Internally Driven and Externally Forced Pacific Decadal Variability in the CESM Last Millennium Ensemble

Anson H. Cheung[1*], Aakash Sane[1†], Baylor Fox-Kemper[1]

[1]Department of Earth, Environmental, and Planetary Sciences, Brown University, Providence, RI 02912, USA

Key Points:

- A last millennium large climate model ensemble was used to characterize the Pacific Decadal Oscillation and Interdecadal Pacific Oscillation
- The external forcing contributions on Pacific decadal variability is small compared to internal variability during years 850–2005 CE
- Multiple marine based proxy records can be used to better constrain Pacific decadal variability compared to a single proxy record

*Current address: Lamont-Doherty Earth Observatory, Columbia University, Palisades, NY 10964

†Current address: Atmospheric and Oceanic Sciences, Princeton University, Princeton, NJ 08544

Corresponding author: Anson H. Cheung, `acheung@ldeo.columbia.edu`

Abstract

Pacific decadal variability are known to drive global and regional climate and ecosystems changes. However, the relative role of internal variability and external forcings in driving PDV and the prospects of obtaining a more accurate PDV reconstruction using a wider marine proxy network remain unclear. Here, we analyze simulations from the Community Earth System Model Last Millennium Ensemble using a information theory metric and find that internal variability dominates PDV, with minor contributions from greenhouse gases and volcanic forcing. Using a more extensive marine proxy network for PDV reconstruction is also shown to outperform PDV reconstruction based on a single proxy record. Collectively, these results offer new insights on the drivers of PDV and pathways to improve PDV reconstruction.

Plain Language Summary

Variations of Pacific sea surface temperature over decades can pose huge influence on global and regional climate. Therefore, we want to understand whether these variations happen naturally in the climate system or can be forced by factors that are not part of the climate system. But, previous studies only provide partial answers because paleoclimate reconstructions do not show a consistent picture of past changes in the Pacific and that there are other factors that can influence results from climate models, but are not sufficiently considered. In this study, we use a climate model with multiple simulations of the past 1000 years to offer a more complete answer. Specifically, we show that decadal scale variations in the Pacific ocean arise mostly from interactions within the climate system. We also show that by using multiple records in the Pacific ocean in a hypothetical scenario, we can obtain a more comprehensive and accurate view of how the Pacific ocean has changed every decades. These results together paint a more complete picture of changes in the Pacific and tell us how we can better understand changes in the Pacific beyond the instrumental record.

1 Introduction

Pacific Decadal Variability (PDV) is a collection of basin-wide phenomena that impose significant impacts on global and regional climate. Based on spatiotemporal patterns used to describe PDV, primarily the Pacific Decadal Oscillation (PDO) (Y. Zhang et al., 1997), Interdecadal Pacific Oscillation (IPO) (Power et al., 1999), and North Pacific Gyre Oscillation (NPGO) (Di Lorenzo et al., 2008), past studies have shown that PDV drives decadal scale global mean temperature variations (Dai et al., 2015; Kosaka & Xie, 2016; Meehl et al., 2016), regional temperature and hydroclimate (B. Dong & Dai, 2015; McCabe et al., 2004), and ecosystems (Mantua & Hare, 2002; Di Lorenzo et al., 2008). The widespread impacts of PDV highlight the importance to understand how it varies internally and changes in response to different external forcings.

Extensive analyses have been done to understand the characteristics, dynamics and drivers of PDV. Multiple review studies have highlighted the role of the atmosphere, ocean, atmosphere-ocean interactions in the Pacific, tropical-extratropical interactions, and inter-basin connections in driving PDV (see Di Lorenzo et al., 2023; Liu, 2012; Newman et al., 2016, and references therein). These processes are shown to occur internally without any external perturbations (Capotondi et al., 2020; L. Zhang & Delworth, 2015). However, more recent studies also noted the influences of greenhouse gases (Bonfils & Santer, 2011; Meehl et al., 2013), volcanic eruptions (Maher et al., 2014), and aerosols (Boo et al., 2015; Dittus et al., 2021) on PDV (L. Dong et al., 2014; Hua et al., 2018).

Despite identifying both the internally driven and externally forced components of PDV, prior studies mostly focus on the instrumental era, which spans only the past ~ 150 years (e.g., Roemmich et al., 2012; Davis et al., 2019; Giese et al., 2016) and contains ~ 15 non-overlapping decadal samples. As such, the short time period considered undermines our ability to accurately characterize internal decadal variability (Stevenson et al., 2010) and quantify the role of external forcing on PDV due to the limited range of magnitude and sources during this limited window (e.g., Jungclaus et al., 2017; Sigl et al., 2015). Hence, our understanding on the spatiotemporal characteristics of PDV and the role of external forcings on PDV remains incomplete.

The last millennium provides a better baseline to more comprehensively characterize PDV that arise through chaotic processes in the climate system and those driven by external perturbations. Proxy reconstructions offer the most direct measurements of past PDV beyond the instrumental record (e.g., d'Arrigo et al., 2001; Felis et al., 2010; O'Mara et al., 2019; Porter et al., 2021). However, most PDV reconstructions to date either rely on a small number of coral records from a localized region or a network of terrestrial based proxy records. Consequently, they might not be able to represent PDV faithfully due to nonstationary teleconnections (Du et al., 2020). Climate model simulations have also been used to understand the spatiotemporal characteristics of PDV (e.g., Fleming & Anchukaitis, 2016; Sun et al., 2022; Wang et al., 2012; Zanchettin et al., 2013; Stevenson et al., 2019). Yet, these analyses either focus on multi-model ensemble or model simulations with a single realization that is subject to both internal and forced variability. Such experimental protocols impede our ability to separate internal variability from differences in model physics and individual forcing effects on PDV.

To better understand the relative roles of internal variability and external forcings on PDV, as well as to determine whether we can improve PDV reconstruction by using a more spatially extensive marine proxy network, we analyze the the Community Earth System Model Last Millennium Ensemble (CESM-LME) (Otto-Bliesner et al., 2016). The CESM-LME contains full forcing and single-forcing simulations, thus offers an opportunity to better distinguish the role of internal variability and external forcings on PDV that is previously unattainable. In addition, efforts to compile proxy records over the Common Era (Walter et al., 2023) now provide adequate information for us to use the CESM-LME as a testbed to understand the sensitivity of PDV reconstructions to a realistic proxy network.

Here, we use CESM-LME and apply an information theory metric (Sane et al., 2021, 2024) to determine (1) the spatiotemporal fingerprint of each external forcing and internally driven decadal sea surface temperature in the Pacific, (2) the contribution of each forcing to total variability of PDV, (3) whether the influence of each forcing on PDV has changed in recent decades, and (4) the prospect of using multiple marine proxy records of those available to reconstruct PDV.

2 Data and Method

2.1 Data

We used the CESM-LME (Otto-Bliesner et al., 2016) for analysis in this study. CESM-LME uses CESM version 1.1, which has a climate sensitivity between that of CCSM4 and CESM2 (Meehl et al., 2013; Gettelman et al., 2019) making it suitable for paleoclimate studies (Zhu et al., 2020). The nominal resolution is $\sim 2^\circ$ for the atmosphere and land components and $\sim 1^\circ$ for the ocean and ice components. CESM-LME contains multiple realizations of full forcing (13), greenhouse gas (GHG) only (3), land use (LULC) only (3), orbital only (3), solar only (4), and volcanic only (5) simulations that span from 850-2005 CE. Similar versions of CESM1 used in the CESM-LME have been shown to capture the spatial pattern (Fasullo et al., 2020) and processes associated with PDV (Newman et al., 2016) more accurately than other climate models that are from the same generation. Nonetheless, CESM1 is also known to exhibit too strong El Niño Southern Oscillation (Capotondi et al., 2020), which can influence PDV, and an inaccurate tropics-PDO linkage (Newman et al., 2016).

We used information from the PAGES CoralHydro2k database (Walter et al., 2023) for pseudoproxy reconstructions. The CoralHydro2k database contains a compilation of coral records with oxygen isotopic composition (δ^{18}O) and strontium-to-calcium ratio (Sr/Ca) measurements from the tropical and subtropical oceans. δ^{18}O in coral is sensitive to ambient temperature and δ^{18}O of seawater (Epstein et al., 1953) whereas changes in Sr/Ca in corals are primarily related to temperature changes (Corrège, 2006, and references therein). Some of these proxy records have previously been used to characterize PDV (e.g., Linsley et al., 2015).

2.2 Method

2.2.1 Spatial and temporal metrics of PDV

We characterized the temporal characteristics of PDV using the PDO index and the IPO tripole index (TPI). PDO is a principal component based index that is commonly used to characterize decadal SST variability in the north Pacific (Y. Zhang et al., 1997), whereas TPI is an index that does not rely on principal components and is used to characterize Pacific basin-wide decadal SST variability (Henley et al., 2015). Details can be found in Text S1. To understand the spatial pattern of forced and internal PDV, we applied a 13-year low pass filter to SST at each grid cell. A 4^{th} order Butterworth filter was used in all the low-pass filtering procedures.

2.2.2 Information theory based metric

Because of the complexity of the indices used and described, as well as their potentially nonlinear responses to different forcings and non-Gaussian statistics, we employ information theory-based metrics to quantify the relative role of internally forced and externally forced variability on PDV. Whereas the traditional approach relies on comparing the variance of model ensemble average with across realizations (e.g., Deser et al., 2020; Leroux et al., 2018; Llovel et al., 2018), this approach builds on the mutual information between time series, not just their correlation, and thus is robust to nonlinear relationships between the forcing and response and does not rely on assumed shapes for their probability distributions (Sane et al., 2024).

Following Sane et al. (2024), we used two information theory metrics to quantify shared variability between two variables and variability of a variable: Shannon entropy and mutual information (Text S2). Shannon entropy characterizes variability of a variable (Carcassi et al., 2021), and can be understood as a quantity that increases with the likelihood of finding a surprising or unprecedented result, or as the number of bits needed to count all of the states that a system visits. On the other hand, mutual information describes how much information entropy within the x signal can be explained by the y signal and vice versa. In the context of this study, the Shannon entropy of all realizations of full forcing simulations can be used to represent total variability, whereas the mutual information between each realization and the ensemble average can represent external variability assuming internal variability of each realization is independent from the other ensemble members.

For our analysis, we first defined f as all full forcing ensemble members, and g_x as the ensemble mean of a model ensemble (either a single-forcing or initial conditions ensemble), where x represents different forcing scenarios. Then, following Sane et al. (2024), we defined a metric γ (Equation 1) to estimate the fraction of information in externally forced variability x within the information in the total variability. Because we are interested in the relative fraction of each forcing x to total variability, we calculated γ using different g_x afterwards. g_x was derived from the initial conditions ensemble (i.e., full forcing) to estimate the fraction of all forcings to total variability, and g_x derived from different single forcing ensembles (i.e., GHG, LULC, orbital, solar, and volcanic forcings) to understand the fraction of forcing x on total variability. By calculating a g_x for each forcing x, $I(f; g_x)$ then can be used to represent the variability driven by forcing x. Consequently, γ of x will then represent the fraction of variability driven by forcing x relative to total variability.

$$\gamma = \frac{I(f; g_x)}{H(f)} \tag{1}$$

As this fraction is a fraction of bits of information entropy explained by external forcing versus bits of information entropy explained by internal and external processes, it does not equal numerically to the ratios of variance in other studies (Llovel et al., 2018). Nonetheless, it monotonically increases from 0 to 1 as more and more of the signal is explained by external forcing. It is also noteworthy that this calculation was done by combining all members of full forcing ensemble into a single vector (f), and by generating multiple copies of the ensemble mean in a single vector (g_x) so that the length of f and g_x would match. More complex methods to compare all ensemble members against each other (Watanabe, 1960) have proven too expensive computationally to be tractable in this type of application (Chen, 2024).

2.2.3 Pseudoproxy reconstruction

To better understand and quantify the advantage of a multiproxy record over a single proxy record based PDV reconstruction, we carried out pseudoproxy reconstructions (Smerdon, 2012) of annual PDO and IPO indices. Pseudoproxies were generated using the CESM-LME and based on information from the HydroCoral2k database (Figure 1), and were then used to reconstruct annual PDO and IPO indices using a nested composite-plus-scaling approach (e.g., Wilson et al., 2006, 2010) (Text S3). Using the CESM-LME as a perfect model framework allows evaluations of the reconstructed PDV against the actual PDV in model. We used Shannon entropy, mutual information, and coefficient of determination to compare each reconstruction's ability to capture the actual PDO and IPO in each nest, outside of the calibration period (1958-1994).

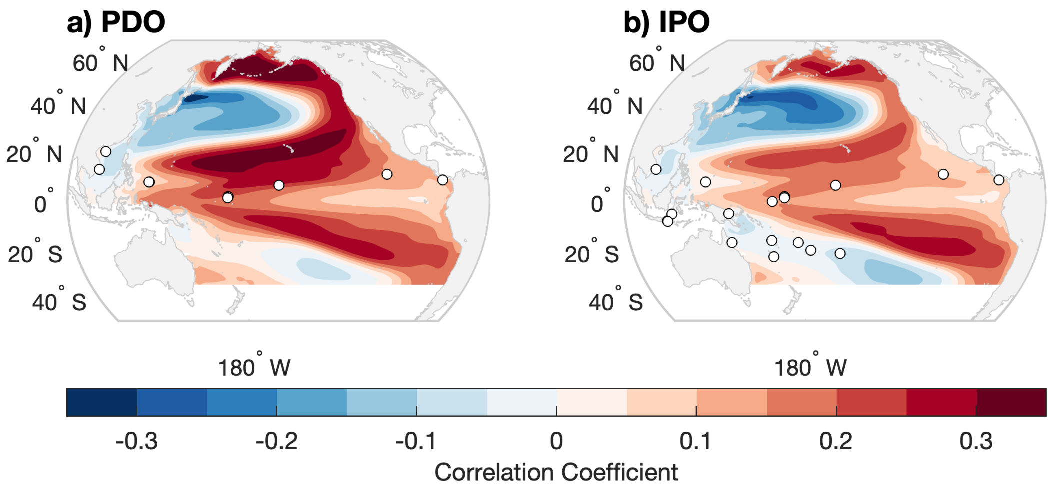

Figure 1. Correlation patterns of PDO and IPO and locations of proxy records. Correlation patterns between SST and (a) the PDO index and (b) the IPO index. The white circles represent proxy record locations used for each index based reconstruction.

3 Results

There are broad scale similarities in the spatial pattern and the strength of γ (fraction of externally forced decadal SST variability in total decadal SST variability) between forcings (Figure 2). The external forcing signal is most notable along the western boundary current in the Northern Hemisphere and off Baja California, whereas the internal forcing is most dominant in the eastern equatorial Pacific. The forced signal is most prominent in full forcing simulations (Figure 2a), with notable signals in GHG and volcanic simulations (Figure 2b,e). However, the overall forced decadal SST signal is relatively weak compared to internally driven decadal SST variability. Indeed, the spatial patterns of decadal total variability, as denoted by Shannon entropy (Figure 2g-l), do not resemble spatial patterns that are associated with forced changes.

The strength of γ in the time domain is similar to results obtained from analyzing the spatial pattern of γ. γ is strongest in the full forcing simulation ($\sim 0.035 - 0.045$) in both PDO and IPO. However, the relative strength between each forcing is dependent on the index used, with GHG and volcanic forcings have stronger roles for PDO, and GHG, LULC and orbital for IPO (Figure 3).

The time dependent information theory based metric suggests external forcing plays a minor role in PDV throughout the last millennium. For both PDO and IPO, γ is relatively low (< 0.1) throughout the last millennium (Figure 4). Interestingly, there are periods where the full forcing and volcanic forcing signal increases in PDO, but less so in IPO. The forced signal in the final 50 year window also does not appear to be anomalous relative to the last millennium.

Incorporating multiple proxy records appears to improve the reconstruction skill of PDO/IPO compared to using a single proxy record (Figure 5). Both γ and the coefficient of determination (R^2) are higher when multiple proxy records are incorporated into the proxy reconstruction.

4 Discussion and Conclusion

Our analysis based on CESM-LME suggests that the imprint of externally forced PDV is small relative to internally driven PDV. Typically, less than 5% of the Shannon entropy can be explained by mutual information. From an information theory perspective, this result can be understood as if files containing the record of PDV were optimally compressed, a

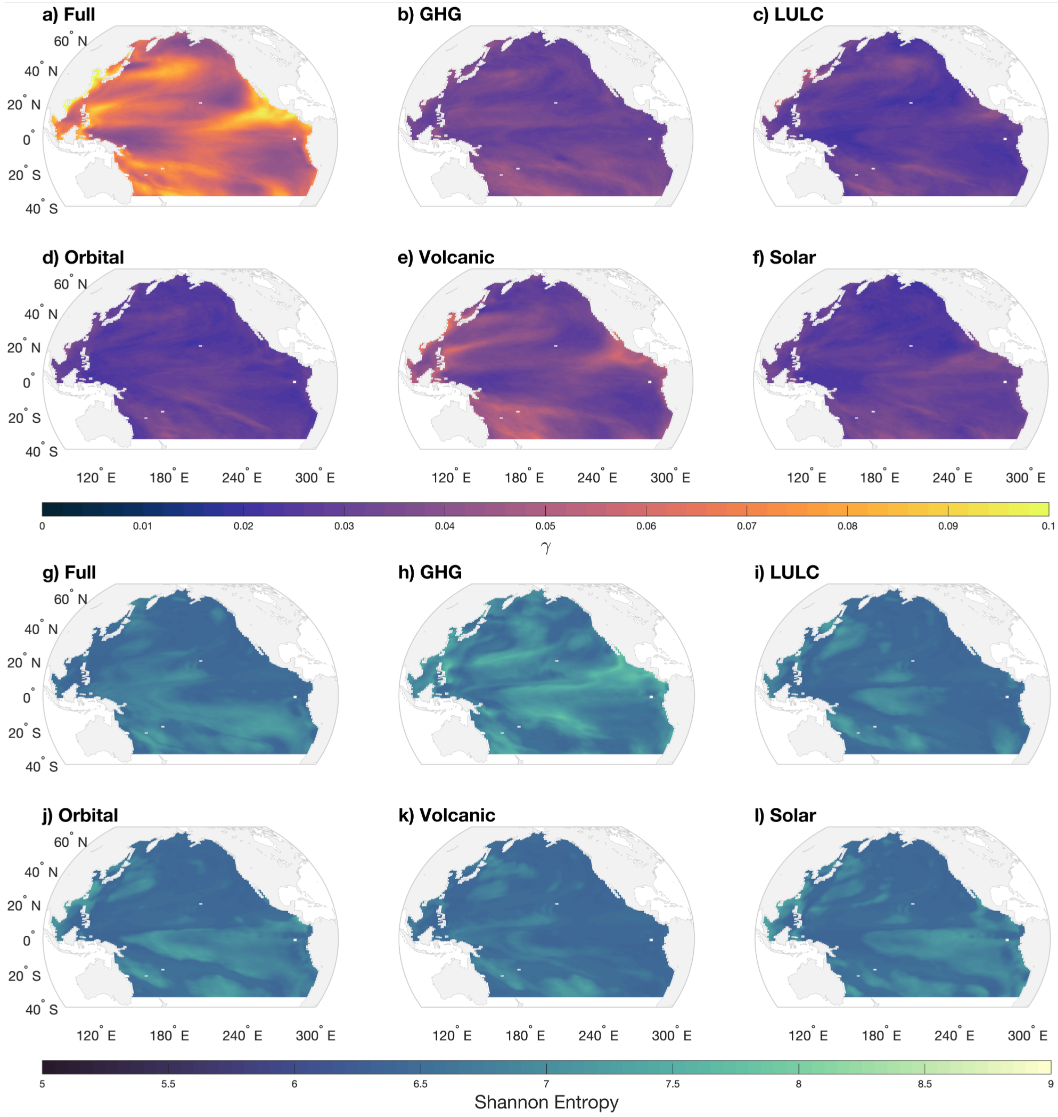

Figure 2. Spatial fingerprint of external forcing and total variability. Spatial patterns of the fractional contribution of external forcing to total information (a-f) and the Shannon entropy (g-l) at each grid point calculated using 13-year low pass filter SST data. Shown are results from (a,g) full forcing, (b,h) GHG forcing, (c,i) LULC forcing, (d,j) orbital forcing, (e,k) volcanic forcing, and (f,l) solar forcing simulations.

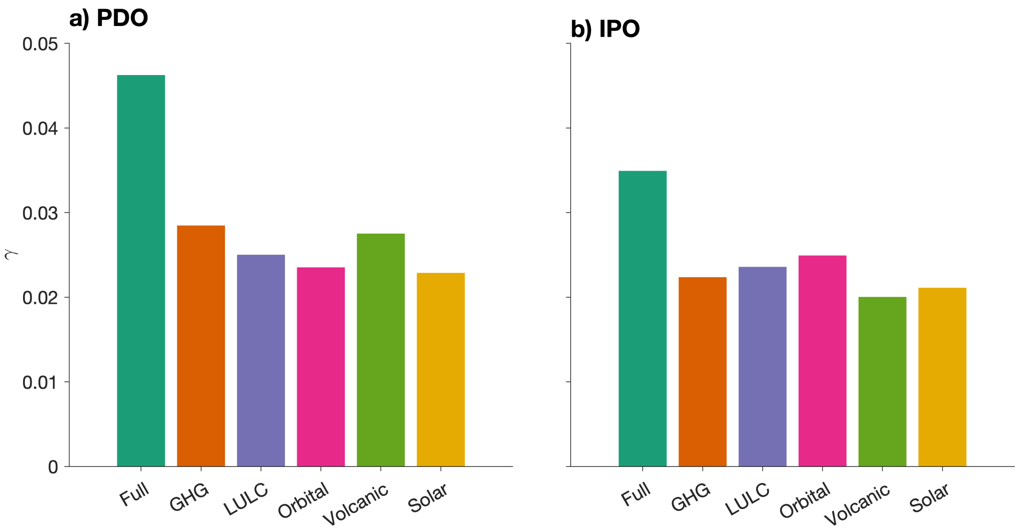

Figure 3. Fraction of information driven by external forcing. The fraction of shared information between ensemble average and all realizations relative to information from all realizations for (a) PDO and (b) IPO indices. Shown are results from full forcing (teal), GHG forcing (brown), LULC forcing (purple), orbital forcing (pink), solar forcing (green), and volcanic forcing (gold) simulations.

file containing fully-forced PDV responses (i.e., g_{full}) would be less than $1/20^{\mathrm{th}}$ the size of a file capturing PDV in any ensemble member's history (i.e., f). This result is analogous to detection and attribution studies where regression coefficients for forced changes based on regression fingerprinting methods are < 0.05 (Hegerl & Zwiers, 2011). Thus, internal variability is more pronounced everywhere in the Pacific basin, and also dominates the variability in indices that represent PDV. This agrees with previous last millennium studies that highlight the internal nature of PDV (Fleming & Anchukaitis, 2016; Zanchettin et al., 2013).

Although the influences of GHG and volcanic forcings on PDV are notable, these forcings only play a minor role in driving PDV. Even though we observe a strengthened forced signal during periods with notable volcanic eruptions (Figure 4), they are weak compared to internal variability. Furthermore, the GHG and volcanic signals in the final 50 years of the simulation (1966-2005) are also not anomalous in the context of the last millennium. Hence, although GHG and volcanic forcings can influence PDV during the historical period (e.g., L. Dong et al., 2014; Hua et al., 2018), our results suggest that they have only played a minor role in driving PDV in comparison to internal variability over the last millennium and have not become more important during the instrumental era. From the emergence of climate

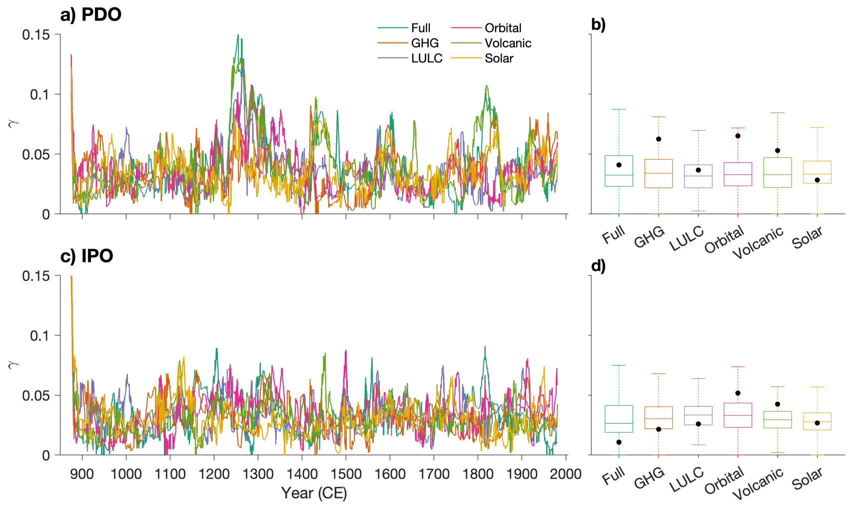

Figure 4. Time dependent fraction of information driven by external forcing. 50 year sliding window of the fraction of shared information between the ensemble average and all full forcing realizations relative to information from all realizations in a) PDO and c) IPO. Box plots of the different forcings' γ values in b) PDO and d) IPO. The black circles indicate the value at the final 50 year window (1966-2005).

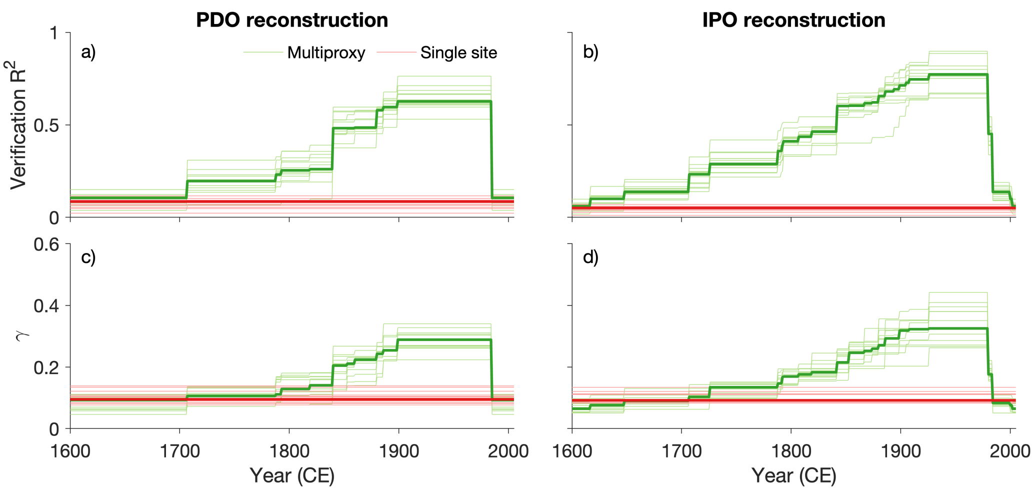

Figure 5. PDV pseudoproxy reconstruction skill. a-b) r^2 and c-d) γ of PDO (left) and IPO (right) pseudoproxy reconstructions using CESM-LME full forcing members. Green denotes multiproxy based reconstruction whereas red denotes single site based reconstruction skill. The thick lines represent median values across 13 ensemble members.

change signal perspective (Hawkins et al., 2020), our results represent small 'signal-to-noise' values in PDV indices over 1966-2005 CE compared to the last millennium. Nonetheless, it is noteworthy that aerosols are not included in our analysis except through volcanism. Prior studies have indicated that aerosols could be partially responsible for PDV shifts in the historical period (Boo et al., 2015; Dittus et al., 2021). However, owing to the lack of constraints of aerosols prior to the historical period, aerosol forcings are kept constant in CESM-LME prior to 1850 CE (Otto-Bliesner et al., 2016). As a result, we are unable to quantify the role of non-volcanic aerosols in the context of the last millennium.

The fact that PDV is primarily internally driven has significant implications on decadal climate predictions. Most importantly, our results imply that future improvements in predicting PDV will be less dependent on how well we can model responses of the Pacific to the forcings analyzed in this study, but more on our understanding of dynamics that can generate PDV and models' ability in representing processes in the Pacific Ocean. As such, future studies should focus on improving our knowledge about mechanisms that generate PDV. This implication is further corroborated by analyses on decadal prediction modeling experiments that suggest external forcing exerts a smaller influence in predicting Pacific than in the Atlantic (Yeager et al., 2018), and that initialization does not provide as much skill in the Pacific as in the Atlantic (Smith et al., 2019; Yeager et al., 2018).

Lastly, our psuedoproxy reconstruction allows us to demonstrate the ability to use multiple marine based proxy records to reconstruct PDV. Prior studies either relied on terrestrial proxy records that were subject to influences of teleconnections or a small number of coral records from a localized region. By using the HydroCoral2k network, we show that there is a potential to improve PDV reconstructions using published proxy records within a pseudoproxy framework. This open avenues to improve our understanding of PDV changes and also calls for more high resolution marine based proxy records to be developed in the Pacific basin.

Understanding the internal and external nature of PDV has significant implications on near term climate predictions. Our work has advanced our understanding about PDV by characterizing spatiotemporal fingerprints of internally driven and externally forced PDV and determining how well current marine based proxies can capture PDV. Although, our study is only based on one climate model, does not include an analysis on the impacts of aerosols, and employs a pseudoproxy framework, these results are helpful to the climate prediction and paleoclimate communities as we continue to try to improve our understanding about decadal variability in the Pacific.

Open Research

CESM-LME data are available on `www.earthsystemgrid.org`. HydroCoral2k data is available at Walter et al. (2022). The code to calculate mutual information and Shannon entropy by our methods will be made available at the Brown Digital Repository with doi provided upon acceptance.

Acknowledgments

We acknowledge the CESM1(CAM5) Last Millennium Ensemble Community Project and supercomputing resources provided by NSF/CISL/Yellowstone. We thank Emanuele Di Lorenzo for helpful comments and suggestions. BFK was supported by NOAA NA19OAR4310366.

References

Bonfils, C., & Santer, B. D. (2011). Investigating the possibility of a human component in various Pacific decadal oscillation indices. *Climate dynamics*, *37*(7), 1457–1468.

Boo, K.-O., Booth, B. B., Byun, Y.-H., Lee, J., Cho, C., Shim, S., & Kim, K.-T. (2015). Influence of aerosols in multidecadal SST variability simulations over the North Pacific. *Journal of Geophysical Research: Atmospheres*, *120*(2), 517–531.

Capotondi, A., Deser, C., Phillips, A., Okumura, Y., & Larson, S. (2020). ENSO and Pacific Decadal Variability in the Community Earth System Model Version 2. *Journal of Advances in Modeling Earth Systems*, e2019MS002022.

Carcassi, G., Aidala, C. A., & Barbour, J. (2021). Variability as a better characterization of Shannon entropy. *European Journal of Physics*, *42*(4), 045102.

Chen, J. J. (2024). *Information-theoretic quantification of predictability in ensemble climate models.* ScB thesis, Applied Math, Brown University, Brown University. Retrieved from `http://fox-kemper.com/pubs/pdfs/Chen24.pdf`

Corrège, T. (2006). Sea surface temperature and salinity reconstruction from coral geochemical tracers. *Palaeogeography, Palaeoclimatology, Palaeoecology*, *232*(2-4), 408–428.

Dai, A., Fyfe, J. C., Xie, S.-P., & Dai, X. (2015). Decadal modulation of global surface temperature by internal climate variability. *Nature Climate Change*, *5*(6), 555–559.

d'Arrigo, R., Villalba, R., & Wiles, G. (2001). Tree-ring estimates of Pacific decadal climate variability. *Climate Dynamics*, *18*(3-4), 219–224.

Davis, R. E., Talley, L. D., Roemmich, D., Owens, W. B., Rudnick, D. L., Toole, J., ... Barth, J. A. (2019). 100 years of progress in ocean observing systems. *Meteorological Monographs*, *59*, 3–1.

Deser, C., Lehner, F., Rodgers, K. B., Ault, T., Delworth, T. L., DiNezio, P. N., ... others (2020). Insights from Earth system model initial-condition large ensembles and future prospects. *Nature Climate Change*, *10*(4), 277–286.

Di Lorenzo, E., Schneider, N., Cobb, K. M., Franks, P., Chhak, K., Miller, A. J., ... others (2008). North Pacific Gyre Oscillation links ocean climate and ecosystem change. *Geophysical Research Letters*, *35*(8).

Di Lorenzo, E., Xu, T., Zhao, Y., Newman, M., Capotondi, A., Stevenson, S., ... others (2023). Modes and mechanisms of pacific decadal-scale variability. *Annual Review of Marine Science*, *15*, 249–275.

Dittus, A. J., Hawkins, E., Robson, J. I., Smith, D. M., & Wilcox, L. J. (2021). Drivers of recent North Pacific decadal variability: the role of aerosol forcing. *Earth's Future*, *9*(12).

Dong, B., & Dai, A. (2015). The influence of the interdecadal Pacific oscillation on temperature and precipitation over the globe. *Climate Dynamics*, *45*(9), 2667–2681.

Dong, L., Zhou, T., & Chen, X. (2014). Changes of Pacific decadal variability in the twentieth century driven by internal variability, greenhouse gases, and aerosols. *Geophysical Research Letters*, *41*(23), 8570–8577.

Du, X., Hendy, I., Hinnov, L., Brown, E., Schimmelmann, A., & Pak, D. (2020). Interannual southern california precipitation variability during the common era and the enso teleconnection. *Geophysical Research Letters*, *47*(1), e2019GL085891. doi: https://doi.org/10.1029/2019GL085891

Epstein, S., Buchsbaum, R., Lowenstam, H. A., & Urey, H. C. (1953). Revised carbonate-water isotopic temperature scale. *Geological Society of America Bulletin*, *64*(11), 1315–1326.

Fasullo, J. T., Phillips, A. S., & Deser, C. (2020, July). Evaluation of leading modes of climate variability in the cmip archives. *Journal of Climate*, *33*(13), 5527–5545. doi: 10.1175/jcli-d-19-1024.1

Felis, T., Suzuki, A., Kuhnert, H., Rimbu, N., & Kawahata, H. (2010). Pacific Decadal Oscillation documented in a coral record of North Pacific winter temperature since

1873. *Geophysical Research Letters*, *37*(14).

Fleming, L. E., & Anchukaitis, K. J. (2016). North Pacific decadal variability in the CMIP5 last millennium simulations. *Climate Dynamics*, *47*(12), 3783–3801.

Gettelman, A., Hannay, C., Bacmeister, J. T., Neale, R. B., Pendergrass, A., Danabasoglu, G., ... others (2019). High climate sensitivity in the community earth system model version 2 (cesm2). *Geophysical Research Letters*, *46*(14), 8329–8337.

Giese, B. S., Seidel, H. F., Compo, G. P., & Sardeshmukh, P. D. (2016). An ensemble of ocean reanalyses for 1815–2013 with sparse observational input. *Journal of Geophysical Research: Oceans*, *121*(9), 6891–6910.

Hawkins, E., Frame, D., Harrington, L., Joshi, M., King, A., Rojas, M., & Sutton, R. (2020). Observed emergence of the climate change signal: From the familiar to the unknown. *Geophysical Research Letters*, *47*(6), e2019GL086259. doi: https://doi.org/10.1029/2019GL086259

Hegerl, G., & Zwiers, F. (2011). Use of models in detection and attribution of climate change. *WIREs Climate Change*, *2*(4), 570-591. doi: https://doi.org/10.1002/wcc.121

Henley, B. J., Gergis, J., Karoly, D. J., Power, S., Kennedy, J., & Folland, C. K. (2015). A tripole index for the interdecadal Pacific oscillation. *Climate Dynamics*, *45*(11-12), 3077–3090.

Hua, W., Dai, A., & Qin, M. (2018). Contributions of internal variability and external forcing to the recent Pacific decadal variations. *Geophysical Research Letters*, *45*(14), 7084–7092.

Jungclaus, J. H., Bard, E., Baroni, M., Braconnot, P., Cao, J., Chini, L. P., ... Zorita, E. (2017). The PMIP4 contribution to CMIP6 – Part 3: The last millennium, scientific objective, and experimental design for the PMIP4 *past1000* simulations. *Geoscientific Model Development*, *10*(11), 4005–4033. Retrieved from `https://gmd.copernicus.org/articles/10/4005/2017/` doi: 10.5194/gmd-10-4005-2017

Kosaka, Y., & Xie, S.-P. (2016). The tropical Pacific as a key pacemaker of the variable rates of global warming. *Nature Geoscience*, *9*(9), 669–673.

Leroux, S., Penduff, T., Bessières, L., Molines, J.-M., Brankart, J.-M., Sérazin, G., ... Terray, L. (2018, January). Intrinsic and atmospherically forced variability of the amoc: Insights from a large-ensemble ocean hindcast. *Journal of Climate*, *31*(3), 1183–1203. doi: 10.1175/jcli-d-17-0168.1

Linsley, B. K., Wu, H. C., Dassié, E. P., & Schrag, D. P. (2015). Decadal changes in South

Pacific sea surface temperatures and the relationship to the Pacific decadal oscillation and upper ocean heat content. *Geophysical Research Letters*, *42*(7), 2358–2366.

Liu, Z. (2012). Dynamics of interdecadal climate variability: A historical perspective. *Journal of Climate*, *25*(6), 1963–1995.

Llovel, W., Penduff, T., Meyssignac, B., Molines, J., Terray, L., Bessières, L., & Barnier, B. (2018, December). Contributions of atmospheric forcing and chaotic ocean variability to regional sea level trends over 1993–2015. *Geophysical Research Letters*, *45*(24). doi: 10.1029/2018gl080838

Maher, N., Gupta, A. S., & England, M. H. (2014). Drivers of decadal hiatus periods in the 20th and 21st centuries. *Geophysical Research Letters*, *41*(16), 5978–5986.

Mantua, N. J., & Hare, S. R. (2002). The pacific decadal oscillation. *Journal of oceanography*, *58*(1), 35–44.

McCabe, G. J., Palecki, M. A., & Betancourt, J. L. (2004). Pacific and Atlantic Ocean influences on multidecadal drought frequency in the United States. *Proceedings of the National Academy of Sciences*, *101*(12), 4136–4141.

Meehl, G. A., Hu, A., Arblaster, J. M., Fasullo, J., & Trenberth, K. E. (2013). Externally forced and internally generated decadal climate variability associated with the Interdecadal Pacific Oscillation. *Journal of Climate*, *26*(18), 7298–7310.

Meehl, G. A., Hu, A., Santer, B. D., & Xie, S.-P. (2016). Contribution of the Interdecadal Pacific Oscillation to twentieth-century global surface temperature trends. *Nature Climate Change*, *6*(11), 1005–1008.

Newman, M., Alexander, M. A., Ault, T. R., Cobb, K. M., Deser, C., Di Lorenzo, E., ... others (2016). The Pacific decadal oscillation, revisited. *Journal of Climate*, *29*(12), 4399–4427.

O'Mara, N. A., Cheung, A. H., Kelly, C. S., Sandwick, S., Herbert, T. D., Russell, J. M., ... Herguera, J. C. (2019). Subtropical Pacific Ocean Temperature Fluctuations in the Common Era: Multidecadal Variability and Its Relationship With Southwestern North American Megadroughts. *Geophysical Research Letters*, *46*(24), 14662-14673.

Otto-Bliesner, B. L., Brady, E. C., Fasullo, J., Jahn, A., Landrum, L., Stevenson, S., ... Strand, G. (2016). Climate variability and change since 850 CE: An ensemble approach with the Community Earth System Model. *Bulletin of the American Meteorological Society*, *97*(5), 735–754.

Porter, S. E., Mosley-Thompson, E., Thompson, L. G., & Wilson, A. B. (2021). Reconstructing an Interdecadal Pacific Oscillation Index from a Pacific Basin–Wide Collection of Ice Core Records. *Journal of Climate*, *34*(10), 3839–3852.

Power, S., Casey, T., Folland, C., Colman, A., & Mehta, V. (1999). Inter-decadal modulation of the impact of ENSO on Australia. *Climate Dynamics*, *15*(5), 319–324.

Roemmich, D., John Gould, W., & Gilson, J. (2012). 135 years of global ocean warming between the Challenger expedition and the Argo Programme. *Nature Climate Change*, *2*(6), 425–428.

Sane, A., Fox-Kemper, B., Ullman, D., Kincaid, C., & Rothstein, L. (2021). Consistent predictability of the Ocean State Ocean Model (OSOM) using information theory and flushing timescales. *Journal of Geophysical Research – Oceans*, *126*(7), e2020JC016875. doi: 10.1029/2020JC016875

Sane, A., Fox-Kemper, B., & Ullman, D. (2024). Internal vs forced variability metrics for geophysical flows using information theory. *Journal of Geophysical Research–Oceans*. Retrieved from `https://fox-kemper.com/pubs/pdfs/SaneFox-Kemper23.pdf` (Accepted)

Sigl, M., Winstrup, M., McConnell, J. R., Welten, K. C., Plunkett, G., Ludlow, F., ... others (2015). Timing and climate forcing of volcanic eruptions for the past 2,500 years. *Nature*, *523*(7562), 543–549.

Smerdon, J. E. (2012). Climate models as a test bed for climate reconstruction methods: pseudoproxy experiments. *WIREs Climate Change*, *3*(1), 63-77. doi: https://doi.org/10.1002/wcc.149

Smith, D., Eade, R., Scaife, A. A., Caron, L.-P., Danabasoglu, G., DelSole, T., ... others (2019). Robust skill of decadal climate predictions. *Npj Climate and Atmospheric Science*, *2*(1), 1–10.

Stevenson, S., Fox-Kemper, B., Jochum, M., Rajagopalan, B., & Yeager, S. G. (2010). Enso model validation using wavelet probability analysis. *Journal of Climate*, *23*(20), 5540–5547.

Stevenson, S., Otto-Bliesner, B., Brady, E., Nusbaumer, J., Tabor, C., Tomas, R., ... Liu, Z. (2019). Volcanic eruption signatures in the isotope-enabled last millennium ensemble. *Paleoceanography and Paleoclimatology*, *34*(8), 1534–1552.

Sun, W., Liu, J., Wang, B., Chen, D., & Gao, C. (2022). Pacific multidecadal (50–70 year) variability instigated by volcanic forcing during the Little Ice Age (1250–1850). *Climate Dynamics*. doi: 10.1007/s00382-021-06127-7

Walter, R. M., Sayani, H. R., Felis, T., Cobb, K. M., Abram, N. J., Arzey, A. K., ... PAGES CoralHydro2k Project Members (2022). *NOAA/WDS Paleoclimatology - CoralHydro2k Database (Common Era coral d18O and Sr/Ca data compilation)* [dataset]. NOAA National Centers for Environmental Information. doi: 10.25921/YP94-V135

Walter, R. M., Sayani, H. R., Felis, T., Cobb, K. M., Abram, N. J., Arzey, A. K., ... the PAGES CoralHydro2k Project Members (2023). The CoralHydro2k database: a global, actively curated compilation of coral δ^{18}O and Sr / Ca proxy records of tropical ocean hydrology and temperature for the Common Era. *Earth System Science Data*, *15*(5), 2081–2116. doi: 10.5194/essd-15-2081-2023

Wang, T., Otterå, O. H., Gao, Y., & Wang, H. (2012). The response of the North Pacific Decadal Variability to strong tropical volcanic eruptions. *Climate Dynamics*, *39*(12), 2917–2936.

Watanabe, S. (1960). Information theoretical analysis of multivariate correlation. *IBM Journal of research and development*, *4*(1), 66–82.

Wilson, R., Cook, E., D'Arrigo, R., Riedwyl, N., Evans, M. N., Tudhope, A., & Allan, R. (2010). Reconstructing ENSO: the influence of method, proxy data, climate forcing and teleconnections. *Journal of Quaternary Science*, *25*(1), 62-78. doi: https://doi.org/10.1002/jqs.1297

Wilson, R., Tudhope, A., Brohan, P., Briffa, K., Osborn, T., & Tett, S. (2006). Two-hundred-fifty years of reconstructed and modeled tropical temperatures. *Journal of Geophysical Research: Oceans*, *111*(C10). doi: https://doi.org/10.1029/2005JC003188

Yeager, S. G., Danabasoglu, G., Rosenbloom, N., Strand, W., Bates, S., Meehl, G., ... others (2018). Predicting near-term changes in the Earth System: A large ensemble of initialized decadal prediction simulations using the Community Earth System Model. *Bulletin of the American Meteorological Society*, *99*(9), 1867–1886.

Zanchettin, D., Rubino, A., Matei, D., Bothe, O., & Jungclaus, J. (2013). Multidecadal-to-centennial SST variability in the MPI-ESM simulation ensemble for the last millennium. *Climate dynamics*, *40*(5-6), 1301–1318.

Zhang, L., & Delworth, T. L. (2015). Analysis of the characteristics and mechanisms of the Pacific decadal oscillation in a suite of coupled models from the Geophysical Fluid Dynamics Laboratory. *Journal of Climate*, *28*(19), 7678–7701.

Zhang, Y., Wallace, J. M., & Battisti, D. S. (1997). ENSO-like interdecadal variability:

1900–93. *Journal of Climate*, *10*(5), 1004–1020.
Zhu, J., Poulsen, C. J., & Otto-Bliesner, B. L. (2020). High climate sensitivity in cmip6 model not supported by paleoclimate. *Nature Climate Change*, *10*(5), 378–379.

Supporting Information for "Internally Driven and Externally Forced Pacific Decadal Variability in the CESM Last Millennium Ensemble"

Anson H. Cheung[1]*, Aakash Sane[1]†, Baylor Fox-Kemper[1]

[1]Department of Earth, Environmental, and Planetary Sciences, Brown University, Providence, RI 02912, USA

Contents of this file

Corresponding author: A. H. Cheung, Lamont-Doherty Earth Observatory, Columbia University, Palisades, NY 10964. (acheung@ldeo.columbia.edu)

*Current address: Lamont-Doherty Earth Observatory, Columbia University, Palisades, NY 10964

†Current address: Atmospheric and Oceanic Sciences, Princeton University, Princeton, NJ 08544

Text S1. PDO and TPI definitions

In this study, we relied on PDO and TPI to understand temporal characteristics of PDV. PDO is a principal component based index that is commonly used to characterize decadal SST variability in the north Pacific (Zhang et al., 1997), whereas TPI is an index that does not rely on principal components and is used to characterize basin-wide decadal SST variability (Henley et al., 2015). The PDO is defined as the first principal component of north Pacific ($20 - 70°$ N; $110 - 260°$ E) SST (Zhang et al., 1997), which was calculated before any filtering beyond the removal of the seasonal cycle. To avoid contamination from external forcings and allow direct comparison of the temporal characteristics of PDO between different forcing experiments, we calculated PDO by first defining their empirical orthogonal functions (EOFs) using the 850 CE pre-industrial simulation of CESM1.1, and then projecting them onto forced simulations. The seasonal cycle and global mean SST anomaly were removed from the preindustrial simulation latitudinally-weighted SST field prior to calculation. For other simulations, the latitudinally-weighted SST field was mean centered and the seasonal cycle was removed consistent with the preindustrial simulation training. The TPI is an index that does not rely on principal components and is used to characterize IPO (Henley et al., 2015). TPI is calculated by first removing the seasonal cycle from the temperature field, then computing the difference between temperature average of the eastern equatorial Pacific and the average temperature of the northwest and southwest Pacific (see Equation (1) in Henley et al. (2015) for details), and applying a 13-year low pass filter to the timeseries.

Text S2. Shannon Entropy and Mutual Information

We used Shannon Entropy and Mutual Information to quantify the relative role of internal variability and external variability on total PDV variability. Shannon Entropy is defined as (Shannon, 1948):

$$H = \sum_{i=1}^{N} p_i \log_2(1/p_i) \tag{1}$$

where H is the Shannon entropy with units of bits, and p_i is the probability of the i^{th} outcome. The factor $\log_2(1/p_i)$ measures the uncertainty shown by a variable or the freedom that a variable has in visiting different combinations of the N bins.

Mutual Information is defined as (Shannon, 1948):

$$I = \sum_{j=1}^{N} \sum_{i=1}^{N} p_{ij} \log_2 \left(\frac{p_{ij}}{p_i p_j} \right) \tag{2}$$

where p_{ij} is the joint probability of i^{th} outcome of x and j^{th} outcome of y. The marginal probability of i^{th} and j^{th} outcomes of x and y respectively are p_i and p_j. If the distributions are statistically independent, then $p_{ij} = p_i p_j$ and $I = 0$. Alternatively, if x and y are identical, then $p_{ij} = p_i = p_j$ and $I = H$.

Because the probability distribution in Shannon entropy and mutual information is estimated based on a binning procedure and the results are somewhat sensitive to binning choices, it is important to use an objective criterion to avoid biasing the probability estimates. Numerous techniques have been proposed to obtain optimal binning for precise measurements of information entropy (Papana & Kugiumtzis, 2008). Here, we derived the histogram for Shannon entropy based on equidistant partitioning and determined the bin width based on the Freedman-Diaconis rule. The Freedman-Diaconis rule estimates the bin width by assuming the underlying distribution is Gaussian (Freedman & Diaconis, 1981). This bin width was then also used in marginal and joint probability distributions when calculating mutual information. Although a previous study showed the information theory metrics used here are sensitive to the uncertainty in bin width (Sane et al., 2024), slight deviation of the bin width from the optimal binning estimate does not appreciably change the ratio shown in Equation 1 in the main text, and the (Freedman & Diaconis, 1981) method seems to have the lowest bin uncertainty for geophysical applications (Chen, 2024). Until further research is done to improve the estimation of various entropies for data in the climate sciences, histogram based estimation is a reasonable approximation for practical purposes.

Text S3. Pseudoproxy Reconstruction

We carried out pseudoproxy reconstructions of annual PDO and IPO indices using the CESM-LME full forcing simulations to understand and quantify the advantage of a multiproxy record over a single proxy record based PDV reconstruction. To do so, we first identified coral records from the HydroCoral2k database (Walter et al., 2023) that are at least 80 years long, have an annual or higher temporal resolution, and are located in the North Pacific basin for PDO and the Pacific basin for IPO (Figure 1). Then, we isolated the model grid cell nearest to the proxy records and truncated these data points based on the temporal length of the proxy records. Since some proxy records do not have monthly resolution, we averaged the monthly timeseries that were isolated into annual timeseries. Afterwards, following prior pseudoproxy experiments, we added white noise (σ=2) into each annual timeseries to mimic characteristics of proxy records. Lastly, we used a nested composite-plus-scaling approach (e.g., Wilson et al., 2006, 2010) to reconstruct annual PDO and annual IPO.

References

Chen, J. J. (2024). *Information-theoretic quantification of predictability in ensemble climate models.* ScB thesis, Applied Math, Brown University, Brown University. Retrieved from http://fox-kemper.com/pubs/pdfs/Chen24.pdf

Freedman, D., & Diaconis, P. (1981). On the histogram as a density estimator: L 2 theory. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, *57*(4), 453–476.

Henley, B. J., Gergis, J., Karoly, D. J., Power, S., Kennedy, J., & Folland, C. K. (2015). A tripole index for the interdecadal Pacific oscillation. *Climate Dynamics*, *45*(11-12), 3077–3090.

Papana, A., & Kugiumtzis, D. (2008). Evaluation of mutual information estimators on nonlinear dynamic systems. *Nonlinear Phenomena in Complex Systems*.

Sane, A., Fox-Kemper, B., & Ullman, D. (2024). Internal vs forced variability metrics for geophysical flows using information theory. *Journal of Geophysical Research–Oceans*. Retrieved from https://fox-kemper.com/pubs/pdfs/SaneFox-Kemper23.pdf (Accepted)

Shannon, C. E. (1948). A mathematical theory of communication. *The Bell system technical journal*, *27*(3), 379–423.

Walter, R. M., Sayani, H. R., Felis, T., Cobb, K. M., Abram, N. J., Arzey, A. K., ... the PAGES CoralHydro2k Project Members (2023). The CoralHydro2k database: a global, actively curated compilation of coral δ^{18}O and Sr / Ca proxy records of tropical ocean hydrology and temperature for the Common Era. *Earth System Science Data*, *15*(5), 2081–2116. doi: 10.5194/essd-15-2081-2023

Wilson, R., Cook, E., D'Arrigo, R., Riedwyl, N., Evans, M. N., Tudhope, A., & Allan, R. (2010). Reconstructing ENSO: the influence of method, proxy data, climate forcing and teleconnections. *Journal of Quaternary Science*, *25*(1), 62-78. doi: https://doi.org/10.1002/jqs.1297

Wilson, R., Tudhope, A., Brohan, P., Briffa, K., Osborn, T., & Tett, S. (2006). Two-hundred-fifty years of reconstructed and modeled tropical temperatures. *Journal of Geophysical Research: Oceans*, *111*(C10). doi: https://doi.org/10.1029/2005JC003188

Zhang, Y., Wallace, J. M., & Battisti, D. S. (1997). ENSO-like interdecadal variability: 1900–93. *Journal of Climate*, *10*(5), 1004–1020.