
<td>29.25°S</td>
<td>1937–2010</td>
</tr>
<tr>
<td>Chatham Island</td>
<td>176.57°W</td>
<td>43.95°S</td>
<td>1951–2010</td>
</tr>
</tbody>
</table>
from 1871 to 2009 (http://www.cgd.ucar.edu/cas/catalog/climind/Nino_3.4_indices.html). The SAM index represents the difference between the normalized zonally averaged sea level pressure anomalies at 40° and 60°S, which were generated from 12 stations near these latitude lines (Marshall 2003). The SAM index was available from 1957 to 2009. Visbeck (2009) applied the Marshall (2003) method with relaxed criteria for station selection and an assumption of atmospheric mass conservation between the Southern Hemisphere subtropics and midlatitudes. This produced an extended station dataset that included subtropical stations, which were used to construct a monthly SAM index from 1884 to 2005. For their overlapping period, the Visbeck SAM index (hereafter referred to as SAMv) correlates strongly with the Marshall SAM index ($r = 0.85$) (Visbeck 2009).

The Niño-3.4 and SAM indices and mean temperature and precipitation data were examined from 1960 to 2009. The precipitation stations were examined from 1900 to 2005 for the SAMv index, and from 1900 to 2009 for the Niño-3.4 index. Hereafter, these data are described as being examined over the period 1900-2005/09 to denote the different data availability of the SAMv and Niño-3.4 indices. A small subset of extended mean temperature stations was also examined from 1920 to 2005/09.

b. Climate model data

Monthly surface temperature and precipitation data from multicentennial control simulations from three different coupled ocean–atmosphere general circulation models (GCMs) were employed, namely the Geophysical Fluid Dynamics Laboratory Climate Model version(s) 2.0 (GFDL CM2.0) and 2.1 (GFDL CM2.1), and the third climate configuration of the Met Office (UKMO) Hadley Centre Unified Model (HadCM3). These three models were selected as they have some of the most realistic dynamical simulations of ENSO from phase 3 of the Coupled Model Intercomparison Project (CMIP3) suite of GCMs (Guilyardi et al. 2009). The Niño-3.4 and SAM indices were generated from model output using the same definitions as for the instrumental data.

The GFDL CM2.0 and CM2.1 control simulations were 500 years long and are described in detail in Delworth et al. (2006). Both are fully coupled atmosphere–ocean models with no flux adjustments and differ primarily in their dynamical treatment of the atmosphere. They have the same grid resolution and the land and atmosphere components utilize a 2.0° latitude by 2.5° longitude grid with 24 vertical levels. The ocean component has 50 vertical levels and is on a 1.0° latitude by 1.0° longitude grid poleward of 30° with resolution increasing to $\frac{1}{2}^\circ$ at the equator.

The control simulation from the UKMO-HadCM3 model spans 341 years. The UKMO-HadCM3 is also a fully coupled model without flux adjustments but has coarser resolution in both atmosphere and ocean components compared to the GFDL models. The resolution of the atmospheric component is 2.75° latitude by 3.75° longitude with 19 vertical levels, and the ocean component has resolution of 1.25° latitude by 1.25° longitude with 20 vertical levels. Detailed descriptions of the model can be found in Johns et al. (1997) and Gordon et al. (2000).

c. Paleoclimate proxy data

Three published Australasian paleoclimate proxy records and reconstructions, and two remote proxy reconstructions of the Niño-3.4 index, were examined.

The first Australasian record was a precipitation reconstruction for tropical northeast Queensland, Australia (Lough 2007). Variations in luminescence from up to 25 coral cores were used to generate four local river flow reconstructions within the approximate domain 17°–23°S, 146°–151°E. Reconstructed river flow was then regressed against precipitation across the same domain, producing a precipitation reconstruction spanning AD 1631–2005.

The second Australasian proxy record was the Agathis australis (kauri) tree-ring master chronology, generated from 196 trees at 14 sites in the northwest of the North Island of New Zealand (Fowler et al. 2008). Nonclimatic trends were removed using a 200-yr spline. The kauri tree-ring width is responsive to local temperature and precipitation (Buckley et al. 2000) but shows a stronger statistical response to the Southern Oscillation index and, therefore, ENSO (Fowler et al. 2008). The chronology used here was uncalibrated and the record spans AD 1580–2002.

The third Australasian record was a temperature reconstruction for Tasmania. The reconstruction was developed from living and subfossil tree-ring records of the Huon Pine from around Mt. Read in western Tasmania (Cook et al. 2000). The well-replicated record is sensitive to warm season (November–April) temperatures and has been filtered to remove outliers associated with summer freeze events. The full record extends from 1600 BC to AD 1991.

The first reconstruction of the Niño-3.4 index was developed from up to eight coral records in the central and eastern equatorial Pacific in a region where the local observed SSTs were strongly, positively correlated with the observed Niño-3.4 index. The reconstruction (hereafter COA) has limited input data prior to 1850, making replication and validation difficult (Wilson et al. 2010). However, we use it because of its proximity to the Niño-3.4 region. The reconstruction extends from AD 1607–1998.

The second reconstruction of the Niño-3.4 index was generated from a network of 404 tree-ring width records in the southwest United States and Mexico (hereafter
called TEXMEX) (Cook et al. 2008). The tree-ring network utilizes the existence of teleconnections between North America and the tropical Pacific Ocean. The tree-ring records were used to develop a field reconstruction of SSTs in the tropical Pacific Ocean over a domain that included the Niño-3.4 region previously defined. The reconstructed SST variations were then used to generate the TEXMEX Niño-3.4 index using the same methodology as in Trenberth (1997).

The relationships between the Niño-3.4 reconstructions and the two Australasian proxy reconstruction data-sets were examined over the period of AD 1631–1998. Those running correlations calculated using the Tasmanian temperature reconstruction were computed to AD 1991 only.

3. Historical relationships between ENSO/SAM and the Australasian climate

The historical relationships between ENSO/SAM and the Australasian climate are well understood. A review of these relationships is now presented, employing data that were described in section 2.

Typically, higher sea level pressures over the west Pacific have occurred during El Niño events, which are associated with changes to the Walker circulation (Troup 1965; Drosdowsky and Williams 1991; Allan et al. 1996; Diaz and Markgraf 2000). The correlations between the Niño-3.4 index and Australian stations confirmed that much of the country was warm and dry during past El Niño events and cool and wet during La Niña events, with the strongest relationships during austral spring and early summer (Fig. 2) (McBride and Nicholls 1983; Drosdowsky and Williams 1991; Nicholls et al. 1996; Risbey et al. 2009). During December–February (DJF) the magnitudes of these correlations were often above 0.5, although inverted for precipitation.

There were inverse correlations between the Niño-3.4 index and mean temperatures in New Zealand. These were strongest on the North Island ($r > 0.53$) and link El Niño (La Niña) events to cooler (warmer) temperatures (Mullan 1995, 1998; Kidson 2000). In New Zealand, the precipitation patterns associated with ENSO are complex owing to orographic influences (Fig. 2). La Niña events have previously been linked to wetter conditions, especially in the North Island, and El Niño events with wetter conditions in the south and west of the South Island and dry conditions elsewhere, particularly during DJF (Mullan 1998; Kidson 2000; Lorrey et al. 2007).

A positive SAM is defined as having stronger than normal high-latitude westerly winds and a more poleward storm track, and vice versa for negative SAM (Marshall 2003). The positive correlations with surface pressure...
over the southern half of the Australasian domain are consistent with this definition (Fig. 3).

The links between SAM and the Australasian climate varied seasonally and spatially (Fig. 3) (Hendon et al. 2007; Meneghini et al. 2007; Kidston et al. 2009; Risbey et al. 2009). Warm conditions in southern Australia and New Zealand were associated with positive SAM during DJF (Kidston et al. 2009). These relationships were inverted during June–August (JJA) in southern Australia only. During DJF almost all stations on the Australian mainland had higher precipitation totals with positive SAM (Hendon et al. 2007; Meneghini et al. 2007; Risbey et al. 2009). However, for Tasmania and much of New Zealand, positive SAM was associated with lower precipitation totals in both seasons.

4. Methods

This study examines multidecadal variations in the strength of interannual relationships between Australasian temperature and precipitation and ENSO and SAM. Given that teleconnections have distinct seasonal signals, monthly data were averaged over the austral summer, December–February (DJF), and winter, June–August (JJA). Stations from the high-quality instrumental datasets were included only if they were more than 80% complete. Seasonal anomalies were generated relative to 1971–2000 for the instrumental data and relative to all years from the model simulations and the paleoclimate proxy data. However, note that, because anomalies are linear transformations, the choice of base period has no effect on the correlation. All data were detrended using linear regression so that there was no influence from long-term trends, including model drift.

The relationships between Australasian and remote climate data were measured as 31-yr running correlations. While correlations do not include any dynamical interpretation, they are useful tools for elucidating climatic relationships. Using 31 years minimized data loss due to the use of running correlations and was particularly important for the instrumental data, which mostly spanned only 50 years. Longer windows of 51 and 71 years were also applied to the model data to examine sensitivity of the results to the sample length (see section 6).

Nonstationarity implies the existence of dynamical (i.e., variable) relationships between local and remote climates, rather than static relationships. Our null hypothesis was that the multidecadal-scale (31–71 yr) variations in the correlations between the Australasian mean temperature or precipitation and the Niño-3.4 and SAM indices are stationary. A nonstationary teleconnection was defined if the running correlation time series fell outside a two-tailed 95% confidence interval generated from synthetic, stochastic data.

We used synthetic running correlations, similar to van Oldenborgh and Burgers (2005) and Sterl et al. (2007), to develop our 95% confidence interval. The synthetic running correlations were developed from stochastic precipitation and mean temperature series that had the same statistical properties as the instrumental, model, or

FIG. 3. As in Fig. 2 but for the SAM index. The time series shows the SAM index from 1960 to 2009.
proxy data and the same relationships with the Niño-3.4 or SAM indices. Thus, the relationship was assumed to be static and to vary because of stochastic variation in local mean temperature or precipitation only.

These variables were defined as

\[ \nu(t) = a_0 + a_1 c(t) + \sigma_v \sqrt{1 - r^2} \left[ \eta_v(t) + \beta \eta_v(t - 1) \right], \]

(1)

where \( \nu(t) \) is the synthetic precipitation or temperature series. The coefficients \( a_0 \) and \( a_1 \) describe the regression relationship between local precipitation or temperature and the Niño-3.4 or SAM index, \( c(t) \). The final term describes the magnitude of the local climate noise that influences the strength of this regression relationship. This was weighted by the standard deviation of the variable \( \sigma_v \) and the proportion of the variance not associated with the regression relationship \( \sqrt{1 - r^2} \), where \( r \) is the correlation between the climate index and precipitation or mean temperature. We used red noise \( [\eta_v(t) + \beta \eta_v(t - 1)] \), defined as the combination of random Gaussian noise \( \eta_v(t) \) with unit standard deviation and zero mean, and the autocorrelation at lag 1 (\( \beta \)).

One thousand variants of \( \nu(t) \) were generated and the running correlations between these and the Niño-3.4 or SAM index, \( c(t) \), were computed. From these, a probability density function (PDF) of the range of running correlations that could be expected from a relationship only influenced by local climate noise was generated and the 95% confidence interval was estimated. Note that a second confidence interval was generated using an estimate of the PDF directly from the data being tested (i.e., an empirical confidence interval). The results were insensitive to the method applied, particularly for the longer GCM and paleoclimate proxy datasets.

In statistics, the variance of strong sampled correlations is implicitly lower than weak sampled correlations due to the finite bounds of correlations between minus one and plus one. So, all running correlations were translated to Fisher z scores (Wilks 2011) in order to stabilize variance for statistical testing using

\[ Z = \frac{1}{2} \ln \left( \frac{1 + r}{1 - r} \right), \]

(2)

where \( Z \) is the Fisher z score and \( r \) is the correlation. The resulting PDFs of \( Z \) for each location were normally distributed, so the 95% confidence interval (CI) was approximated as CI = \( Z \pm 1.96 \text{SD}_Z \), where \( Z \) was the mean from all simulations and \( \text{SD}_Z \) the standard deviation.

The mean of the 31-yr running correlations (\( \bar{r} \)) was used to define the typical strength of the teleconnection. The mean was calculated as an inverse transform of the mean Fisher z score (\( \bar{Z} \)) given by

\[ \bar{r} = \frac{e^{\bar{Z}} - e^{-\bar{Z}}}{e^{\bar{Z}} + e^{-\bar{Z}}}. \]

(3)

Spatiotemporal coherence in any nonstationary time series was examined by computing empirical orthogonal functions (EOFs) and their related principal components (PCs) from the running z scores (Wilks 2011). All stations were weighted using Thiessen polygons (Thiessen 1911) so that the EOFs were not biased toward areas with higher station density.

Field significance was computed to determine whether the number of stations registering nonstationarity was greater than that expected by random chance, as in Livezey and Chen (1983). Field significance levels were adjusted to account for the reduced number of degrees of freedom associated with this spatial coherence between stations and grid boxes. Field significance was determined as in Sterl et al. (2007) using adjusted degrees of freedom \( N \), defined as

\[ N = \sin^{-2} \frac{L}{2}. \]

(4)

where \( L \) is the decorrelation length scale: \( L \) was computed for each station \( i \) (\( L_i \)) and was the median e-folding distance in degrees along a great circle arc, that is, the distance at which the correlation between station \( i \) and all other stations reached 1/e. The distance \( L_i \) varied between stations and was dependent on aspects such as topography and remoteness from other stations. For example, New Zealand stations generally had smaller values of \( L_i \) compared to Australian stations. Thus, we used \( L \) as the median of \( L_i \).

To assess the ability of the GCMs to simulate the relationships between local and remote data, irregular station observations had to be translated to the model grids to allow for direct comparison. The instrumental grids were computed as area averages from nearby stations where the station weights were defined using Thiessen polygons. A station was only included in a gridbox average if it was within a minimum distance of influence from the grid box. Given the differences between \( L_i \) as previously described, the distance of influence was defined as the conservative estimate of \( L = \min(L_i) \). The grid boxes where no station was within distance \( L \) were ignored.

5. Testing stationarity in the relationships between ENSO/SAM and the Australasian climate

Stationarity in the variations of the 31-yr running correlations between Australasian temperature and precipitation
and the Niño-3.4 and SAM indices were tested using instrumental (section 5a), climate model (section 5b), and paleoclimate proxy (section 5c) data. The results highlight the existence of a significant number of non-stationary local–remote climate relationships.

a. Instrumental data

Nonstationarities were detected in the 31-yr running correlations between the observed Niño-3.4 and SAM indices and Australasian meteorological stations. These were most evident in the longer time series beginning in 1900 or 1920. In the shorter time series spanning 1960–2009, there was a total of 49 nonstationary time series (Fig. 4). These relationships were mostly associated with precipitation (8% of stations) and mean temperature (26% of stations) and the SAM index during DJF and were field significant at the 5% level.

Nonstationarities were detected in eight cases for temperature stations with data available from 1920. Five of these were in southeast Australia or on the east coast of the North Island of New Zealand. These stations had nonstationary relationships with the SAMv during DJF or JJA and represented 33% and 43% of all available temperature stations, respectively. The remaining three cases represented 50% of all available temperature stations and all occurred on the southeast Australian coast for the DJF relationships between mean temperature the Niño-3.4 index.

Using precipitation data from 1900, nonstationary correlations with the Niño-3.4 and SAMv indices were detected at 61 and 99 stations (21% and 34% of all available stations) during DJF and 68 and 108 stations (23% and 37% of all available stations) during JJA (Fig. 5). This represents a 4-fold to 68-fold increase compared to the data from 1960. The number of nonstationary locations was field significant at the 5% level in every case.

The increases in the number of nonstationarities using the longer time series suggest that variations in local–remote climate relationships are inconsistent with stochastic variations in the local climate on time scales longer than 50 yr. Provided the running correlations represent a real change in the relative influence of the ENSO or SAM on the Australasian climate, then the station time series should display some spatiotemporal coherence, which was examined using EOF analysis.

Figure 5 identifies those stations that most strongly loaded the first two EOFs of the nonstationary 31-yr running correlations between precipitation and the Niño-3.4 and SAMv indices from 1900 to 2005/09 and shows their corresponding PCs. The strongest EOF loadings were commonly concentrated in eastern and/or southwest Australia. Furthermore, the temporal coherence of these modes was consistent between those relationships based on the Niño-3.4 and SAMv indices. The strongest similarities were between the variations of the PCs computed using the Niño-3.4 and SAMv indices during DJF and the SAMv index during JJA. The magnitudes of the correlations between these PCs were between 0.84 and 0.94, although they were sometimes out of phase (Fig. 5). Accounting for sample sizes
and large autocorrelations as in Bretherton et al. (1999), these 31-yr running correlations were statistically significant at the 3% level.

Figure 6 shows examples of the similarities in the variations of the 31-yr running correlation time series at the stations most strongly associated with the first PCs from all precipitation-based relationships from 1900–2005/09. Note that these precipitation-based variations also resemble the variations in the nonstationary relationships based on temperature from 1920–2005/09. During DJF, the common signals were manifest at stations in western and eastern Australia but were out of phase in the two regions (Fig. 6). During JJA, the variations associated with SAMv were out of phase compared to those associated with the Niño-3.4 index, but were in phase at all stations regardless of their location.

This temporal coherence between stations suggests that there might be a common modulator of local–remote ENSO- and SAM-based climatic relationships in Australasia. Note that the SAMv index is partially based on station data from the subtropics early in the record and so may be displaying some dependence on ENSO. However, the similarities in the data from 1960 with the Marshall (2003) SAM index, which uses high-latitude stations only, suggests this is not necessarily the case.

Any mechanism causing the larger-than-expected variations in the teleconnected relationships would fall into one or both of two general categories. The first is that there is a common, but external, process affecting the relative influence that ENSO and SAM have on local climatic variations. The second is that there is a direct change in the characteristic behavior of ENSO and/or SAM (Kestin et al. 1998; Power et al. 1999; Verdon et al. 2004) perhaps resulting from, or at least including, an interaction between the two (Fogt et al. 2011).

To investigate, we identified significant associations between the changes in the characteristics of the Niño-3.4 or SAM indices and the timing of nonstationarities,
defined as the time when the running correlation fell outside the two-tailed 95% confidence interval. This identified how likely the nonstationarity was to be directly associated with a change in the behavior of the driving mechanism(s) compared to an injection of external noise or forcing into the system. Two statistics representing the characteristics of ENSO and the SAM were calculated: (i) the mean state, defined as the 31-yr mean value of the Niño-3.4 and SAM indices, and (ii) the magnitude of the interannual variations, defined as the standard deviation of each index over running 31-yr periods. These statistics were averaged over the $n$ nonstationary periods identified in each 31-yr running correlation time series. The nonstationary values were then compared to a 95% confidence interval generated from up to 120 randomized permutations of the remaining $n$-averaged stationary periods, defined as those times where the 31-yr correlation did not exceed the 95% confidence interval. Both successive and nonsuccessive stationary periods were used to generate the confidence interval to ensure that any temporal dependence was retained. We defined a statistically significant association between the state of the remote climate driver and nonstationary local–remote climatic relationships when the mean value calculated over the nonstationary periods was outside this 95% confidence interval.

There were significant relationships between nonstationarities and the statistical characteristics (i.e., the mean and variance) of the Niño-3.4 and SAM indices at 23% and 40% of nonstationary precipitation stations, respectively, during JJA. There were fewer significant relationships during DJF, at 16% and 18% of nonstationary precipitation stations. Likewise, more nonstationary temperature stations had statistically significant links to the statistical characteristics of the climate indices during JJA (two stations) than in DJF (one station). The significant changes described were more likely to be associated with a change in variance than the mean and more likely to occur in JJA. These results suggest that internal changes in the ENSO and SAM explain the observed behavior, rather than changes in the teleconnections.

**b. Climate model data**

We do not attempt to directly compare the GCM output described in section 2 to the instrumental data. As shown below, the GCMs are inconsistent in their ability to simulate the spatiotemporal variations in the relationships between the Australasian climate and remote climate drivers realistically. Instead, our assessment examines whether we should expect nonstationarity in an environment solely regulated by the internal dynamics of the atmosphere–ocean system. In other words, is it possible that internal dynamical interactions alone can cause significant changes to the relationships between the Australasian climate and the remote processes important to its regulation?

The ability of the simulated Niño-3.4 and SAM indices to represent ENSO- and SAM-like dynamics in the models was first tested. Figure 7 shows that the indices produce ENSO-like and SAM-like patterns in sea surface temperatures (SSTs) and sea level pressure (SLP) (e.g., Figs. 2 and 3). However, there were some differences compared to the observed patterns that influenced their subsequent relationships with the Australasian climate in the simulations. For example, in all models, the canonical coupled atmosphere–ocean ENSO pattern over the tropical Pacific Ocean extended too far to the west. Also, the correlations between modeled SAM and SLP over Australasia were stronger than observed and extended too far north compared to Figs. 2 and 3, particularly in the GFDL CM2.0 simulation. These differences in the dynamical representations of the Niño-3.4 and SAM indices mean that the teleconnections have different strengths and/or spatial patterns compared to reality.
The abilities of the GCMs used in this study to simulate the observed 31-yr running correlations between the Niño-3.4 and SAM indices and Australasian temperature and precipitation were assessed using Student’s t test. The test was applied to establish whether there were statistically significant differences between the modeled and observed 31-yr running correlations given their respective spreads and sample sizes. If the t statistic was significant at less than the 1% level, then the simulated local–remote climate relationship was defined as inconsistent with the observations.

The means of the simulated 31-yr running correlations were first compared across all grid points to determine whether the models were able to simulate the observed magnitudes of the local–remote climatic relationships independent of their spatial patterns. The t tests showed skill in 15 of 24, or 63%, of cases (2 variables × 2 climate indices × 2 seasons × 3 models). The models were weaker at simulating the correlations based on temperature compared to precipitation, where 50% and 83% of cases were skillful, respectively. The GFDL CM2.1 model showed the most inconsistencies between the modeled and observed 31-yr running correlations.

Figures 8 and 9 show the ability of each model to simulate the spatial patterns, magnitude, and spread of the observed 31-yr running correlations. Unsurprisingly, all models showed less skill for this more stringent test. The simulations were comparable to the observations for between 8% and 56% of the domain, but no model was systematically more proficient than the others. The spatial patterns computed using precipitation were more skillful than those based on temperature. The best simulated relationship was between the Australasian climate and the SAM index during JJA, for which 38%–56% of the grid boxes were skillful (Fig. 9).

The results show that the GCM control simulations display teleconnection patterns between the Australasian climate and the Niño-3.4 and SAM indices but that these differ from the observations. Therefore, the models cannot be directly compared to the observations. However, there is utility in examining the model-based teleconnection patterns to determine whether nonstationarities should

FIG. 7. The shading shows the correlations between the Niño-3.4 index generated from each model and the simulated SST anomalies. The red contours show the correlations between the simulated sea level pressure anomalies and Niño-3.4 (red) and SAM (black) indices; dashed contours indicate inverse correlations.
be expected from an environment with internally forced climate variations alone. Between 76% and 98% of the grid boxes across the Australasian domain were defined as nonstationary when 31-yr running correlations were computed from the internally driven, multicentennial GCM simulations described in section 2b. However, as we used a 95% confidence interval, approximately 5% of the nonstationary periods detected in any long time series might be false positives. So, for a more conservative estimate we show only grid boxes with a proportion of nonstationary periods in a time series of more than 5% in Figs. 10 and 11. Using this definition, between 22% and 54% of model grid boxes were defined as nonstationary, some for over a quarter of their total length. This is unlikely to be solely an artifact of the method. There was little consistency as to where each model produced nonstationarities.

Nonstationary periods denote the times at which nonstationarities occurred. The spatial patterns of the prevalence of nonstationarities reflected the spatial patterns in the number of independent nonstationary periods, defined as a consecutive block of 1 to n nonstationary periods. In other words, the locations that were more frequently nonstationary were because there were more independent events, not because each independent event lasted longer. The average number of independent nonstationary periods was between three and six in the GFDL CM2.0 and GFDL CM2.1 simulations and between two and four in the UKMO-HadCM3 simulation. The former models have 500-yr simulations and the latter a 341-yr simulation. Assuming the events are relatively uniformly spaced in time, on average, a nonstationary period occurred approximately every 80 to 160 years at most locations, consistent with the instrumental data.

The links between the 31-yr variations in the states of the Niño-3.4 and SAM indices and nonstationary correlations with the Australasian climate were compared for each model using the same technique as for the instrumental data (see section 5a). Figures 10 and 11 show that an average of 25%, and up to 40%, of the
nonstationary grid boxes had nonstationary periods that coincided with abnormally high or low interannual variations in the Niño-3.4 and SAM indices. Tables 2 and 3 show that these nonstationarities were similarly likely to be associated with both smaller and larger than normal interannual variability in the index representing the climatic driver. The mean of the index was important to fewer grid boxes, an average of 10% and up to 23%. The greater likelihood that a nonstationarity would coincide with a change in the magnitude of the interannual variations in the climate driver, rather than the mean, was consistent between the results produced from the GCMs and the observations.

The prevalence of nonstationarities in the GCM simulations is evidence that a change in local–remote climate relationships that is not driven by stochastic variation in the local precipitation or mean temperature can occur from internally driven dynamics alone. While these simulated nonstationarities were sometimes linked to fluctuations in the state of the climate driver, many grid boxes also showed no association, suggesting that external processes are also responsible. Note that what we describe as external mechanisms are not necessarily independent of the climate driver (i.e., ENSO or SAM). It is possible that they are the result of indirect and/or nonlinear responses to the climate driver but are not easily distinguished from noise.

c. Paleoclimate proxy data

Examination of the observational record and GCM simulations (sections 5a and 5b) suggested that nonstationary relationships between the Australian climate and ENSO and SAM that are distinct from stationary interannual stochastic variation occur on near-centennial time scales. However, the short length of the instrumental series (and the shortfalls in the abilities of the GCMs to simulate the observed teleconnection patterns) prompts examination in a longer-term context by another means. Paleoclimate analogs might be useful tools for such a comparison. However, their application to the methodology described in section 4 must be tested for robustness, given the extraneous nonclimatic noise present in these reconstructions.

The TEXMEX Niño-3.4 reconstruction uses data outside the Niño-3.4 region to reconstruct tropical Pacific
SSTs, although it retains temporal coherence with that region (Cook et al. 2008). The COA Niño-3.4 reconstruction was developed from coral records within the Niño-3.4 region. However, it was deemed nonrobust prior to 1850 as it used one coral record only to represent the index, and this record is known to contain dating errors (Wilson et al. 2010). This issue was confirmed through a comparison of the TEXMEX and COA records before and after 1850 when the correlations between the two were 0.02 and 0.55, respectively. Wilson et al. reported similar correlations of 0.02 from 1800–70 and 0.59 from 1871–1980. Therefore, we examined the COA record from 1850 only.

Each Niño-3.4 proxy was correlated against the Queensland precipitation reconstruction (hereafter defined as Pairs 1C and 1T, where the C and T subscripts indicate that the COA or TEXMEX records were used as the Niño-3.4 reconstruction), the kauri tree-ring chronology from the North Island of New Zealand (hereafter defined as Pairs 2C and 2T), and the Tasmanian temperature tree-ring reconstruction (hereafter defined as Pairs 3C and 3T). These correlations act as analogs for the relationships between the Niño-3.4 region and the climates of tropical northeast Australia, northern New Zealand, and southeast Australia. Although each proxy spans a different seasonal window, the strongest correlations between the two pairs of data were at lag 0. Thus, seasonal differences in the data did not require adjustments on the interannual scale.

Despite a statistically significant relationship between TEXMEX and COA after 1850, the pairs of 31-yr running correlations produced using each of these Niño-3.4 reconstructions (i.e., Pairs 1C and 1T, Pairs 2C and 2T, and
Pairs $3_C$ and $3_T$ did not exhibit any statistically significant correlations with each another ($-0.15 \leq r \leq +0.10$). This suggests that either one of the ENSO proxies, or both, does not contain a strong enough ENSO signal to adequately assess variations in the relationship between the Niño-3.4 region and the Australasian climate.

To test this further we assumed that Pairs $1_C$, $2_C$, and $3_C$ were “truth”; that is, they realistically estimated the relationships between the variations in the Niño-3.4 index and the Australasian climate. White noise was then added to the COA data so that a new, noisy COA proxy ($COA_N$) had the same correlation as that found between...
the COA and TEXMEX data. This effectively assumes that the TEXMEX data represents the true Niño-3.4 index with an additional noise component associated with changes in the teleconnection pattern between ENSO and the North American climate, local climate variability, or biological noise. Five thousand pairs of 31-yr running correlations between the Australasian proxies and the COAN (new Pairs 1CN, 2CN, and 3CN) were then correlated with running correlations of Pairs 1C, 2C, and 3C. We use the term \( r_N \) to define the correlation between each pair. The distribution of \( r_N \) values indicates whether we should expect Pairs 1C, 2C, and 3C to strongly covary with Pairs 1T, 2T, and 3T, given the magnitude of the extraneous noise in the signal that we estimated from the correlation between the COA and TEXMEX data. Given that we prescribed the TEXMEX proxy as the noisy dataset, we can robustly assess nonstationarities using Pairs 1T, 2T, and 3T if there is a small spread and no weak or out of phase relationships in the \( r_N \) distribution. Note that the results from this synthetic testing are interchangeable; that is, they would be the same if we had defined TEXMEX as truth and COA as the noisy signal. In reality, it is impossible to know which, or if either, proxy is truth and, therefore, we cannot make an assessment about which might be better to employ for testing nonstationarity. If Pairs 1T, 2T, and 3T are outliers in this synthetic noise distribution, then there is further evidence that the true relationship between the COA and TEXMEX data is either weaker or stronger than the 30% common variance \( (r = 0.55) \) computed using data from 1850.

It is unknown which of the COA or TEXMEX Niño-3.4 reconstructions should be used to estimate preinstrumental variations in the Niño-3.4 index for our purposes or, indeed, if either should. Noise in the Australasian proxies, not examined here, will further decrease the robustness of these estimates. Therefore,

<table>
<thead>
<tr>
<th></th>
<th>DJF</th>
<th>JJA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>GFDL CM2.0</td>
<td>GFDL CM2.1</td>
</tr>
<tr>
<td>Niño-3.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low SD</td>
<td>11%</td>
<td>6%</td>
</tr>
<tr>
<td>High SD</td>
<td>22%</td>
<td>1%</td>
</tr>
<tr>
<td>Mean</td>
<td>1%</td>
<td>3%</td>
</tr>
<tr>
<td>SAM</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low SD</td>
<td>6%</td>
<td>13%</td>
</tr>
<tr>
<td>High SD</td>
<td>4%</td>
<td>15%</td>
</tr>
<tr>
<td>Mean</td>
<td>10%</td>
<td>6%</td>
</tr>
</tbody>
</table>

Table 3. Percentage of model grid boxes, by season and model, showing how nonstationary correlations between mean temperature and Niño-3.4 and between mean temperature and SAM relate to low or high index interannual variation (SD) or to the mean level of the driving index.

Figure 12 shows a large spread in the \( r_N \) distributions computed for each pair of proxies. The minimum bound of the 95% confidence interval for each distribution ranged between \(-0.32\) and \(-0.08\) and the maximum bound was 0.85 for all three pairs. Inclusion of weak correlations in this interval indicates that the noise distorting the relationship between the TEXMEX and COA data is too large to robustly compare any nonstationarities recorded between Pairs 1C, 2C, and 3C and Pairs 1T, 2T, and 3T. The results also suggest that, if there is an equivalent or larger noise component in the paleoclimate data when compared to more precise measurements, then paleoclimate data cannot produce robust assessments of variations in interannual local-remote climatic relationships. Finally, Pairs 1T, 2T, and 3T were at or below the 9th percentile of the 5000-member \( r_N \) distributions (Fig. 12), indicating that the common variations between the COA and TEXMEX proxies may, in reality, be less than the 30% as measured from 1850 to 1998.

It is unknown which of the COA or TEXMEX Niño-3.4 reconstructions should be used to estimate preinstrumental variations in the Niño-3.4 index for our purposes or, indeed, if either should. Noise in the Australasian proxies, not examined here, will further decrease the robustness of these estimates. Therefore,
we cannot say whether Pairs 1T, 2T, and 3T or Pairs 1C, 2C, and 3C will provide a more realistic representation of the true relationships between the Niño-3.4 index and the Australasian climate. However, given the amount of noise in paleoclimate proxy data, it is also possible that neither can. With these caveats in mind, we investigated nonstationarity in both pairs but did not draw any conclusions about which might be more representative of reality.

Figure 13 shows the variations in the 31-yr running correlations between the Pairs 1C, 2C, and 3C and Pairs 1T, 2T, and 3T. In all cases, the changes over time were large, showing a range of variations similar to the magnitudes computed from the instrumental data (e.g., Fig. 6). Nonstationary relationships were detected for Pairs 1T and 3T but not for Pairs 1C, 2C, 2T, and 3C. These discrepancies further highlight the differences between the relationships computed using the COA and TExMex data. Wilson et al. (2010) also identified periods with very different correlations between the COA and TExMex indices on the centennial scale. Given that one of the COA utilizes data very near or within the Niño-3.4 region and the TExMex assumes ENSO variations from a North American teleconnection, their work supports our results of nonstationarity.

Ultimately, it is unknown which, if either, of the sets of the 31-yr running correlations based on the COA and TExMex Niño-3.4 reconstructions is representing reality. However, importantly, the TExMex reconstruction is fundamentally based on teleconnections. So, extraneous noise causing the differences between the two sets may be a result of a nonstationary teleconnection between the Niño-3.4 region and the U. S. Southwest and Mexico, from which the TExMex reconstruction was developed. Such discrepancies further highlight that nonstationary relationships exist as part of an internally varying system.

6. The sensitivity of nonstationary relationships to sample length

The results in section 5 were tested for nonstationarity in local–remote climate relationships using 31-yr running z scores in order to retain as much of the instrumental data as possible. The sensitivity of these results to the choice of window length used to compute the running correlations is now described for the model data only, which had a long enough time series that robust results could still be generated using longer windows.

Nonstationary local–remote climate relationships were evident regardless of the window used to determine the running correlations. Figure 14 shows an example of the differences in the nonstationarities computed using 31-yr, 51-yr, and 71-yr running correlations for the UKMO-HadCM3 simulation of relationships between the SAM and Australasian precipitation and mean temperatures during JJA. Although differences exist, the general patterns of nonstationarity were similar regardless of the length of the window applied. For example, there was consistently a region of more frequent nonstationarities in northern Australia for mean temperature, and a line from southwest Western Australia over the Great Australian Bight and Tasmania for precipitation. Pattern correlations computed for those data, in Fig. 14, ranged between 0.66 and 0.88, with the largest discrepancies between the 31-yr and 71-yr smoothed data.

7. Nonstationary local–remote climate relationships and implications for paleoclimate reconstructions

Paleoclimate reconstructions for the Australasian region have utilized teleconnection patterns to infer remote (local) climate variations from a local (remote) source of proxy data (D’Arrigo et al. 2005; Shen et al. 2006; Fowler et al. 2008; van Ommen and Morgan 2010; Gallant and Gergis 2011; Fowler et al. 2012; Gergis et al. 2012). The evidence of nonstationary teleconnections presented here has implications for the robustness of
these reconstructions and suggests that current methodological practices require reassessment.

Fowler et al. (2012) highlight that using teleconnected relationships to discern past variations in remote climate data is imperfect. In a comparison between the kauri record and other ENSO paleoclimate reconstructions, they propose significant changes to the ENSO teleconnection with northern New Zealand in the 1300 and 1400s. Moreover, Wilson et al. (2010) compared three ENSO reconstructions and identified centennial-scale periods with vastly different correlations between the ENSO indices derived from the different regions. In some cases the correlations were near zero.

Most paleoclimate studies use subcentennial-scale data to estimate the strength of the teleconnection on which their reconstruction is based. Moreover, a single multidecadal period is often used for either calibration or verification of the relationship between the reconstruction and a remote climate variable. Given the large variations in the relationships between local and remote data, using a single calibration period may bias results (Gallant and Gergis 2011; Gergis et al. 2012). This is a particular problem for Australasia, where limited instrumental data in some areas means that the paleoclimate proxies are calibrated to short station time series. For these cases, climate models are an alternative and useful tool for establishing a more robust estimate of the teleconnection (Phipps et al. 2013). However, establishing the ability of climate models to simulate not only the strength of the relationships but the dynamical mechanisms causing the local–remote climate relationship is crucial prior to their use.

We now provide a simple example of how the variations in teleconnected relationships can affect the perceived skill
of paleoclimate reconstructions. Note that this example greatly simplifies the process of generating reconstructions from proxy data and does not perform any of the rigorous testing involved. However, similar sensitivities have been highlighted in previously published reconstructions (Gallant and Gergis 2011; Gergis et al. 2012).

As previously described, Fowler et al. (2008) used the Kauri tree-ring record as an ENSO proxy as it agreed well with variations in the Southern Oscillation index. Using this known agreement, we regressed the Kauri record against the November–April Niño-3.4 index to develop a very simplified ENSO reconstruction. A series of traditional verification statistics commonly used to assess the skill of paleoclimate reconstructions were computed. These statistics include the Pearson correlation and consequent variance explained, reduction of error statistic (RE), coefficient of efficiency statistic (CE), and root-mean-square error (RMSE) (Cook and Kairiukstis 1990). The RE and CE are often used to define whether a reconstruction is skillful, with values larger than zero denoting skill. We used all possible 30-yr, 50-yr, and 70-yr sequences of the 99 years of overlapping Kauri and Niño-3.4 data to develop the transfer function. The verification statistics previously described were then calculated using the remaining years not used for calibration. Table 4 shows the sensitivity of these skill scores to the choice of calibration and verification period and, hence, to the state of the relationship between the New Zealand climate and the Niño-3.4 index. The magnitudes of the statistics varied by up to 50% in some cases, and some choices of the calibration period were defined as unskilled (i.e., a CE or RE of 0.0).

This simple example illustrates that caution should be exercised when producing reconstructions associated with teleconnected relationships between local and remote climates, particularly when calibrating to short observational records. However, it does not necessarily preclude the practice. Instead, the uncertainties must be included and acknowledged in reconstructions. Trying to capture the full range of variation in the strength of teleconnections, including any nonstationary states, will achieve this.

8. Discussion

The simple statistical definitions and indices used here do not capture the dynamical complexities of teleconnection patterns. Furthermore, there were shortcomings in instrumental, model, and proxy data, as previously discussed. However, despite these issues, all three datasets consistently identified variable, rather than static, relationships between the Australasian climate and ENSO and SAM indices on near-centennial time scales. However, the results stemming from the paleoclimate data were not robust.

The evidence presented here suggests that many Australasian local–remote climate relationships vary. However, recall that Eq. (1), which was used to test our hypothesis, included a term describing stochastic variation in local Australasian mean temperature and precipitation only. The consistency of nonstationarities on near-centennial time scales from all data sources suggests that an important component has been omitted. For example, low-frequency oscillations or stochastic variation in the local–remote climatic relationships might be associated with stochastic or dynamical fluctuations in the remote climate driver.

Variations in the relationships between ENSO and the Australasian climate occur with large-scale changes to the Pacific Ocean, commonly referred to as the interdecadal Pacific oscillation (IPO) (Power et al. 1999) or the Pacific decadal oscillation (PDO) (Mantua et al. 1997). However, the IPO/PDO oscillates at a higher frequency (15–30 yr) compared to the near-centennial-scale nonstationary relationships identified in this study (Zhang et al. 1997). The statistical stationarity on time scales less than 30 yr reported by previous studies

Table 4. A suite of verification statistics computed from an ensemble of reconstructions of the Niño-3.4 index from the Kauri tree ring record (Fowler et al. 2008). The range and maximum variation from the mean value of each verification statistic, computed from a distribution of possibilities, is shown. The distributions of the metrics were computed using the independent data that remained after transfer functions were calculated using 30-, 50-, and 70-yr periods for calibration. The metrics computed were the Pearson correlation (Pearson $r$), the reduction error statistic (RE), the coefficient of efficiency statistic (CE), and the root-mean-square error (RMSE) (Cook and Kairiukstis 1990).

<table>
<thead>
<tr>
<th>Metric</th>
<th>30-yr calibration</th>
<th>50-yr calibration</th>
<th>70-yr calibration</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Range (min–max)</td>
<td>Max variation</td>
<td>Range (min–max)</td>
</tr>
<tr>
<td>Pearson $r$</td>
<td>0.34–0.52</td>
<td>16%</td>
<td>0.41–0.61</td>
</tr>
<tr>
<td>Pearson $r^2$</td>
<td>14–27</td>
<td>34%</td>
<td>17–37</td>
</tr>
<tr>
<td>RE statistic</td>
<td>0.04–0.20</td>
<td>33%</td>
<td>0.10–0.22</td>
</tr>
<tr>
<td>CE statistic</td>
<td>0.00–0.19</td>
<td>50%</td>
<td>0.09–0.21</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.73–0.90</td>
<td>10%</td>
<td>0.75–0.93</td>
</tr>
</tbody>
</table>
and SAM-based nonstationary relationships with local common variations associated with both ENSO-based spatially distinct climate patterns (Jiang et al. 2012). The where multiple climate drivers interact to generate interactions are well grounded in physics. Such process-guidance from dynamical climate models so that the in-
to remote paleoclimate proxy variability, ideally with unique circulation patterns can be established and linked the role of the remote climate driver as a regulator of patterns (Lorrey et al. 2007, 2008, 2012). Subsequently, the paradigm, investigating nonstationary teleconnections of past climatic fluctuations. For the paleoclimate research connections can provide a better physical understanding simulations in section 5b.

Centennial-scale fluctuations in the Pacific atmosphere–ocean system, with a signature that is strongest in the west Pacific warm pool region, have been identified in control simulations that span at least 1000 years from three coupled GCMs (Karnauskas et al. 2012). These internal variations induced large tropical and extratropical circulation anomalies and their temporal scale is consistent with that of the nonstationary periods identified here. Whether such an oscillation could induce nonstationary teleconnection patterns between the Australasian climate and ENSO or SAM on the 31-yr to 71-yr time scale is unknown. However, if this was the case, the implication is that variable teleconnection patterns are an inherent part of natural climate variability. This notion is supported by our results derived from the GCM control simulations in section 5b.

Characterization of the variability in observed teleconnections can provide a better physical understanding of past climatic fluctuations. For the paleoclimate research paradigm, investigating nonstationary teleconnections should first focus on the likely circulation patterns associated with the localized variability in paleoclimate proxies. For example, circulation-based approaches using multiproxy networks and regional climate regime classification can be used to investigate teleconnection patterns (Lorrey et al. 2007, 2008, 2012). Subsequently, the role of the remote climate driver as a regulator of unique circulation patterns can be established and linked to remote paleoclimate proxy variability, ideally with guidance from dynamical climate models so that the interactions are well grounded in physics. Such process-based studies are particularly advantageous for situations where multiple climate drivers interact to generate spatially distinct climate patterns (Jiang et al. 2012). The common variations associated with both ENSO-based and SAM-based nonstationary relationships with local climates in section 5a indicate this might be particularly important for the Australasian region.

9. Conclusions

We have provided evidence that some Australasian teleconnection patterns change over time in a manner inconsistent with stochastic variation in the local climate on interannual time scales; that is, they can be nonstationary on near-centennial time scales. Statistical nonstationary relationships between interannual variations in seasonal Australasian temperature or precipitation and indices representing ENSO and the SAM were robustly identified. In the instrumental record, the running correlations associated with both ENSO and SAM strongly covaried, suggesting common modulation. Sometimes, the statistically significant variations in local–remote climate correlations coincided with distinct states of ENSO or SAM in the instrumental and model data, but sometimes not. The nonstationarities identified from model simulations suggest that nonstationary local–remote climate relationships can be forced solely by internal climate variability. The paleoclimate proxies employed here contained sufficient extraneous noise to prohibit extraction of a consistent signal, suggesting caution should be applied when assessing teleconnections in a paleoclimatic context.

Nonstationary teleconnection patterns have implications for paleoclimate research that exploits these links to generate a reconstruction. This is particularly problematic for calibrating and assessing paleoclimate proxy records if instrumental records, which are mostly less than 100 years long, contain a nonstationary relationship with remote climates on the near-centennial time scale. Every reconstruction utilizing relationships between local and remote climates should test stationarity in rigorous detail in order to capture the associated uncertainty. This should include interpreting nonstationarities in paleoclimate proxies using concepts and methods grounded in physics.

Acknowledgments. Funding for this research comes in part from the Australian Research Council (Grants FF0668679 and DP1092945) and the ARC Centre of Excellence for Climate System Science (Grant CE110001028). ABM and AML are supported by the NIWA core-funded project “Climate Present and Past.” Thank you to NIWA for providing the New Zealand temperature and precipitation records and to Georgina Griffiths, Joelle Gergis, Tas van Ommen, Stewart Allen, and three anonymous reviewers for their useful discussion and constructive comments. We thank NOAA for supporting the production of this special issue of Journal of Climate.
REFERENCES


