

Multi-proxy summer and winter precipitation reconstruction for southern Africa over the last 200 years

Raphael Neukom* · David J. Nash · Georgina H. Endfield · Stefan W. Grab · Craig A. Grove · Clare Kelso · Coleen H. Vogel · Jens Zinke

R. Neukom (*)

Oeschger Centre for Climate Change Research, University of Bern, 3012 Bern, Switzerland e-mail: neukom@giub.unibe.ch

D. J. Nash

School of Environment and Technology, University of Brighton, Lewes Road, Brighton BN2 4GJ, UK

D. J. Nash & S. W. Grab

School of Geography, Archaeology and Environmental Studies, University of the Witwatersrand (Wits), Private Bag 3, Johannesburg 2050, South Africa

G. H. Endfield

School of Geography, University of Nottingham, University Park, Nottingham NG7 2RD, UK

C. A. Grove

Department of Marine Geology, NIOZ Royal Netherlands Institute for Sea Research, P.O. Box 59, 1790 AB Den Burg, Texel, The Netherlands

C. Kelso

Department of Geography, Environmental Management and Energy Studies, University of Johannesburg, PO Box 524, Auckland Park 2006, South Africa

C. H. Vogel

Department of Geography, Geoinformatics and Meteorology, University of Pretoria, Hatfield, Private Bag X20, Pretoria 0028, South Africa

J. Zinke

School of Earth and Environment, University of Western Australia Oceans Institute, 35 Stirling Highway, Crawley, WA 6009, Australia

J. Zinke

Australian Institute of Marine Science, Nedlands, WA 6009, Australia

Abstract This study presents the first consolidation of palaeoclimate proxy records from multiple archives to develop statistical rainfall reconstructions for southern Africa covering the last two centuries. State-of-the-art ensemble reconstructions reveal multi-decadal rainfall variability in the summer and winter rainfall zones. A decrease in precipitation amount over time is identified in the summer rainfall zone. No significant change in precipitation amount occurred in the winter rainfall zone, but rainfall variability has increased over time. Generally synchronous rainfall fluctuations between the two zones are identified on decadal scales, with common wet (dry) periods reconstructed around 1890 (1930). A strong relationship between seasonal rainfall and sea surface temperatures (SSTs) in the surrounding oceans is confirmed. Coherence among decadal-scale fluctuations of southern African rainfall, regional SST, SSTs in the Pacific Ocean and rainfall in south-eastern Australia suggest SST-rainfall teleconnections across the southern hemisphere. Temporal breakdowns of the SST-rainfall relationship in the southern African regions and the connection between the two rainfall zones are observed, for

example during the 1950s. Our results confirm the complex interplay between large-scale teleconnections, regional SSTs and local effects in modulating multi-decadal southern African rainfall variability over long timescales.

Keywords Precipitation reconstruction · Southern Africa · Palaeoclimate · Climate dynamics · Southern hemisphere

1 Introduction

Interannual variations in rainfall across southern Africa have major consequences for human livelihoods through their impacts upon water supply and vegetation. Rainfall distribution is highly seasonal over most of the subcontinent and is influenced by a variety of atmospheric and oceanic circulation systems (see Tyson 1986; Chase and Meadows 2007 for reviews). Rainfall in southwestern South Africa (approximately 31–34°S, 17–21°E) occurs mainly during the austral winter months, with more than 66 % of mean annual precipitation falling from April to September. Here, rainfall is associated with the passage of westerly cold frontal systems (Reason and Jagadheesha 2005). In contrast, > 66 % of mean annual precipitation over the interior regions (including northern and eastern South Africa, plus Angola, Botswana, Lesotho, Malawi, Mozambique, Namibia, Swaziland, Zambia and Zimbabwe) falls between the summer months of October and March. Here, the most significant contributors to rainfall are tropical-temperate troughs and their associated cloud bands (Mason and Jury 1997). In recognition of the antiphase nature of the annual rainfall cycle, these two regions are traditionally referred to as the southern African winter rainfall zone (WRZ) and summer rainfall zone (SRZ), respectively (Tyson 1986). The narrow belt of South Africa between the SRZ and WRZ receives year-round precipitation (Chase and Meadows 2007).

Rainfall levels in southern Africa are modulated by the interplay of large-scale climate modes, including the El Niño–Southern Oscillation (ENSO), Indian Ocean Dipole (IOD) and Southern Annular Mode (SAM). ENSO warm events may be preceded by periods of anomalously high rainfall and are often followed by drought in the SRZ (Nicholson et al. 2001). Positive phases of the SAM are usually associated with anomalously wet conditions over the SRZ and weak decreases in rainfall in the WRZ (Gillett et al. 2006; Watterson 2009). Rainfall in the northeast of the subcontinent is modulated by interactions of ENSO with the IOD, with positive IOD phases generally associated with enhanced rainfall (Abram et al. 2008). In addition to large scale modes of climate variability, SRZ and WRZ

rainfall are also linked to variability of sea surface temperatures (SSTs) in the oceans surrounding southern Africa (Goddard and Graham 1999; Behera and Yamagata 2001; Reason 2001; Reason and Jagadheesha 2005).

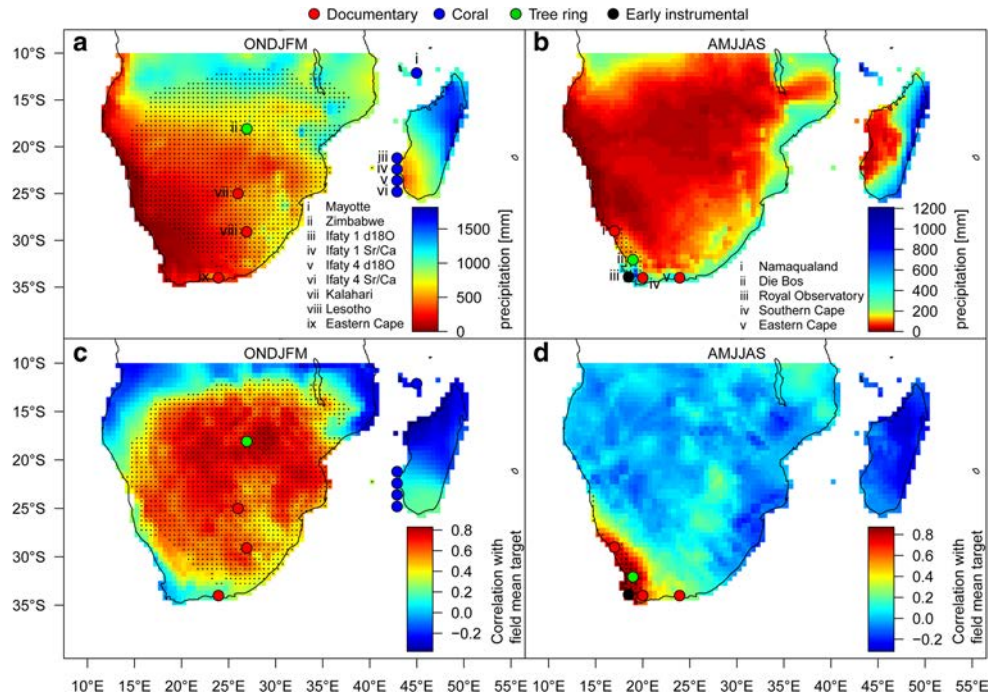
Understanding long-term rainfall variability is dependent upon the availability of extended precipitation time-series. Instrumental data, however, are not available until the end-nineteenth century for most areas beyond coastal South Africa (Nicholson et al. 2012). Recent efforts to develop proxy records, including reconstructions of annual rainfall variability from historical documents, tree-rings and speleothems, and SSTs from isotopic analyses of corals, have extended the instrumental dataset (Neukom and Gergis 2012). To date, no attempt has been made to consolidate these records for southern Africa. In this paper, we use state-of-the-art statistical analyses to combine all available published rainfall proxies and reconstruct summer and winter rainfall variations over the last 200 years. We describe the reliability of the reconstructions and explore the influence of the important modes of Southern Hemisphere (SH) climate variability such as ENSO, SAM and IOD on long-term rainfall variability. Further details, relating mainly to the methodologies used in our reconstructions, and additional illustrations are provided in the supplementary material (SM; sections S1-S6).

2 Data and methods

2.1 Instrumental target

As our instrumental rainfall dataset we used rainfall totals for the October–March (ONDJFM) and April–September (AMJJAS) seasons in the CRU TS 3.0 grid (updated from Mitchell and Jones 2005). The spatial distributions of seasonal rainfall totals are shown in Fig. 1a (ONDJFM) and b (AMJJAS). The SRZ and WRZ areas were defined as follows. First, we separated the area south of 10°S into the regions where $\geq 66\%$ of annual precipitation falls in the respective (ONDJFM and AMJJAS) seasons (see Chase and Meadows 2007). The resulting area of $\geq 66\%$ winter (AMJJAS) rainfall is relatively small and spatially coherent (Fig. 1d, see also Fig. S3.2 in the SM). Therefore, we defined this area as the WRZ and used its spatial average as our winter rainfall reconstruction target. In contrast, the area with $\geq 66\%$ summer (ONFJDM) rainfall is very diverse in terms of rainfall amounts and spatial coherence. We performed a Principal Component (PC) analysis to identify the dominant modes of rainfall variability over this area. The loadings and spatial correlations of the resulting PCs are presented in Figs. S1.1-S1.2 and Table S1.1. The first PC explains 26 % of the total rainfall variability and has the strongest loadings and spatial rainfall correlations over the area traditionally described as the SRZ (see Sect. 1).

Fig. 1 *Top row* seasonal southern African rainfall totals for the austral summer (ONDJFM; panel **a**) and winter (AMJJAS; panel **b**) seasons in the CRU TS3.0 grid averaged over the 1901–2006 period. *Bottom row* correlation of rainfall with the SRZ (panel **c**) and WRZ (panel **d**) spatial average over the period 1901–2006. The grid cells that were averaged to create the SRZ (WRZ) spatial averages are highlighted with *black dots* in the *left (right) panels*. Locations of the proxy records used for the rainfall reconstructions are shown with *circles*. The latitudinal positions of the Ifaty coral records are adjusted to improve visibility and avoid overlapping circles. Note the non-linear colour scheme in panel **b**



We defined the SRZ as the area including all grid-cells with significant ($p < 0.05$) and positive correlations with this first PC. For all correlations calculated herein, significance was tested using a t-distribution, with the degrees of freedom adjusted for auto-correlation at lag 1 using the method of Bretherton et al. (1999). As our summer rainfall reconstruction target we used the spatial average of the SRZ. Figure 1c shows the strong and coherent relationship of the individual SRZ grid cells with the field mean reconstruction target (see also Fig. S3.2 in the SM).

2.2 Proxy data processing

As predictors for our rainfall reconstruction, we used climate proxy records of at least annual resolution from southern Africa and the surrounding oceanic areas. Only proxies within the domain 10°E – 50°E / 10°S – 35°S were used, resulting in a total of 13 records from 11 sites (Fig. 1; Table 1; see Neukom and Gergis 2012 for a detailed review of proxy data availability). Since the documentary records all end in 1900 CE, it was necessary to extend these records to the present-day to allow calibration within the multi-proxy dataset. The extension was performed using a “pseudo-documentary” approach (Neukom et al. 2009), which degrades instrumental data to match the statistical properties of an index time series based on historical documents. This approach has been applied successfully in South America (Neukom et al. 2009), where the availability of documentary and instrumental data in terms of quality, length and type of sources is very similar to southern Africa. Further details of the process of extending the documentary records are provided in section

S2 of the SM. In addition to palaeoclimate proxy data, we also included a very long instrumental record from the Royal Observatory in Cape Town as a rainfall predictor. Records from 1850 to 2008 were provided by the South African Weather Service, with information dating back to 1842 digitised directly from meteorological ‘day books’ held at the Royal Observatory.

The proxies were allocated to the SRZ and WRZ according to their locations, resulting in 9 (5) rainfall proxies used for the SRZ (WRZ) reconstructions (Fig. 1; Table 1). SRZ proxies include coral records from Mayotte (Zinke et al. 2009) and Ifaty (Zinke et al. 2004), a tree-ring record from Zimbabwe (Therrell et al. 2006), and documentary records from the Kalahari (Nash and Endfield 2002, 2008), Lesotho (Nash and Grab 2010) and Eastern Cape (Vogel 1989). The WRZ rainfall proxies include a tree-ring record from Die Bos (Dunwiddie and Lamarche 1980), documentary records from Namaqualand (Kelso and Vogel 2007) and the Southern and Eastern Cape (Vogel 1989), and the Royal Observatory instrumental record. Correlations of the records with the instrumental SRZ and WRZ reconstruction targets are provided in Table S3.1. Missing values in the proxy matrices in the 1911–1995 period (0.52 % for the SRZ; 4.7 % for the WRZ) were infilled using composite plus scaling (Mann et al. 2008; Neukom et al. 2011).

2.3 Reconstruction method description

We used an ensemble-based nested principal component regression (PCR; Cook et al. 1994; Luterbacher et al. 2002,

Table 1 Proxy records used in the southern African rainfall reconstruction

Name	Archive/Proxy	Resolution	Lon E	Lat S	Start	End	Reference	Zone
Mayotte	Coral d18O	Bimonthly	44.98	12.11	1865	1993	Zinke et al. (2009)	SRZ
Ifaty 1	Coral d18O, Sr/Ca	Annual	42.90	23.03	1882	1994	Unpublished	SRZ
Ifaty 4	Coral d18O, Sr/Ca	Bim./ann.	42.90	23.03	1660	1995	Zinke et al. (2004)	SRZ
Zimbabwe rainfall reconstruction	Tree-ring width	Annual	26.94	18.08	1796	1996	Therrell et al. (2006)	SRZ
Die Bos	Tree-ring width	Annual	18.96	32.07	1762	1976	Dunwiddie and LaMarche (1980)	WRZ
Kalahari rainfall indices	Historical documents	Annual	26.00	25.00	1815	2002	Nash and Endfield (2002, 2008)	SRZ
Namaqualand rainfall indices	Historical documents	Annual	17.00	29.00	1817	1997	Kelso and Vogel (2007)	WRZ
Lesotho rainfall indices	Historical documents	Annual	26.92	29.08	1824	1995	Nash and Grab (2010)	SRZ
Eastern Cape rainfall indices ^a	Historical documents	Annual	23.92	34.00	1821	2007	Vogel (1989)	SRZ/WRZ
Southern Cape rainfall indices	Historical documents	Annual	20.00	34.00	1821	1996	Vogel (1989)	WRZ
Royal observatory	Early instruments	Monthly	18.50	33.90	1842	2008	This study	WRZ

^a The area covered by the Eastern Cape documentaries includes the transition zone between the SRZ and WRZ (Vogel 1989) and the record is therefore used for both regions

Table 2 Correlation of instrumental SRZ and WRZ rainfall averages with SH climate modes using the full period of overlap for each index

	SOI	NINO3.4	PDO	IPO	SAM	IOD
SRZ	0.51 ($p < 0.001$)	-0.52 ($p < 0.001$)	-0.24 ($p = 0.02$)	-0.41 ($p < 0.001$)	0.09 ($p = 0.52$)	-0.25 ($p = 0.02$)
WRZ	0.01 ($p = 0.95$)	0.10 ($p = 0.33$)	0.02 ($p = 0.81$)	-0.02 ($p = 0.81$)	-0.1 ($p = 0.50$)	0.11 ($p = 0.28$)

Bold values indicate significant correlations ($p < 0.05$)

2004; Neukom et al. 2010, 2011; Wahl and Smerdon 2012; Trouet et al. 2013) to reconstruct southern African rainfall. A description of the PCR reconstruction technique is provided by Luterbacher et al. (2002), and also outlined in section S4. We used the period 1921–1995 for calibration and verification. In addition, the 1911–1920 period, which has reduced rainfall station data availability, was used for independent early verification. These periods were chosen because before 1911 and after 1995 the coverage of instrumental data within the Global Historical Climatology Network database (GHCN; Peterson and Vose 1997) is limited (Fig. S3.1).

2.4 Independent datasets for comparison and verification

Instrumental and reconstructed data of important modes of SH climate variability were used to assess large-scale teleconnections of southern African rainfall and their temporal stability. A number of independent palaeoclimate datasets were also used to perform an additional

verification of our results. In this section we describe these datasets and the methods used to process them.

2.4.1 Instrumental climate indices

To evaluate the influence of large-scale climate modes on southern African rainfall, we used the following climate indices: The Southern Oscillation index (SOI; 1867–2006; CRU dataset; Allan et al. 1991), NINO 3.4 tropical Pacific SST index (1871–2008; HadISST data set; Rayner et al. 2003), Pacific Decadal Oscillation index (PDO, 1901–2010; Mantua et al. 1997), Inter-decadal Pacific Oscillation (IPO; 1872–2005; Power et al. 1999; Folland et al. 2002), Southern Annular Mode (SAM; 1949–2009; NCEP data; Thompson and Wallace 2000) and Indian Ocean Dipole (IOD; 1872–1997; Dipole Mode Index calculated from the HadISST data set (Rayner et al. 2003) using the Saji et al. (1999) definition). Correlations between the climate indices and our instrumental rainfall targets using the full period of overlap for each index are shown in Table 2. Note that the high p value in the SAM correlation is influenced by the

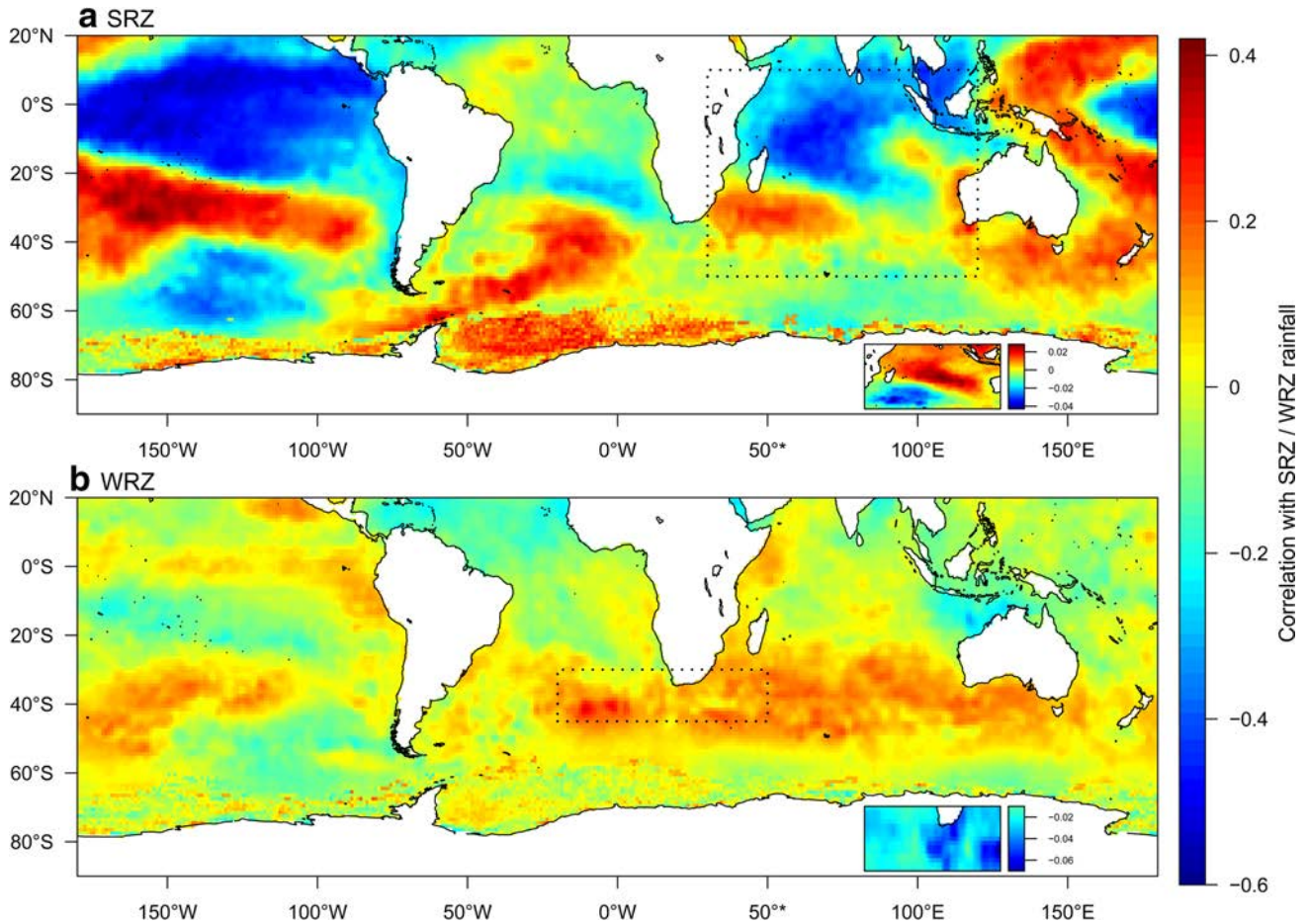


Fig. 2 **a** Spatial correlations of instrumental SRZ rainfall with SSTs (HadISST; Rayner et al. 2003) over the period 1901–2006 in the ONDJFM season. *Dashed line* demarcates the area that was used to calculate IOPC2; *inset map* displays the loadings of IOPC2. **b** As lower number of years available for analysis (1949–2006) compared to the other indices.

2.4.2 Reconstructions of climate indices

To allow for a long-term comparison of our reconstruction with modes of SH climate variability, we used a suite of existing reconstructions of these modes. For ENSO, we used a tree-ring based reconstruction of the SOI (Stahle et al. 1998) covering 1706–1977, the multi-proxy Unified ENSO Proxy (UEP; McGregor et al. 2010) covering 1650–1977 and the Li et al. (2011) tree-ring ENSO reconstruction covering 900–2002. For the SAM, we used a tree-ring-based reconstruction (Villalba et al. 2012) covering 1592–2010, and a shorter reconstruction based on instrumental station data (Jones et al. 2009) covering 1851–2006. Additionally, we used reconstructions of IPO (corals; 1650–2004; Linsley et al. 2008) and two PDO reconstructions based on tree-rings (1565–2004; D’Arrigo and Wilson 2006) and historical documents (1470–1998; Shen et al. 2006) for

above but for instrumental WRZ rainfall in the AMJJAS season. *Dashed line* demarcates the area that was used to calculate SOPC1, with the *inset map* displaying the loadings of SOPC1

comparison with our reconstructions. Among the variety of existing PDO reconstructions, recent work (Grove et al. 2013) has identified the D’Arrigo and Wilson (2006) and Shen et al. (2006) PDO reconstructions, which are based on Asian proxies, as most coherent with eastern African climate. Finally, we use the coral-based IOD reconstruction of Abram et al. (2008). Note that the tree-ring based reconstructions represent the austral summer (SOI and SAM) and boreal spring (PDO) and summer (Li et al. 2011 ENSO reconstruction) seasons. The documentary-based PDO reconstruction also represents the boreal summer season, whereas the UEP, IPO and IOD reconstructions represent annual mean values. The Jones et al. (2009) SAM reconstruction is available at seasonal resolution.

2.4.3 Regional SSTs

Given that southern African rainfall is not only linked to large-scale SH circulation but also related to the SSTs of

the surrounding oceans (Nicholson and Kim 1997; Nicholson 1997; Goddard and Graham 1999; Behera and Yamagata 2001; Reason 2001; Reason and Jagadheesha 2005), we also compared our reconstructions to modes of regional variability. We extracted modes of regional SST variability using the HadISST dataset (Rayner et al. 2003). Figure 2a, b shows the spatial correlation of instrumental SRZ (WRZ) spatial mean rainfall with ONDJFM (AMJJAS) SSTs. Whereas 40 % of all SST grid cells south of 20°N show significant ($p < 0.05$) correlations with SRZ rainfall, only 2 % are significantly correlated to WRZ rainfall. Apart from the clear ENSO signal in the tropical Pacific, SRZ rainfall is also strongly linked to Indian Ocean SSTs, whereas WRZ rainfall has the strongest correlations with SSTs in the southern Atlantic and Indian oceans around 40°S.

Next we performed a PC analysis on the instrumental SST record to extract regional modes displaying strong associations with SRZ and WRZ rainfall, respectively. For the SRZ, we selected ONDJFM SSTs in the area of the Indian Ocean within 30°E–120°E and 10°N–50°S (area demarcated by a dashed line in Fig. 2a). The first four PCs explain 48, 15, 7 and 5 percent of the total variance in the basin over the period 1871–2008, respectively. PC2 shows the highest correlations with instrumental SRZ rainfall ($r = -0.40$, $p < 0.01$). The loadings of this Indian Ocean PC2 (IOPC2) are illustrated in the inset of Fig. 2a, showing a dipole between the latitudes around 20°S and 40°S in the western Indian Ocean.

For the WRZ, we calculated the PCs of AMJJAS mean SSTs within the 20°W–50°E, 30°S–45°S domain of the Southern Ocean (dashed line in Fig. 2b). SSTs in this domain show a significant positive trend over the instrumental period, in contrast to SRZ/WRZ rainfall and the selected Indian Ocean SST data. Therefore the Southern Ocean SSTs were linearly detrended prior to calculating the PCs. The four dominant PCs explain 22, 21, 12 and 8 % of total SST variance within the domain. Southern Ocean PC1 (SOPC1) has clearly the highest correlations with instrumental WRZ rainfall ($r = -0.23$, $p = 0.02$). The loadings of PC1 shown in the inset to Fig. 2b show a consistent signal with all values of the same (negative) sign and the strongest loadings south and southeast of the WRZ.

2.4.4 Independent palaeoclimate records from southern Africa

Comparison of our rainfall reconstructions with independent records is problematic since all of the annually-resolved proxies available for southern Africa were used in our rainfall reconstruction. Nevertheless, a recently published synthesis of interannual rainfall variability across Africa during the nineteenth century (Nicholson et al.

2012) offers the opportunity for semi-independent validation. The Nicholson et al. (2012) study uses published documentary evidence and gauge data to reconstruct rainfall variability for a series of zones with statistically coherent contemporary precipitation patterns. Statistical inference from adjacent zones is used to infer conditions in zones with sparse documentary and/or instrumental data. Four zones from this study were selected for the validation process, two from the SRZ (zone 76 centred on Kimberley, South Africa; and 72 in Swaziland) and two from the WRZ (zone 81 in the western Karoo, South Africa and 82 in the adjacent Atlantic coast areas). Care was taken to select zones that did not overlap with the geographical coverage of documentary and early instrumental proxies used in our reconstruction. Zones 72 and 76 in the SRZ fully meet this criterion. Zone 81 is not an ideal candidate for validation of the WRZ reconstruction as it falls partly within the year-round rainfall zone. However, as the remaining three zones within the WRZ overlap with proxies used in our reconstruction it was the best possible option. Zone 82 is not fully independent from our reconstruction due to some proxy-overlap, but is included to investigate the agreement between two reconstructions with overlapping data but independent methodologies.

Given that southern African rainfall is strongly related to large-scale circulation over the surrounding oceans, we used wind speed measurements recorded in ships' logbooks from the CLIWOC database (García-Herrera et al. 2005) as rainfall predictors to perform additional independent reconstructions. Wind speed measurements were taken from the combined CLIWOC/ICOADS dataset (Küttel et al. 2010), consisting of gridded ($8^\circ \times 8^\circ$ spatial resolution) u- and v-wind vector data and covering the period 1750–2002. Only grid cells within a search radius of 1,000 km of the SRZ and WRZ regions were used. Not all grid cells around southern Africa could be used due to insufficient data coverage in the pre-1850 CLIWOC period (Küttel et al. 2010). Grid cells with significant correlations ($p < 0.05$) with the reconstruction targets were selected as rainfall predictors (details in SM section S6). This resulted in eight grid cells being identified as suitable for SRZ rainfall reconstruction and only one cell for the WRZ. Correlations of each candidate cell are shown in Table S6.1 (SM) and the locations of the selected cells are shown in Fig. S6.1. For the SRZ, the selected cells were then used as predictors in an ensemble PCR rainfall reconstruction using the method described in Sect. 2.3. For the WRZ, the independent rainfall reconstruction was developed by scaling the single selected cell to the mean and standard deviation of the instrumental data in the 1921–1990 period.

Finally, we also compared our SRZ reconstruction with an additional coral record from Madagascar. This "MASB" coral record (Grove et al. 2013) is affected by a

human deforestation signal in the twentieth century and was therefore not included in our reconstruction proxy database. However, the record is potentially useful for comparison with our results in the nineteenth century.

3 Rainfall variability over the last two centuries

3.1 Summer rainfall zone

Our proxy networks allow rainfall reconstructions for the SRZ with positive ensemble mean reduction of error (RE; Cook et al. 1994) skill and early verification RE skill back to 1796. Results suggest that the SRZ was significantly wetter in the nineteenth than the twentieth century (students t -test, $p \ll 0.01$). Average reconstructed rainfall anomalies in the nineteenth (twentieth) century are 60.1 mm (11.5 mm) wrt 1921–1995. A 30-year loess-filtered reconstruction (Fig. 3a) reveals persistent decadal rainfall anomalies relative to the 1921–1995 mean. Relatively drier periods are reconstructed around 1845 and in the early 1860s, whereas the last 30 years of the nineteenth century are the wettest of our reconstruction. The twentieth century is characterised by dry phases around 1930, 1945 and in the most recent period since the early 1970s. These decadal scale fluctuations mask considerable interannual

variability. The driest years of our reconstruction are 1983, 1973, 1992 and 1995 (consistent with instrumental data), whilst the driest years of the nineteenth century are 1851, 1862 and 1842. In the nineteenth (twentieth) century, 4 (10) years exceed -1 standard deviations (s.d.) from the 1921–95 mean. The wettest reconstructed years are 1871, 1830 and 1925, and an additional 34 (21) years of the nineteenth (twentieth) century exceed the $+1$ s.d. threshold. The agreement of the reconstruction with instrumental data is weaker for extreme wet years than for droughts.

Figure 3b shows the 10-year running standard deviation of the SRZ rainfall reconstruction. This reveals periods of reduced rainfall variability around 1825, 1960 and 1985. Increased rainfall variability is reconstructed around 1835 and 1975. The instrumental record identifies an additional period of increased rainfall variability around 1920; this is not reflected by the proxy data and may be an artefact of the instrumental rainfall data during these early years of observation. There is no evidence for a loss in variability back in time, as may be caused by a regression artefact within the reconstruction (e.g. Smerdon 2011, and references therein). Figure 3c shows the evolution of reconstruction skill over time. Whereas the RE values of the individual ensemble members and their median are close to zero or even negative over large parts of the reconstruction period, the early verification RE (average of 0.55 over

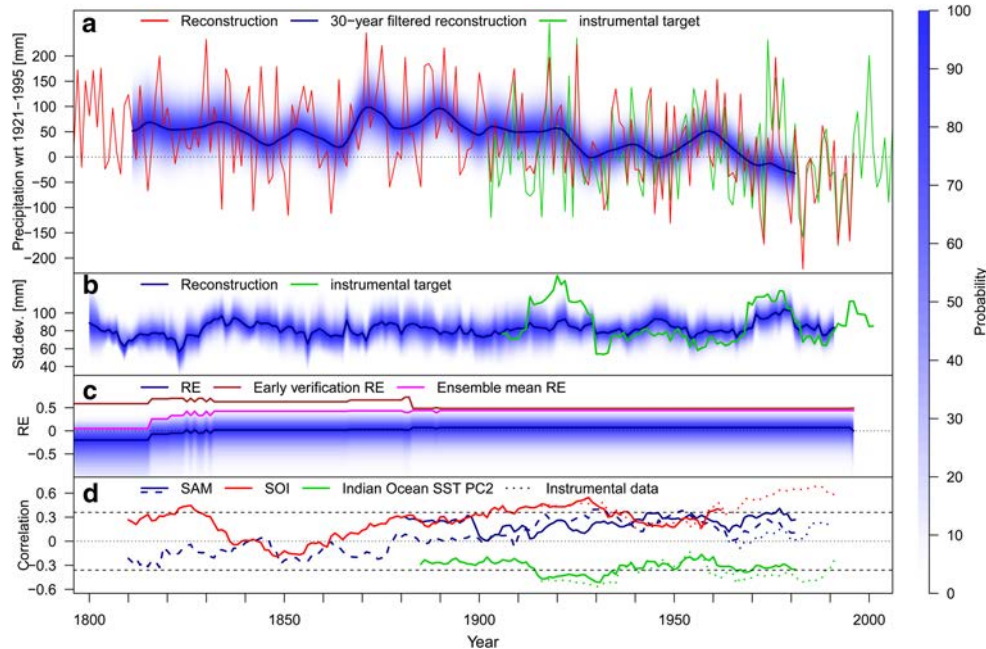


Fig. 3 **a** Summer rainfall zone precipitation reconstruction for 1796–1996 on interannual (red) and 30-year filtered (blue, with uncertainty bands shaded) timescales. Unfiltered instrumental data are shown in green. **b** 30-year running standard deviation of the unfiltered reconstruction (blue) and instrumental (green) data. **c** Distribution of RE values of the reconstruction ensemble members (blue shaded, dark blue line is the median), early verification RE (brown) and RE of

the ensemble mean (pink). **d** 30-year running correlations of the SRZ reconstruction with reconstructions of the SAM index (blue dashed; Villalba et al. 2012; blue; Jones et al. 2009); SOI index (red; Stahle et al. 1998) and IOPC2 (green; HadISST data; Rayner et al. 2003). The dotted lines represent the correlations between the corresponding instrumental data

1796–1996) and the RE of the ensemble mean (0.39) are clearly higher (section S4.3). This indicates reasonable skill of the rainfall reconstruction and emphasises the importance of using ensemble reconstruction approaches to achieve more accurate and robust results, particularly in regions with limited proxy data availability and short overlap periods with instrumental data. There is no shift in mean SRZ rainfall between 1911–1920 and 1921–1995 that could explain the high early verification RE compared to the RE of the ensemble mean. However, due to the short time window available for early verification (10 years), the early verification needs to be interpreted cautiously. Single years, such as the very wet 1918, which is well captured by the reconstruction, can have a strong influence on the early verification RE.

Correlations between reconstructed SH circulation indices, instrumental and reconstructed SRZ rainfall over the full respective periods of overlap are shown in Table 3. Significant correlations are identified with the Stahle et al. (1998) and Li et al. (2011) ENSO reconstructions, but not with the UEP. In subsequent analyses, we use the Stahle et al. (1998) SOI reconstruction because it shows the strongest relationship to instrumental and reconstructed SRZ rainfall. The instrumental SAM reconstruction of Jones et al. (2009) shows a significant relationship with SRZ rainfall in contrast to the tree-ring based SAM reconstruction of Villalba et al. (2012). Despite the significant correlations of instrumental PDO and IPO indices with SRZ rainfall (Table 2), the reconstructions of these indices do not show a significant relationship with our rainfall reconstructions over their period of overlap (Table 3) or during 1921–1995 (not shown) and are therefore not used for further comparisons. The IOD reconstruction (Abram et al. 2008) is significantly correlated to our SRZ reconstruction but not to the instrumental target. Given that our Indian Ocean SST extraction IOPC2 has a clearly stronger correlation with instrumental rainfall ($r = -0.40$, $p < 0.01$) and it is not much

shorter than the Abram et al. (2008) IOD reconstruction (start date 1871 vs. 1846), we use IOPC2 for further analyses of the relationship between Indian Ocean SSTs and SRZ rainfall.

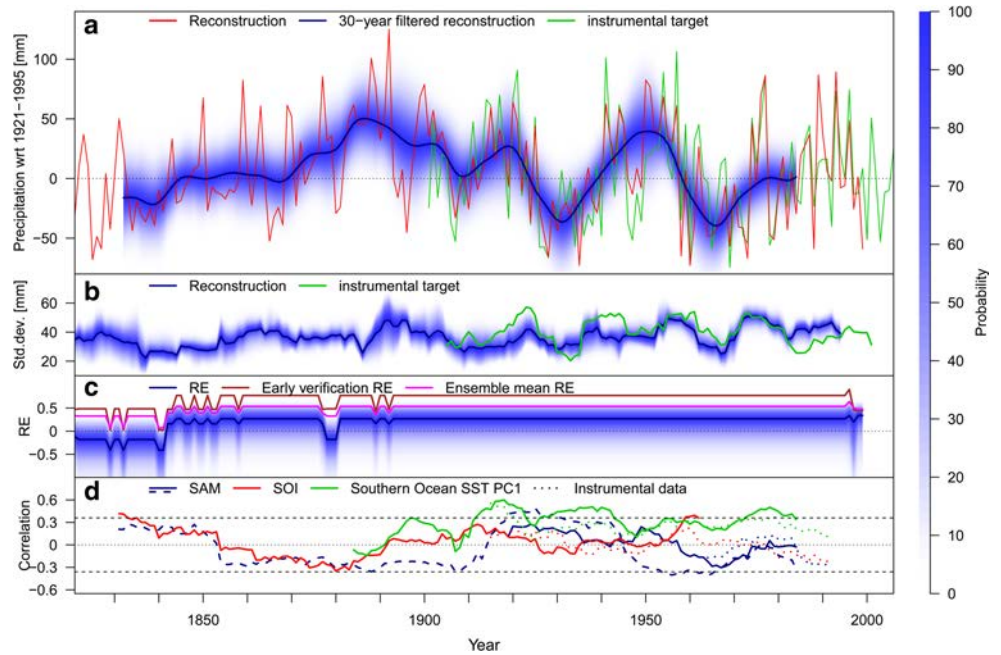
Figure 3d shows the 30-year running correlations of our SRZ rainfall reconstruction with climate indices affecting southern African rainfall, allowing the stability of the teleconnections to be assessed (visual comparison of the time series is provided in Fig. S5.1, alternative illustration in Fig. S5.3). While the relationship between SRZ rainfall and the Southern Oscillation Index (SOI) is significantly positive over the full period of data availability (Tables 2 and 3), the running correlations show a clear breakdown in the mid-nineteenth century. Apart from this period, the correlations oscillate around 0.3. The running correlations with the SAM are weaker, unstable and rarely significant. The most stable relationship is found with IOPC2, confirming that Indian Ocean SSTs are a main driver of SRZ rainfall. The absolute correlations of the SOI and IOPC2 fluctuate in phase ($r = -0.60$), demonstrating the influence of ENSO on regional Indian Ocean SSTs. Both SOI and IOPC2 correlations are relatively weak around 1950, a period where correlations with the SAM reconstruction of Jones et al. (2009) are above average. This suggests a stronger high-latitude influence on SRZ rainfall and a breakdown of tropical influence during this period. The much stronger and longer lasting breakdown of the relationship with the SOI around 1850 does not appear to be associated with a strengthening of the SAM teleconnection. It must be noted, however, that the relationship between SRZ rainfall and the long SAM reconstruction of Villalba et al. (2012) is weak (Table 3). Therefore, interpretations of these correlations in the nineteenth century need to be viewed cautiously. It should also be noted that the temporal instability of running correlations may be caused partly by the various sources of noise potentially biasing this analysis. These uncertainties include noise in the proxy-based

Table 3 Correlations of ENSO, SAM, PDO, IPO and IOD reconstructions with SRZ and WRZ instrumental and reconstructed rainfall data

	SRZ		WRZ	
	Instrumental target	Reconstruction	Instrumental target	Reconstruction
SOI Stahle	0.46 ($p < 0.01$)	0.26 ($p < 0.01$)	0.13 ($p = 0.24$)	0.09 ($p = 0.26$)
UEP	0.20 ($p = 0.09$)	0.11 ($p = 0.15$)	0.11 ($p = 0.35$)	-0.09 ($p = 0.28$)
ENSO Li	-0.46 ($p < 0.01$)	-0.23 ($p < 0.01$)	-0.11 ($p = 0.26$)	-0.07 ($p = 0.38$)
SAM Villalba	0.17 ($p = 0.12$)	-0.12 ($p = 0.12$)	0.06 ($p = 0.57$)	-0.11 ($p = 0.22$)
SAM Jones	0.21 ($p = 0.03$)	0.19 ($p = 0.04$)	-0.07 ($p = 0.51$)	0.05 ($p = 0.60$)
PDO D'Arrigo	-0.18 ($p = 0.09$)	-0.19 ($p = 0.08$)	-0.07 ($p = 0.47$)	-0.05 ($p = 0.68$)
PDO Shen	-0.18 ($p = 0.09$)	-0.12 ($p = 0.10$)	-0.10 ($p = 0.35$)	-0.06 ($p = 0.43$)
IPO Linsley	-0.25 ($p = 0.02$)	-0.09 ($p = 0.24$)	-0.08 ($p = 0.45$)	-0.03 ($p = 0.74$)
IOD Abram	-0.17 ($p = 0.13$)	-0.26 ($p < 0.01$)	0.00 ($p = 1$)	-0.17 ($p = 0.07$)

Correlations with $p < 0.05$ are bolded

Fig. 4 Same as Fig. 3 but for the winter rainfall zone reconstruction covering 1817–1999, and with Southern Ocean SST PC1 instead of Indian Ocean SST PC2. SOPC1 correlations are inverted



reconstructions of SRZ rainfall and the reconstructed climate indices, and the potentially noisy coupling between rainfall and large-scale circulation. Temporal fluctuations of the noise can lead to breakdowns in the correlations. However, the very similar running correlation values of the corresponding instrumental data (dotted lines in Fig. 3d) indicate that the noise introduced by the proxies plays a minor role, at least during the twentieth century.

3.2 Winter rainfall zone

Due to the shorter time period covered by proxies from the WRZ, rainfall can only be reconstructed with positive skill back to 1817 (ensemble mean RE and early verification RE). Reconstructed WRZ rainfall anomalies are shown in Fig. 4a. There is no significant offset in mean rainfall between the nineteenth and twentieth centuries (t-test), suggesting that rainfall levels have not changed significantly through time. Decadal-scale dry conditions are reconstructed around 1835, 1930 and 1965. Wet conditions prevailed around 1890, 1920 and 1950. The driest years of the nineteenth (twentieth) century include 1825, 1827 and 1865 (1935, 1960 and 1973), whereas the wettest are 1892, 1888 and 1889 (1993, 1989 and 1977).

In contrast to the SRZ reconstruction, there is a small but significant (Mann–Kendall test, $p < 0.01$) increase in reconstructed rainfall variability over the reconstruction period (Fig. 4b). The twentieth century instrumental WRZ data show no such trend; however, a positive trend is evident in three of the WRZ proxy records, including the instrumental Royal Observatory data. This suggests that the positive trend is not an artefact of the reconstruction

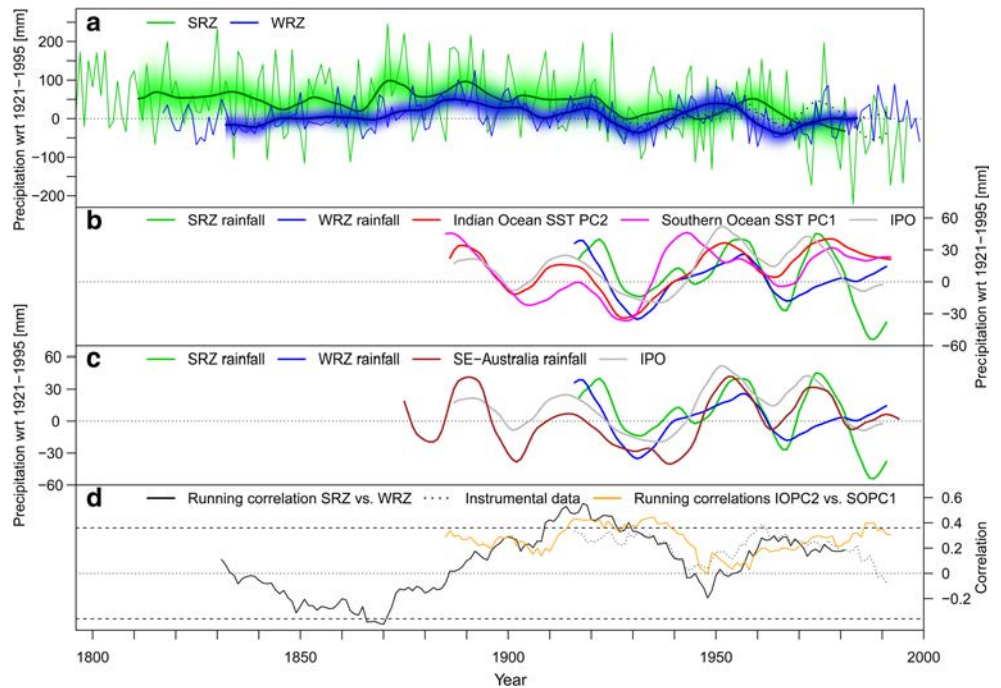
method but instead is reflecting real fluctuations of rainfall variability. RE skills of the individual ensemble members are relatively low (Fig. 4c) but the early verification (average 0.69 for 1817–1999) and ensemble mean RE (0.49) values indicate a skilful WRZ reconstruction. In contrast to the SRZ, WRZ mean rainfall amounts are significantly larger ($p < 0.05$; t-test) during the 1911–1920 early verification period than during the 1921–1995 calibration period in both the reconstructed and instrumental data. This may in part explain the high early verification RE compared to the RE of the ensemble mean and indicates to some extent that our reconstruction is able to detect shifts in mean rainfall.

Running correlations with climate circulation indices (Fig. 4d) show an unstable and mostly weak relationship between WRZ rainfall and the SOI and SAM over time (see also Figs. S5.2 and S5.4). This is unsurprising given the non-significant correlations of SH climate circulation indices with WRZ rainfall (Tables 2 and 3). The regional SST index SOPC1 shows the most stable relationship with WRZ rainfall, similar to IOPC2 for the SRZ. The relationship with SOPC1 is weaker prior to 1911, potentially due to the decrease in SST data quality caused by fewer available observations.

3.3 Comparison of summer and winter rainfall zones

Rainfall levels in the SRZ and WRZ are only weakly correlated at interannual time scales ($r = 0.22$, $p = 0.06$ over 1921–1995 and $r = 0.10$, $p = 0.19$ over 1817–1996 in the instrumental and reconstruction data, respectively). On decadal timescales, the relationship is much stronger

Fig. 5 a Comparison of interannual and 30-year filtered SRZ (*green*) and WRZ (*blue*) rainfall reconstructions. Filtered instrumental data are *dotted*. **b** Comparison of 30-year filtered instrumental rainfall data with IOPC2 (*red*), SOPC1 (*pink*) and IPO index (*grey*). Indices are scaled to the mean and variance of the rainfall data, and IPO data are inverted for better visibility. **c** Same as **b** but compared against instrumental south-eastern Australian rainfall (1860–2008) (*brown*; Gergis and Ashcroft 2012) and IPO (*grey*). **d** 30-year running correlations between reconstructed (*black solid*) and instrumental (*black, dotted*) SRZ and WRZ rainfall and between IOPC2 and SOPC1 (*orange*)



($r = 0.41$, $p = 0.56$ and $r = 0.46$, $p = 0.51$, respectively) but Fig. 5a shows that the fluctuations are not always synchronous. While some very dry (e.g. around 1930) and wet (e.g. around 1890) decadal-scale periods are present across both rainfall regions, there are periods of contrasting anomalies. For example a wet SRZ and dry WRZ occurs during the 1830s and 1960s, while very wet conditions were experienced in the SRZ (WRZ) around 1873 (1950), when the WRZ (SRZ) has only moderate anomalies. On decadal timescales, SRZ, WRZ, IOPC1 and IOPC2 show synchronous fluctuations and a negative relationship with low frequency variability of the Pacific Ocean, illustrated in Fig. 5b by the IPO (Folland et al. 2002). Compared to the high frequency indices of ENSO variability (SOI and NINO3.4), IPO represents Pacific SST variability on decadal timescales and has similar and higher correlations with 30-year loess-filtered SRZ and WRZ rainfall, respectively (SRZ: $r = -0.54$; WRZ: $r = -0.34$). Unfortunately, existing reconstructions of the IPO and our rainfall reconstructions do not share significant fractions of variance to allow an assessment of this relationship further back in time (Table 3).

The late nineteenth century is the wettest period identified in the 30-year filtered southern African rainfall reconstructions for both seasons. Significantly, the period around 1890 was also one of the wettest of the last 200 years in south-eastern Australia, as identified by long instrumental data (Gergis and Ashcroft 2012), natural palaeoclimate archives (Gergis et al. 2012) and documentary sources (Gergis and Ashcroft 2012). Rainfall in south-eastern Australia is also strongly linked to low frequency SST

variability in the Pacific Ocean (correlation with the IPO: $r = -0.77$; Gergis et al. 2012). Figure 5c shows the close agreement between decadal-scale rainfall fluctuations in the SRZ, WRZ, south-eastern Australia and the IPO, suggesting large-scale rainfall teleconnections in the SH through low frequency SST variability in the Pacific Ocean.

Around 1950, there is a breakdown in the running correlations between SRZ and WRZ rainfall (Fig. 5d). This decrease in correlation between the rainfall zones may also be related to the reduced connectivity of the SRZ with the tropical Pacific around this time (Fig. 3d) and the suggested stronger influence of high latitude variability. During the same period around 1950, correlations between IOPC2 and SOPC1 SST fluctuations are also close to zero (Fig. 5d), again highlighting the importance of regional SST fluctuations on rainfall in southern Africa. Due to the lack of long SST data, we can only hypothesise that a similar disconnection between the two ocean basins occurred between c.1840 and 1885, where our SRZ and WRZ reconstructions also exhibit negative correlations. Alternatively, the various sources of noise in the running correlations, as discussed in Sect. 3.1, may be responsible for the unstable relationship in the nineteenth century.

4 Comparison against independent palaeoclimate datasets

The SRZ and WRZ rainfall reconstructions described above show close statistical agreement with the majority of

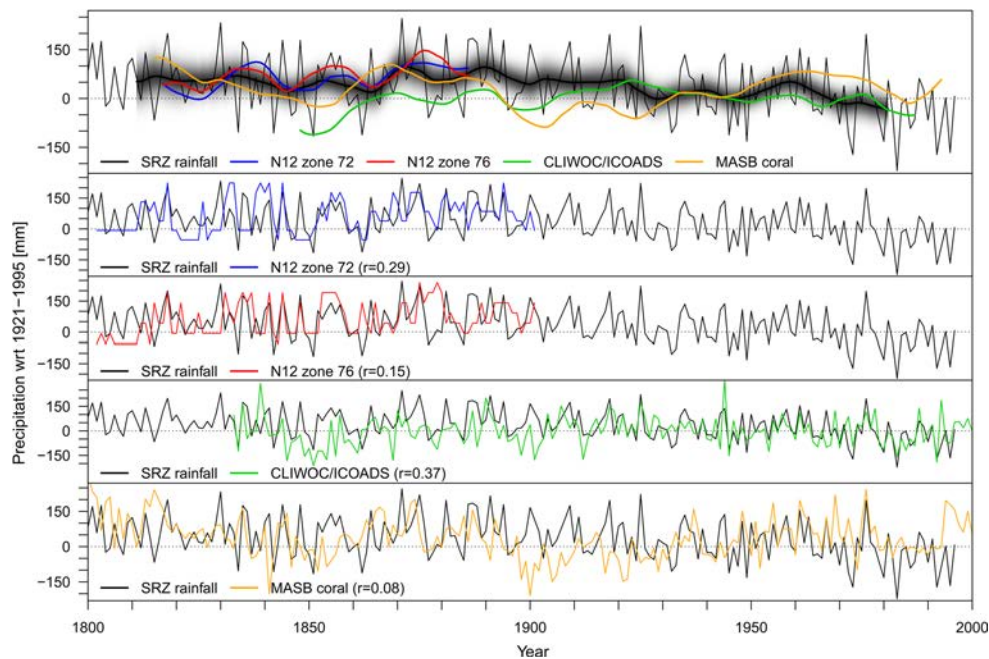
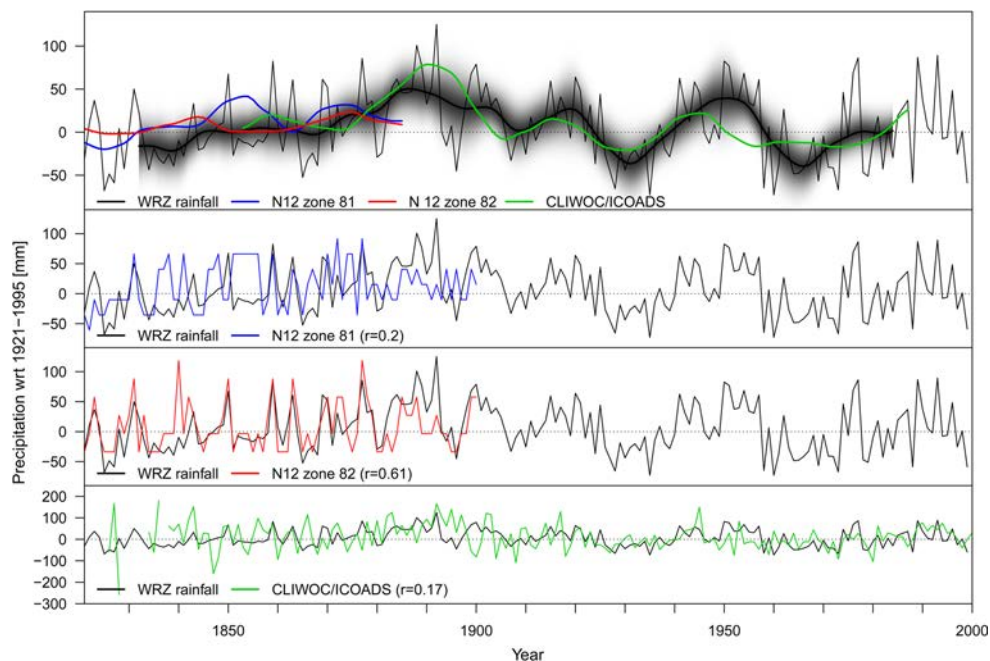


Fig. 6 Comparison of the SRZ rainfall reconstruction with independent reconstructions. *Top* interannual and 30-year loess-filtered reconstruction (*black with shaded uncertainties*) and 30-year filtered rainfall reconstructions for Nicholson et al. (2012) zone 72 (*blue*) and zone 76 (*red*), the CLIWOC/ICOADS reconstruction (*green*) and the MASB coral record (*orange*; Grove et al. 2013). The lower four

panels compare the unfiltered SRZ reconstruction with the unfiltered independent reconstructions (*same colours as above*, correlations over the full period of overlap are indicated in *parentheses* for each reconstruction). Data in the independent reconstructions were scaled to the mean and standard deviation of the SRZ reconstruction over the period 1921–1995 (1801–1900 for the Nicholson et al. (2012) data)

Fig. 7 Same as Fig. 6 but for the WRZ and using Nicholson et al. (2012) zones 81 and 82 as well as the CLIWOC/ICOADS WRZ reconstruction for comparison



selected independent palaeoclimate time-series, suggesting that they are robust. Correlations between our reconstructed annual SRZ rainfall anomalies and independent palaeoclimate data are shown in Fig. 6. The SRZ reconstruction is significantly correlated with Nicholson et al.

(2012) zone 72 ($r = 0.29$, $p < 0.01$) but only weakly correlated with zone 76 ($r = 0.15$, $p = 0.15$). For zone 76, the agreement is good between 1835 and 1880 ($r = 0.37$, $p = 0.02$) but weak in the years before and after this period. A strong and significant correlation is identified between SRZ

rainfall and the CLIWOC/ICOADS wind vector-based reconstruction ($r = 0.37$, $p < 0.001$), whereas no relation with the MASB coral record could be identified (Fig. 6). The agreement of our reconstruction with the Nicholson et al. (2012) zones 72 and 76 is good on decadal scales (Fig. 6a), where all three datasets identify the wet (dry) periods around 1835, 1855 and 1875 (1845 and 1865). The CLIWOC/ICOADS reconstruction suggests drier conditions in the nineteenth century, particularly around 1850, suggesting a weaker influence of oceanic conditions on WRZ rainfall during this period, which is in agreement with low correlations with teleconnection indices (Fig. 3d). However, it is important to note that data coverage in the CLIWOC/ICOADS database is significantly reduced prior to c.1860 (Kutiel et al. 2010), indicating that this reduction in the wind vector-based rainfall amount may also be a data artefact.

Our WRZ rainfall reconstruction is correlated with Nicholson et al. (2012) zones 81 ($r = 0.20$, $p = 0.08$) and 82 ($r = 0.61$, $p < 0.01$) and the CLIWOC/ICOADS reconstruction ($r = 0.17$, $p = 0.02$). The very good agreement with Nicholson et al. (2012) zone 82 (Fig. 7) can be explained by the partial overlap in rainfall predictor data, but also confirms the ability of our multi-proxy approach to identify extremes on various timescales. The CLIWOC/ICOADS data show good agreement with our reconstruction on decadal timescales (Fig. 7a), confirming the very wet phase around 1890. In contrast, the Nicholson et al. (2012) data suggest only moderately wet conditions during this period. Within our WRZ rainfall predictor matrix, this wet phase is predominantly present in the instrumental and tree ring proxies, whereas the documentary records used in our reconstruction show anomalies similar to the Nicholson et al. (2012) zones. This may be explained by the different ability of proxies from different archives to capture very wet phases, or the spatial variability of rainfall within the WRZ over the 1880–1895 period.

5 Conclusions

This paper has presented the first consolidation of terrestrial and marine rainfall proxy records from multiple archives to investigate long-term rainfall variability over southern Africa. The extended rainfall series presented here should provide a baseline for future water-allocation and climate change adaptation strategies in southern Africa. We have demonstrated that our results are skilful and verify well with independent data. We identify decadal-scale rainfall variations in both the SRZ and WRZ. Our results suggest that the SRZ was wetter in the nineteenth than the twentieth century. Rainfall variability in the WRZ shows a small increase over time. SRZ and

WRZ rainfall fluctuations are generally synchronous on decadal scales, but interrupted by periods of contrasting anomalies. Our results confirm that southern African rainfall is influenced by local effects, regional SSTs and large scale teleconnections. The latter also appear to influence rainfall co-variability between the SRZ, WRZ and south-eastern Australia on decadal timescales. The relationship between SRZ rainfall and ENSO has been relatively stable over the last 200 years, although it may have broken down at times, particularly around 1850 and 1950. Running correlations show the most stable relationship between SRZ (WRZ) rainfall and Indian (Southern) Ocean SSTs. Additional rainfall-proxy records are now required to extend this analysis beyond the early nineteenth century.

Acknowledgments RN and DN would like to thank the University of Brighton for funding to support this study. JZ was supported by an AIMS/CSIRO/UWA collaborative post-doctoral Fellowship.

References

- Abram N, Gagan M, Cole J, Hantoro W, Mudelsee M (2008) Recent intensification of tropical climate variability in the Indian Ocean. *Nat Geosci* 1:849–853. doi:[10.1038/ngeo357](https://doi.org/10.1038/ngeo357)
- Allan R, Nicholls N, Jones P, Butterworth I (1991) A further extension of the Tahiti-Darwin SOI, early SOI results and Darwin pressure. *J Clim* 4(7):743–749
- Behera SK, Yamagata T (2001) Subtropical SST dipole events in the southern Indian Ocean. *Geophys Res Lett* 28(2):327–330
- Bretherton CS, Widmann M, Dymnikov VP, Wallace JM, Bladé I (1999) The effective number of spatial degrees of freedom of a time-varying field. *J Clim* 12(7):1990–2009
- Chase BM, Meadows ME (2007) Late quaternary dynamics of southern Africa’s winter rainfall zone. *Earth Sci Rev* 84:103–138
- Cook ER, Briffa KR, Jones PD (1994) Spatial regression methods in dendroclimatology—a review and comparison of two techniques. *Int J Climatol* 14(4):379–402
- D’Arrigo R, Wilson R (2006) On the Asian expression of the PDO. *Int J Climatol* 26(12):1607–1617. doi:[10.1002/joc.1326](https://doi.org/10.1002/joc.1326)
- Dunwiddie PW, Lamarche VC (1980) A Climatically Responsive Tree-Ring Record from *Widdringtonia cedarbergensis*, Cape-Province, South-Africa. *Nature* 286(5775):796–797
- Folland C, Renwick J, Salinger J, Mullan A (2002) Relative influences of the interdecadal Pacific oscillation and ENSO on the south Pacific convergence zone. *Geophys Res Lett* 29(13):21/21–21/24
- García-Herrera R, Konnen GP, Wheeler DA, Prieto MR, Jones PD, Koek FB (2005) CLIWOC: a climatological database for the world’s oceans 1750–1854. *Clim Change* 73(1–2):1–12
- Gergis J, Ashcroft L (2012) Rainfall variations in south-eastern Australia part 2: a comparison of documentary, early instrumental and palaeoclimate records, 1788–2008. *Int J Climatol*. doi:[10.1002/joc.3639](https://doi.org/10.1002/joc.3639)
- Gergis J, Gallant A, Braganza K, Karoly DJ, Allen K, Cullen L, D’Arrigo R, Goodwin I, Grierson P, McGregor S (2012) On the long-term context of the 1997–2009 ‘Big Dry’ in south-eastern Australia: insights from a 206-year multi-proxy rainfall reconstruction. *Clim Change* 111:923–944

- Gillett NP, Kell TD, Jones PD (2006) Regional climate impacts of the Southern Annular Mode. *Geophys Res Lett* 33(23):L23704
- Goddard L, Graham NE (1999) Importance of the Indian Ocean for simulating rainfall anomalies over eastern and southern Africa. *J Geophys Res Atmos* 104(16):19099–19116
- Grove CA, Zinke J, Peeters F, Park W, Scheufen T, Kasper S, Randriamanantsoa B, McCulloch MT, Brummer GJA (2013) Madagascar corals reveal a multidecadal signature of rainfall and river runoff since 1708. *Clim Past* 9(2):641–656
- Jones J, Fogt R, Widmann M, Marshall G, Jones P, Visbeck M (2009) Historical SAM variability. Part I: century-length seasonal reconstructions. *J Clim* 22(20):5319–5345. doi:10.1175/2009JCLI2785.1
- Kelso C, Vogel CH (2007) The climate of Namaqualand in the nineteenth century. *Clim Change* 83(3):357–380. doi:10.1007/s10584-007-9264-1
- Küttel M, Xoplaki E, Gallego D, Luterbacher J, García-Herrera R, Allan R, Barriendos M, Jones P, Wheeler D, Wanner H (2010) The importance of ship log data: reconstructing North Atlantic, European and Mediterranean sea level pressure fields back to 1750. *Clim Dyn* 34:1115–1128
- Li J, Xie SP, Cook ER, Huang G, D'Arrigo R, Liu F, Ma J, Zheng XT (2011) Interdecadal modulation of El Niño amplitude during the past millennium. *Nature Clim Change* 1:114–118
- Linsley BK, Zhang PP, Kaplan A, Howe SS, Wellington GM (2008) Interdecadal-decadal climate variability from multicoralline oxygen isotope records in the South Pacific Convergence Zone region since 1650 A.D. *Paleoceanography* 23(2):PA2219
- Luterbacher J, Xoplaki E, Dietrich D, Rickli R, Jacobeit J, Beck C, Gyalistras D, Schmutz C, Wanner H (2002) Reconstruction of sea level pressure fields over the Eastern North Atlantic and Europe back to 1500. *Clim Dyn* 18(7):545–561
- Luterbacher J, Dietrich D, Xoplaki E, Grosjean M, Wanner H (2004) European seasonal and annual temperature variability, trends, and extremes since 1500. *Science* 303(5663):1499–1503
- Mann M, Zhang Z, Hughes M, Bradley R, Miller S, Rutherford S, NI F (2008) Proxy-based reconstructions of hemispheric and global surface temperature variations over the past two millennia. *Proc Natl Acad Sci USA* 105(36):13252–13257
- Mantua N, Hare S, Zhang Y, Wallace J, Francis R (1997) A Pacific interdecadal climate oscillation with impacts on Salmon Production. *Bull Am Meteorol Soc* 78(6):1069–1079
- Mason SJ, Jury MR (1997) Climatic variability and change over southern Africa: a reflection on underlying processes. *Prog Phys Geogr* 21(1):23–50
- McGregor S, Timmermann A, Timm O (2010) A unified proxy for ENSO and PDO variability since 1650. *Clim Past* 6(1):1–17
- Mitchell TD, Jones PD (2005) An improved method of constructing a database of monthly climate observations and associated high-resolution grids. *Int J Climatol* 25(6):693–712
- Nash DJ, Endfield GH (2002) A 19th century climate chronology for the Kalahari region of central southern Africa derived from missionary correspondence. *Int J Climatol* 22(7):821–841. doi:10.1002/joc.753
- Nash DJ, Endfield GH (2008) 'Splendid rains have fallen': links between El Niño and rainfall variability in the Kalahari, 1840–1900. *Clim Change* 86(3–4):257–290. doi:10.1007/s10584-007-9274-z
- Nash DJ, Grab SW (2010) "A sky of brass and burning winds": documentary evidence of rainfall variability in the Kingdom of Lesotho, Southern Africa, 1824–1900. *Clim Change* 101(3–4):617–653. doi:10.1007/s10584-009-9707-y
- Neukom R, Gergis J (2012) Southern Hemisphere high-resolution palaeoclimate records of the last 2000 years. *The Holocene* 22(5):501–524. doi:10.1177/0959683611427335
- Neukom R, Prieto MD, Moyano R, Luterbacher J, Pfister C, Villalba R, Jones PD, Wanner H (2009) An extended network of documentary data from South America and its potential for quantitative precipitation reconstructions back to the 16th century. *Geophys Res Lett* 36:L12703
- Neukom R, Luterbacher J, Villalba R, Küttel M, Frank D, Jones PD, Grosjean M, Esper J, Lopez L, Wanner H (2010) Multi-centennial summer and winter precipitation variability in southern South America. *Geophys Res Lett* 37:L14708. doi:10.1029/2010GL043680
- Neukom R, Luterbacher J, Villalba R, Küttel M, Frank D, Jones PD, Grosjean M, Wanner H, Aravena JC, Black DE, Christie DA, D'Arrigo R, Lara A, Morales M, Soliz-Gamboa C, Srur A, Urrutia R, von Gunten L (2011) Multiproxy summer and winter surface air temperature field reconstructions for southern South America covering the past centuries. *Clim Dyn* 37(1–2):35–51. doi:10.1007/s00382-010-0793-3
- Nicholson SE (1997) An analysis of the ENSO signal in the tropical Atlantic and western Indian Oceans. *Int J Climatol* 17:345–375
- Nicholson S, Kim J (1997) The relationship of the El Niño-Southern oscillation to African rainfall. *Int J Climatol* 17:117–135
- Nicholson SE, Leposo D, Grist J (2001) The relationship between El Niño and drought over Botswana. *J Clim* 14:323–335
- Nicholson SE, Klotter D, Dezfuli AK (2012) Spatial reconstruction of semi-quantitative precipitation fields over Africa during the nineteenth century from documentary evidence and gauge data. *Quat Res* 78(1):13–23
- Peterson TC, Vose RS (1997) An overview of the global historical climatology network temperature database. *Bull Am Meteorol Soc* 78(12):2837–2849
- Power S, Casey T, Folland C, Colman A, Mehta V (1999) Interdecadal modulation of the impact of ENSO on Australia. *Clim Dyn* 15:319–324
- Rayner NA, Parker DE, Horton EB, Folland CK, Alexander LV, Rowell DP, Kent EC, Kaplan A (2003) Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. *J Geophys Res* 108(D14):4407
- Reason CJC (2001) Subtropical Indian Ocean SST dipole events and southern African rainfall. *Geophys Res Lett* 28(11):2225–2227
- Reason CJC, Jagadheesha D (2005) Relationships between South Atlantic SST variability and atmospheric circulation over the South African region during austral winter. *J Clim* 18(16):3339–3355
- Saji N, Goswami G, Vinayachandran P, Yamagata T (1999) A dipole mode in the tropical Indian Ocean. *Nature* 401:360–363
- Shen C, Wang WC, Gong W, Hao Z (2006) A Pacific decadal oscillation record since 1470 AD reconstructed from proxy data of summer rainfall over eastern China. *Geophys Res Lett* 33(L03702). doi:10.1029/2005GL024804
- Smerdon JE (2011) Climate models as a test bed for climate reconstruction methods: pseudoproxy experiments. *Wiley Interdisciplinary Reviews: Clim Change*. doi:10.1002/wcc.1149
- Stahle DW, D'Arrigo RD, Krusic PJ, Cleaveland MK, Cook ER, Allan RJ, Cole JE, Dunbar RB, Therrell MD, Gay DA, Moore MD, Stokes MA, Burns BT, Villanueva-Diaz J, Thompson LG (1998) Experimental dendroclimatic reconstruction of the Southern oscillation. *Bull Am Meteorol Soc* 79(10):2137–2152
- Therrell MD, Stahle DW, Ries LP, Shugart HH (2006) Tree-ring reconstructed rainfall variability in Zimbabwe. *Clim Dyn* 26(7–8):677–685. doi:10.1007/s00382-005-0108-2
- Thompson DWJ, Wallace JM (2000) Annular modes in the extratropical circulation. Part I: month-to-month variability. *J Clim* 13(5):1000–1016
- Trouet V, Diaz HF, Wahl ER, Viau AE, Graham R, Graham N, Cook ER (2013) A 1500-year reconstruction of annual mean temperature for temperate North America on decadal-to-multidecadal time scales. *Environ Res Lett* 8:024008. doi:10.1088/1748-9326/8/2/024008

- Tyson PD (1986) *Climate change and variability in Southern Africa*. Oxford University Press, Capetown
- Villalba R, Lara A, Masiokas M, Urrutia R, Cook ER, Christie D, Mundo IA, Boninsegna J, Fenwick P, Neukom R, Allen K, Morales M, Araneo DC, Marshall G, Srur A, Aravena JC, Palmer J (2012) Unusual southern hemisphere tree growth patterns induced by changes in the southern annular mode. *Nat Geosci* 5:793–798. doi:[10.1038/NGEO1613](https://doi.org/10.1038/NGEO1613)
- Vogel CH (1989) A documentary-derived climatic chronology for South Africa, 1820–1900. *Clim Change* 14(3):291–307. doi:[10.1007/bf00134967](https://doi.org/10.1007/bf00134967)
- Wahl ER, Smerdon JE (2012) Comparative performance of paleoclimate field and index reconstructions derived from climate proxies and noise-only predictors. *Geophys Res Lett* 39(6):L06703
- Watterson IG (2009) Components of precipitation and temperature anomalies and change associated with modes of the Southern Hemisphere. *Int J Climatol* 29(6):809–826
- Zinke J, Dullo WC, Heiss GA, Eisenhauer A (2004) ENSO and Indian Ocean subtropical dipole variability is recorded in a coral record off southwest Madagascar for the period 1659 to 1995. *Earth Planet Sci Lett* 228(1–2):177–194
- Zinke J, Pfeiffer M, Timm O, Dullo W, Brummer G (2009) Western Indian Ocean marine and terrestrial records of climate variability: a review and new concepts on land-ocean interactions since AD 1660. *Int J Earth Sci* 98:115–133. doi:[10.1007/s00531-008-0365-5](https://doi.org/10.1007/s00531-008-0365-5)

Supplementary material

Table of Contents

S1.	Definition of SRZ.....	3
S2.	Documentary data: extension to present	5
S2.1.	Rationale	5
S2.2.	Instrumental records used for the extension	6
S2.3.	Methods to calculate the pseudo-documentaries	9
S3.	Correlation between proxy and instrumental data	17
S4.	Description of reconstruction method.....	19
S4.1.	Rainfall reconstruction.....	19
S4.2.	Reconstruction uncertainties.....	20
S4.3.	Reconstruction skill	20
S5.	Climate index and SST comparison plots	23
S5.1.	Reconstruction vs. indices	23
S5.2.	Running correlations regional SSTs vs. ENSO indices.....	24
S6.	CLIWOC/ICOADS-based rainfall reconstructions	26
S7.	References	29

S1. Definition of SRZ

Fig. S1.1 shows the loadings of first three the rainfall Principal Components (PC 1-3) over all grid cells south of 10°S where $\geq 66\%$ of annual precipitation falls in the ONDJFM season. Fig. S1.2 shows the spatial correlations of PC 1-3. Table S1.1 lists the fraction of total variance explained by PC 1-3, together with the fraction of grid cells that display significant correlations with PC 1-3.

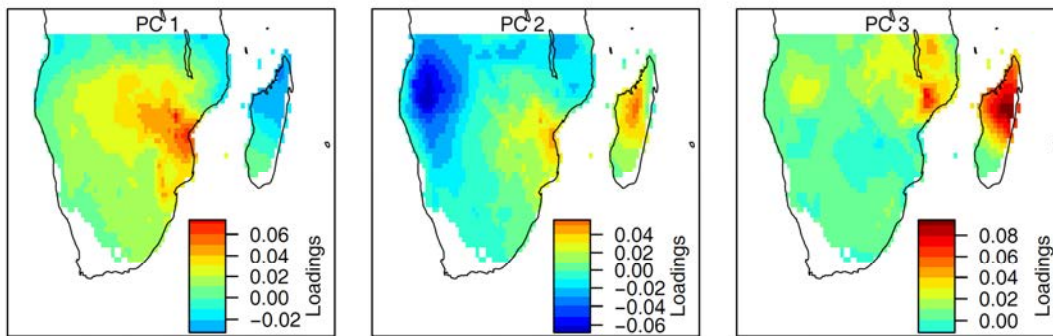


Fig. S1.1: Spatial distributions of the loadings to the first three PCs of ONDJFM rainfall. Only grid cells with $\geq 66\%$ of annual precipitation falling during the ONDJFM season are used.

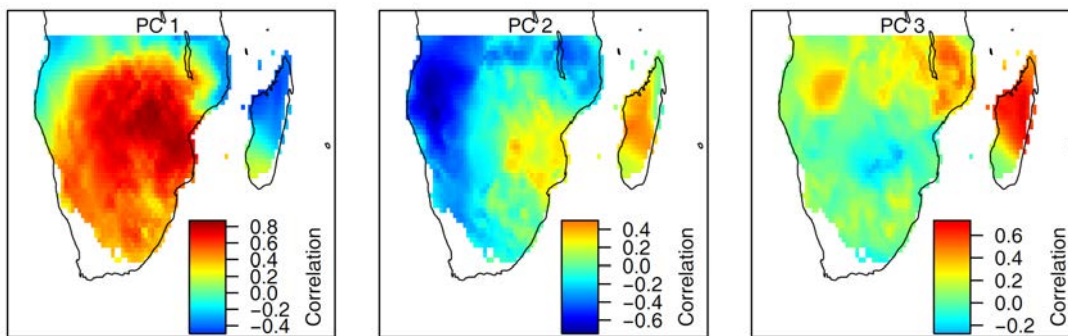


Fig. S1.2: Spatial distributions of ONDJFM rainfall correlations with the first three PCs. Only grid cells with $\geq 66\%$ of annual precipitation falling during the ONDJFM season are used.

Table S1.1: ev: Explained variance of total rainfall variability over the area with $\geq 66\%$ of annual precipitation falling during the ONDJFM season by PC 1-3. Fsig: Fraction of grid cells with significant correlation with PC 1-3. Fsig.pos: Same as Fsig but counting only significant positive correlations.

	ev	Fsig	Fsig.pos
PC 1	26	77	67
PC 2	10	51	13
PC 3	9	26	26

S2. Documentary data: extension to present

S2.1. Rationale

The five documentary-based rainfall index-time series available from southern Africa (see main text and Table S2.2 below) cover a period of 76-85 years within the 19th century. None of these records extend into the 20th century, making calibration with instrumental data impossible. To allow inclusion in our multi-proxy rainfall reconstruction, these records needed to be extended to the present day. There are two ways in which it is possible to perform such an extension:

- a) ‘Real documentaries’: Using the same or comparable sources that were used to derive the 19th-century documentary records to obtain rainfall indices for the 20th century;
- b) ‘Pseudo-documentaries’: Degrading instrumental data to generate pseudo-documentary indices, which mimic the properties of a ‘real’ documentary time series.

Both approaches are described in further detail in Neukom et al. (2009). Whereas approach (a) would yield more robust records of true documentary nature, it would be extremely time consuming and would require years of additional archive work to derive full-length 20th century documentary time series for all five regions. Given that 20th century instrumental data are available for southern Africa, the outcome of this approach would not provide any new information or benefit except for the possibility of calibrating the documentary records against instrumental data. Neukom et al. (2009) have shown that the results of a rainfall reconstruction using ‘real’ and ‘pseudo’ documentaries are very similar. Furthermore, pseudo-documentaries are able to mimic realistically the statistical properties of a documentary time series, if the correlation between documentary and instrumental data can be estimated (Neukom et al. 2009). Therefore, we applied approach (b) here, to extend the southern Africa rainfall records to the present day. The methods we used are based on Neukom et al. (2009) and are described in detail in the following paragraphs.

S2.2. Instrumental records used for the extension

First, we needed to identify which instrumental rainfall station data adequately represented the five regional documentary records. Candidate stations needed to extend further back than AD 1900 in order to have an overlap period with the documentary data. Table S2.1 summarises the available data for each region.

For all calculations, anomalies relative to the 1901-1990 period were used for the instrumental data. For larger regions, the available station data were grouped into sub-groups using cluster analysis (k-means method). Various combinations of composite time-series among individual stations and clusters were tested to obtain the optimal target data to derive the pseudo-documentaries for each region. We used the temporal coverage of the stations and Spearman correlation coefficients with the documentary data in the overlap period as selection criteria. Details for each region are provided in the following paragraphs. Table S2.2 summarises the instrumental target data for each region.

Kalahari (original annual indices and all calculations based on July-June rain-years)

We used the mean of the stations Barkly West and Kimberley as the instrumental target composite.

Namaqualand (calendar years)

We used the Springbok time series. Two years with missing values (1880 and 1993) were infilled with linear regression from the Port Nolloth station.

Lesotho (July-June rain-years)

We used a composite of all available stations.

Southern Cape (calendar years)

Cluster analysis resulted in a cluster from the inland/north region (Clanwilliam, Malmesbury and Worcester), separated from the other records. This cluster has a correlation of 0.34 with the documentary record in the overlap period. Given that the Clanwilliam station alone has higher correlations and the best temporal coverage of the three records from the cluster, Clanwilliam was

used as the target time series. Missing values (seven years between 1880 and 1910) were infilled using multiple linear regression from Malmesbury and Worcester.

Eastern Cape (calendar years)

Cluster analysis yielded four clearly separated clusters, which also matched the geographical distribution of the stations: northeastern stations (Steynsburg and Tarkastad), eastern stations (Grahamstown, Fort Beaufort, King William's Town, Somerset East and Alicedale) southeastern/coastal stations (Port Elizabeth, Uitenhage and Humansdorp) and western stations (Oudtshoorn, Prince Albert, Calitzdorp and Willowmore). Only the eastern ($r=0.49$, $p<0.01$) and southeastern/coastal ($r=0.42$, $p<0.01$) clusters showed significant correlations with the documentary data in the overlap period; the northeastern ($r=0.10$, $p=0.71$) and western ($r=0.34$, $p=0.12$) show a clearly weaker relationship to the documentary data. This result also reflected the distribution of locations from which historical information used in the construction of documentary time series was derived; predominantly sites within 100km of the coast. The average of the eastern and southeastern/coastal composites has a correlation of 0.51 ($p<0.001$) with the documentary time series, a higher value than the individual composites from the two regions, suggesting that both regions contribute to the data derived from the historical archives. We used this average of the eastern and southeastern/coastal composites as the target time series. All composites were calculated after standardising all stations, in order to account for differences in variance among the stations.

Table S2.1: Instrumental data available for each region (Peterson and Vose 1997). Station names, coordinates, time period covered, Spearman correlation with documentary data in the overlap period (*cor*) and length of the overlap period (*n*). Temporal coverage of the documentary data for each region in brackets.

Region/station name	Lat (°S)	Long (°E)	Start	End	<i>cor</i>	<i>n</i>
<i>Kalahari (1815-1900)</i>						
Barkly West	28.30	24.55	1884	1998	0.67	11
Kimberley	28.80	24.77	1877	2008	0.61	23
<i>Namaqualand (1817-1900)</i>						
Pella	29.00	19.10	1878	1981	0.08	14
Port Nolloth	29.23	16.87	1879	1997	0.51	13
Springbok	29.67	17.88	1878	1997	0.62	20
<i>Lesotho (1824-1900)</i>						
Leribe	28.88	28.05	1886	1996	0.67	12
Ladybrand	29.17	27.43	1879	1993	0.75	11
Maseru Airport	29.45	27.57	1888	1996	NA ^a	2
Quthing	30.42	27.72	1887	1996	0.58	13
Aliwal North	30.80	26.90	1866	1993	0.37	29
<i>Southern Cape (1821-1900)</i>						
Clanwilliam	32.18	18.90	1869	1997	0.52	24
Malmesbury	33.45	18.73	1877	1995	0.27	15
Wellington	33.65	19.00	1876	2000	0.09 ^b	23
Worcester	33.65	19.43	1880	1995	0.32	21
Groot Drakenstein	33.80	19.00	1865	1975	0.31 ^c	22
Royal Observatory	33.93	18.48	1850	2000	0.27	50
Bredasdorp	34.53	20.03	1869	1997	0.22	22
<i>Eastern Cape (1821-1900)</i>						
Steynsburg	31.30	25.83	1877	1997	-0.02	13
Tarkastad	32.00	26.27	1877	1995	-0.01	15
Somerset East	32.72	25.58	1870	1997	0.24	25
Fort Beaufort	32.78	26.63	1877	1997	0.43	16
King William's Town	32.80	27.40	1868	1997	0.49	29
Prince Albert	33.22	22.03	1877	1997	0.43	19
Willowmore	33.28	23.50	1877	2000	0.39	18
Grahamstown	33.30	26.53	1854	1997	0.43	32
Alicedale	33.32	26.08	1877	1995	0.17	11
Calitzdorp	33.53	21.68	1877	1995	0.36	17
Oudtshoorn	33.58	22.20	1878	1997	0.28	17
Uitenhage	33.80	25.40	1859	1999	0.42	37
Port Elizabeth	34.00	25.60	1867	2008	0.32	17
Humansdorp	34.03	24.77	1878	1997	-0.03	17

^a Only two overlapping years, both with a documentary index of 0;

^b This time series has variance issues and many missing values and is therefore not used;

^c This time series has an artificial shift in its mean between the 19th and 20th centuries (change in mean of 2.45 standard deviations) and is therefore not used.

Table S2.2: Properties of the instrumental target data used for each region: Number and period of the overlapping years with the documentary record and Spearman correlation over this period (*cor*); start and end years of the combined documentary/pseudo-documentary time series.

Region	Years overlap	<i>cor</i>	Start	End
Kalahari	23 (1877-1899)	0.66	1815	2002
Namaqualand	21 (1878-1900)	0.65	1817	1996
Lesotho	33 (1866-1899)	0.54	1824	1995
Southern Cape	29 (1870-1900)	0.56	1821	1996
Eastern Cape	46 (1854-1900)	0.51	1821	2007

S2.3. Methods to calculate the pseudo-documentaries

The selected instrumental target rainfall series of all regions were degraded to pseudo-documentaries using the approach of Neukom et al. (2009), as illustrated in the workflow shown in Fig. S2.1. First, the overlap period correlation between the documentary data and instrumental rainfall series was calculated and used to derive the signal-to-noise ratio for the degrading process (Mann et al. 2007). The corresponding amount of white noise was then added to the instrumental data. This was needed to obtain a time series with realistic amounts of noise, to avoid over- or under-weighting the documentary data in the climate reconstruction relative to the proxy records from other archives, and to obtain realistic reconstruction verification and uncertainty measures.

Next, the degraded instrumental data needed to be categorised to the index categories of the documentary data. All southern African documentary rainfall records consist of five index categories: -2 (very dry), -1 (dry), 0 (normal), +1 (wet), +2 (very wet). A major difficulty in quantitative analyses of documentary data is allocating these categories to absolute rainfall values (Brazdil et al. 2005; Pfister et al. 2008; Dobrovolny et al. 2009). Pfister (1999) uses thresholds based on deviations from the climatological mean: ± 0.65 standard deviations (std.dev.) for the category ± 1 , and ± 1.3 std.dev. for the category ± 2 . Note that the definition of the zero value (a normal year) is also not trivial as it depends strongly on the reference period chosen. Given the

nature of documentary sources, the distribution of the index categories is often skewed, and ‘normal’ years (index 0) are often less frequent than wet (+1) or dry (-1) years.

The use of such std.dev.-based and symmetric thresholds for the categories may not be ideal, leading to pseudo-documentaries with very different distributions of categories compared to the ‘original’ documentary data, most probably not reflecting real climatic shifts. We therefore used a slightly different approach to define the categories, making use of the available but short overlap periods of instrumental and documentary data. We defined the thresholds in a way that the fraction of years allocated to each category remained the same for the original documentary data and the categorised instrumental data. For example, in the case of Lesotho, three years of the overlap period have the category -2 in the documentary data, reflecting a fraction of 9%. The other categories -1, 0, +1 and +2 have fractions of 35%, 26%, 21% and 9%, respectively. The percentiles of each category in the overlap periods of all regions are shown in Table S2.3. In the example of Lesotho, all years with rainfall anomalies smaller than -52mm (with respect to 1901-1990; representing the 9th percentile of rainfall amounts in the 1866-1899 overlap period) were allocated to the category -2. Years with anomalies between -52mm and 55mm (representing the 44th percentile (9th + 35th percentile)) were allocated to an index of -1. Accordingly, the thresholds of 121mm, representing the 70th percentile (100-21-9) and 238mm, representing the 91st percentile (100-9) were used as thresholds for the categories +1 and +2, respectively.

Note that we used the raw and non-degraded instrumental data to define the category-thresholds, whereas the allocation to categories for the pseudo-documentaries was calculated after adding white noise to the instrumental data. The disadvantage of this percentile-based categorisation approach is that it depends strongly on the overlap periods, which are relatively short, so that each individual year can affect the outcome (such as outliers or problematic years in terms of amount or coherence of the available documentary information). The categorisation can have a substantial influence on the relation of the pseudo-documentaries to the original instrumental

data. To avoid biases introduced by this step, we repeated the white noise degrading and categorisation until three criteria were fulfilled:

- i) Realistic target correlation - The correlation of the pseudo-documentaries with the original instrumental data must be within ± 0.05 of the target correlation calculated within the overlap period;
- ii) No artificial trend - The residuals between the original instrumental data and the white noise-degraded rainfall amounts must not have a significant ($p < 0.05$) trend over the full period of instrumental data availability, as assessed by the Mann-Kendall test (Mann 1945);
- iii) No variance bias - The ratio of the standard deviations between the overlap period (sdo) and the full period of instrumental data availability (sdl) must remain stable. For the documentary data we used the standard deviation of the original documentary data in the overlap period (sdo_{docu}) and the standard deviation of the pseudo-documentaries over the full period of instrumental data availability (sdl_{pseudo}). Their ratio must not exceed 1 ± 0.15 times the same ratio of the instrumental data (the value of 0.15 reflects a subjective choice based on the trade-off between calculating too many iterations and allowing too large variance biases):

$$0.85 < \frac{\frac{sdl_{pseudo}}{sdo_{docu}}}{\frac{sdl_{instr}}{sdo_{instr}}} < 1.15$$

To assess the importance of the definition of the categories, we repeated the calculation of the pseudo-documentaries using two other approaches. First, we used the method of Pfister (1999) using ± 0.65 and ± 1.3 standard deviations from the “normal” value for the categories ± 1 and ± 2 , respectively. Instead of an average to define “normal” years, we used the maximum value of the distribution function of rainfall amounts over the full period of instrumental data availability (Neukom et al. 2009). Note that none of the instrumental composites has a significant trend, except for Lesotho with a negative trend of -0.9mm/year , making the choice of the reference period less important. Second, we applied a scaling approach where the thresholds were identified by scaling

the documentary data to the mean and standard deviation of the instrumental data over the overlap period, and with the resulting rainfall amounts used as thresholds for each category.

Fig. S2.2 compares the reconstruction outcomes of the three approaches. The reconstructions are very similar but the approach of Pfister (1999) yields systematically lower rainfall amounts during the 19th century (Fig. S2.2a). This is because the method of Pfister (1999) does not account for the unequal distribution of positive and negative indices in the original documentary data. These are all slightly skewed towards dry extremes (Figs. S2.3-S2.7; Table S2.3). As a consequence, the method of Pfister (1999) yielded a too large fraction of positive indices in the 20th century pseudo-documentaries relative to the original data from the 19th century (Fig. S2.2b), leading to a relatively drier 19th century reconstruction. The other two approaches are able to deal with the unequal distribution of indices in the documentaries and yielded more realistic results. The use of standard deviation-based methods for categorisation is therefore not ideal because of the non-normal distribution of most documentary-based time series.

The pseudo-documentaries (starting in the first year after the overlap period) were then spliced together with the original historical indices to obtain the documentary time series used in the rainfall reconstructions. The entire routine was repeated 1000 times to obtain an ensemble of pseudo-documentaries. For each reconstruction ensemble member, one of the 1000 pseudo-documentary records was selected for each region. Figs S2.3-S2.7 show the documentary records (grey bars), instrumental composites (black) and one example of a pseudo-documentary time series (red bars) for each region.

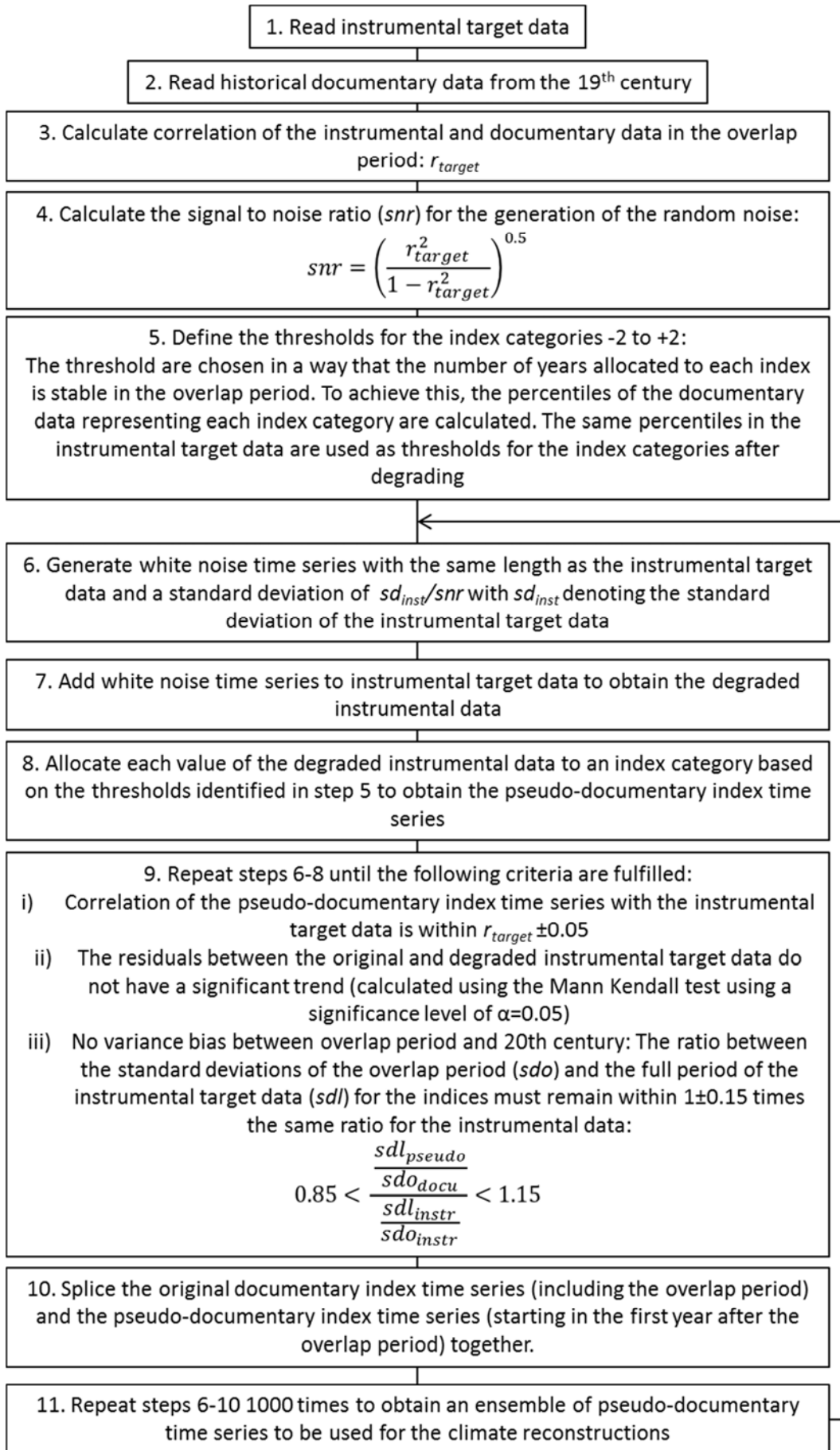


Fig. S2.1: Workflow of the process of generating pseudo-documentaries

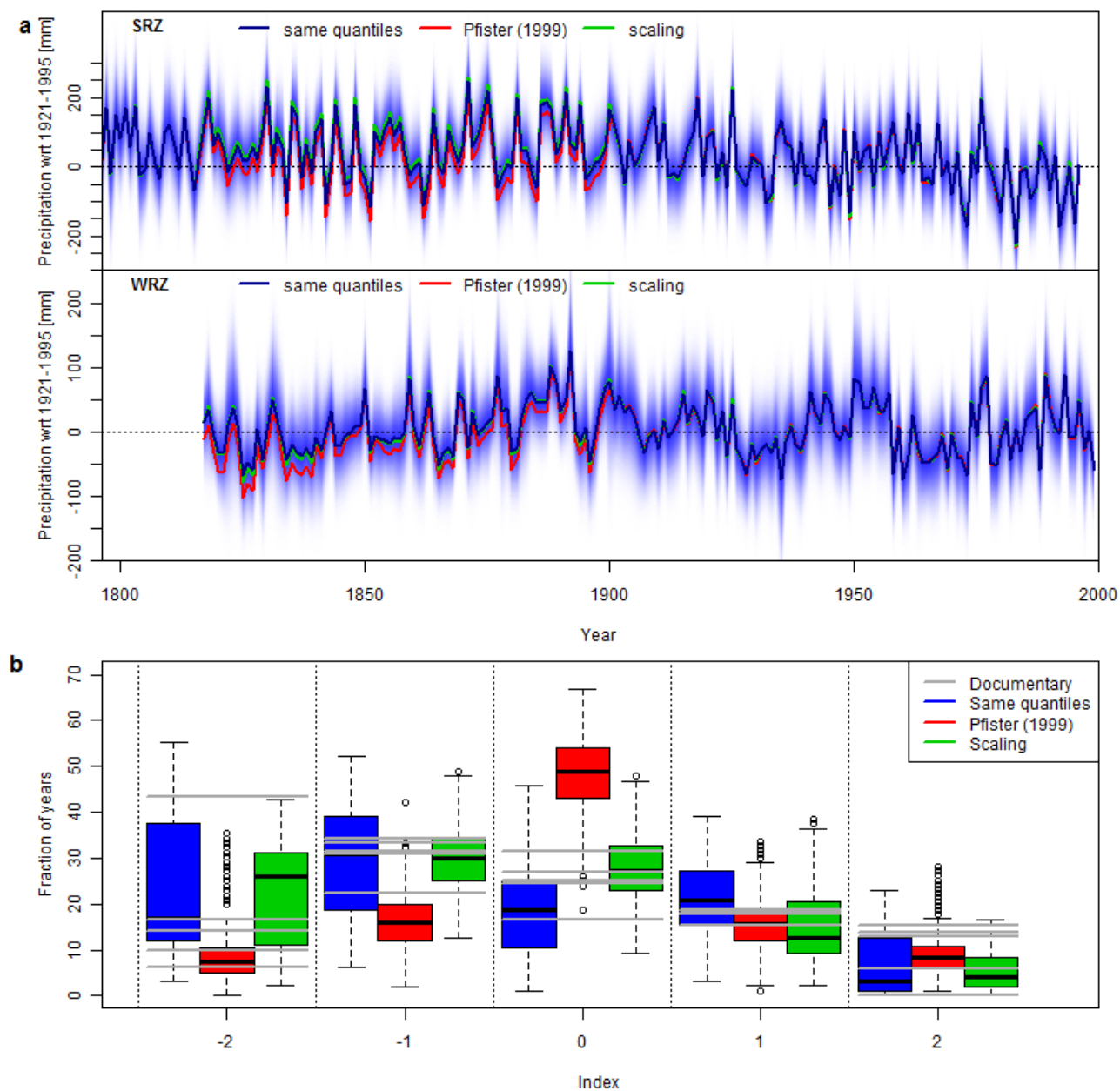


Fig. S2.2: (a) Comparison of SRZ (top) and WRZ (bottom) rainfall reconstructions based on different methods to allocate rainfall data to index categories. Blue: Same fraction of years in each category during the overlap period. This approach was used for the final reconstructions (see text). Blue shading represents the reconstruction uncertainties as in Figs 2 and 3 in the main text. Red: Standard deviation-based thresholds as defined by Pfister (1999). Green: Scaling of indices to mean and standard deviation of instrumental data in the overlap period. (b) Fraction of years allocated to each index category. Grey lines: Values from the original documentary data of the 19th century for each of the five regions. Coloured boxplots: Fraction of years allocated to each category over all regions and noise-iterations (1000 per region) in the 20th century pseudo-documentaries. Colours same as in (a)

Table S2.3: Percentiles of each index category of the documentary data in the overlap period with instrumental data.

Category	-2	-1	0	1	2
Kalahari	26	22	13	30	9
Namaqualand	38	24	14	24	0
Lesotho	9	35	26	21	9
Southern Cape	13	35	26	13	13
Eastern Cape	17	30	24	15	13

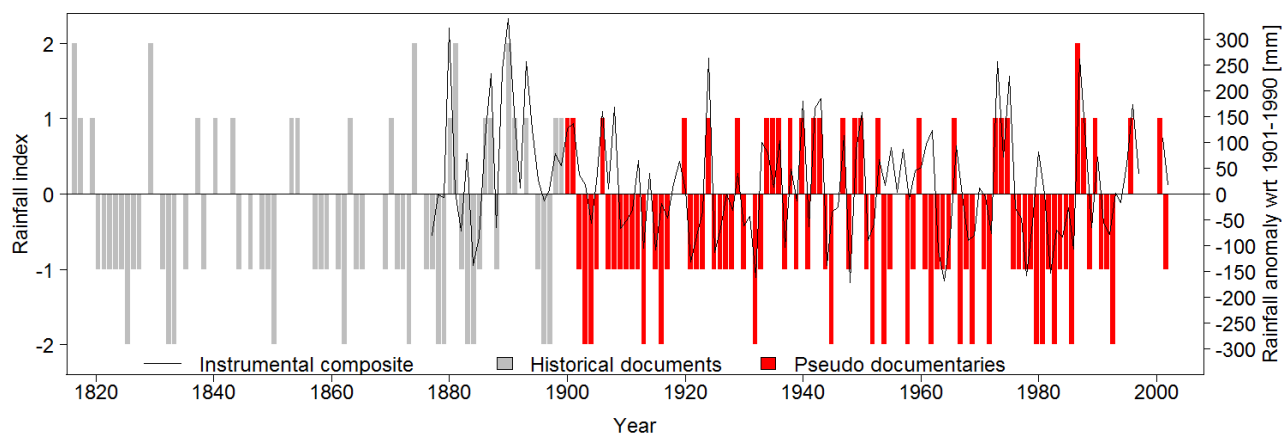


Fig. S2.3: Kalahari documentary record (grey bars), instrumental composite (black line) and one realisation of pseudo-documentary indices (red bars)

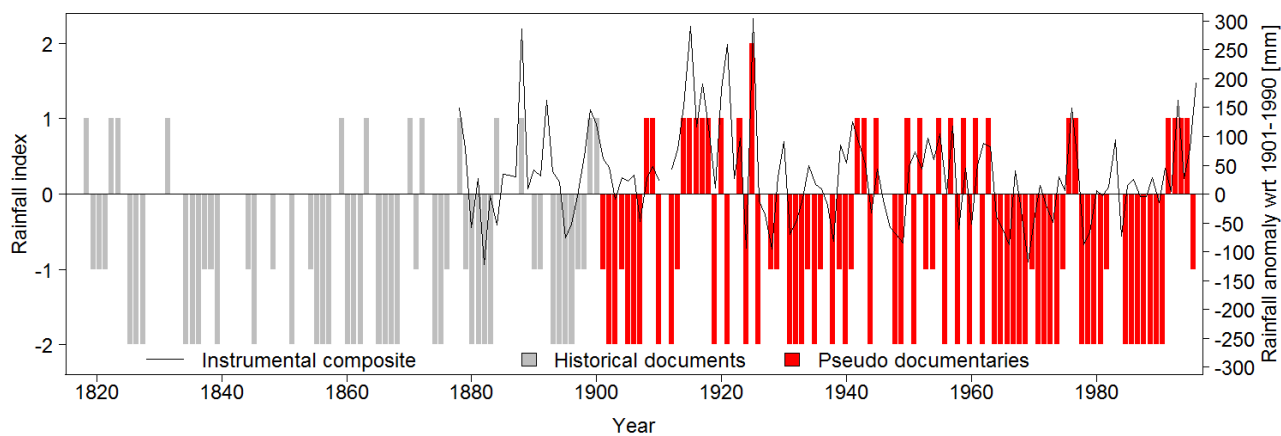


Fig. S2.4: Same as Fig. S2.3 but for the Namaqualand region.

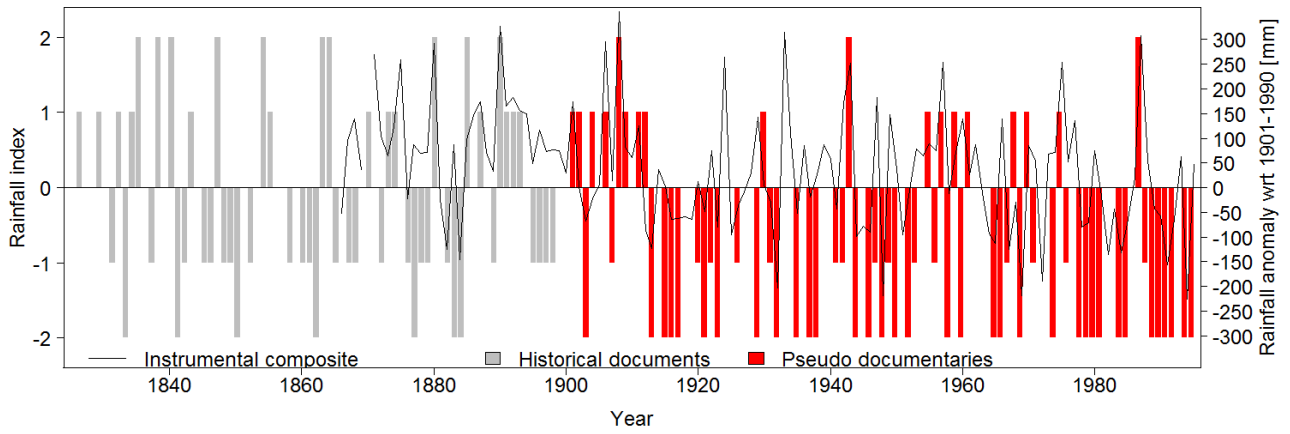


Fig. S2.5: Same as Fig. S2.3 but for the Lesotho region

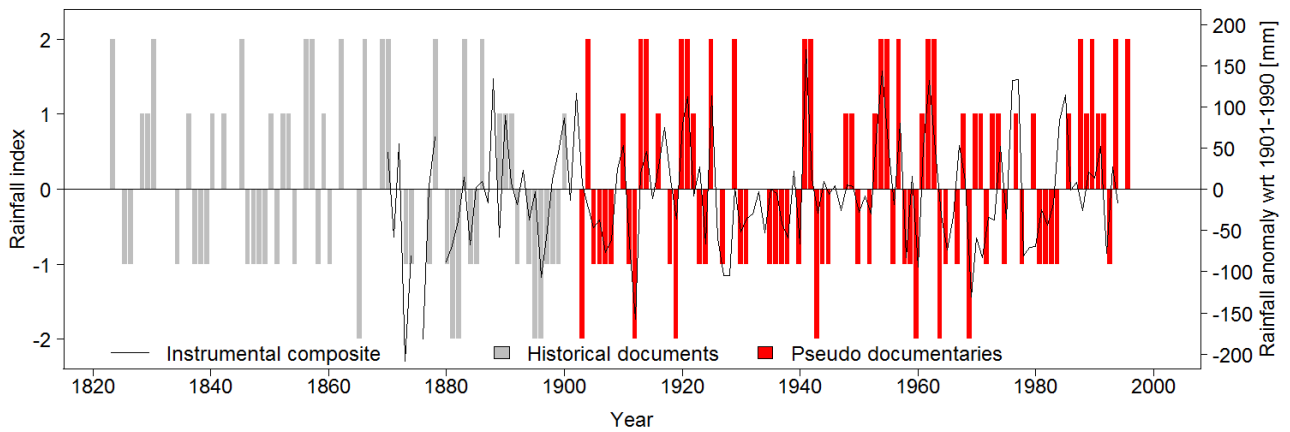


Fig. S2.6: Same as Fig. S2.3 but for the Southern Cape region

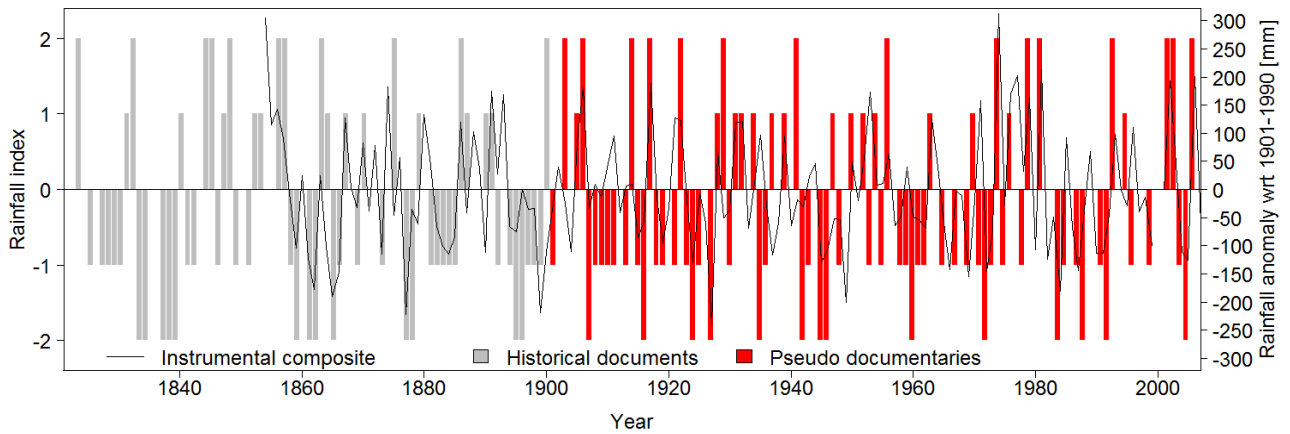


Fig. S2.7: Same as Fig. S2.3 but for the Eastern Cape region

S3. Correlation between proxy and instrumental data

Table S3.1 shows the correlations of our SRZ and WRZ rainfall proxies with the corresponding reconstruction targets. Correlations were calculated over the 1911-1995 period. This period was chosen because before and after these years the coverage with instrumental station data in the Global Historical Climatology Network (GHCN) database (Peterson and Vose 1997) is clearly reduced (Fig. S3.1).

Fig. S3.2 shows an alternative illustration of Fig.1 in the main text but showing only significant ($p < 0.05$) correlations in panels c and d.

Table S3.1: Pearson correlation coefficients with the instrumental targets (*cor*) and an indicator of selection (*sel*) are provided for all proxy records and both seasons.

Name	ONDJFM		AMJJAS	
	<i>cor</i>	<i>sel</i>	<i>cor</i>	<i>sel</i>
Mayotte	0.21	1	0.11	
Ifaty1 d18O	0.25	1	0.11	
Ifaty1 Sr/Ca	-0.21	1	0.06	
Ifaty4 d18O	-0.28	1	-0.25	
Ifaty4 Sr/Ca	-0.20	1	-0.21	
Zimbabwe precipitation reconstruction	0.47	1	0.10	
Die Bos	0.12		0.27	1
Kalahari precipitation reconstruction	0.35	1	0.12	
Namaqualand precipitation reconstruction	0.21		0.26	1
Lesotho precipitation reconstruction	0.35	1	0.11	
Eastern Cape precipitation reconstruction ^a	0.20	1	0.25	1
Southern Cape precipitation reconstruction	0.10		0.41	1
Royal Observatory	0.20		0.68	1

^a The area covered by the Eastern Cape documentaries includes the transition zone between the SRZ and WRZ and the record is therefore used for both regions;

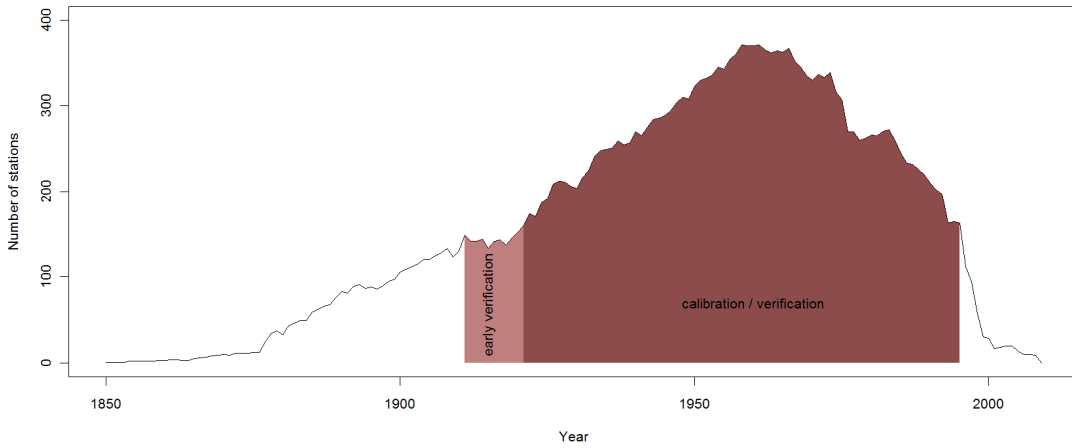


Fig. S3.1: Temporal evolution of rainfall station data coverage in southern Africa. The curve shows the number of rainfall stations within GHCNv2 (Peterson and Vose 1997), with measurements available for each year within the domain 10°S - 35°S / 10°E - 55°E . The period used for calibration/verification in our reconstruction is dark brown shaded (1921-1995, at least 160 stations with data available for each year), the early verification period is light brown shaded (1911-1920, 133-152 stations available)

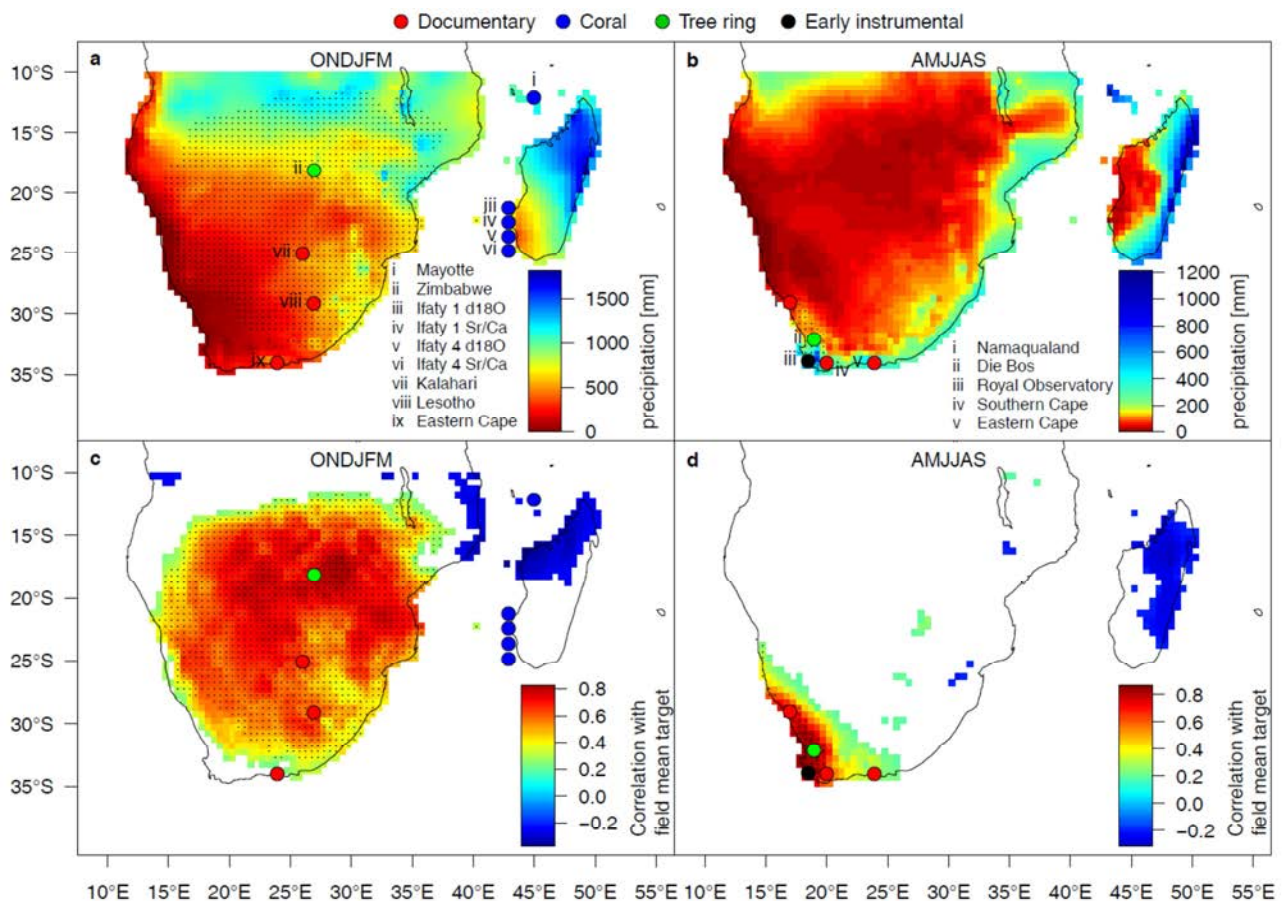


Fig. S3.2: Same as Fig. 1 in the main text, but showing only significant ($p < 0.05$) correlations in panels c and d.

S4. Description of reconstruction method

S4.1. Rainfall reconstruction

We used an ensemble-based nested principal component regression (PCR) to reconstruct southern African rainfall (Luterbacher et al. 2002; Neukom et al. 2010; Neukom et al. 2011). In this approach, principal components of the rainfall predictor data were regressed against the instrumental target over the calibration period to establish a transfer function. This transfer function was then applied to the proxy-PCs over their full temporal coverage to obtain reconstructed rainfall values. This step was repeated for each existing combination of proxy records during the reconstruction period (nested approach). Given that objective selection criteria for a number of reconstruction parameters are largely missing, we performed an ensemble of 3000 reconstructions with randomly chosen parameters for each ensemble member (Neukom et al. 2010). Randomisation was applied to the following reconstruction parameters for each member:

- One of the nine (five) proxies for SRZ (WRZ) rainfall was removed from the proxy nest;
- The percentage of total variance of the predictor matrix explained by the retained PCs was varied between 50% and 95% by varying the number of PCs included in the regression analysis;
- A calibration period of 35-50 (non-successive) years between 1921-1995 was selected, and the remaining 25-40 years used for verification;
- The weight of each proxy record in the PC analysis was scaled with a factor of 0.67 to 1.5;
- For the documentary proxy records, one of the 1000 realisations of pseudo-documentary data was chosen (see section S2).

Of the resulting 3000 reconstructions, we considered the ensemble mean the most probable outcome and used it as our best estimate reconstruction. Fig. S3.1 shows a workflow with all steps of the reconstruction.

S4.2. Reconstruction uncertainties

An advantage of the ensemble approach is that it makes it possible to quantify not only the traditional regression residual-based uncertainties, referred to as ‘calibration error’ (e.g. Cook and Kairiukstis 1990), but also the spread of the ensemble members generated from the random selection of the reconstruction parameters, described as the ‘ensemble error’. The total reconstruction confidence interval was defined as the combined calibration and ensemble standard error (SE), calculated as $SE = \sqrt{\sigma_{res}^2 + \sigma_{ens}^2}$ with σ_{res} denoting the standard deviation of the regression residuals (calibration error) and σ_{ens} the standard deviation of the ensemble members (ensemble error). Uncertainties of the filtered curves were calculated the same way using the residuals of the filtered data and standard deviation between the filtered ensemble members.

S4.3. Reconstruction skill

We assessed the reconstruction skill based on different variations of the reduction of error measure (RE; Cook et al. 1994). The RE is defined as $RE = 1 - \frac{\sum_{i=1}^n (x_{inst_i} - x_{recon_i})^2}{\sum_{i=1}^n (x_{inst_i} - x_{calib})^2}$, with x_{inst} denoting instrumental values and x_{recon} the reconstructed values for each year i of the verification period. x_{calib} is the calibration period mean of the instrumental data. The aim of the RE is to test whether the reconstruction has more predictive potential than the calibration period climatology (x_{calib}). If the reconstruction has skill, then its RE values lie between 0 and 1 (a hypothetically perfect reconstruction), while RE values <0 have no predictive skill. While we calculated an RE value for each ensemble member over its individual verification period (blue shading in Figs 3c and 4c), the skill of the ensemble mean, which we considered our best estimate reconstruction, was assessed using two variations of the RE.

Ensemble mean RE

The ensemble mean RE was calculated over the 1921-1995 period based on verification years only. This means that for each ensemble member, we used only the years that were withheld for verification and were therefore independent from the calibration period. Because the years used for calibration and verification were randomly chosen for each of the 3000 ensemble members and between 33% and 53% of the years were used for verification, for each year of the 1921-1995 period around 1000-1600 values were available for verification. These values were averaged to obtain a validation ensemble mean, which was then used together with the instrumental target data to calculate the ensemble mean RE (see above). This RE value is slightly different from the traditional RE statistic because the verification years (covered by x_{inst} and x_{recon}) and the calibration years (represented by x_{calib}) are both drawn from the same 1921-1990 period.

Early verification RE

To calculate the early verification RE, the ensemble mean reconstruction (in this case based on the full 3000 member ensemble) was verified against the instrumental target over the 1911-1920 pre-calibration period.

S5. Climate index and SST comparison plots

S5.1. Reconstruction vs. indices

In addition to the running correlations shown in Figs 3d (4d) in the main text, Fig. S5.1 (Fig. S5.2) shows a direct comparison of our SRZ (WRZ) reconstruction with the selected climate indices.

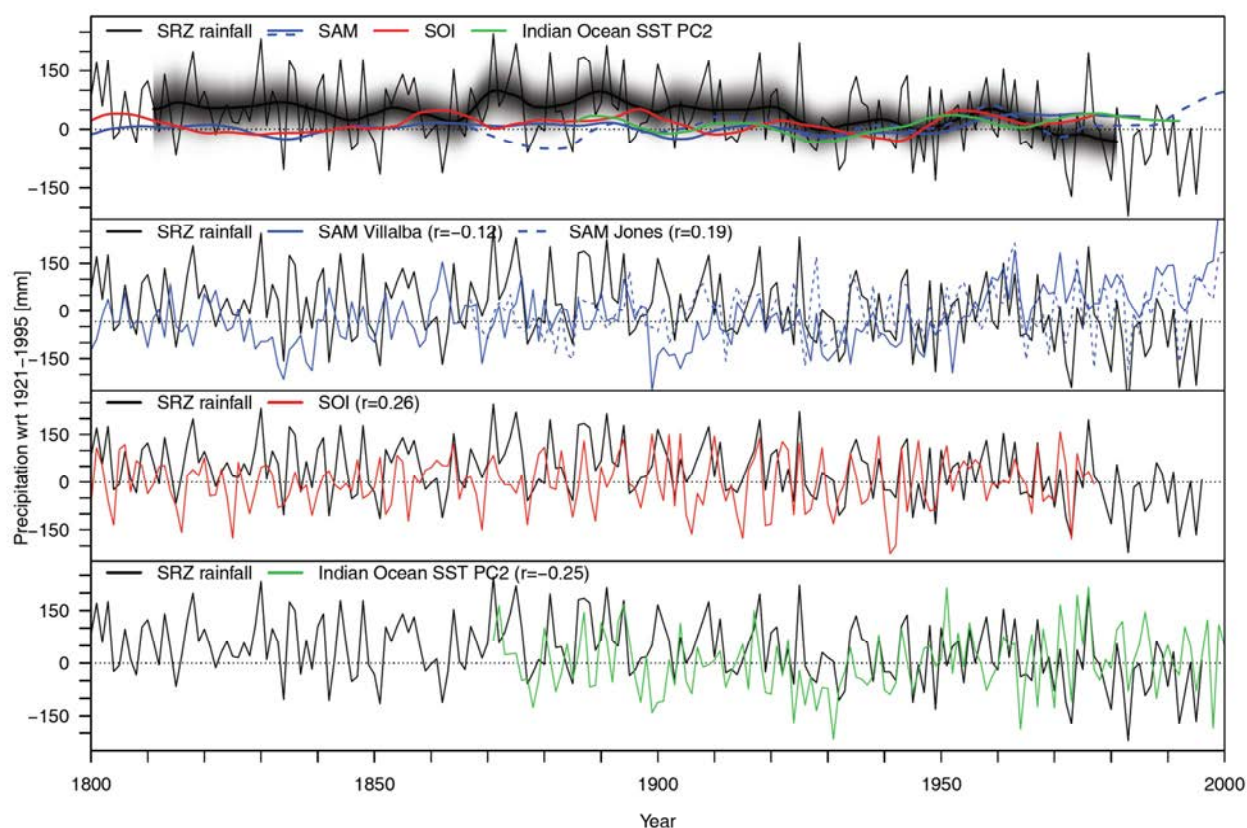


Fig. S5.1: Comparison of the SRZ rainfall reconstruction with independent climate indices. Top: interannual and 30-year loess-filtered reconstruction (black with shaded uncertainties) and 30-year filtered SAM reconstructions (solid blue; Villalba et al. 2012 and dashed blue; Jones et al. 2009), SOI reconstruction (red; Stahle et al. 1998) and Indian Ocean PC2 (green). The lower three panels compare the unfiltered reconstruction with the unfiltered indices (same colours as above, correlations over the full period of overlap are indicated in parentheses for each index). Data for the indices were scaled to the mean and standard deviation of the SRZ reconstruction over the period 1921-1995. Indian Ocean SST PC2 curves are inverted for clearer comparison

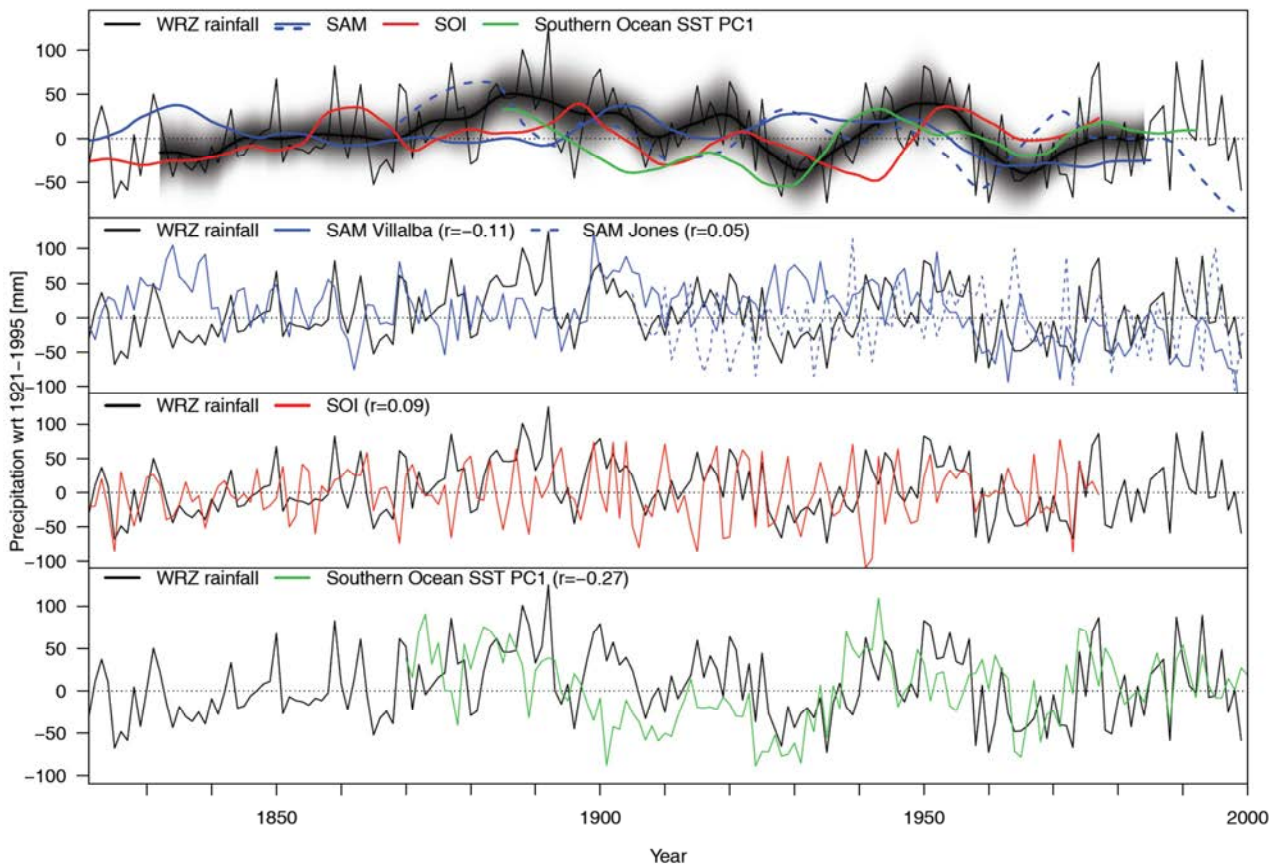


Fig. S5.2: Same as Fig. S5.1 but for the WRZ and Southern Ocean SST PC1 instead of Indian Ocean SST PC2. SAM and Southern Ocean SST PC1 curves are inverted for clearer comparison

S5.2. Running correlations regional SSTs vs. ENSO indices

Alternative running correlations to Figs. 3d and 4d in the main text, are provided in Figs. S5.3 and S5.4. They include also the instrumental and reconstructed NINO3.4 running correlations with SRZ and WRZ rainfall, respectively. Fig. S5.3 illustrates the very similar relation of SRZ rainfall to the SOI, NINO3.4 and Indian Ocean SSTs.

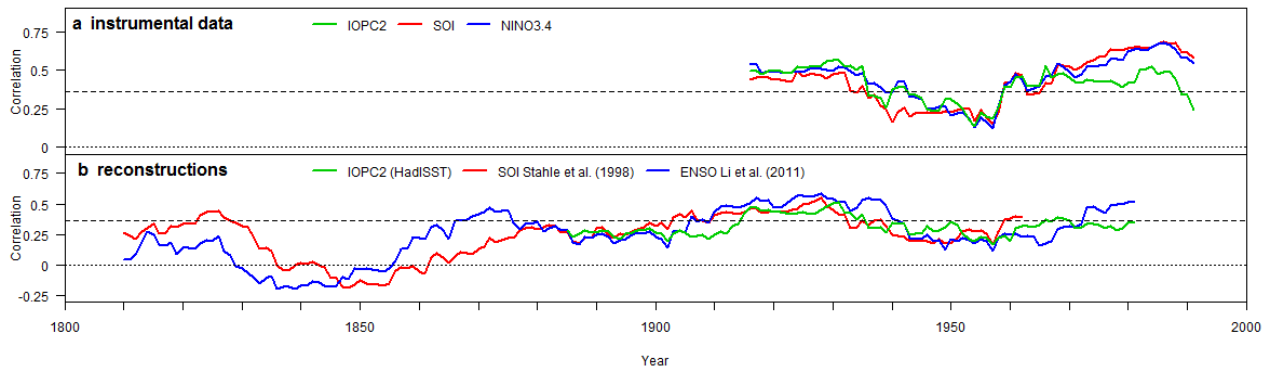


Fig. S5.3: 30-year running correlations between the SRZ rainfall and SST indices. a) Running correlations between instrumental SRZ (reconstruction target) and instrumental Indian Ocean SST PC2 (IOPC2; green), SOI (red) and NINO3.4 (blue) indices. b) Running correlations between reconstructed SRZ rainfall and reconstructed indices. For the IOPC2 correlations, instrumental (HadISST) data are used. NINO3.4 and IOPC2 correlations are inverted for improved comparison. The red and green lines in panel a (b) correspond to the solid (dotted) red and green lines in Fig. 3 in the main text.

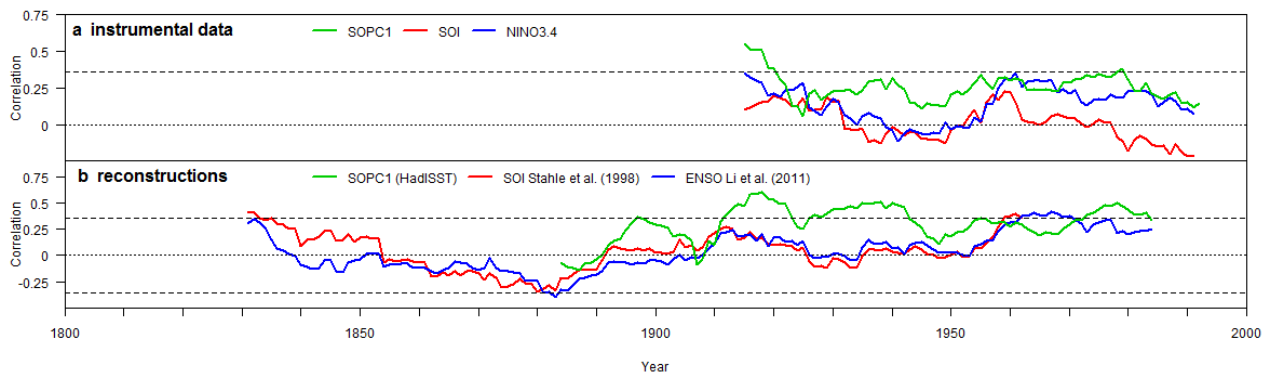


Fig. S5.4: Same as Fig. S5.3 but for the WRZ and using Southern Ocean PC1 (SOPC1) as regional SST index (green). NINO3.4 and SOPC1 correlations are inverted. The red and green lines in panel a (b) correspond to the solid (dotted) red and green lines in Fig. 4 in the main text.

S6. CLIWOC/ICOADS-based rainfall reconstructions

Prior to the reconstructions, the gridded CLIWOC/ICOADS wind speed data series were screened for suitability as rainfall predictors by correlating them against the seasonal reconstruction targets over the 1911-1995 period. Only grid cells with a maximum distance of 1000 km from the SRZ and WRZ areas were considered.

Table S6.1 lists the cells of the gridded wind speed data that are available in the southern African domain of the CLIWOC/ICOADS data set (Küttel et al. 2010) and their correlation with SRZ and WRZ rainfall. The cells that were selected for the reconstructions are indicated in Table S6.1 and shown in Fig. S6.1.

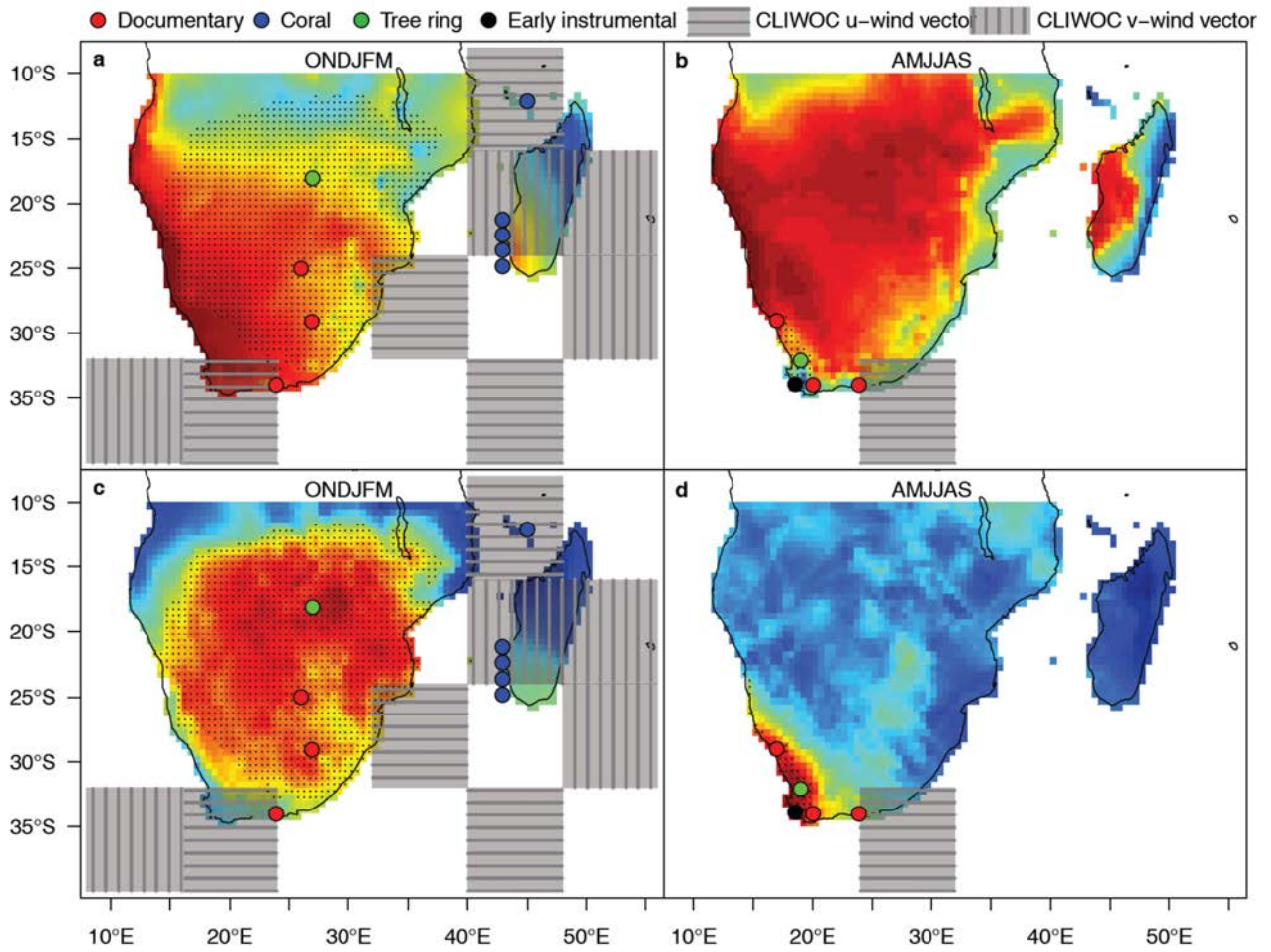


Fig. S6.1: Same as Fig. 1 in the main text but including the CLIWOC/ICOADS cells (Küttel et al. 2010) that were selected for the independent reconstruction

Table S6.1: Results of the proxy screening procedure for the SRZ (ONDJFM) and WRZ (AMJJAS) reconstructions for the CLIWOC/ICOADS data. Pearson correlations coefficients (*cor*) and an indicator of selection (*sel*) are provided for both seasons. ‘U’ (‘V’) denotes the zonal (meridional) direction of the wind vectors and the coordinates indicate the centre of the 8°x8° grid cells. No correlations are provided if the record is outside the search radius of 1000km (from the closest grid cell corresponding to the SRZ and WRZ respectively). Some of the grid cells surrounding southern Africa are not considered because they do not have sufficient data coverage in the respective season (Küttel et al. 2010). Note that the degrees of freedom for the correlations are dependent on the data coverage and lag-1 autocorrelation of each individual record.

Name	ONDJFM		AMJJAS		Name	ONDJFM		AMJJAS	
	<i>cor</i>	<i>sel</i>	<i>cor</i>	<i>sel</i>		<i>cor</i>	<i>sel</i>	<i>cor</i>	<i>sel</i>
CLIWOC U 4°E / 4°S	-0.16				CLIWOC U 60°E / 28°S	0.13			
CLIWOC U 12°E / 4°S	0.04				CLIWOC V 20°E / 28°S				-0.24
CLIWOC U 44°E / 4°S	0.29				CLIWOC V 28°E / 28°S	-0.14			0.12
CLIWOC V 12°E / 4°S	-0.15				CLIWOC V 36°E / 28°S	0.18			
CLIWOC V 44°E / 4°S	-0.13				CLIWOC V 44°E / 28°S	0.03			
CLIWOC U 12°E / 12°S	0.22		-0.26		CLIWOC V 52°E / 28°S	-0.22	1		
CLIWOC U 44°E / 12°S	-0.26	1			CLIWOC U 4°E / 36°S	-0.12			0.19
CLIWOC V 12°E / 12°S	-0.09		0.24		CLIWOC U 12°E / 36°S	-0.24			0.16
CLIWOC V 44°E / 12°S	0.12				CLIWOC U 20°E / 36°S	-0.22	1		0.30
CLIWOC U 4°E / 20°S	0.22				CLIWOC U 28°E / 36°S	0.09			0.29
CLIWOC U 36°E / 20°S	0.07				CLIWOC U 36°E / 36°S	0.06			1
CLIWOC U 44°E / 20°S	-0.16				CLIWOC U 44°E / 36°S	-0.23	1		
CLIWOC U 52°E / 20°S	-0.09				CLIWOC V 4°E / 36°S	0.09			-0.06
CLIWOC V 4°E / 20°S	-0.23				CLIWOC V 12°E / 36°S	0.27	1		0.15
CLIWOC V 36°E / 20°S	-0.12				CLIWOC V 20°E / 36°S	0.03			-0.17
CLIWOC V 44°E / 20°S	-0.37	1			CLIWOC V 28°E / 36°S	0.20			0.16
CLIWOC V 52°E / 20°S	-0.30	1			CLIWOC V 36°E / 36°S	0.14			
CLIWOC U 4°E / 28°S	-0.17		0.14		CLIWOC V 44°E / 36°S	0.20			
CLIWOC U 12°E / 28°S	-0.13		-0.08		CLIWOC U 12°E / 44°S				-0.27
CLIWOC U 20°E / 28°S			-0.08		CLIWOC U 20°E / 44°S	-0.15			-0.17
CLIWOC U 28°E / 28°S	-0.15		-0.15		CLIWOC U 28°E / 44°S	0.09			0.38
CLIWOC U 36°E / 28°S	0.28	1			CLIWOC U 36°E / 44°S	-0.13			
CLIWOC U 44°E / 28°S	0.09				CLIWOC V 12°E / 44°S				-0.21
CLIWOC U 52°E / 28°S	0.15				CLIWOC V 20°E / 44°S	0.19			0.39
CLIWOC V 4°E / 28°S	0.15		0.05		CLIWOC V 28°E / 44°S	0.25			0.30
CLIWOC V 12°E / 28°S	-0.07		-0.17		CLIWOC V 36°E / 44°S	0.11			

S7. References

- Brazdil R, Pfister C, Wanner H, Von Storch H, Luterbacher J (2005) Historical climatology in Europe - The state of the art. *Clim Change* 70 (3):363-430
- Cook E, Kairiukstis L (1990) *Methods of Dendrochronology*. Kluwer Academic Publishers, Dordrecht, The Netherlands
- Cook ER, Briffa KR, Jones PD (1994) Spatial regression methods in dendroclimatology –a review and comparision of two techniques. *Int J Climatol* 14 (4):379-402
- Dobrovolny P, Brazdil R, Valasek H, Kotyza O, Mackova J, Halickova M (2009) A standard paleoclimatological approach to temperature reconstruction in historical climatology: an example from the Czech Republic, AD 1718-2007. *Int J Climatol* 29 (10):1478-1492. doi:10.1002/joc.1789
- Jones J, Fogt R, Widmann M, Marshall G, Jones P, Visbeck M (2009) Historical SAM Variability. Part I: Century-Length Seasonal Reconstructions. *J Clim* 22 (20):5319-5345. doi:10.1175/2009JCLI2785.1
- Kittel M, Xoplaki E, Gallego D, Luterbacher J, Garc ía-Herrera R, Allan R, Barriendos M, Jones P, Wheeler D, Wanner H (2010) The importance of ship log data: reconstructing North Atlantic, European and Mediterranean sea level pressure fields back to 1750. *Clim Dynam* 34:1115-1128
- Luterbacher J, Xoplaki E, Dietrich D, Rickli R, Jacobeit J, Beck C, Gyalistras D, Schmutz C, Wanner H (2002) Reconstruction of sea level pressure fields over the Eastern North Atlantic and Europe back to 1500. *Clim Dynam* 18 (7):545-561
- Mann HB (1945) Nonparametric tests against trend. *Econometrica* 13:245-259
- Mann ME, Rutherford S, Wahl E, Ammann C (2007) Robustness of proxy-based climate field reconstruction methods. *J Geophys Res* 112 (D12):D12109. doi:10.1029/2006jd008272
- Neukom R, Prieto MD, Moyano R, Luterbacher J, Pfister C, Villalba R, Jones PD, Wanner H (2009) An extended network of documentary data from South America and its potential for

quantitative precipitation reconstructions back to the 16th century. *Geophys Res Lett* 36:L12703

Neukom R, Luterbacher J, Villalba R, Küttel M, Frank D, Jones PD, Grosjean M, Esper J, Lopez L, Wanner H (2010) Multi-centennial summer and winter precipitation variability in southern South America. *Geophys Res Lett* 37:L14708. doi:10.1029/2010GL043680

Neukom R, Luterbacher J, Villalba R, Kuttel M, Frank D, Jones PD, Grosjean M, Wanner H, Aravena JC, Black DE, Christie DA, D'Arrigo R, Lara A, Morales M, Soliz-Gamboa C, Srur A, Urrutia R, von Gunten L (2011) Multiproxy summer and winter surface air temperature field reconstructions for southern South America covering the past centuries. *Clim Dynam* 37 (1-2):35-51. doi:10.1007/s00382-010-0793-3

Peterson TC, Vose RS (1997) An overview of the global historical climatology network temperature database. *Bull Amer Meteorol Soc* 78 (12):2837-2849

Pfister C (1999) *Wetternachhersage. 500 Jahre Klimavariationen und Naturkatastrophen (1496-1995)*. Verlag Paul Haupt, Bern, Stuttgart, Wien

Pfister C, Luterbacher J, Wanner H, Wheeler D, Brázdil R, Ge Q, Hao Z, Moberg A, Grab S, Prieto MR (2008) Documentary evidence as climate proxies. Paper presented at the Proxy Uncertainty Workshop, Trieste, 9-11 June 2008

Stahle DW, D'Arrigo RD, Krusic PJ, Cleaveland MK, Cook ER, Allan RJ, Cole JE, Dunbar RB, Therrell MD, Gay DA, Moore MD, Stokes MA, Burns BT, Villanueva-Diaz J, Thompson LG (1998) Experimental dendroclimatic reconstruction of the Southern Oscillation. *Bull Amer Meteorol Soc* 79 (10):2137-2152

Villalba R, Lara A, Masiokas MH, Urrutia R, Luckman BH, Marshall GJ, Mundo IA, Christie DA, Cook ER, Neukom R, Allen K, Fenwick P, Boninsegna JA, Srur AM, Morales MS, Araneo D, Palmer JG, Cuq E, Aravena JC, Holz A, LeQuesne C (2012) Unusual Southern Hemisphere tree growth patterns induced by changes in the Southern Annular Mode. *Nat Geosci* doi:10.1038/ngeo1613