nature geoscience

ARTICLES

https://doi.org/10.1038/s41561-021-00819-9

ENSO-driven coupled megadroughts in North and South America over the last millennium

Nathan J. Steiger[1,2]✉, Jason E. Smerdon[1], Richard Seager[1], A. Park Williams[1,3] and Arianna M. Varuolo-Clarke[1]

Geological evidence from the last millennium indicates that multidecadal megadroughts may have occurred simultaneously in California and Patagonia at least once. However, it is unclear whether or not megadroughts were common in South America, whether or not simultaneous megadroughts in North and South America occurred repeatedly, and what would cause their simultaneous occurrence. Here we use a data-assimilation-based global hydroclimate reconstruction, which integrates palaeoclimate records with constraints from a climate model, to show that there were about a dozen megadroughts in the South American Southwest over the last millennium. Using dynamical variables from the hydroclimate reconstruction, we show that these megadroughts were driven by the El Niño/Southern Oscillation (ENSO). We also find that North American Southwest and South American Southwest megadroughts have occurred simultaneously more often than expected by chance. These coincident megadroughts were driven by an increased frequency of cold ENSO states relative to the last millennium-average frequency. Our results establish the substantial risk that exists for ENSO-driven, coupled megadroughts in two critical agricultural regions.

Over 15 years ago, tree-ring records revealed prolonged megadroughts that occurred repeatedly in western North America during the mediaeval period and that dwarfed any prolonged droughts observed during the 19th and 20th centuries[1,2]. These historical megadroughts profoundly impacted human and ecological communities and the landscape across western North America[3]. In the 1990s it was discovered that at least one mediaeval megadrought in California coincided with prolonged drought conditions in Patagonia[4]. It was proposed that these droughts—referred to as 'Stine droughts' after their discoverer—were connected through a decades-long polewards shift of the storm tracks in both hemispheres. Since that time, several climate drivers have been implicated in causing North American megadroughts, including modes of variability in the Pacific and Atlantic Oceans as well as historical radiative forcing[5–10]. In southwestern South America, Pacific Ocean variability has been linked to interannual fluctuations in precipitation and streamflow[11,12]. Both palaeo- and observation-based analyses suggest that the fluctuations in the interannual hydroclimates of North and South America can be partially explained by a common Pacific Ocean driver[13–17]. There is also evidence of multidecadal coherence between precipitation-sensitive proxies in central Chile and ENSO over the past 400 years[17]. Taken together, this evidence indicates the existence of an ENSO-driven, interhemispheric link between the hydroclimates of North and South America. Despite this progress in understanding the occurrence and causes of megadroughts in North America and shorter duration droughts in South America, little is known about the existence and characteristics of multidecadal megadroughts in South America, let alone their connection to North American megadroughts.

Understanding the possible connection between megadroughts in North and South America is critical given the substantial impacts of simultaneous severe droughts in both regions. The past two decades have, perhaps coincidentally, seen long-term drying develop in both the North American Southwest[18] and central Chile[19], which has had considerable economic and ecological impacts. Drying in the North American Southwest has reduced the flow of the Colorado River[20] (which irrigates approximately 22,000 km² of land and provides water for over 40 million people[21]) and has promoted rapid increases in ground-water extraction[22]. Drying has also contributed to several massive forest mortality events[23] and increased the area of forest fires throughout the region[24]. Chile is among the top ten agricultural exporters in the world, with central Chile being the economic heart of the country: it is home to approximately three-quarters of the country's population, contains over 75% of the country's total irrigated agriculture and has the majority of the country's water reservoirs[25,26]. During the most recent severe drought (2010 until now), central Chile has experienced precipitation deficits of up to 45% of the 1970–2000 mean[19,27]. This has resulted in substantial depletion of water resources for both agriculture and human consumption as well as an extension of forest-fire seasons[19,27].

Megadroughts in North and South America

Here we use a newly available reconstruction product to test the possibility and potential causes of coupled megadroughts in the Americas (we define megadroughts as severe droughts that last longer than a decade, see Methods for details). The Paleo Hydrodynamics Data Assimilation product (PHYDA) is an internally consistent reconstruction that provides, among other climate variables, globally gridded hydroclimate and sea surface temperatures (SSTs) over the Common Era. PHYDA combines information from both a global climate model and a global collection of proxy time series to produce a probabilistic reconstruction of the past 2,000 years (see Methods and ref. [28] for more information about PHYDA). We focus our analyses on two regions in North and South America: the North American Southwest (NASW) and the South American Southwest (SASW) (Methods, see Extended Data Fig. 1 for boundaries and Extended Data Fig. 2 for reconstruction verification). These two regions are important agricultural centres and

---

[1]Lamont-Doherty Earth Observatory, Columbia University, Palisades, NY, USA. [2]Institute of Earth Sciences, Hebrew University of Jerusalem, Jerusalem, Israel. [3]Department of Geography, University of California Los Angeles, Los Angeles, CA, USA. ✉e-mail: nsteiger@ldeo.columbia.edu

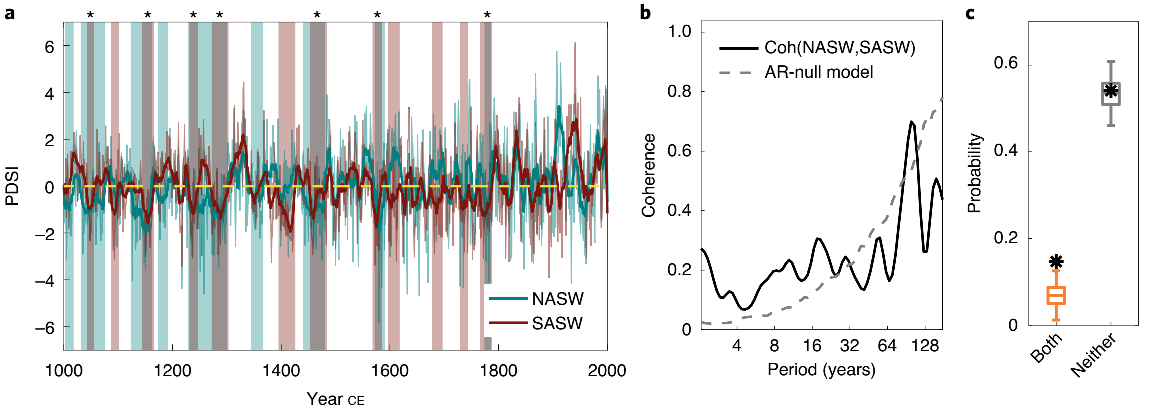


Fig. 1 | Megadrought in the NASW and the SASW. a, Reconstructed PDSI of the NASW and SASW for the summer season (June–August (JJA) for NASW; December–February (DJF) for SASW) from PHYDA[28] (note that the period of our analyses and the baseline of the plotted PDSI is 1000–1925 CE, the baseline is indicated by the dotted yellow line, see Methods). Lighter coloured time series are annually resolved while darker coloured time series are an 11-year lowpass filter (moving average), corresponding to the timescale used in the megadrought definition (Methods). Megadrought periods for each region are highlighted in green (NASW) and maroon (SASW), with the overlap represented by the darkest shading and an asterisk symbol above the plot. **b**, Spectral coherence between the NASW and the SASW time series shown in **a** along with a Monte Carlo-based AR-null (significance level, $\alpha = 0.05$)[47]. Coh, coherence. **c**, Black asterisks show the probability of the occurrence of 'both' NASW and SASW experiencing a megadrought and for 'neither' experiencing a megadrought (where the probability is the number of years experiencing both or neither conditions divided by the total number of years). Box plots show the distribution of corresponding probabilities for a best-fit AR-null model (Methods); the central box mark of the distributions indicates the median, the bottom and top edges of the box indicate the 25th and 75th percentiles, respectively, and the whiskers extend to the most extreme data points not considered outliers, approximately $\pm 2.7\sigma$.

they are climatologically similar: both regions receive the majority of their precipitation in the cold season (which supplies the soil moisture for the following growing season), and both regions contain semi-arid lands and the only Mediterranean-type climate zones in the Americas[29]. These two regions have also been observed to experience large variability in long-term hydroclimate throughout the Holocene[3,4,11,19,30].

We find that megadroughts occurred repeatedly in the SASW (Fig. 1a) and that they often co-occurred with megadroughts in the NASW (Fig. 1a). Using a specific megadrought metric (Methods), we find nine NASW and 12 SASW events, with seven events occurring simultaneously in the two regions (one NASW drought spans two SASW droughts). NASW droughts tend to be longer than SASW droughts, with a mean length of 31.8 years compared with 21.2 years; this additional length in NASW droughts also makes them more severe, with the average sum of the Palmer drought severity index (PDSI) over the droughts being −25.7 PDSI compared with −15.0 PDSI for the SASW. Time series of the averaged summer PDSI[31] in these two regions show an unusual degree of coherence at interannual to decadal timescales (Fig. 1b). Among the seven identified periods of coincident drought, both the NASW and the SASW experienced megadroughts during the classic Stine megadrought periods (Fig. 1a): circa 1050–1200 CE and 1250–1400 CE (the approximate, wide ranges account for the dating uncertainties of the relic tree stumps presented in ref. [4]). Although ref. [4] identified only the first of these events as a period of simultaneous megadrought in California and southern Patagonia, we find evidence for much more frequent simultaneous megadroughts in the NASW and the SASW. The association between megadroughts in the two regions is not by chance: the occurrence of simultaneous events exceeds the frequency that would be expected in two unrelated time series with the same auto-regressive (AR) statistical properties. Specifically, the co-occurrence exceeds 2.7σ of the AR-null distribution (Fig. 1c). Moreover, the variability and drought periods in the SASW PDSI time series reflect the local proxy data: they are not imparted by North American proxy sources (Extended Data Fig. 3) nor by the spatial covariance information used by the PHYDA reconstruction methodology (Extended Data Figs. 4 and 5). Relatively continuous historical documentary evidence overlaps with the last of our identified megadroughts in the SASW. This evidence indicates a severe drought in central Chile from 1769 to 1784 CE (refs. [32,33]), approximately coincident with the megadrought we identify in our proxy-based reconstructions from 1773 to 1785 CE.

Causes of coupled megadroughts

Composites of global PHYDA reconstructions of SSTs and the PDSI for all megadroughts in the NASW and all megadroughts in the SASW show a cold Pacific SST pattern similar to La Niña (Fig. 2). Consistent with previous work[10], the NASW is associated with both a La Niña-like Pacific and a warm North Atlantic (basin-wide mostly positive SST anomalies). SASW megadroughts lack the warm North Atlantic and are characteristic of a typical La Niña pattern through the tropical Atlantic. Consistently, the years in which neither location experienced a megadrought are characteristic of El Niño conditions in the Pacific and neutral conditions in the Atlantic (Fig. 2). When there are megadrought conditions in both regions, the global SST pattern is an enhanced version of the SASW megadrought conditions, typical of a stronger La Niña pattern (Fig. 2). These results hold even if SASW megadrought periods are determined using an alternative reconstruction of the SASW PDSI that relies only on the local proxy data (Extended Data Fig. 6); this confirms that the SST–megadrought associations are not imposed through a far-field influence of the large North American proxy network on the PHYDA reconstruction over South America.

The ENSO-dominant result established in Fig. 2 is also supported by the full probabilistic ensemble of the NINO3.4 index from PHYDA: individual megadroughts in either the NASW or the SASW correspond with colder NINO3.4 conditions, the lack of megadrought in either location corresponds with a warm NINO3.4 shift, and simultaneous megadroughts in both regions occur with a large negative shift in the NINO3.4 index (Fig. 3a). An increased frequency of cold Pacific SST conditions may therefore partially explain megadroughts and their co-occurrence.

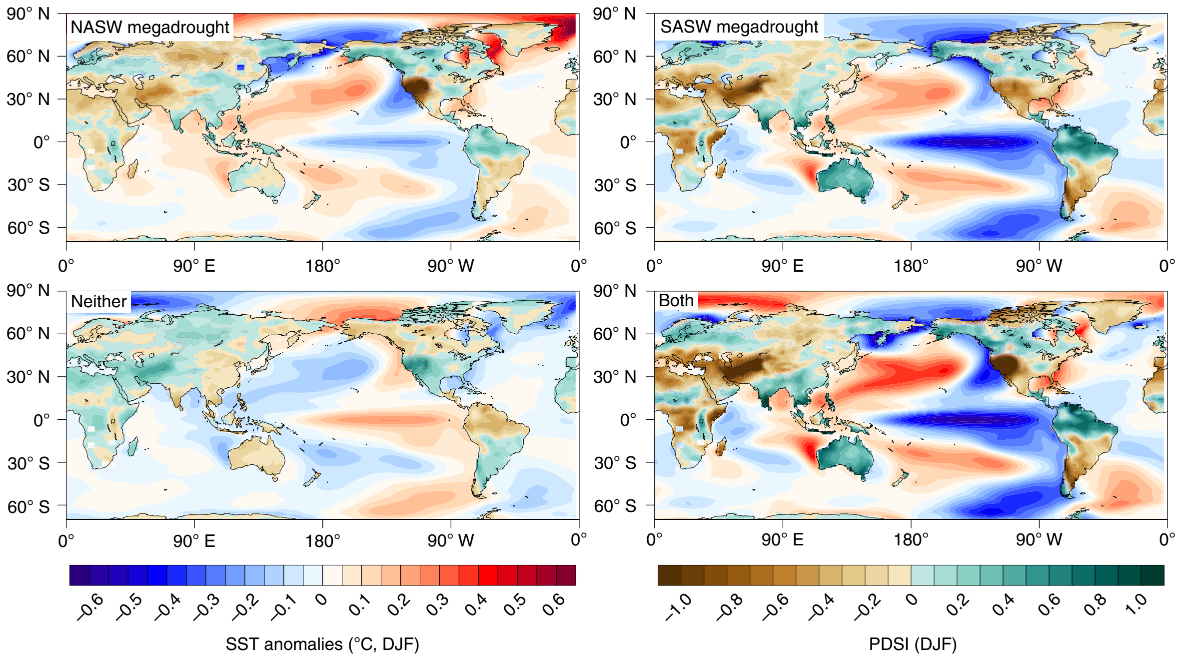


Fig. 2 | Mean climate states during megadrought and non-megadrought years. Composites of DJF SST and PDSI from PHYDA for all years corresponding to NASW megadrought conditions, SASW megadrought conditions, when megadroughts do not exist in either location (Neither) and when megadroughts exist in both locations (Both). Temperature and PDSI data are anomalies with respect to the analysis period 1000–1925 CE.

Using self-organizing map (SOM) analysis[34], we partition PHYDA's DJF Pacific SST reconstruction into its fundamental flavours, similar to the observations-based analysis of ref. [35] and previous PHYDA analyses[10,36] (Methods). We find that when megadroughts occur in both the NASW and the SASW, there is a robust decrease in warm, El Niño-like ENSO flavours (Fig. 3b, patterns 1–4). There is also an unusual increase in the frequency of flavours of ENSO that have cold SST anomalies: in Fig. 3b, pattern numbers 5 and 7 increase 40% and 138%. These pattern increases exceed the 97.5th percentile of a Markov chain Monte Carlo-based null distribution (Methods). These results are robust to the choice of the number of SOM patterns (Extended Data Fig. 7). While the La Niña-like flavours of ENSO with a broad latitudinal extent may be connected to longer timescale expressions of ENSO, such as the Pacific Decadal Oscillation or the Interdecadal Pacific Oscillation[37], these ENSO flavours are also present in interannual analyses[35].

To assess the relative importance of ocean-dynamical factors and radiative forcing in explaining the hydroclimate in these regions, a decadal-scale linear regression analysis is performed over the period 1000–1925 CE using detrended indices (Methods). Consistent with ref. [10], we find that the hydroclimate in the NASW is driven most strongly by ENSO, with the Atlantic Multidecadal Oscillation (AMO) index and radiative forcing explaining less than half the NASW PDSI variance explained by NINO3.4 (Fig. 4a). Including NINO3.4, AMO and forcing as predictors, however, improves the regression model relative to NINO3.4 alone (Fig. 4a; the distribution of adjusted the coefficient of determination, r^2, for 'N + A + F' is distinct from the distribution for 'NINO3.4', via the Kolmogorov–Smirnov test with distribution size $n=999$ and $\alpha=0.05$). In contrast to these NASW results, the NINO3.4 index has r^2 values that are approximately two times higher for the SASW (Fig. 4a). In addition, there is no discernible impact of radiative forcing on the decadal-scale PDSI in the SASW and it does not improve the r^2 values when included with the NINO3.4 index (Fig. 4a; Kolmogorov–Smirnov test with $n=999$ and $\alpha=0.05$). This result holds if the radiative forcing model (see Methods) is driven solely by volcanic events in the Southern Hemisphere (and thus more local to the SASW).

We do not find conclusive evidence for whether or not radiative forcing influenced the simultaneous occurrence of megadroughts in the NASW and the SASW (Fig. 4b), despite the fact that radiative forcing has been shown to influence the occurrence of NASW megadroughts[10]. Among the seven simultaneous megadrought events, three experience predominantly positive forcing anomalies, two experience predominantly negative forcing anomalies, one experiences average forcing anomalies and the single remainder experiences a mix of positive and negative forcing anomalies (Fig. 4b). This assessment is hindered in large part by the relatively small number of simultaneous megadrought events, making a more sophisticated analysis (for example, logistic regression as in ref. [10]) impossible.

All our analyses rely on palaeoclimate reconstructions, which are estimates of the past climate based on sparse and noisy measurements from climatic proxies. These endemic uncertainties necessitate the probabilistic analyses that we have presented herein. In addition, although PHYDA benefits from the dynamical information of a global climate model, it is also constrained by that model. While PHYDA's input variables have undergone bias correction[28] and PHYDA itself (including its ENSO covariances and spectra) has been shown to validate well in multiple contexts[10,28,36], it likely still inherits some of the fundamental inaccuracies of the climate model that was used in its construction[38]. Such inescapable features of data-assimilation products therefore require diligence and caution regarding interpretation, hence the additional validation exercises that we have provided herein. Moreover, documentary data[32,33] and the new South American Drought Atlas[39] allow for some independent validations of PHDYA, as discussed previously for our proxy-identified drought from 1773–1785 CE. However, documentary data are not available before the 16th century and the South American Drought Atlas is only available back to 1400 CE, leaving the critical mediaeval megadrought period unsampled by these products.

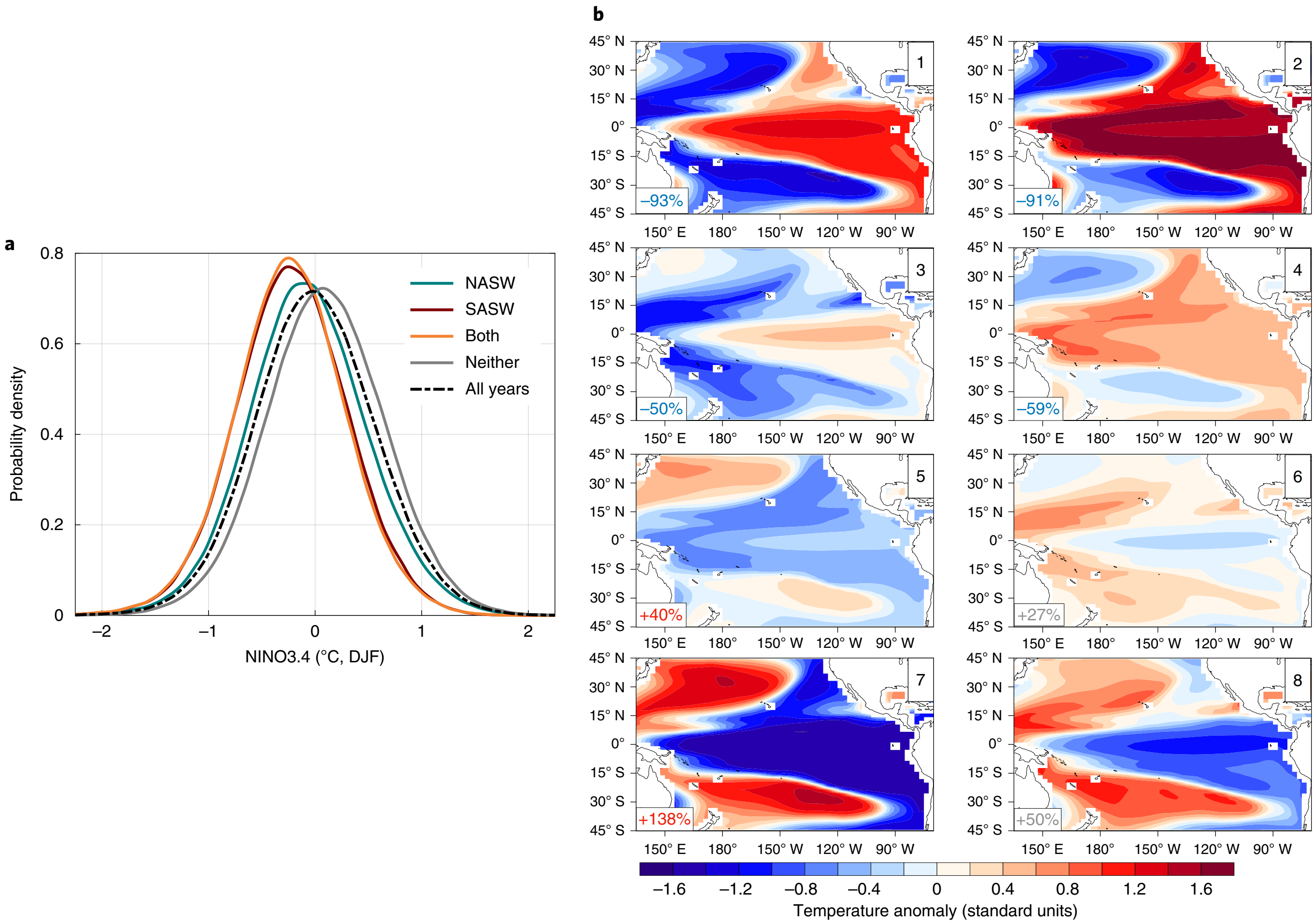


Fig. 3 | Frequency changes in tropical SST conditions. a, Probability density functions for the NINO3.4 index from PHYDA, including the full probabilistic ensemble, for periods of NASW megadrought, SASW megadrought, neither location in megadrought, both locations in megadrought and a reference distribution for all years of the analysis, 1000–1925 CE. **b**, SOM patterns based on detrended, standardized DJF SST data from PHYDA for the years 1000–1925 CE (see Methods). Boxes in the lower left corner of each panel show the percentage change in the frequency of occurrence of that pattern during simultaneous megadrought years relative to the average frequency over 1000–1925 CE; grey-coloured frequency changes indicate those percentage changes that fall within the 2.5th and 97.5th percentiles of a Markov chain Monte Carlo null distribution for each pattern (see Methods).

A last-millennium perspective

Despite the existing caveats, the global and last-millennium perspective provided by PHYDA has allowed a more definitive assessment of the existence and causes of megadroughts in the SASW than has previously been possible. Unlike their counterparts in the NASW, which are also influenced by Atlantic SSTs and radiative forcing, past megadroughts in the SASW appear to have been driven by cold Pacific SSTs[10]; ENSO is approximately twice as important a driver for the SASW as it is for the NASW. The lack of these additional forcings may underlie the slightly reduced persistence and severity of the SASW megadroughts relative to those in the NASW.

This work has also shown that megadroughts have appeared simultaneously in the NASW and the SASW throughout the past millennium. The co-occurrence of these events cannot be explained by chance, by shared proxy information or as an artefact of the reconstruction methodology, but instead probably have physical origins. In the pioneering work of ref. [4] it was suggested that a single simultaneous megadrought period was driven by a persistent contraction of both polar vortices that would directly decrease precipitation delivery to California and purportedly decrease precipitation in Patagonia through an enhanced rain-shadow effect. Contrary to this suggestion, we find that the more likely cause of multiple simultaneous megadroughts was an increase in the frequency of cold Pacific SSTs. These cold tropical Pacific Ocean states shift the jet streams and storm tracks polewards over both North and South America, thus inducing the coincident droughts in the NASW and the SASW[40].

Our results suggest that radiative forcing did not play a major role in past simultaneous megadroughts. However, because there is evidence that past NASW megadroughts were partially explained by a positive radiative forcing (Fig. 4a and ref. [10]), it is reasonable to expect that forcing may have some role in increasing the probability of their co-occurrence with SASW megadroughts. The causes of past megadroughts also provide only a partial insight into current and future events, given the large magnitude of contemporary radiative forcing. Indeed, both the NASW and central Chile have experienced severe and prolonged droughts over the past decade or longer, both of which have been explained partially by radiative forcing and partially by unforced atmosphere–ocean dynamics[18,27]. For both droughts, SST anomaly patterns distinct from ENSO and including warming in the western Pacific have been invoked as possible drivers[27,41] but it is not clear whether these patterns arise from natural variability or as a forced response to rising greenhouse gases[29]. Either way, the large contemporary radiative forcing can in part explain why drought has persisted even without a dominant La Niña driver over the same time period[27,29].

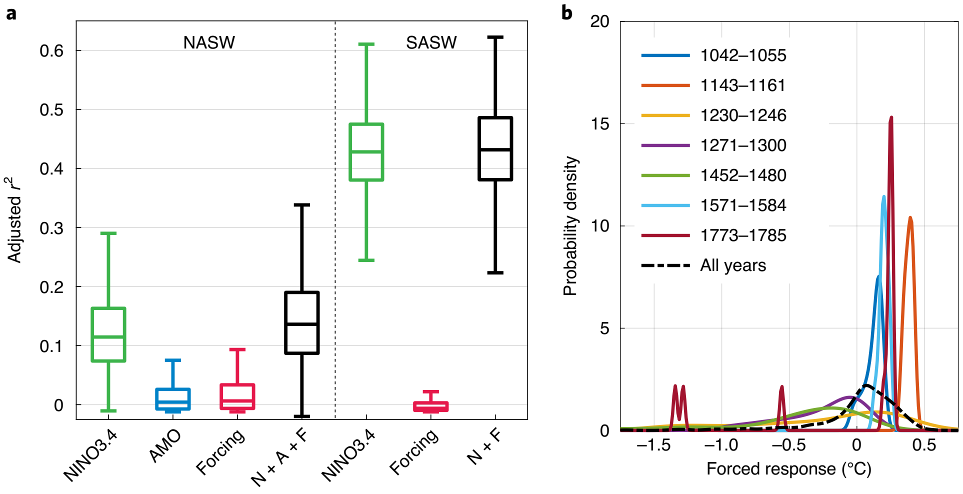


Fig. 4 | Hydroclimate dynamics and forcing analysis. **a**, The bootstrap distributions of adjusted r^2 values for the decadal linear regression of the predictands of the NASW PDSI and SASW PDSI and the predictors of NINO3.4 (or N), AMO (or A) and total radiative forcing (Forcing or F) from ref. [10]; see Methods. The central box mark of the distributions indicates the median, the bottom and top edges of the box indicate the 25th and 75th percentiles, respectively, and the whiskers extend to the most extreme data points not considered to be outliers, approximately $\pm 2.7\sigma$. **b**, Probability density functions of the estimated total forced response (the same underlying radiative forcing data as in **a**) for each of the simultaneous megadrought periods ($n = 7$; dates in years CE) compared against all years of the analysis, 1000–1925 CE. Spikes in the probability density functions indicate a narrow range of forcing values during a particular drought period.

Implications for future climate

Our results imply that ENSO variability has the capacity to induce megadroughts in both North and South America at any time in the future. It may also be that global warming could make coupled megadroughts more frequent, given the known influence of greenhouse-gas-driven drying in both areas[27,42–45]. Projecting future changes in these correlated extremes is nevertheless complicated by the fact that state-of-the-art global climate models have difficulty reproducing important features of ENSO[46] and appear to disagree with observations about the response of the tropical Pacific to historical forcings[42]. It is therefore essential that we improve our understanding and modelling of the tropical Pacific in response to increasing radiative forcing. This will improve our capacity to predict and to prepare for simultaneous megadroughts in the critical agricultural centres of the NASW and the SASW.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41561-021-00819-9.

Received: 13 April 2020; Accepted: 19 July 2021;

References

1. Cook, E. R., Woodhouse, C. A., Eakin, C. M., Meko, D. M. & Stahle, D. W. Long-term aridity changes in the western United States. *Science* **306**, 1015–1018 (2004).
2. Meko, D. M. et al. Medieval drought in the upper Colorado River Basin. *Geophys. Res. Lett.* **34**, L10705 (2007).
3. Cook, B. I. et al. North American megadroughts in the Common Era: reconstructions and simulations. *Wiley Interdiscip. Rev. Clim. Change* **7**, 411–432 (2016).
4. Stine, S. Extreme and persistent drought in California and Patagonia during mediaeval time. *Nature* **369**, 546–549 (1994).
5. Seager, R. et al. Tropical Pacific forcing of North American medieval megadroughts: testing the concept with an atmosphere model forced by coral-reconstructed SSTs. *J. Clim.* **21**, 6175–6190 (2008).
6. Burgman, R., Richard, S., Amy, C. & Celine, H. Role of tropical Pacific SSTs in global medieval hydroclimate: a modeling study. *Geophys. Res. Lett.* **37**, L06705 (2010).
7. Hunt, B. G. Global characteristics of pluvial and dry multi-year episodes, with emphasis on megadroughts. *Int. J. Climatol.* **31**, 1425–1439 (2011).
8. Stevenson, S., Timmermann, A., Chikamoto, Y., Langford, S. & DiNezio, P. Stochastically generated North American megadroughts. *J. Clim.* **28**, 1865–1880 (2015).
9. Coats, S., Smerdon, J. E., Karnauskas, K. B. & Seager, R. The improbable but unexceptional occurrence of megadrought clustering in the American West during the Medieval Climate Anomaly. *Environ. Res. Lett.* **11**, 074025 (2016).
10. Steiger, N. J. et al. Oceanic and radiative forcing of medieval megadroughts in the American Southwest. *Sci. Adv.* **5**, eaax0087 (2019).
11. Le Quesne, C. et al. Ancient *Austrocedrus* tree-ring chronologies used to reconstruct central Chile precipitation variability from AD 1200 to 2000. *J. Clim.* **19**, 5731–5744 (2006).
12. Urrutia, R. B. et al. Multicentury tree ring reconstruction of annual streamflow for the Maule River watershed in south central Chile. *Water Resour. Res.* **47**, W06527 (2011).
13. Evans, M. N. et al. Support for tropically-driven pacific decadal variability based on paleoproxy evidence. *Geophys. Res. Lett.* **28**, 3689–3692 (2001).
14. Villalba, R., D'Arrigo, R. D., Cook, E. R., Jacoby, G. C. & Wiles, G. in *Interhemispheric Climate Linkages* (ed. Markgraf, V.) 155–172 (Academic, 2001).
15. Seager, R. et al. Mechanisms of ENSO-forcing of hemispherically symmetric precipitation variability. *Q. J. R. Meteorol. Soc.* **131**, 1501–1527 (2005).
16. Seager, R. et al. Blueprints for Medieval hydroclimate. *Quat. Sci. Rev.* **26**, 2322 – 2336 (2007).
17. Villalba, R. et al. in *Dendroclimatology: Progress and Prospects* Vol. 11 (eds Hughes M. K. et al.) 175–227 (Springer, 2011).
18. Williams, A. P. et al. Large contribution from anthropogenic warming to an emerging North American megadrought. *Science* **368**, 314–318 (2020).
19. Garreaud, R. et al. The 2010–2015 megadrought in central Chile: impacts on regional hydroclimate and vegetation. *Hydrol. Earth Syst. Sci.* **21**, 6307–6327 (2017).
20. Xiao, M., Udall, B. & Lettenmaier, D. P. On the causes of declining Colorado River streamflows. *Water Resour. Res.* **54**, 6739–6756 (2018).
21. Bureau of Reclamation *Colorado River Basin Water Supply and Demand Study* (US Department of the Interior, 2012).
22. Faunt, C. C., Sneed, M., Traum, J. & Brandt, J. T. Water availability and land subsidence in the Central Valley, California, USA. *Hydrogeol. J.* **24**, 675–684 (2016).
23. Williams, A. P. et al. Temperature as a potent driver of regional forest drought stress and tree mortality. *Nat. Clim. Change* **3**, 292–297 (2013).
24. Abatzoglou, J. T. & Williams, A. P. Impact of anthropogenic climate change on wildfire across western US forests. *Proc. Natl Acad. Sci. USA* **113**, 11770–11775 (2016).

25. Masiokas, M., Villalba, R., Luckman, B., LeQuesne, C. & Aravena, J. C. Snowpack variations in the central Andes of Argentina and Chile, 1951–2005: large-scale atmospheric influences and implications for water resources in the region. *J. Clim.* **19**, 6334–6352 (2006).
26. Bozkurt, D., Rojas, M., Boisier, J. P. & Valdivieso, J. Projected hydroclimate changes over Andean basins in central Chile from downscaled CMIP5 models under the low and high emission scenarios. *Clim. Change* **150**, 131–147 (2018).
27. Garreaud, R. et al. The central Chile mega drought (2010–2018): a climate dynamics perspective. *Int. J. Climatol.* **40**, 421–439 (2019).
28. Steiger, N. J., Smerdon, J. E., Cook, E. R. & Cook, B. I. A reconstruction of global hydroclimate and dynamical variables over the Common Era. *Sci. Data* **5**, 180086 (2018).
29. Seager, R. et al. Climate variability and vhange of Mediterranean-type climates. *J. Clim.* **32**, 2887–2915 (2019).
30. Jenny, B. et al. Early to mid-Holocene aridity in central Chile and the southern westerlies: the Laguna Aculeo record (34° S). *Quat. Res.* **58**, 160–170 (2002).
31. Palmer, C. W. *Meteorological Drought* (US Department of Commerce, 1965).
32. Vicuña-Mackenna, B. *Ensayo Histórico Sobre el Clima de Chile: Desde los Tiempos Prehistóricos Hasta el Gran Temporal de Julio de 1877* (Imprenta del Mercurio, 1877).
33. Jana, P., Torrejón, F., Araneda, A. & Stehr, A. Drought periods during 18th century in central Chile (33° S): a historical reconstruction perspective revisiting Vicuña Mackenna's work. *Int. J. Climatol.* **39**, 1748–1755 (2019).
34. Kohonen, T. The self-organizing map. *Neurocomputing* **21**, 1–6 (1998).
35. Johnson, N. C., Feldstein, S. B. & Tremblay, B. The continuum of Northern Hemisphere teleconnection patterns and a description of the NAO shift with the use of self-organizing maps. *J. Clim.* **21**, 6354–6371 (2008).
36. Baek, S. H., Steiger, N. J., Smerdon, J. E. & Seager, R. Oceanic drivers of widespread summer droughts in the United States over the Common Era. *Geophys. Res. Lett.* **46**, 8271–8280 (2019).
37. Newman, M. et al. The Pacific Decadal Oscillation, revisited. *J. Clim.* **29**, 4399–4427 (2016).
38. Otto-Bliesner, B. L. et al. Climate variability and change since 850 CE: an ensemble approach with the Community Earth System Model. *Bull. Am. Meteorol. Soc.* **97**, 735–754 (2016).
39. Morales, M. S. et al. Six hundred years of South American tree rings reveal an increase in severe hydroclimatic events since mid-20th century. *Proc. Natl Acad. Sci. USA* **117**, 16816–16823 (2020).
40. Seager, R., Harnik, N., Kushnir, Y., Robinson, W. & Miller, J. Mechanisms of hemispherically symmetric climate variability. *J. Clim.* **16**, 2960–2978 (2003).
41. Seager, R. et al. Causes of the 2011–14 California drought. *J. Clim.* **28**, 6997–7024 (2015).
42. Seager, R. et al. Strengthening tropical Pacific zonal sea surface temperature gradient consistent with rising greenhouse gases. *Nat. Clim. Change* **9**, 517–522 (2019).
43. Cook, B. I., Ault, T. R. & Smerdon, J. E. Unprecedented 21st century drought risk in the American Southwest and Central Plains. *Sci. Adv.* **1**, e1400082 (2015).
44. Ault, T. R., Mankin, J. S., Cook, B. I. & Smerdon, J. E. Relative impacts of mitigation, temperature, and precipitation on 21st-century megadrought risk in the American Southwest. *Sci. Adv.* **2**, e1600873 (2016).
45. Cook, B. I. et al. Twenty-first century drought projections in the CMIP6 forcing scenarios. *Earths Future* **8**, e2019EF001461 (2020).
46. Bellenger, H., Guilyardi, E., Leloup, J., Lengaigne, M. & Vialard, J. ENSO representation in climate models: from CMIP3 to CMIP5. *Clim. Dyn.* **42**, 1999–2018 (2014).
47. Grinsted, A., Moore, J. C. & Jevrejeva, S. Application of the cross wavelet transform and wavelet coherence to geophysical time series. *Nonlinear Process. Geophys.* **11**, 561–566 (2004).

Methods

Paleo Hydrodynamics Data Assimilation product. We use PHYDA for the basis of this study[28,48]. PHYDA was constructed using palaeoclimate data assimilation, a reconstruction method that probabilistically fuses proxy information with the dynamical constraints of climate models; this method simultaneously estimates both hydroclimate fields and corresponding atmosphere–ocean states[49]. PHYDA incorporates information from a network of 2,978 annually resolved proxy-data time series[50] together with the Community Earth System Model (CESM) Last Millennium Ensemble of climate model simulations[38]; as a data-assimilation product, PHYDA therefore represents a statistically optimal amalgamation of both model and proxy information (see ref. [28] for details about PHYDA). PHYDA covers the years 1–2000 CE, targeting three different temporal windows: the boreal growing season of JJA, the austral growing season of DJF and the April to the next calendar year March annual means.

PHYDA has been extensively validated against observational data in addition to being compared against other palaeoclimate reconstructions[10,28]. It has also been shown that the ENSO and other ocean dynamics of PHYDA are not merely a reflection of the estimated background, that is, the CESM climate model, but have power spectra that are closer to observations than to the CESM simulation used as the basis of PHYDA[10]. Additionally, it has been verified that PHYDA has realistic ENSO–hydroclimate teleconnections, and it has been successfully used to elucidate the dynamical causes of pancontinental droughts in North America[36] and megadroughts in the NASW region[10]. Extended Data Figure 2 shows a specific verification of PHYDA's DJF reconstruction of PDSI over the central SASW region.

For the megadrought specifications, time series analyses and regression analyses, we use summer PDSI data for both regions, JJA averages in the NASW and DJF averages in the SASW. The summer dry season reflects the time when droughts are most acute. However, we note that both regions have similar hydroclimatologies and receive the majority of their precipitation in the preceding winter. Both the tree-ring data and the time-lagged PDSI variable therefore incorporate information about both the summer growing season and the amount of winter precipitation; this has been shown explicitly for the NASW[51,52] and is expected to apply to the SASW given the climatological conditions (Extended Data Fig. 1) and similarities to the NASW.

For the megadrought composites (Fig. 2 and Extended Data Fig. 6), we use the DJF SST anomalies (the peak season of ENSO) and the DJF PDSI globally. PHYDA has been shown to skillfully reconstruct both of these seasonal variables[28] and has lower uncertainties through the last millennium for the Pacific DJF SSTs than the JJA or annual SSTs. Despite the complexities of seasonal SST associations with annual droughts[53], these complexities appear negligible at the decadal timescale analysed herein: the composite results are very similar for other seasonal choices (for example, the JJA or annual mean).

Analysis period. We limit the beginning of the analysis to the year 1000 CE because of the relative dearth of South American proxy information before that time[28]. As in ref. [10], we end the analysis in 1925 CE to exclude the period of most pronounced anthropogenic warming that is likely to introduce complicating trends in the data.

North American Southwest region. Following previous work[10], the NASW region is defined as land in the regional box bounded by 125–105° W and 31–42° N, which wholly includes the US states of California, Nevada, Utah and Arizona, a southern portion of Wyoming, and both Colorado and New Mexico west of the Rocky Mountains. This regional box also includes some of the northernmost portions of Mexico where the international border deviates from a straight line at 31° N. For nearly all of this region, the majority of the precipitation falls during the cold season.

South American Southwest region. We use the South American land region bounded by 75–69° W and 26–45° S, excluding the size of a single PHYDA grid box near 28° S at the northeastern corner of the region encompassing part of the Chilean Altiplano (see Extended Data Fig. 1). The boxed region contains all of central Chile and a section of western Argentina. The boundaries were selected based on two factors: the precipitation climatology[54] and the proxy availability. The eastern boundary of this region follows the sharp gradient in the precipitation climatology that divides areas with a majority of JJA versus DJF precipitation (Extended Data Fig. 1a,b). The eastern and southern boundaries enclose a region with a majority JJA precipitation signal (Extended Data Fig. 1a); like the NASW, this corresponds to a cold-season-dominated precipitation regime. This bounded region also contains a high density of proxy locations used in PHYDA[28]. These proxy locations covary strongly with other locations in the regional box, with the DJF PDSI at proxy locations correlating at $r^2 > 0.6$ with other locations in the box and mean annual temperature at $r^2 > 0.9$ (Extended Data Fig. 1c,d).

Megadrought specification. Drought periods in each of the two time series were identified as the persistently negative PDSI values relative to an 11-year moving average[55]. Of these drought events, we analysed those that were both at least ten years long[3] and were more severe than the 50th percentile of each region's droughts (where severity is the sum of the PDSI over the drought period, with more severe being more negative); these criteria isolate the most severe 'megadroughts' in these regions. Nevertheless, the results presented here are insensitive to the precise choice of length and percentile thresholds (for example, for the NASW, changing the percentile threshold to the 67th and 75th adds only one and two drought events, respectively).

Radiative forcing response. We use an estimate of the global mean temperature response to radiative forcing from the megadrought analysis of ref. [10]. This global mean temperature response is based on a coupled energy balance model and estimated global-scale radiative forcings over the past two millennia. The coupled model has been shown to capture both the fast and slow components of atmosphere–ocean global climate models as they respond to time varying radiative forcings[56]. The model simulates the forced global mean temperature response to historical estimates of greenhouse gases (CO_2, N_2O and CH_4)[57], solar irradiance[58] and global volcanic forcing[59]. Whereas this model is not spatially resolved, we argue that it is a conservative use of the palaeoclimate forcing time series that are most reliable on hemispheric or global scales. See ref. [10] for further specifications of the model, its validation over the instrumental period and the resulting forced time series over the past two millennia.

Pairwise comparisons reconstruction. We use the PaiCo (or 'pairwise comparisons') reconstruction methodology[60] as an alternative to PHYDA for the mean SASW time series reconstruction (Extended Data Figs. 4 and 5). PaiCo is a robust index reconstruction methodology that has been widely used in palaeoclimatology and palaeoclimate reconstructions of the Common Era (for example, ref. [61]). PaiCo compares all pairs of sample values within each proxy record and produces a time series that best matches (maximum likelihood) all of the pairwise comparisons. The reconstruction method is loosely similar to finding the time series that has the best Kendall's tau correlation coefficient with all proxy records. The resulting unitless composite time series is then scaled to match the instrumental record, here the mean SASW PDSI from ref. [62] from 1903–1995 CE. The input proxy data for the PaiCo reconstruction are all of the moisture-sensitive tree-ring time series ($n = 47$) within the SASW region that were used in PHYDA. Note that we used the moisture sensitivity as determined in the construction of PHYDA: if the local growing season PDSI correlated more strongly with the proxy than the local growing season temperature (as derived from ref. [63]), the proxy was flagged as moisture sensitive and modelled as such in the reconstruction (see ref. [28]). A verification[64] of the PaiCo reconstruction over the SASW is shown in Extended Data Fig. 2.

AR-null model for probability of co-occurrence. We constructed an AR-null model to test whether or not the co-occurrence of megadroughts in both the NASW and the SASW regions (Fig. 1a) can be explained by chance. We first calculated[65] that a second-order AR model best fitted both the unsmoothed NASW and SASW time series shown in Fig. 1a. We then used the coefficients of the best-fit AR models to simulate 1,000 noise time series representing each location. We used the same megadrought specification as above to calculate the amount of megadrought overlap expected for these unrelated time series. The null distribution of the probability of overlap, as well as for the absence of drought in both locations, is shown as the box plots in Fig. 1c ('Both' and 'Neither').

Self-organizing map analysis. SOMs are a neural-network-based cluster analysis[34] that have been used in climate science to describe the continuum of patterns within a dataset (for example, refs. [35,66]). Conceptually, an analysis of climate modes using SOMs is similar to traditional principal component analysis in that the 'nodes' of a SOM represent the primary modes of the underlying climate field. In contrast to a principal component analysis, SOMs do not impose linearity and orthogonality on states that may be related in non-linear, non-orthogonal ways; in practice this makes SOMs more likely to correspond to physically plausible patterns[67].

Here we use the SOM algorithm to assign DJF SST anomaly fields from the Pacific to spatial patterns of a pre-set number. The SOM algorithm creates a set of spatial patterns that maximize their similarity with the underlying SST fields by minimizing their Euclidian distance. The algorithm then assigns each DJF SST field to the best-matching pattern. The SOM patterns themselves are approximately a composite of the assigned SST fields[35]. In addition, the SOM analysis organizes the patterns such that similar patterns are set in nearby locations within a regular two-dimensional grid. Thus, this full process allows for the visualization of a reduced-space continuum of patterns in the dataset (Fig. 3 and Extended Data Fig. 7).

For the SOM analysis, we used the mean DJF SSTs from PHYDA in the largely Pacific sector bounded by 45° N to 45° S and 135° E to 70° W. The SST fields were detrended by removing the linear trend at each grid point over the full analysis period, 1000–1925 CE; this detrending prevents spurious trends in the pattern occurrence[66] due to the long-term cooling trend over the pre-industrial Common Era. Because PHYDA is on a regular latitude–longitude grid, we apply an area weighting of the SST fields according to the cosine of the latitude[35].

As a test of the robustness of the findings in Fig. 3 we repeated the analysis with six nodes (Extended Data Fig. 7). We did not include more SOM nodes than the eight presented in the main text because eight modes of tropical Pacific variability are already at or very near the upper limit of physical plausibility[68].

We constructed a Markov chain Monte Carlo null distribution to assess whether or not the pattern frequency changes under 'both' megadrought conditions are unusual. We first found the SOM-state transition matrix given the best-matching unit time series. This transition matrix was then used to generate a Markov chain-based 10,000 year SOM-state time series (that is, a synthetic 10,000 year time series, composed of SOM node numbers, that matches the statistical characteristics of the PHYDA-based best-matching unit time series). We built up a null distribution by sampling from this time series the same number of 'both' years as observed ($n = 136$) 1,000 times. For the SOM node frequency changes shown in Fig. 3b, the coloured 'unusual' frequency changes were those that exceeded the 2.5th or 97.5th percentiles of the null distribution. Note that in Fig. 3b, node 5 is an unusual frequency change while node 8 is not because node 5 is more common than node 8 (17% frequency versus 11% frequency). The same null-generating process was followed for the supplementary SOM analyses shown in Extended Data Fig. 7.

Linear regression analysis. As in ref. [10], we performed a decadal-scale linear regression analysis to assess the relative importance of ENSO and forcing drivers in explaining the decadal hydroclimate over the period 1000–1925 CE for both the NASW and the SASW (Fig. 4a). Here, decadal-scale PDSI was the predictand and decadal-scale NINO3.4, AMO (North Atlantic SST anomalies relative to the period 1000–1925 CE) and the modelled temperature response to radiative forcing (see the 'Radiative forcing response' section above) were the predictors. We performed the univariate and multivariate regressions within a bootstrap sampling framework that sampled both the values in the time series and the choice of averaging timescale: the unsmoothed time series were averaged at nine-, ten- and eleven-year block averages and randomly sampled with replacement such that 999 (3×333) total regressions were performed for each predictor/predictand(s) option. This bootstrap sampling allowed us to incorporate the uncertainty associated with the choice of time averaging, as well as account for the relatively small sample size ($n \approx 92$ at decadal timescales). Before computing the regressions, the averaged time series were detrended with respect to the full analysis period 1000–1925 CE, in part to remove autocorrelation (the removal of which was confirmed). The adjusted r^2 values (adjusted for the number of predictors) for all 999 bootstrap regressions are summarized by the box plots in Fig. 4a.

Data availability

All data needed to evaluate the conclusions in the paper are present in public repositories. PHYDA and the palaeoclimate records used in its creation are publicly available in Zenodo data repositories at https://doi.org/10.5281/zenodo.1154913 (ref. [48]) and https://doi.org/10.5281/zenodo.1189006 (ref. [50]).

Code availability

The code for computing the SOMs is publicly available from https://github.com/ilarinieminen/SOM-Toolbox, while the code for computing PHYDA is publicly available from https://github.com/njsteiger/PHYDA. All additional code related to this paper may be requested from the authors.

References

48. Steiger, N. J. *Paleo Hydrodynamics Data Assimilation product (PHYDA)* (Zenodo, 2018); https://doi.org/10.5281/zenodo.1154913
49. Steiger, N. J. & Smerdon, J. E. A pseudoproxy assessment of data assimilation for reconstructing the atmosphere–ocean dynamics of hydroclimate extremes. *Climate of the Past* **13**, 1435–1449 (2017).
50. Steiger, N. et al. *A global Collection of Paleoclimate Proxy Time Series Over the Common Era* (Zenodo, 2018); https://doi.org/10.5281/zenodo.1189006
51. St. George, S., Meko, D. M. & Cook, E. R. The seasonality of precipitation signals embedded within the North American Drought Atlas. *Holocene* **20**, 983–988 (2010).
52. Baek, S. H. et al. Precipitation, temperature, and teleconnection signals across the combined North American, Monsoon Asia, and Old World Drought Atlases. *J. Clim.* **30**, 7141–7155 (2017).
53. Anderson, W., Seager, R., Baethgen, W. & Cane, M. Life cycles of agriculturally relevant ENSO teleconnections in North and South America. *Int. J. Climatol.* **37**, 3297–3318 (2017).
54. Harris, I., Jones, P., Osborn, T. & Lister, D. Updated high-resolution grids of monthly climatic observations – the CRU TS3.10 dataset. *Int. J. Climatol.* **34**, 623–642 (2014).
55. Meehl, G. A. & Hu, A. Megadroughts in the Indian monsoon region and southwest North America and a mechanism for associated multidecadal Pacific sea surface temperature anomalies. *J. Clim.* **19**, 1605–1623 (2006).
56. Held, I. M. et al. Probing the fast and slow components of global warming by returning abruptly to preindustrial forcing. *J. Clim.* **23**, 2418–2427 (2010).
57. Meinshausen, M. et al. Historical greenhouse gas concentrations for climate modelling (CMIP6). *Geosci. Model Dev.* **10**, 2057–2116 (2017).
58. Vieira, L. E. A., Solanki, S. K., Krivova, N. A. & Usoskin, I. Evolution of the solar irradiance during the Holocene. *Astron. Astrophys.* **531**, A6 (2011).
59. Sigl, M. et al. Timing and climate forcing of volcanic eruptions for the past 2,500 years. *Nature* **523**, 543–549 (2015).
60. Hanhijärvi, S., Tingley, M. & Korhola, A. Pairwise comparisons to reconstruct mean temperature in the Arctic Atlantic region over the last 2,000 years. *Clim. Dyn.* **41**, 2039–2060 (2013).
61. PAGES 2k Consortium. Continental-scale temperature variability during the past two millennia. *Nat. Geosci.* **6**, 339–346 (2013).
62. Schrier, G., Barichivich, J., Briffa, K. & Jones, P. A scPDSI-based global data set of dry and wet spells for 1901–2009. *J. Geophys. Res. Atmos.* **118**, 4025–4048 (2013).
63. Rohde, R. et al. Berkeley Earth temperature averaging process. *Geoinform. Geostat. Overview* https://doi.org/10.4172/2327-4581.1000103 (2013).
64. Gneiting, T., Raftery, A. E., Westveld III, A. H. & Goldman, T. Calibrated probabilistic forecasting using ensemble model output statistics and minimum CRPS estimation. *Mon. Weather Rev.* **133**, 1098–1118 (2005).
65. Schneider, T. & Neumaier, A. Algorithm 808: ARfit—a Matlab package for the estimation of parameters and eigenmodes of multivariate autoregressive models. *ACM Trans. Math. Softw.* **27**, 58–65 (2001).
66. Horton, D. E. et al. Contribution of changes in atmospheric circulation patterns to extreme temperature trends. *Nature* **522**, 465–469 (2015).
67. Reusch, D. B., Alley, R. B. & Hewitson, B. C. Relative performance of self-organizing maps and principal component analysis in pattern extraction from synthetic climatological data. *Polar Geogr.* **29**, 188–212 (2005).
68. Johnson, N. C. How many ENSO flavors can we distinguish? *J. Clim.* **26**, 4816–4827 (2013).

Acknowledgements

We thank E. Tejedor for helpful discussions about historical droughts in Chile. This work was supported in part by the National Science Foundation (NSF) under grants NSF OCE 1657209, OISE-1743738, AGS-1602581, AGS-1703029, AGS-1602920 and AGS-1805490. This work was also supported in part by the National Oceanic and Atmospheric Administration NA20OAR4310379 and in part by the Israel Science Foundation grant 2654/20.

Author contributions

N.J.S. designed the experiments with J.E.S., A.P.W. and R.S. The analysis code was developed by N.J.S., who also ran the experiments, performed the statistical analyses of the experiments and produced the figures. N.J.S. and J.E.S. prepared the manuscript with assistance from A.P.W., R.S. and A.V.-C.

Competing interests

The authors declare no competing interests.

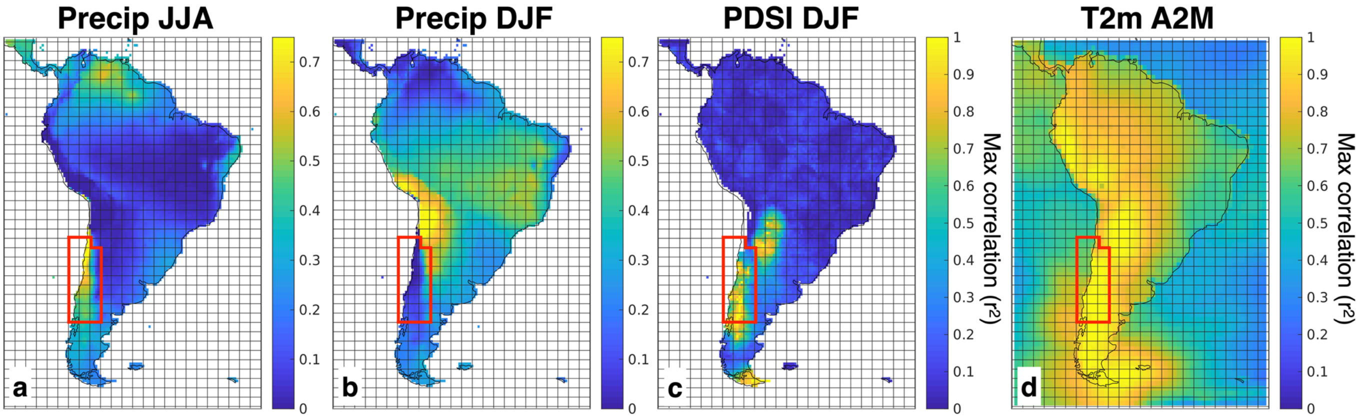
Additional information

Extended data is available for this paper at https://doi.org/10.1038/s41561-021-00819-9.

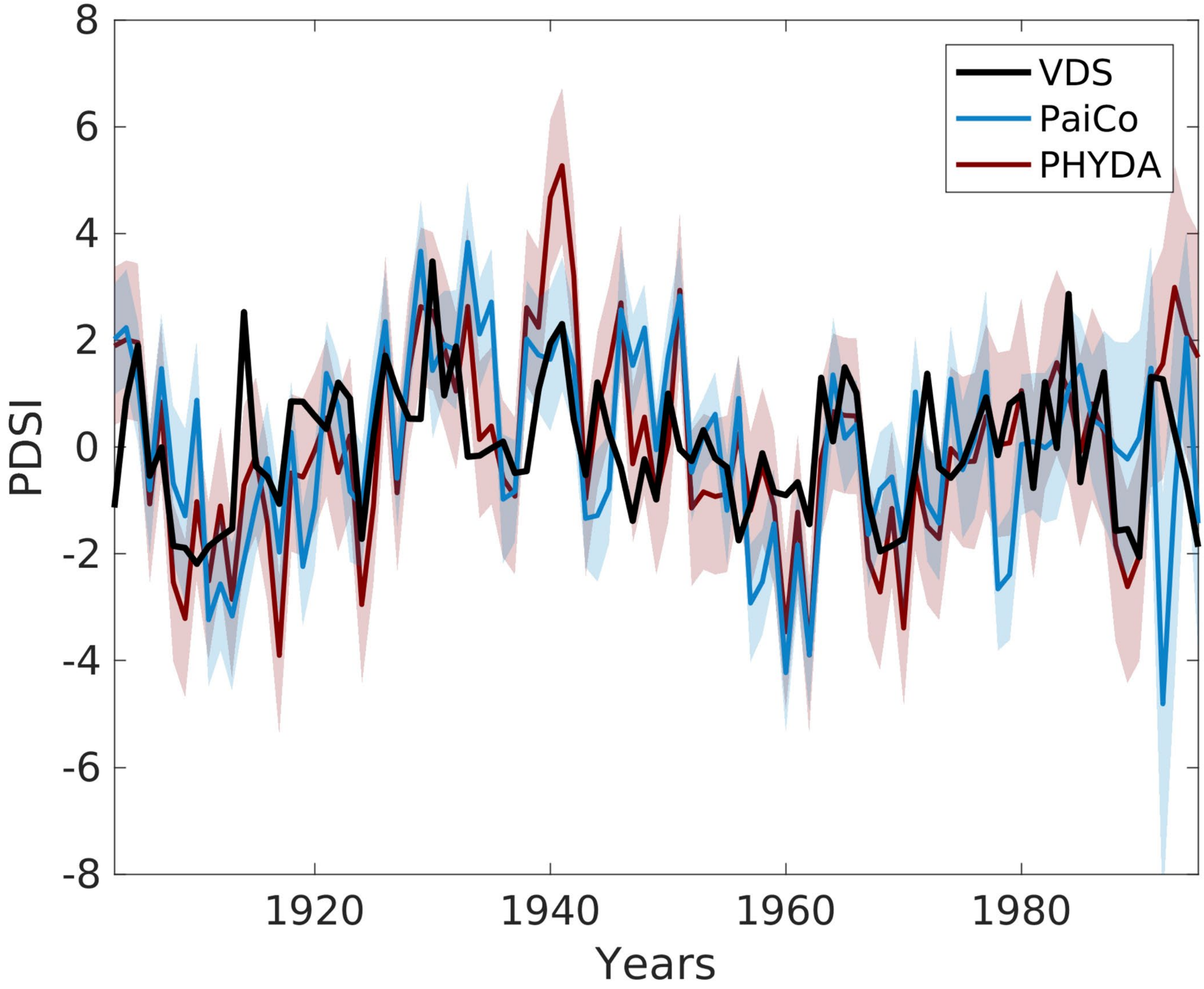
Correspondence and requests for materials should be addressed to N.J.S.

Peer review information *Nature Geoscience* thanks Roberto Rondanelli, Matthieu Lengaigne and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Primary Handling Editor: James Super.

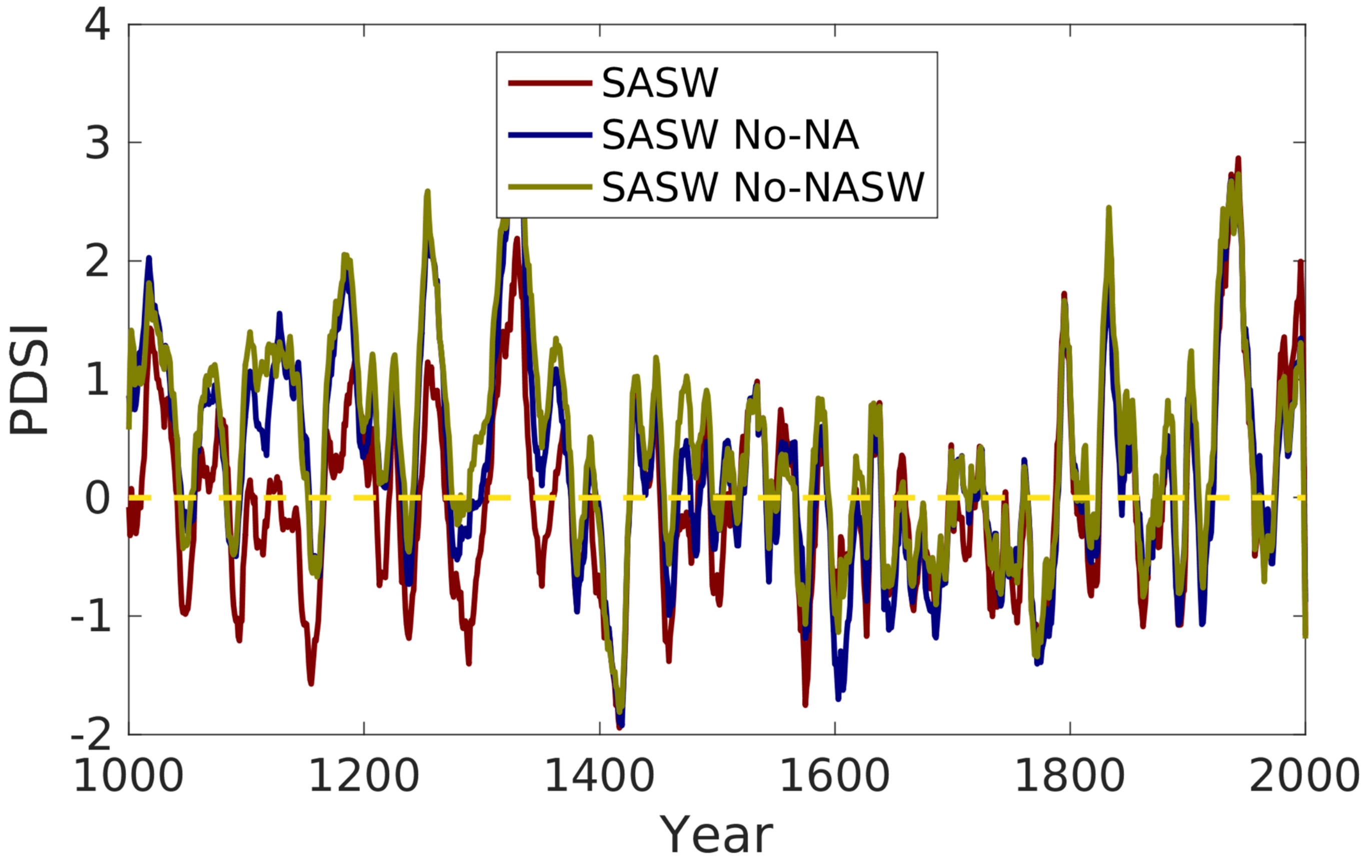
Reprints and permissions information is available at www.nature.com/reprints.



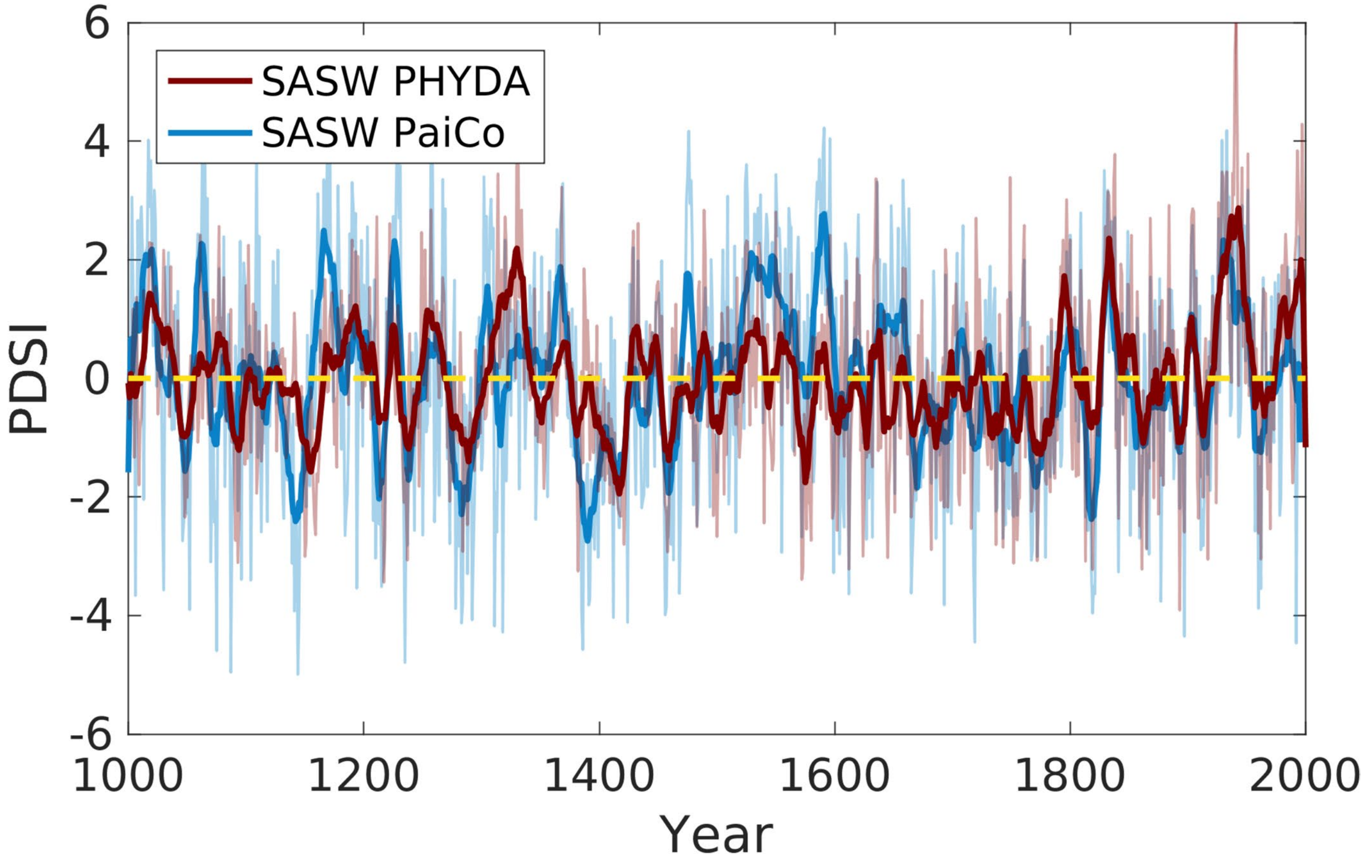
Extended Data Fig. 1 | South American analysis region. (**a**) The fraction of annual precipitation that falls within the JJA season, based on a climatology of CRUTS3.23 from 1950 to the present[54]. The SASW region analyzed here is the land area enclosed within the red bounding box. In this and all panels, the approximately 2 degree grid of PHYDA is overlaid on South America. (**b**) The fraction of annual precipitation that falls within the DJF season, similar to (a). (**c**) The maximum r^2 correlation of observations-based PDSI[62] at the South American moisture-sensitive proxy locations used in PHYDA ($n = 60$). This is constructed by first computing the r^2 value between each proxy location and all other South American locations, using only the observations-based PDSI data. The maximum value of all of these point correlation maps is displayed here. The moisture-sensitive proxy locations are those locations associated with proxy time series that were modeled with PDSI (and not temperature) in the data assimilation framework of PHYDA. (**d**) Same as in (c), but using an observations-based temperature product[63] and the South American temperature-sensitive proxy locations used in PHYDA ($n = 116$). The correlations were computed using annual mean data, defined as April to the next calendar year March (A2M).



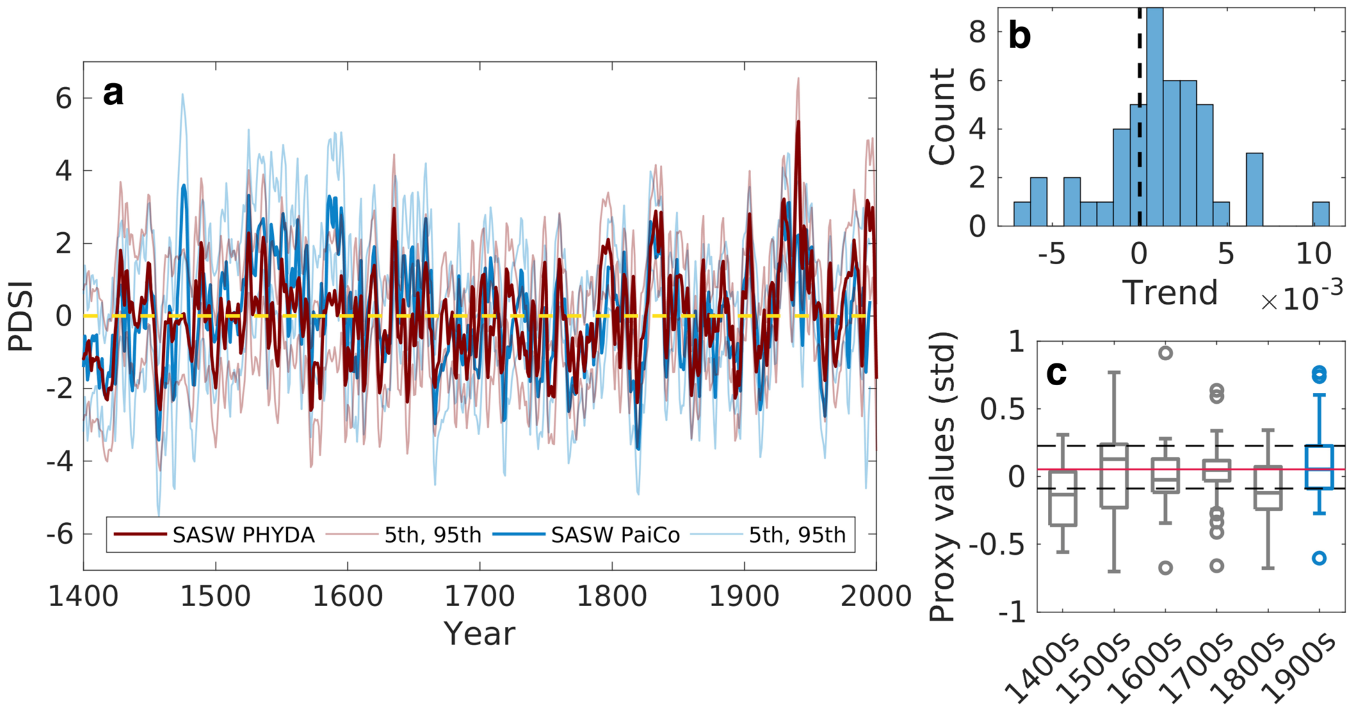
Extended Data Fig. 2 | Verification of the SASW reconstructions. Verification of PHYDA's reconstruction of the SASW along with a reconstruction of this same region using only the underlying proxy data (PaiCo; using a method of 'Pairwise Comparisons'[60], see Methods and Materials) and the DJF PDSI over this region from the 'van der Schrier' (VDS) PDSI product[62]. PHYDA and PaiCo are shown with error bounds extending to $\pm 2\sigma$ of their respective reconstruction ensembles. All PDSI time series show a DJF average. The PDSI baseline and period of comparison is 1903–1995, bounded on the oldest date by the availability of observations-based PDSI from Ref. [62] and on the newest date by the availability of proxy data for the PaiCo reconstruction. Verification metrics for PHYDA against VDS over this time period are: $r = 0.60$, mean absolute error = 1.14, continuous ranked probability score = 0.87. Corresponding verification metrics for PaiCo are: $r = 0.29$, mean absolute error = 1.46, continuous ranked probability score = 1.20. Note that for mean absolute error and continuous ranked probability score[64], lower values are better.



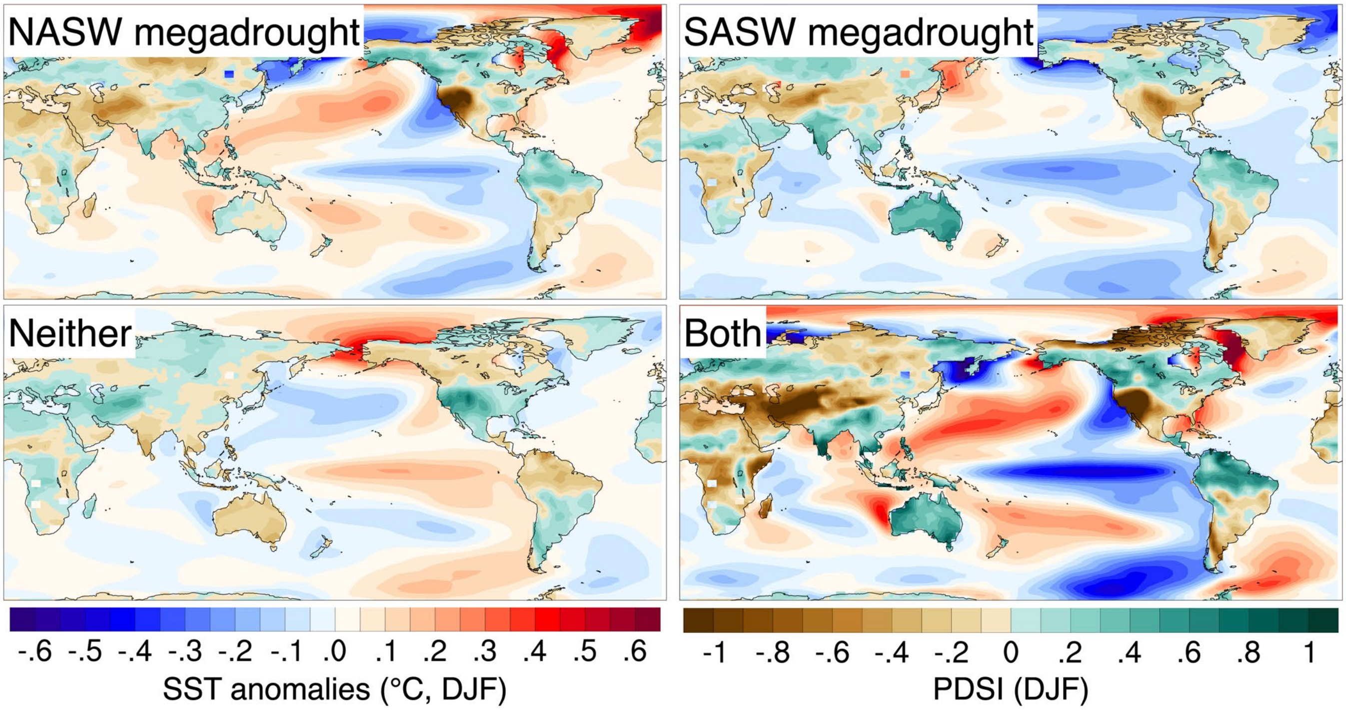
Extended Data Fig. 3 | Alternative reconstructions of the SASW from PHYDA. Reconstruction of the SASW region from the published PHYDA along with alternative PHYDA-based reconstructions that were made with no proxies in all of North America (SASW No-NA) and no proxies in all of the North American Southwest (SASW No-NASW). The reconstructions are displayed with an 11-year smoothing (same as in Fig. 1) for clarity of comparison at the decadal-scale. Note that the No-NA and No-NASW reconstructions reflect not only the lack of proxies in NA and the NASW, but also the increased influence of other non-local proxies in the absence of NA and NASW proxies. The PDSI series are shown relative to the mean of the SASW time series over the analysis period (1000-1925).



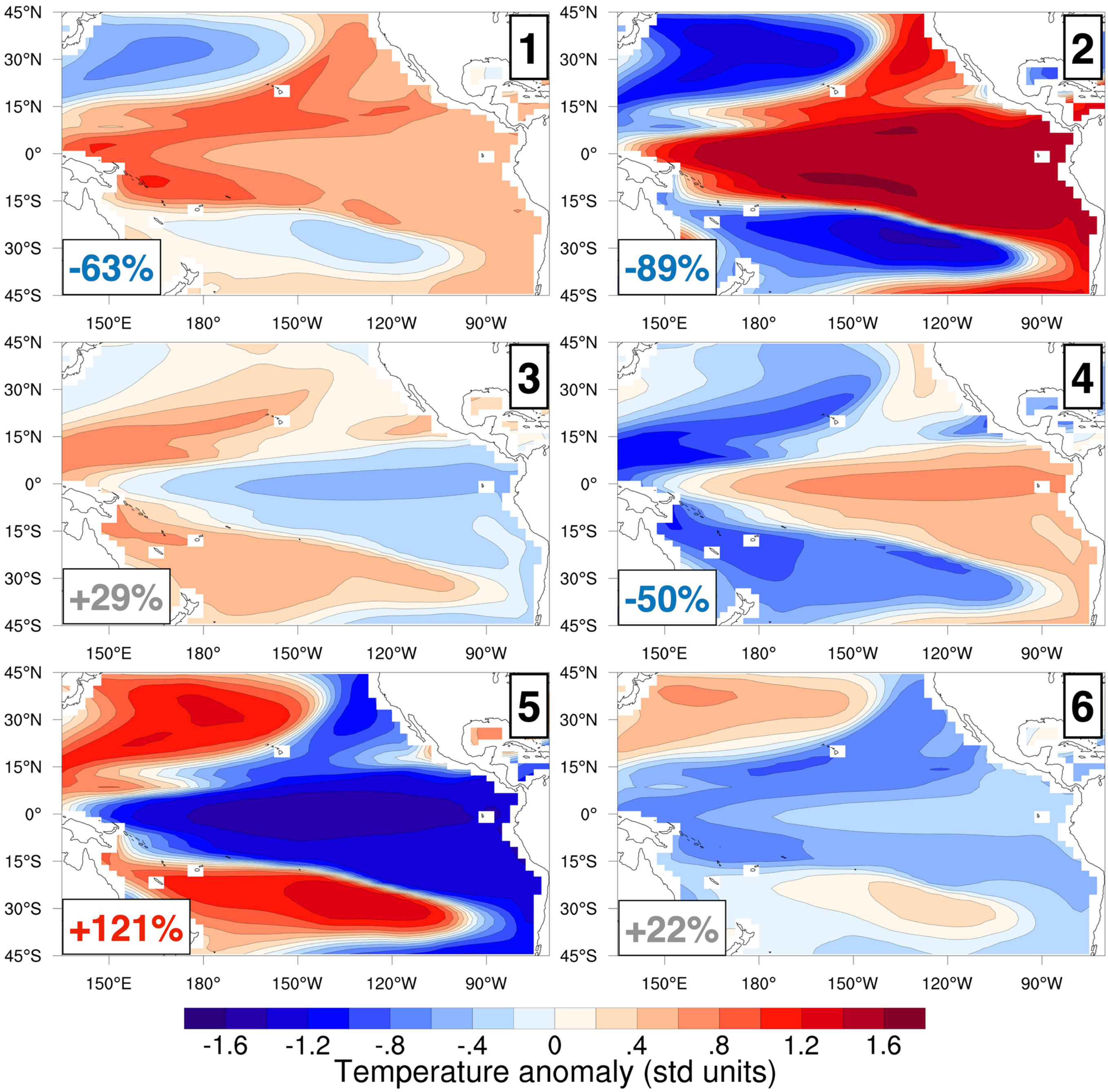
Extended Data Fig. 4 | Comparison of SASW reconstructions. Comparison of PHYDA's reconstruction of the SASW (the same as presented in Fig. 1) with a reconstruction of this same region using only the underlying proxy data (SASW PaiCo). SASW PaiCo is a reconstruction based on a method of 'Pairwise Comparisons' (PaiCo)[60] and all of the moisture-sensitive tree ring time series ($n=47$) within the SASW region (Fig. S1) used in PHYDA, see Methods. The PDSI series are shown relative to the analysis period mean (1000-1925).



Extended Data Fig. 5 | SASW time series and proxy data analysis. (**a**) Comparison of PHYDA and PaiCo reconstructions for the SASW, as in Fig. S4 but for the time series over the years 1400–1800, a probable relative dry period (c.f. Fig. 1). For clarity, each annual time series and uncertainties have been smoothed using a 5-year LOWESS filter. Uncertainties extend to the 5th and 95th percentiles of the probabilistic ensemble generated by each reconstruction method. This panel illustrates that a reconstruction method based on the local proxy data alone, PaiCo, agrees with the PHYDA reconstruction in the result of a generally drier period from approximately 1400–1800; though this period is also marked by high interannual and multi-year variance, with many wet years interspersed within the dry period. The PDSI anomalies are relative to the mean over 1000-1925 (c.f. Fig. 1). (**b**) Trend in the standardized moisture sensitive proxy data from the years 1800–2000 in the SASW. The sign of the proxy time series have been aligned so as to correlate positively with PDSI for this particular calculation. This panel shows that a majority of proxy data have a wetting trend from 1800–2000. (**c**) Box plot summaries of the distribution of 100 year means of each proxy time series in the SASW [proxy data were aligned as in panel (b)]. For reference, the century of the 1900s is highlighted in blue and horizontal lines across the panel indicate the top (black), middle (red), and bottom (black) box edges of the distribution for this century. The central box mark of the distributions indicates the median, the bottom and top edges of the box indicate the 25th and 75th percentiles, the whiskers extend to the most extreme data points not considered outliers (approximately ± 2.7σ), with outliers indicated by circles. This panel illustrates that the distribution of proxy values are generally shifted towards drier conditions for all centuries during the 1400–1800s, with the exception of the 1500s; this relatively wetter century is also visible in panel (a).



Extended Data Fig. 6 | Alternative SST and PDSI composite analysis. Composites of the original PHYDA DJF SST and DJF PDSI, as in Fig. 2, except that the SASW megadrought years are calculated according to the alternative PaiCO reconstruction of SASW DJF PDSI (which is based solely on the underlying proxy data, see Methods) shown in Figs. S2 and S4. The years of NASW megadrought are determined through the original PHYDA, thus the top left 'NASW megadrought' panel in this figure and Fig. 2 are identical. The years of SASW megadrought use the alternative PaiCo reconstruction of SASW PDSI, thus the remaining panels differ in their constituent years from Fig. 2. Temperature and PDSI data are anomalies with respect to the analysis period 1000–1925.



Extended Data Fig. 7 | SOM analysis using six nodes. This is the same analysis as shown in Fig. 3 but for six SOM nodes instead of eight. The SOM patterns are based on detrended, standardized DJF SST data from PHYDA for the years 1000–1925 (see Methods). Boxes in the lower left corner of each panel show the percent change in the frequency of occurrence of that pattern during simultaneous megadrought years relative to the average frequency over 1000–1925; gray colored frequency changes indicate those percent changes that fall within the 2.5th and 97.5th percentiles of a Markov chain Monte Carlo null distribution for each pattern (see Methods).