AGU ADVANCING EARTH AND SPACE SCIENCE

Geophysical Research Letters®

RESEARCH LETTER
10.1029/2021GL094707

Detecting Climate Teleconnections With Granger Causality

Filipi N. Silva[1], Didier A. Vega-Oliveros[2,3], Xiaoran Yan[4], Alessandro Flammini[3], Filippo Menczer[1,3], Filippo Radicchi[3], Ben Kravitz[5,6], and Santo Fortunato[1,3]

[1]Indiana University Network Science Institute (IUNI), Bloomington, IN, USA, [2]Institute of Computing, University of Campinas, Campinas, SP, Brazil, [3]Center for Complex Networks and Systems Research, Luddy School of Informatics, Computing, and Engineering, Indiana University, Bloomington, IN, USA, [4]Artificial Intelligence Research Institute, Zhejiang Lab, Zhejiang, China, [5]Department of Earth and Atmospheric Sciences, Indiana University, Bloomington, IN, USA, [6]Atmospheric Sciences and Global Change Division, Pacific Northwest National Laboratory, Richland, WA, USA

Abstract Climate system teleconnections are crucial for improving climate predictability, but difficult to quantify. Standard approaches to identify teleconnections are often based on correlations between time series. Here we present a novel method leveraging Granger causality, which can infer/detect relationships between any two fields. We compare teleconnections identified by correlation and Granger causality at different timescales. We find that both Granger causality and correlation consistently recover known seasonal precipitation responses to the sea surface temperature pattern associated with the El Niño Southern Oscillation. Such findings are robust across multiple time resolutions. In addition, we identify candidates for unexplored teleconnection responses.

Plain Language Summary Teleconnections, or climate responses far away from a perturbation, are difficult to study. We have developed a new method of quantifying teleconnection strengths and identifying new teleconnections. When testing this method on global precipitation responses to El Niño surface temperature changes, not only does our method do a good job of recovering previous findings, but we identify several new possible teleconnections. Our method can be applied to any two climate variables and could be useful for attribution studies, for example changes in frequency of extreme events.

Didier A. Vega-Oliveros and Filipi Nascimento Silva contributed equally to this work.

Key Points:

- We propose a new method for identifying significant teleconnections between any two climate fields
- We test this method on Niño 3.4 temperature teleconnections to precipitation anomalies everywhere
- The method recovers previously known El Niño teleconnections and discovers potential new ones

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

S. Fortunato and B. Kravitz, santo.fortunato@gmail.com; bkravitz@iu.edu

Citation:

Silva, F. N., Vega-Oliveros, D. A., Yan, X., Flammini, A., Menczer, F., Radicchi, F., et al. (2021). Detecting climate teleconnections with Granger causality. *Geophysical Research Letters*, *48*, e2021GL094707. https://doi.org/10.1029/2021GL094707

Received 9 JUN 2021
Accepted 22 AUG 2021

1. Introduction

Climate system teleconnections are responses in one region of the globe to perturbations far away. The most well-known and well-studied teleconnection stems from the El Niño Southern Oscillation (ENSO): A warmer or cooler anomaly condition in the tropical Pacific Ocean that results in changes in seasonal weather patterns in (for example) North America and Asia (Yuan et al., 2018). Other well-studied teleconnections include correlations between midlatitude temperature features that are mediated by Rossby waves (Hoskins & Karoly, 1981) and potential links from Arctic sea ice loss to midlatitude winter storms (Cohen et al., 2014).

In all cases, there must be physical mechanisms that link the forcing in one region to a response in another. In the case of the non-regular and complex ENSO fluctuations, equatorially propagating Kelvin waves trigger the event, leading to numerous downstream effects that are mediated by other modes of variability, including the Pacific Meridional Mode, the Pacific Decadal Oscillation, and the Indian Ocean Dipole (Chang et al., 2007; Chiang & Vimont, 2004; Hare & Mantua, 2002; Lorenzo et al., 2015; McPhaden, 2015; Stuecker, 2018; Stuecker et al., n.d.; Timmerman et al., 2018; Trenberth, 2019; Yuan et al., 2018). In addition, there is evidence of ENSO teleconnections via stratospheric mechanisms (Domiesen et al., 2019). Uncovering these mechanisms can be a time-consuming, research-intensive process. Methods that have been explored in previous studies include empirical orthogonal functions (Liu et al., 2002), Green's function approaches (Harrop et al., 2018; Hill & Ming, 2012; Liu et al., 2020), the fluctuation-dissipation theorem (Fuchs et al., 2015), and system identification. Each of these methods has advantages and disadvantages and situations in which they are most useful.

A standard technique to detect teleconnections (prior to uncovering the underlying mechanisms) relies on computing correlations between time series of climate variables in different locations (Donner et al., 2017). Typically the world map is turned into a latitude-longitude grid, and each grid cell is represented by time series of variables measured in that location, for example, temperature, precipitation, or pressure. Then, correlations between time series in separate locations are computed (Boers et al., 2019; Fan et al., 2017; Wang et al., 2013; Yamasaki et al., 2008). Most methods in the literature rely on computing linear correlation coefficients. Nonlinear scores, like mutual information (Kantz & Schreiber, 2004a), do not alter the picture substantially (Donges et al., 2009). The ensuing relationships of strongly correlated time series can be interpreted as the links of a climate network, with the grid cells or locations as nodes (Boers et al., 2019; De Castro Santos et al., 2020; Donges et al., 2009; Donner et al., 2017; Fan et al., 2017; Guez et al., 2012; Steinhaeuser et al., 2011; Tsonis & Swanson, 2008; Wang et al., 2013; Yamasaki et al., 2008). Structural features of climate networks have been associated to critical climate phenomena, like ENSO (Fan et al., 2017; Yamasaki et al., 2008) and Rossby waves (Boers et al., 2019; Wang et al., 2013).

Although correlation-based methods have yielded useful results, they do not necessarily provide causal relationships among phenomena. They also suffer from problems related to interpretability (Fenton & Neil, 2012); for instance, low correlation values can lead to significant p-values depending on the length of the time series (for example, see Text S1). When cross-correlation is employed, a peak in a correlogram may be spurious (Damos, 2016) — for example, it can be a consequence of processes with feedback mechanisms (Chatfield & Xing, 2019), which are common in climate science. Moreover, one can observe high computed cross-correlation peaks even among any two time series due to serial autocorrelation within each time series, which is another common feature in geophysical time series (Falasca et al., 2019; McGraw & Barnes, 2018).

Here we explore a new method of measuring the strength of teleconnections by employing Granger causality (Granger, 1969), a concept that has gained increasing currency in climate science (McGraw & Barnes, 2018; Papagiannopoulou et al., 2017; Triacca, 2005) and differs from standard correlation in several important ways. First, Granger causality is defined explicitly as a statistical test to assess the predictive relationship of one time series with another. Thus, it has a clear interpretation that depends only on the choice of the null models. It also satisfies the two traditional assumptions for causal relationships: given two events such that A causes B, A must happen before B; and A must provide unique information needed to predict B, thus implying an information flow from A to B. Finally, Granger causality can be extended to a multivariate framework to incorporate multiple time series into the model.

In this work we employ Granger causality to uncover potential causal relationships between temperature anomalies in the tropical Pacific and precipitation anomalies in the rest of the globe (Figure 1). We show that our approach is able to recover known ENSO teleconnections and to identify candidates for new ones. Our methodology is robust and represents a powerful new tool to investigate how remote locations influence each other's climate.

2. Methods

The steps in the methodology, which we describe in this section, are illustrated in Figure 1.

2.1. Data Set

We used hourly near-surface air temperature and precipitation data from the MERRA-2 reanalysis project (Molod et al., 2015). The data set covers the period from 1980 to 2018, with a resolution of 0.5° latitude × 0.625° longitude, for a total of 207,936 grid cells. Of these, we only consider grid cells between 80°N and 80° S. We then reduced the spatial resolution and sample frequency by transforming the data set to a resolution of 1.0° × 1.25° and temporally averaging to produce daily mean values, resulting in 365 measurements of each variable for each of the resulting 45,792 cells for each year. In the case of leap-years, we averaged the daily values of the 28th and 29th of February so all years have the same length. Precipitation data tends to be highly skewed, with many very small values and some values orders of magnitude larger. Therefore we re-scaled the precipitation time-series by applying a logarithmic function to all values.

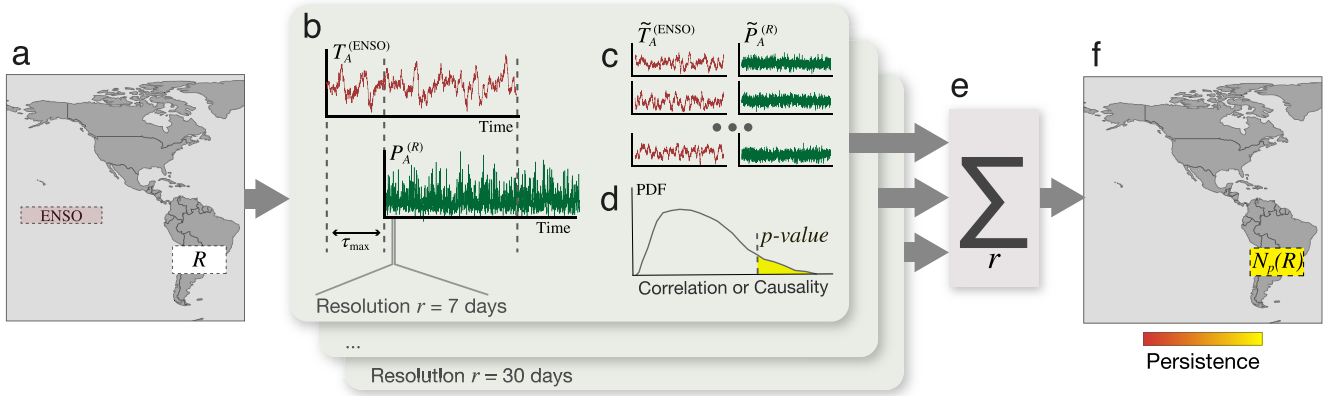

Figure 1. Graphical description of the methodology used in this study (also see the Methods section). Surface air temperature data in the Niño 3.4 region (a) is used to perform lagged correlation or Granger causality analysis against precipitation data throughout the world, up to a maximum lag τ_{max}. For each considered time resolution r (varying from 7 to 30 days), we calculate the lagged correlation or Granger causality between anomalies in precipitation ($P_A^{(R)}$) in a particular region R and the El Niño Southern Oscillation (ENSO) temperature ($T_A^{(ENSO)}$) (b). The significance of the result is found considering surrogates of both time series, that is, $\tilde{T}_A^{(ENSO)}$ and $\tilde{P}_A^{(R)}$ (c), and then calculating the respective p-value for the R region (d). For a given threshold $p_{thres.}$ we define that region R has significant relationship with the ENSO if the corresponding p-value is smaller than $p_{thres.}$. Significant regions are aggregated and counted across the considered resolution values to measure persistence (e). Finally, colors on the maps represent regions with significant p-values that are persistent across r (f).

We removed the seasonal cycle from the data by averaging temperature and precipitation values for each of the 365 calendar days. We considered the interval from January 1, 1980 to February 28, 2018 as the climatic period for which the long-term averages were computed. The anomalies are then obtained by subtracting for each day the respective average temperature or precipitation from the climatic period. For example, the January 1, 1998 anomaly is computed as the value for that day minus the average of all January 1 values between 1980 and 2018.

Due to climate change over this period, the obtained anomaly time series are not sufficiently stationary in time, in that the averages in specific periods are not stable. Since our methods require approximate stationarity, we detrended the anomaly time series at each grid point by employing ordinary least squares regression of order 2 on the time series at each point. We then subtract the fitted curve from the anomaly data. We finally smoothed the results across time by taking weekly averages of the temperature or precipitation anomalies to remove high frequency weather noise.

2.2. Lagged Correlation

Cross-correlation (also sometimes called lagged correlation) methods have been commonly used in climate science to find relationships between time series (Boers et al., 2019; Falasca et al., 2019; Fan et al., 2017). These techniques are based on calculating the Pearson correlation ρ_τ between two time series $x(t)$ and $y(t)$ shifted by a time lag τ. One approach to find the highest correlation across all the lags is by defining an upper-bound time lag τ_{max} and calculating the maximum Pearson correlation across the lags as

$$C_{\tau_{max}} = \max\{|\rho_\tau(x(t), y(t+\tau))| : 0 \leq \tau \leq \tau_{max}\}. \quad (1)$$

This measure has been used as a way to construct a network of teleconnections among different regions and time series around the globe (Donner et al., 2017). We employ this cross-correlation as one of our methodologies to measure the relationships between temperature anomalies in the ENSO region and precipitation in each grid box.

Correlation (and cross-correlation) can be greatly influenced by the presence of outliers and by the overall distributions of the time series across the time and frequency domains. Because of that, we employ a null-model of surrogate instances generated for each pair of the original time series to serve as the basis to calculate the significance of their correlations (e.g., Arizmendi et al., 2014), an approach also known in statistics as bootstrapping (Kantz & Schreiber, 2004b). This is accomplished by calculating the p-value as a measure of significance among the surrogate pairs. Here, we obtained the surrogates using the refined

Amplitude Adjusted Fourier Transform (AAFT) method (Theiler et al., 1992). This technique shuffles the time series by randomizing the phases while preserving the distribution of magnitudes across the frequency spectrum. It is based on applying the Fourier transform to the data and separating the phase from the spectrum of magnitudes. Next the phase spectrum is replaced by uniformly distributed values. These surrogate data are then reassembled by means of the inverse Fourier transform. An extra iterative correction step is used to also preserve the probability distribution of the time series, for example, that average or extreme temperature and precipitation events are occurring with the same frequency in all realizations. Next, the p -value is calculated from the survival function of the surrogate distribution, more specifically, the area under the surrogate distribution curve considering all correlation values higher than $C_{\tau_{\max}}$.

2.3. Granger Causality

The bivariate Granger causality (Damos, 2016; Hamilton, 1994) is defined as a causality test between two time series $x(t)$ and $y(t)$ according to a linear autoregressive model. If the inclusion of $y(t-\tau)$ to a linear predictive model significantly improves the prediction of $x(t)$ we say that y Granger-causes (G-causes) x. More specifically, the test is defined in terms of the linear relationships between $x(t)$ and the lagged time series $x(t-\tau)$ and $y(t-\tau)$ for lags τ varying from 1 to a maximum value $\tau_{\max}$. Two linear models are taken into consideration: the complete model, defined by

$$x(t) = \alpha_0 + \sum_{\tau=1}^{\tau_{\max}} \alpha_\tau x(t-\tau) + \sum_{\tau=1}^{\tau_{\max}} \beta_\tau y(t-\tau) + \epsilon_c(t), \tag{2}$$

and the restricted model,

$$x(t) = \gamma_0 + \sum_{\tau=1}^{\tau_{\max}} \gamma_\tau x(t-\tau) + \epsilon_r(t), \tag{3}$$

where α_τ, β_τ, and γ_τ are constants that can be determined from the data by using ordinary least squares, and ϵ_c and ϵ_r correspond to the residuals of the complete and restricted models respectively.

The restricted model can be regarded as a null model for the hypothesis of $x(t)$ having no dependence on $y(t-\tau)$ (i.e., $\beta_\tau = 0$ for any τ) in the bivariate analysis. The performance of the two models can then be compared in terms of their residuals. Thus, y G-causes x ($y \rightarrow x$) when the sum of the squared residuals for the complete model $R_c = \sum_t [\epsilon_c(t)]^2$ is significantly smaller than those observed for the restricted model, $R_r = \sum_t [\epsilon_r(t)]^2$.

For the Granger causality analysis, we employed a similar pipeline to the lagged correlation analysis. However, to accelerate its computation, instead of applying the significance test based on surrogate time series, we directly performed an F-test with χ^2 asymptotic approximation (Hamilton, 1994) based on the sum of the residuals. In this test, a $\chi^2(\tau_{\max})$ distribution is used to approximate the null model distributions of

$$F = (N - \tau_{\max}) \frac{R_r - R_c}{R_c} \overset{\text{null}}{\approx} \chi^2(\tau_{\max}), \tag{4}$$

where N is the number of points in the time series. Finally, the p-values can be calculated from the survival distribution of $\chi^2(\tau_{\max})$. As for lagged correlation, we select the more significant relationships by filtering the p-values in the higher percentile of the distribution. This way, we obtain a set with the most significant locations (grid cells) impacted by ENSO.

2.4. Persistence

It is possible to obtain spurious results that appear to be statistically significant but only for a single time resolution. To determine the robustness of our results to changes in the resolution of the time series, we introduce a persistence metric. For each method (lagged correlation and Granger causality), we compute at each grid point whether there is a statistically significant relationship between ENSO temperature and precipitation in that grid box, for different resolutions r. We considered values of r ranging between 7 and 30 days to remove short-term weather noise while retaining a sufficient amount of data. We define persistence as the number of instances (choices of resolutions) that result in a statistically significant relationship (minimum of 5). Figure 2 compares the persistence of the two methods: Yellow values indicate equal persistence for lagged

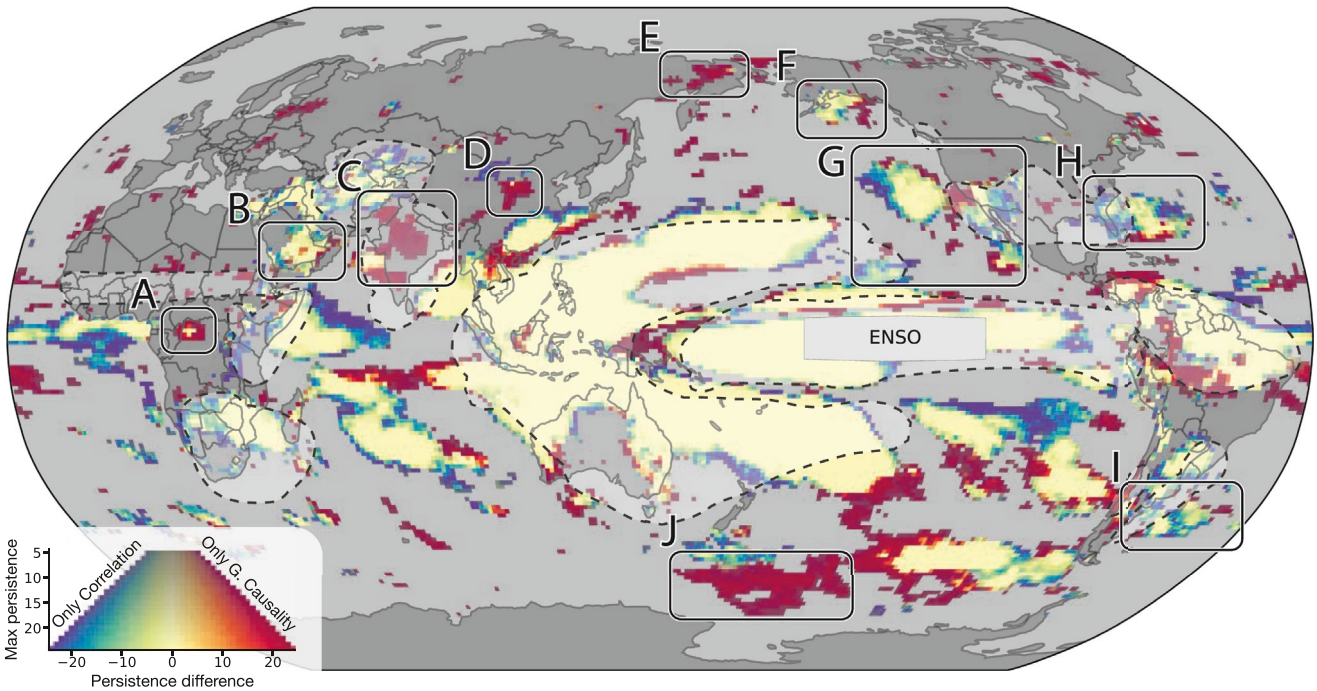

Figure 2. Teleconnections between temperature in the Niño 3.4 region and precipitation everywhere, following the methodology described in Figure 1 for Granger causality and lagged correlation, each with 90-day maximum lags. Colors indicate the persistence (see text): differences in the number of instances (choices of time resolution between 7 and 30 days) for which there is a statistically significant result ($p \leq 0.01$) at each location, with a minimum of 5. Yellow values indicate that lagged correlation and Granger causality are statistically significant for an equal number of instances. Blue values indicate that lagged correlation has more statistically significant instances than Granger causality, and vice versa for red values. The boxes inside the panels show some detected regions that are different between the methods. Light shading indicates ground-truth regions (IRI, 2021a; IRI, 2021b). Maps for each considered decade can be found in Figure S7 as well as a discussion on the major events across the decades (Medhaug et al., 2017; Robock, 2000; Slingo & Annamalai, 2000) in Text S3.

correlation and Granger causality. Blue/red values indicate higher persistence for lagged correlation/Granger causality, respectively. See Text S1 for a comparison between the Granger Causality persistence calculated using the surrogate-based null-model and the approximated distribution.

3. Results

We consider as "ground truth" a number of regions where teleconnections have been reported in prior literature; the regions were digitized from IRI (2021a, 2021b), which are composites of results beginning with Ropelewski and Halpert (1987). Most of these regions (shaded areas in Figure 2) are well captured by both Granger causality (90-day maximum lag) and cross-correlation (also for 90-day maximum lag). Qualitatively, the two methods perform equally well in many of the ground-truth regions, such as the Equatorial Pacific, the horseshoe-shaped pattern in the Western and Central Pacific, the Australian Coasts, the Amazon, Southern Africa, and the Horn of Africa. Both methods tend to miss ground-truth ENSO features in the Sahel, but such relationships are variable and not robust (Janicot et al., 2001). There are several regions (boxes in Figure 2) where performance differs between lagged correlation and Granger causality (Also see Text S1).

The response over the Indian subcontinent (Region C) is different for the two methods, with only Granger causality recovering a statistically significant, persistent result for $p \leq 0.01$. Indian Summer Monsoon Rainfall (ISMR) and ENSO are known to be weakly anticorrelated; the magnitude of correlation is often much less than 0.5 (Ashok et al., 2019; Hrudya et al., 2020). Moreover, the correlation between ISMR and ENSO may have been weakening in recent years, possibly due to effects from increased greenhouse gas concentrations (Kumar et al., 1999), although stochastic variability could also cause an apparent weakening in correlation (Yun & Timmermann, 2018). There is little consensus about whether there is a strong relationship between ISMR and ENSO (Hrudya et al., 2020). Nevertheless, the Granger causality method does reveal a strong relationship between Niño 3.4 temperature and precipitation in central India. We hypothesize that

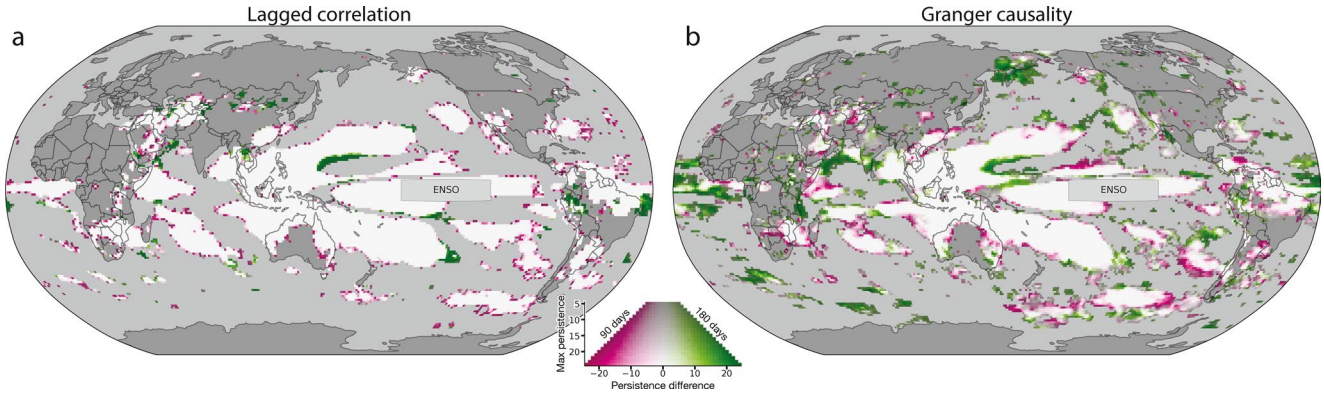

Figure 3. Maps of persistence differences between maximum lags of 90 and 180 days for lagged correlation (a) and Granger causality (b). These maps were obtained for regions with $p < 0.01$.

part of the difficulty in determining whether there is a robust relationship between ISMR and ENSO is because correlation appears not to be an effective diagnostic tool in this case.

For Western North America and the nearby Eastern Pacific (Region G in Figure 2), the picture is more mixed, with some regions showing a stronger response under lagged correlation and some under Granger causality. While there is a long history relating El Niño events and precipitation in western North America, especially California (Anrews et al., 2004; Schonher & Nicholson, 1989), as well as changes in California precipitation due to changing ENSO characteristics in response to climate change (Wang et al., 2014; Yoon et al., 2015), recent studies have cast doubt on the robustness of that relationship (O'Brien et al., 2019). Kumar and Chen (2020) argue that predictability of the ENSO response in California using common methods is inhibited by low signal-to-noise ratios. Nevertheless, Region G does show a region in southern California that is almost exclusively identified by Granger causality, potentially indicating an advantage for Granger causality in low signal-to-noise ratio regimes. Further study is warranted.

Kumar and Chen (2020) argue that Eastern North America and the nearby Western Atlantic (Region H) are known to be influenced by ENSO in boreal winter (DJF) (Ropelewski & Halpert, 1987). Both methods show some statistically significant teleconnections in this region, and cross-correlation has a robust response in more grid boxes. Nevertheless, neither method recovers the robust DJF precipitation response described by Kumar and Chen (2020). It is difficult to pinpoint a reason behind this discrepancy. It could be due to different performance metrics (Kumar and Chen (2020) used the anomaly correlation coefficient, which would not make sense to compute in our current study), as well as a different observational set; MERRA-2 is known to have substantial dry biases in this region, which could obscure teleconnections (Reichle et al., 2017). Further study is warranted before we make firm conclusions about this region.

It is perhaps not surprising that many regions identified by cross correlation fall within the ground truth, as many past analyses of teleconnections used correlation-based methods (Mason & Goddard, 2001). In addition, both Granger causality and cross-correlation have identified several regions that are predominantly picked up by only one method. Regions B and I in Figure 2 are more persistent with cross-correlation but do not substantially overlap with the ground truth. For these regions, Granger causality does detect at least some portion of the areas that the correlation-based method does, but for different maximum lag windows (Figure 3b), indicating that the maximum lag window is an important parameter in this methodology. Conversely, the maximum lag window does not appear to be important for cross-correlation (Figure 3a), which is counter-intuitive: there should be evolving seasonal dynamics of precipitation. This suggests a potential shortcoming of the cross-correlation method; further discussion of this issue can be found below.

Several regions appear to be persistent only for Granger causality (Figure 2), and with few exceptions, do not appear in the ground truth. We are reluctant to make any conclusions about whether the regions indicate real features of ENSO teleconnections. We highlight this as an area of future work, not only for understanding teleconnections, but also as further validation of our methodology.

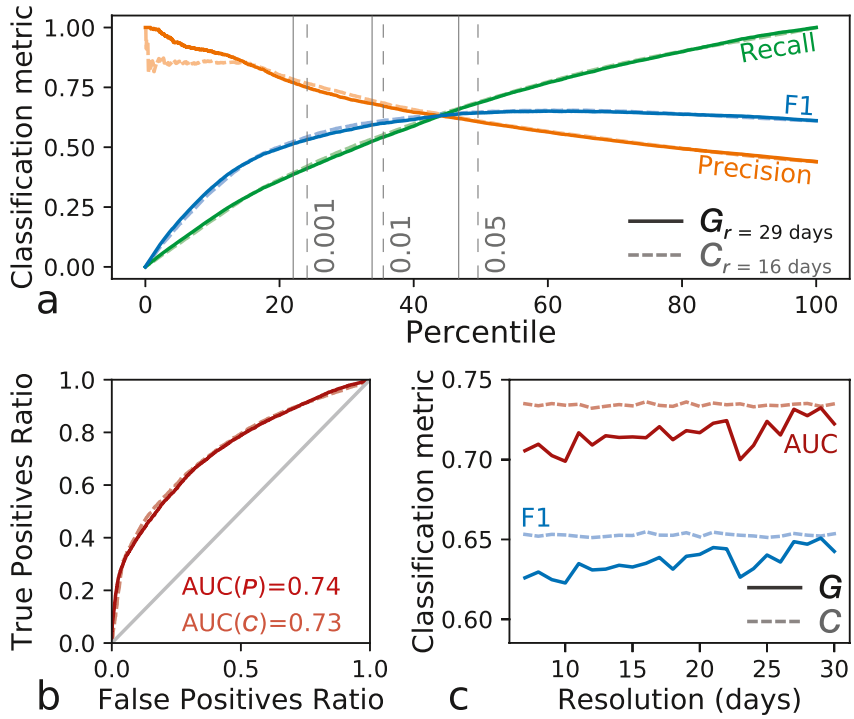

Figure 4. Classification metrics for lagged correlation (solid lines) and Granger causality (dashed lines). (a) Precision, recall, and F1 score versus percentile of the top-most significant relationships, with vertical lines indicating the percentile positions corresponding to p-values 0.05, 0.01, and 0.001. (b) Receiver operating characteristic (ROC) curves. The value of the area under the curve (AUC) is shown for both methodologies. (c) The AUC and maximum F1 scores for Granger causality (solid) and lagged correlation (dashed) as a function of the time resolution in days. See Text S4 for further description of these metrics.

It may appear strange that there are regions that satisfy Granger causality but do not appear to have a significant lagged correlation with El Niño 3.4 temperature (e.g., Regions C or J in Figure 2) according to our methods. Despite the fact that this can occur in real-world time series (Altman & Krzywinski, 2015), we attribute these phenomena to specific shortcomings of lagged correlation that are overcome by Granger causality (see Texts S1 and S2). Thresholds of statistical significance for cross-correlation are sensitive to the choice of the null model, whereas that is not the case for Granger causality. This effect can be easily demonstrated for a simple synthetic autoregressive time series (see Supporting Information S1) that shares many characteristics with large-scale geophysical data (e.g., MacMynowski et al., 2011): Cross-correlation only chooses a single lag of maximum correlation, whereas Granger causality incorporates information across all lags (up to $\tau_{\max}$), resulting in a more robust metric. Also, while the results of the two methods are often consistent, in some cases the cross-correlation method may select a value with a slightly higher correlation with a lag that is temporally inconsistent with neighboring grid points, which is a physically spurious result.

Figure 4 shows globally aggregated metrics gauging how well cross-correlation (C_τ) and Granger causality (G_τ) perform against the ground truth for a given maximum lag τ (in days). For this comparison we adopted only the best results of each model, corresponding to C_{16} and G_{29} (a more complete set of plots is available in Figure S7). While the curves are nearly identical, Granger causality has slightly worse quantitative performance than cross-correlation as measured by the AUC score. This is likely because most of the errors in cross-correlation result from false negatives (misses), whereas in addition to false negatives, Granger causality also identifies several features (e.g., regions D, E, and J in Figure 2) that do not appear in the ground truth, which register as false positives. This effect is more clear by looking at precision and recall metrics (Text S4). Granger causality displays higher precision and recall than cross-correlation if we keep only the areas of lower p-values. This indicates that the most significant Granger causal relationships occur inside the ground-truth regions, lending confidence that the lower AUC scores for Granger causality are indeed primarily because that method identifies new features that are not in the ground truth.

To assess the similarity of the maps obtained by employing the two methodologies, we computed the Spearman rank correlation between $C_{\tau_{max}}$ (cross-correlation) and $G_{\tau_{max}}$ (Granger causality) for different values of the maximum time lag τ_{max} (see Table S1). There is not much difference in the ranks for the cross-correlation method across the considered values of τ_{max} (rank correlation values range between 0.77 and 0.94). On the other hand, Granger causality seems to be more sensitive to the maximum time lag because the autoregressive model encompasses all the information of previous lags (up to τ_{max}) to predict the time series.

4. Discussion and Conclusions

We have developed a new method of quantifying climate system teleconnections using Granger causality that complements existing methods like cross-correlation. In addition to accurately reproducing the results found in the ground truth (as cross-correlation also does), this new method has several advantages:

1. Granger causality identifies statistically causal relationships with high precision and recall when focusing on high-confidence regions (low p; Figure 4).
2. Cross-correlation appears to not vary substantially with the maximum lag window. Based on the information in Figure 3, we cannot conclude whether cross-correlation is indeed picking up longer-term modes of variability, as we would want for evaluating ENSO teleconnections. Conversely, relationships found by Granger causality differ with lag (Figure 3 and Text S2), which not only matches our physical intuition that relationships evolve over time, but also allows us to determine the timescale of different precipitation responses to Niño 3.4 temperature and, in principle, the spatiotemporal evolution of teleconnections.
3. Cross-correlation methods can cause spurious results (e.g., McGraw & Barnes, 2018), appearing as false positives due to serial autocorrelation among the time series. As was seen for the Indian Summer Monsoon, there is also potential for false negatives due to a relative insensitivity. We also encountered a U-shaped distortion for the distribution of maximum lags (Dean & Dunsmuir, 2016; Yule, 1926), which depends on the window size (see Figure S5, as well as an extended discussion in Text S1). Thus the selected best lags for cross-correlation, in some cases, are artifacts of the method.

The potentially novel results highlighted by the regions in Figure 2 need further investigation. Our method is useful for discovering statistically robust relationships, but these explanations are incomplete without connecting physical mechanisms. To aid in this process, one promising attribute of Granger causality is the ability to construct long causal chains, that is, the discovered relationships indicated in Figure 2 may be mediated by several intermediate steps. Other proposed techniques can detect indirect relationships or retrieve the causal graph (e.g., Runge et al., 2019; Zhou et al., 2015). Future work could explore indirect associations through a multivariate Granger causality framework, which is more suited to be compared to the mentioned techniques used to retrieve causal graphs. This could be useful for uncovering physical mechanisms and signal propagation through the Earth system.

Several aspects of our method require further exploration. An example is the maximum lag — for Granger causality, this appears to be an important parameter. We have not yet explored the ability of this parameter to reveal useful properties of relationships, such as signal propagation. Moreover, the lags we chose had an upper limit of 180 days, which is a reasonable limit for many of the effects of ENSO. However, that upper limit may need to be altered for other applications, which could introduce difficulties in obtaining clear signals, such as if the annual cycle interferes with the calculations or if other modes of variability begin to play important roles; these potential effects have not been explored. Other aspects of our method that need to be investigated include time averaging or filtering the precipitation results by percentile. These potential effects may impact the robustness of our method, but we have not explored them in detail.

Perhaps one of the most important aspects of our method is that it provides additional testable predictions about thus far unverified ENSO teleconnection regions. The ground truth we used for ENSO teleconnections (shaded regions in Figure 2) is a composite of results obtained over years of research. Our methodology has identified new regions, which need to be reconciled with that composite, resulting in rejection of our findings or an update to the suite of known ENSO teleconnection responses. This process could apply to teleconnection studies more generally, providing a systematic way of identifying teleconnections between

any two climate fields; this could be especially useful for features in which the response is difficult to understand or simply unknown, for example, for attribution of changes in extreme weather event frequency.

Data Availability Statement

Code to reproduce the analysis presented in this manuscript is available at https://doi.org/10.5281/zenodo.5266377.

Acknowledgments
This research was developed using the computational resources from the Center for Mathematical Sciences Applied to Industry (CeMEAI) funded by the Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) under Grant No. 2013/07375-0. D.A.Vega-Oliveros acknowledges FAPESP Grants 2016/23698-1, 2018/24260-5, and 2019/26283-5. This material is based upon work supported by the National Science Foundation under Grant No. CNS-0521433. Support for B. Kravitz was provided in part by the National Science Foundation through agreement CBET-1931641, the Indiana University Environmental Resilience Institute, and the Prepared for Environmental Change Grand Challenge initiative. This work was based on research supported by the U.S. Department of Energy (DOE), Office of Science, Biological and Environmental Research, as part of the Regional and Global Model Analysis program. The Pacific Northwest National Laboratory is operated for the US Department of Energy by Battelle Memorial Institute under contract DE-AC05-76RL01830.

References

Altman, N., & Krzywinski, M. (2015). Association, correlation and causation. *Nature Methods*, *12*, 899–900. https://doi.org/10.1038/nmeth.3587

Anrews, E. D., Antweiler, R. C., Neiman, P. J., & Martin Ralph, F. (2004). Influence of ENSO on flood frequency along the California coast. *Journal of Climate*, *17*, 337–348. https://doi.org/10.1175/1520-0442(2004)017<0337:IOEOFF>2.0.CO;2

Arizmendi, F., Martí, A. C., & Barreiro, M. (2014). Evolution of atmospheric connectivity in the 20th century. *Nonlinear Processes in Geophysics*, *21*(4), 825–839. https://doi.org/10.5194/npg-21-825-2014

Ashok, K., Feba, F., & Tejavath, C. T. (2019). The Indian summer monsoon rainfall and ENSO. *Mausam*, *3*, 443–452.

Boers, N., Goswami, B., Rheinwalt, A., Bookhagen, B., Hoskins, B., & Kurths, J. (2019). Complex networks reveal global pattern of extreme-rainfall teleconnections. *Nature*, *566*(7744), 373–377. https://doi.org/10.1038/s41586-018-0872-x

Chang, P., Zhang, L., Saravanan, R., Vimont, D. J., Chiang, J. C. H., Ji, L., et al. (2007). Pacific meridional mode and el niño—Southern oscillation. *Geophysical Research Letters*, *34*, L16608. https://doi.org/10.1029/2007GL030302

Chatfield, C., & Xing, H. (2019). *The analysis of time series: An introduction with r*. CRC press.

Chiang, J. C., & Vimont, D. (2004). Analogous pacific and Atlantic meridional modes of tropical atmosphere–ocean variability. *Journal of Climate*, *17*, 4143–4158. https://doi.org/10.1175/JCLI4953.1

Cohen, J., Screen, J. A., Furtado, J. C., Barlow, M., Whittleston, D., Coumou, D., et al. (2014). Recent arctic amplification and extreme mid-latitude weather. *Nature Geoscience*, *7*, 627–637. https://doi.org/10.1038/ngeo2234

Damos, P. (2016). Using multivariate cross correlations, granger causality and graphical models to quantify spatiotemporal synchronization and causality between pest populations. *BMC Ecology*, *16*(1), 33. https://doi.org/10.1186/s12898-016-0087-7

Dean, R. T., & Dunsmuir, W. T. M. (2016). Dangers and uses of cross-correlation in analyzing time series in perception, performance, movement, and neuroscience: The importance of constructing transfer function autoregressive models. *Behavior Research Methods*, *48*(2), 783–802. https://doi.org/10.3758/s13428-015-0611-210.3758/s13428-015-0611-2

De Castro Santos, M. A., Vega-Oliveros, D. A., Zhao, L., & Berton, L. (2020). Classifying el niño-southern oscillation combining network science and machine learning. *IEEE Access*, *8*, 55711–55723. https://doi.org/10.1109/access.2020.2982035

Domiesen, D. I. V., Garfinkel, C. I., & Butler, A. H. (2019). The teleconnection of el niño southern oscillation to the stratosphere. *Reviews of Geophysics*, *57*, 5–47. https://doi.org/10.1029/2018RG000596

Donges, J. F., Zou, Y., Marwan, N., & Kurths, J. (2009). Complex networks in climate dynamics. *The European Physical Journal - Special Topics*, *174*(1), 157–179. https://doi.org/10.1140/epjst/e2009-01098-2

Donner, R. V., Wiedermann, M., & Donges, J. F. (2017). Complex network techniques for climatological data analysis. In C. L. E. Franzke, & T. J. O'Kane (Eds.), *Nonlinear and stochastic climate dynamics* (pp. 159–183). Cambridge University Press. https://doi.org/10.1017/9781316339251.007

Falasca, F., Bracco, A., Nenes, A., & Fountalis, I. (2019). Dimensionality reduction and network inference for climate data using δ-maps: Application to the cesm large ensemble sea surface temperature. *Journal of Advances in Modeling Earth Systems*, *11*(6), 1479–1515. https://doi.org/10.1029/2019ms001654

Fan, J., Meng, J., Ashkenazy, Y., Havlin, S., & Schelllnhuber, H. J. (2017). Network analysis reveals strongly localized impacts of El Niño. *Proceedings of the National Academy of Sciences*, *114*, 7543–7548. https://doi.org/10.1073/pnas.1701214114

Fenton, N., & Neil, M. (2012). *Risk assessment and decision analysis with bayesian networks*. CRC Press.

Fuchs, D., Sherwood, S., & Hernandez, D. (2015). An exploration of multivariate fluctuation dissipation operators and their response to sea surface temperature perturbations. *Journal of the Atmospheric Sciences*, *72*, 472–486. https://doi.org/10.1175/JAS-D-14-0077.1

Granger, C. W. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica: journal of the Econometric Society*, *37*, 424–438. https://doi.org/10.2307/1912791

Guez, O., Gozolchiani, A., Berezin, Y., Brenner, S., & Havlin, S. (2012). Climate network structure evolves with north Atlantic oscillation phases. *Europhysics Letters*, *98*(3), 38006. https://doi.org/10.1209/0295-5075/98/38006

Hamilton, J. (1994). *Time series analysis*. Princeton, NJ: Princeton University Press.

Hare, S., & Mantua, N. J. (2002). The pacific decadal oscillation. *Journal of Oceanography*, *58*, 35–44. https://doi.org/10.1023/A:1015820616384

Harrop, B. E., Lu, J., Liu, F., Garuba, O., & Leung, L. (2018). Sensitivity of the itcz location to ocean forcing via q-flux green's function experiments. *Geophysical Research Letters*, *45*, 13116–13123. https://doi.org/10.1029/2018GL080772

Hill, S., & Ming, Y. (2012). Nonlinear climate response to regional brightening of tropical marine stratocumulus. *Geophysical Research Letters*, *39*, L15707. https://doi.org/10.1029/2012GL052064

Hoskins, B. J., & Karoly, D. J. (1981). The steady linear response of a spherical atmosphere to thermal and orographic forcing. *Journal of the Atmospheric Sciences*, *38*, 1179–1196. https://doi.org/10.1175/1520-0469(1981)038⟨1179:TSLROA⟩2.0.CO;2

Hrudya, P. H., Varikoden, H., & Vishnu, R. (2020). A review on the Indian summer monsoon rainfall, variability and its association with ENSO and IOD. *Meteorology and Atmospheric Physics*, *133*, 1–14. https://doi.org/10.1007/s00703-020-00734-5

IRI. (2021a). What changes in rainfall are typical during El Niño? Retrieved from http://iridl.ldeo.columbia.edu/maproom/IFRC/FIC/elninorain.html

IRI. (2021b). What changes in rainfall are typical during La Niña? Retrieved from http://iridl.ldeo.columbia.edu/maproom/IFRC/FIC/laninarain.html

Janicot, S., Trzaska, S., & Poccard, I. (2001). Summer Sahel-ENSO teleconnection and decadal time scale SST variations. *Climate Dynamics*, *18*, 303–320. https://doi.org/10.1007/s003820100172

Kantz, H., & Schreiber, T. (2004a). Nonlinear time series analysis (Vol. 7). Cambridge university press.

Kantz, H., & Schreiber, T. (2004b). Nonlinear time series analysis (Vol. 7). Cambridge university press.
Kumar, A., & Chen, M. (2020). Understanding skill of seasonal mean precipitation prediction over California during boreal winter and role of predictability limits. *Journal of Climate*, *33*, 6141–6163. https://doi.org/10.1175/JCLI-D-19-0275.1
Kumar, K. K., Rajagopalan, B., & Cane, M. A. (1999). On the weakening relationship between the Indian monsoon and ENSO. *Science*, *284*, 2156–2159. https://doi.org/10.1126/science.284.5423.2156
Liu, F., Lu, J., Huang, Y., Leung, L., Harrop, B., & Luo, Y. (2020). Sensitivity of surface temperature to oceanic forcing via q-flux green's function experiments. Part iii: Asymmetric response to warming and cooling. *Journal of Climate*, *33*, 1283–1297. https://doi.org/10.1175/JCLI-D-19-0131.1
Liu, J., Yuan, X., Rind, D., & Martinson, D. G. (2002). Mechanism study of the ENSO and southern high latitude climate teleconnections. *Geophysical Research Letters*, *29*, 24-1–24-4. https://doi.org/10.1029/2002GL015143
Lorenzo, E. D., Liguori, G., Schneider, N., Furtado, J. C., Anderson, B. T., & Alexander, M. A. (2015). Enso and meridional modes: A null hypothesis for pacific climate variability. *Geophysical Research Letters*, *42*, 9440–9448. https://doi.org/10.1002/2015GL066281
MacMynowski, D. G., Shin, H.-J., & Caldeira, K. (2011). The frequency response of temperature and precipitation in a climate model. *Geophysical Research Letters*, *38*, L16711. https://doi.org/10.1029/2011GL048623
Mason, S. J., & Goddard, L. (2001). Probabilistic precipitation anomalies associated with ENSO. *Bulletin of the American Meteorological Society*, *82*, 619–638. https://doi.org/10.1175/1520-0477(2001)082<0619:ppaawe>2.3.co;2
McGraw, M. C., & Barnes, E. A. (2018). Memory matters: A case for granger causality in climate variability studies. *Journal of Climate*, *31*, 3289–3300. https://doi.org/10.1175/jcli-d-17-0334.1
McPhaden, M. J. (2015). Playing hide and seek with el niño. *Nature Climate Change*, *5*, 791–795. https://doi.org/10.1038/nclimate2775
Medhaug, I., Stolpe, M., Fischer, E., & Knutti, R. (2017). Reconciling controversies about the 'global warming hiatus'. *Nature*, *545*, 41–47. https://doi.org/10.1038/nature22315
Molod, A., Takacs, L., Suarez, M., & Bacmeister, J. (2015). Development of the geos-5 atmospheric general circulation model: Evolution from merra to merra2. *Geoscientific Model Development*, *8*, 1339–1356. https://doi.org/10.5194/gmd-8-1339-2015
O'Brien, J. P., O'Brien, T. A., Patricola, C. M., & Wang, S.-Y. (2019). Metrics for understanding large-scale controls of multivariate temperature and precipitation variability. *Climate Dynamics*, *53*, 3805–3823. https://doi.org/10.1007/s00382-019-04749-6
Papagiannopoulou, C., Miralles, D., Decubber, S., Demuzere, M., Verhoest, N., Dorigo, W. A., & Waegeman, W. (2017). A non-linear granger causality framework to investigate climate-vegetation dynamics. *Geoscientific Model Development*, *10*, 1945–1960. https://doi.org/10.5194/gmd-10-1945-2017
Reichle, R. H., Liu, Q., Koster, R. D., Draper, C. S., Mahanama, S. P., & Partyka, G. S. (2017). Land surface precipitation in MERRA-2. *Journal of Climate*, *30*, 1643–1664. https://doi.org/10.1175/JCLI-D-16-0570.1
Robock, A. (2000). Volcanic eruptions and climate. *Reviews of Geophysics*, *38*(2), 191–219. https://doi.org/10.1029/1998rg000054
Ropelewski, C. F., & Halpert, M. S. (1987). Global and regional scale precipitation patterns associated with the El Niño/Southern oscillation. *Monthly Weather Review*, *115*(8), 1606–1626. https://doi.org/10.1175/1520-0493(1987)115<1606:GARSPP>2.0.CO;2
Runge, J., Nowack, P., Kretschmer, M., Flaxman, S., & Sejdinovic, D. (2019). Detecting and quantifying causal associations in large nonlinear time series datasets. *Science Advances*, *5*(11), eaau4996. https://doi.org/10.1126/sciadv.aau4996
Schonher, T., & Nicholson, S. E. (1989). The relationship between California rainfall and ENSO events. *Journal of Climate*, *2*, 1258–1269. https://doi.org/10.1175/1520-0442(1989)002<1258:trbcra>2.0.co;2
Slingo, J. M., & Annamalai, H. (2000). 1997: The el niño of the century and the response of the Indian summer monsoon. *Monthly Weather Review*, *128*(6), 1778–1797. https://doi.org/10.1175/1520-0493(2000)128<1778:TENOOT>2.0.CO;2
Steinhaeuser, K., Chawla, N. V., & Ganguly, A. R. (2011). Complex networks as a unified framework for descriptive analysis and predictive modeling in climate science. *Statistical Analysis and Data Mining: The ASA Data Science Journal*, *4*(5), 497–511. https://doi.org/10.1002/sam.10100
Stuecker, M. (2018). Revisiting the pacific meridional mode. *Scientific Reports*, *8*, 3216. https://doi.org/10.1038/s41598-018-21537-0
Stuecker, M. F., Timmermann, A., Jin, F.-F., Chikamoto, Y., Zhang, W., Wittenberg, A. T., & Zhao, S. (n.d.) Revisiting enso/indian ocean dipole phase relationships. *Geophysical Research Letters*, *44*(5), 2481–2492. https://doi.org/10.1002/2016GL072308
Theiler, J., Eubank, S., Longtin, A., Galdrikian, B., & Farmer, J. D. (1992). Testing for nonlinearity in time series: The method of surrogate data. *Physica D: Nonlinear Phenomena*, *58*, 77–94. https://doi.org/10.1016/0167-2789(92)90102-S
Timmerman, A., An, S.-I., Kug, J.-S., Jin, F.-F., Cai, W., Capotondi, A., & Zhang, X. (2018). El niño–southern oscillation complexity. *Nature*, *559*, 535–545. https://doi.org/10.1038/s41586-018-0252-6
Trenberth, K. E. (2019). El niño southern oscillation (enso). In Encyclopedia of ocean sciences (3rd ed., Vol. 6, pp. 420–432). https://doi.org/10.1016/B978-0-12-409548-9.04082-3
Triacca, U. (2005). Is granger causality analysis appropriate to investigate the relationship between atmospheric concentration of carbon dioxide and global surface air temperature? *Theoretical and Applied Climatology*, *81*, 133–135. https://doi.org/10.1007/s00704-004-0112-1
Tsonis, A. A., & Swanson, K. L. (2008). Topology and predictability of el nino and la nina networks. *Physical Review Letters*, *100*(22), 228502. https://doi.org/10.1103/physrevlett.100.228502
Wang, S., Hipps, L., Gillies, R. R., & Yoon, J.-H. (2014). Probable causes of the abnormal ridge accompanying the 2013–2014 California drought: ENSO precursor and anthropogenic warming footprint. *Geophysical Research Letters*, *41*, 3220–3226. https://doi.org/10.1002/2014GL059748
Wang, Y., Gozolchiani, A., Ashkenazy, Y., Berezin, Y., Guez, O., & Havlin, S. (2013). Dominant imprint of rossby waves in the climate network. *Physical Review Letters*, *111*(13), 138501. https://doi.org/10.1103/physrevlett.111.138501
Yamasaki, K., Gozolchiani, A., & Havlin, S. (2008). Climate networks around the globe are significantly affected by el nino. *Physical Review Letters*, *100*(22), 228501. https://doi.org/10.1103/physrevlett.100.228501
Yoon, J.-H., Wang, S.-Y., Gillies, R. R., Kravitz, B., Hipps, L., & Rasch, P. J. (2015). Increasing water cycle extremes in California and in relation to ENSO cycle under global warming. *Nature Communications*, *6*, 8657. https://doi.org/10.1038/ncomms9657
Yuan, X., Kaplan, M., & Cane, M. (2018). The interconnected global climate system–a review of tropical-polar teleconnections. *Journal of Climate*, *31*, 5765–5792. https://doi.org/10.1175/jcli-d-16-0637.1
Yule, G. U. (1926). Why do we sometimes get nonsense-correlations between time-series?–A study in sampling and the nature of time-series. *Journal of the Royal Statistical Society*, *89*(1), 1–63. https://doi.org/10.2307/2341482
Yun, K.-S., & Timmermann, A. (2018). Decadal monsoon-ENSO relationships reexamined. *Geophysical Research Letters*, *45*, 2014–2021. https://doi.org/10.1002/2017GL076912
Zhou, D., Gozolchiani, A., Ashkenazy, Y., & Havlin, S. (2015). Teleconnection paths via climate network direct link detection. *Physical Review Letters*, *115*(26), 268501. https://doi.org/10.1103/physrevlett.115.268501