ARTICLE OPEN

Check for updates

Global land drought hubs confounded by teleconnection hotspots in equatorial oceans

Tejasvi Chauhan [1], Vikram Chandel [2] and Subimal Ghosh [1,2]✉

Emerging data-driven techniques, such as Complex Networks (CNs), can identify spatial linkages between droughts on a global scale and can improve early warning systems. Recent studies used CNs to identify hotspots of global drought teleconnections as land drought hubs; however, these studies excluded the ocean regions in CN, an oversight that can upend the insights gained thus far. Here, using a comprehensive global CN analysis on drought onsets, we show that oceanic regions harbor significantly larger drought hubs than land regions. The Indo-Pacific Warm pool (IPWP) in the Maritime continent emerges as the most significant drought hub having the farthest teleconnections. We show that IPWP, together with a few sub-tropical land and ocean regions, exhibit a 'rich club phenomenon' in CN. Further, using a causal network learning algorithm, we demonstrate the confounding role of oceans in modulating drought onsets on land regions. Our study reveals insights on the spatiotemporal linkages of global drought onsets and highlights the role of oceans in driving global drought teleconnections and modulation of land drought hubs.

INTRODUCTION

With their large spatiotemporal scales and high intensities, droughts leave an indelible mark on the global socioeconomy and are thus regarded as one of the most devastating natural disasters[1]. Persistent droughts have profound impacts on water availability[2,3], plant health[4,5], wildfires[6], and human well-being[7–9], the repercussions from which can cascade into critical areas such as water resources management[10], agricultural and food security[11–14], ecosystem services[15–17], and the global economy[18–20]. The sixth Assessment Report (AR6) by the Intergovernmental Panel on Climate Change (IPCC) projects an increase in the intensity and frequency of droughts across the globe, fuelled by global warming, potentially reducing agricultural yields by up to 20% per decade[14,21,22]. Human-induced climate change has intensified agricultural and ecological droughts across the globe, which threatens a global food crisis[19,23]. Globally, around 55 million people are impacted by droughts every year[24], and in the past century, droughts were responsible for the deaths of around 10 million people and have left multiple economies in dire straits[25]. According to United Nations, around 129 countries are projected to have increased exposure to droughts within the next few decades[25].

While spatial variability of droughts is projected to increase[26], our knowledge about the dynamics of droughts, their prediction, and their propagation in a warming world is limited[27–29]. Droughts are complex phenomena driven by both local processes and large-scale circulation of the atmosphere; hence, onsets, intensity, spatial extent of large-scale droughts have significant distant teleconnections and local dependence. Once set up, a large-scale drought inhibits precipitation due to land-atmosphere feedback[30,31] and thus can persist for longer durations. This internal feedback is not completely known in earth systems, and increasing anthropogenic perturbations make predicting drought onsets, durations, and intensities challenging.

Complex network (CN) analysis has gained prominence in many domains, including earth science[32–38] as it can help delineate internal feedback and external teleconnections from the observed earth science data. Due to its ability to capture the spatial embedding of physical processes that current models do not completely resolve, many attempts have been made to understand the climate extremes (droughts and precipitation extremes) using network-based approaches[39–46]. For example, Mondal et al. 2023[42] delineated CN between onsets of droughts on land to identify regions across the globe with very high connectivity called 'drought hubs'. These regions were argued to be directly linked to multiple droughts across the globe as hotspots of global drought teleconnections through their links observed in spatially embedded CN. However, they have overlooked the influence of external confounders in the CN framework, a known cause of spurious links in network analysis[47–53]. It is widely recognized that network delineation approaches can produce misleading links when variables have common drivers[52,53]. In this study, we contend that the exclusion of oceanic sources in CN analysis on drought onsets can give rise to critical oversights.

Droughts in the ocean regions may exhibit a higher degree of connectivity compared to land regions and hence, may hold higher predictive information for droughts across the globe. It is also intuitive because ocean regions are responsible for rainfall in many land regions simultaneously. For example, the Pacific Ocean is a source of moisture for parts of Australia, East Asia, North and South America[54–57]. Similarly, the Indian ocean is also responsible for rainfall in parts of South Asia, Australia, and East Africa[56,58,59]. The presence of ocean regions in a CN can thus modify the existence and characteristics of these teleconnection hotspots which can also alter the identification of "drought hubs" and "Rich Club Phenomenon"[42]. In addition, previous studies have failed to acknowledge inflation in network estimates caused by the projection system. Many projection systems have increasing grids per unit physical area as we move to higher latitudes. This leads to a significantly higher number of nodes near poles, which consequently generates misleading estimates of node importance measures. By incorporating ocean nodes into our global CN

[1]Department of Civil Engineering, Indian Institute of Technology Bombay, Mumbai, India. [2]Interdisciplinary Programme in Climate Studies, Indian Institute of Technology Bombay, Mumbai, India. ✉email: subimal@iitb.ac.in
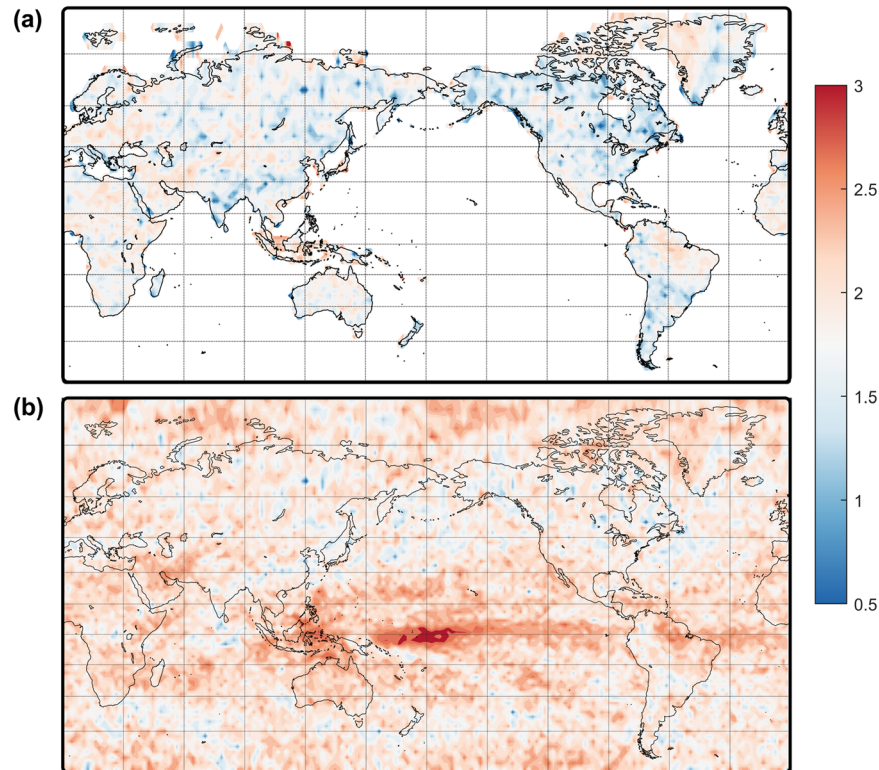
Fig. 1 Degree of connectivity of drought onsets after including ocean regions in CN. Logarithm ($\log_{10}$) of Degree centrality (DC) of grids from complex networks (CN) generated using drought onsets on land (**a**) and both land and ocean (**b**). Links in CN were tested for statistical significance (methods), and the Figure shows DC for only statistically significant links at 95% confidence.

analysis and accounting for the inflation caused by the projection systems, we can better capture the characteristics of teleconnections between drought onsets across the globe.

In this study, by including droughts in ocean regions as nodes in CN, we address these caveats and generate insights, which can better our understanding of drought teleconnections across the globe. We conducted a comprehensive global CN analysis on drought onsets that incorporates nodes on the ocean surface in addition to land. We use the Standardized Precipitation Index (SPI), Standardized Precipitation-Evapotranspiration Index (SPEI), and the self-calibrating Palmer Drought Severity Index (sc-PDSI) to classify droughts. We employ the event synchronization technique and a causal network learning algorithm – PCMCI (see methods) – to demonstrate that oceans have a significantly higher number of connections across the globe than land regions; hence, oceans have bigger 'drought hubs' than on land.

RESULTS AND DISCUSSIONS

Figure 1a shows the degree centrality (DC) of each grid from a complex network (CN) generated using event synchronization (ES) technique (Eqs. 1–5) on onsets of droughts on land grids. Here, DC represents the degree of connectivity of any grid in CN measuring the number of grids on which the onset of a 12-month cumulative drought (Supplementary Fig. 1 and Eq. 7), measured by SPI-12 from ERA-5 data (hereafter called SPI), can occur within a lag (or lead) of up to 6 months. Hence, drought onsets between any two grids can be synchronized in either direction (up to 6 months, see methods for details), leading to an undirected link between each grid pair. While computing DC, we consider only those links of the network on which ES was found statistically significant at 95% confidence. Hence, computed DC gives a robust estimate of the degree of connectivity of each grid. Drought onsets in a few land regions – namely Sahel, South Africa, the Middle East, western

North America, South America, and northern Australia – have a very high DC (Fig. 1, a). A higher DC of a region indicates that the drought onsets in the region are synchronized with many drought onsets across the globe. Using various CN metrics on the synchronized grids across the globe, these regions are declared as 'drought hubs' in the literature[42] which are argued to be hotspots of global teleconnections of droughts.

The DC estimated in this study has been corrected against error caused by the projection system (Eq. 6), a vital correction that is missing in many previous studies using CN on spatial data-sets[39–42,45]. Polar regions have significantly more nodes per unit area than the equator. Hence, we get an increasing number of links for the same physical area as we move from the equator to the poles, which can lead to an overestimation of network measures and misleading inferences, as shown in Supplementary Fig. 2. The Supplementary Fig. 2a shows DC generated using sc-PDSI on land-only grids as done in the literature[42], while Supplementary Fig. 2b shows the same plot after correcting for the projection system. It can be clearly seen that applying correction (methods) reduces the estimated degree of higher latitudes. The amount of overestimation of the degree will be directly proportional to the resolution of the data. A higher resolution data will have a higher number of nodes to penalize as we move away from the equator (Eq. 6). Hence, we used the corrected plots of DC for further analysis.

We have also generated Fig. 1a using SPI and sc-PDSI indices from the CRU dataset (methods), and DC from these datasets is shown in Supplementary Fig. 2b, c. We observe a general agreement of hotspot regions of high DC using multiple datasets and drought metrics, though they differ slightly in the estimates of DC of these regions which could be because of different processes captured by the index. PDSI depends on land-atmospheric interactions and involves a greater role of soil moisture and evapotranspiration, which also depend on a region's land-use and

land cover types. On the other hand, drought onsets from SPI are based on only precipitation with no consideration of soil moisture or land atmospheric interactions. PDSI is known to correlate well with SPI and SPEI at time scales higher than around 18 months[60] which is higher than the timescale considered in this study. Given the difference in what these indices represent, SPI at time scales shorter than 18 months may show droughts when sc-PDSI doesn't. Hence, slight differences between drought onsets from SPI and PDSI and their synchronization are expected. Regardless, the results from all the datasets point towards similar highly connected drought regions over land and are also similar to that reported in the previous studies[42].

To verify the efficacy of ERA-5 for this study, we compare it with monthly precipitation data from Global Precipitation Climatology Project (GPCP)[61] from 1979-2022 (Supplementary Fig. 3). GPCP provides combined estimates of precipitation using observations from ground stations and satellite products and hence can be used to validate the ERA5 data. Supplementary Fig. 3 shows statistically significant correlation (at 95% confidence) between monthly precipitation from ERA5 and GPCP. The ERA5 has a positive correlation over both ocean and land regions, however, the correlation of ERA5 with land regions is higher than the ocean regions. Within ocean regions, the highest correlation of up to 0.9 exists in the equatorial oceans. Similar results have also been reported using other datasets in the literature[62]. As an additional check on the applicability of the ERA5 dataset, we have also generated Event Synchronization networks using SPI from GPCP and results are presented in Supplementary Fig. 4. We find that while the degree centrality (DC, Supplementary Fig. 4) is systematically overestimated in the ERA-5 dataset (Fig. 1b) across the globe, it well captures the spatial embedding of the global network as well as the relative node importance and drought hubs. The correlation and consistency of network metric (DC) between ERA5 and GPCP demonstrate the validity of ERA5 in this study.

Drought hubs in ocean
Concurrent regional droughts on land are known to be driven by ocean-atmosphere interaction[63]. While the precipitation deficits responsible for the onsets of meteorological droughts are brought upon by large scale ocean-atmosphere interactions, land-droughts like hydrological droughts, are complex, and their duration and intensity are also modulated by land-atmosphere interactions[64]. Temperature also plays a major role in driving the land precipitation and vapor pressure deficit – important factors controlling the land-atmospheric feedback. Ocean-atmospheric interactions, for example, ENSO plays a major role in temperature variability in many regions around the globe[57]. This leads to a hypothesis that the onsets of droughts on the ocean might be synchronized with drought onsets on land regions. Considering ocean regions along with land regions in a CN framework of drought onsets may lead to a bigger hubs outside land regions (Fig. 1a). To test this, we generate the CN by including drought onsets over the oceans, and the updated DC is shown in Fig. 1b using SPI and Supplementary Fig. 5 using SPEI. Networks from both datasets show similar results. The oceanic regions in the tropics have DC around two to three times the DC of land regions (Fig. 1b, maximum DC of land regions on log scale is around 2.5 while that in the ocean it is around 3). The wider spatial spread of oceanic hubs is also apparent from Fig. 1, which might also be because of high interconnectedness of these regions. Physically, this happens because ocean regions have ability to simultaneously modulate droughts on multiple land regions at varying time scales (El Niño Southern Oscillation, Atlantic Niño, Indian Ocean dipole), which translates into synchronization of their drought onsets. The land hotspots in Fig. 1a have also gained more connectivity in Fig. 1b, which is expected because total number of nodes in the

network have increased from 'land only' to 'land+ocean' case. While the regions with high connectivity remain apparent on land, the DC on the ocean regions– primarily the Indo-Pacific warm pool (IPWP) in the Maritime continent, the central Pacific, and the Atlantic Ocean – is manifold higher than the DC of land regions. Hence, the oceans can also be hotspots of drought teleconnections or regions that drive global land droughts simultaneously. These regions of high connectivity in the ocean might contain more predictive information for droughts on land than the land drought hubs from Fig. 1a.

From Fig. 1, we can see that the inclusion of the ocean region reveals teleconnection hotspots in equatorial oceans. The betweenness centrality (BC, Eq. 8) generated using networks from SPI and SPEI, which indicates the importance of a node as an intermediary between any two nodes in the network, is shown in Supplementary Fig. 6. Both datasets show a consistency in the spatial pattern of BC. We observe the highest BC in the Pacific Ocean and the IPWP followed by the Atlantic Ocean, Western Indian Ocean, and some parts of the Middle East and South Africa. It should be noted here that we haven't corrected BC for the errors due to the projection system, and this correction is out of the scope of the current study. The regions in the equatorial ocean regions with high DC and BC than land regions indicate the presence of manifold bigger drought hubs in the ocean than on land. This result is also intuitively related to the dependence of all major droughts on ocean-atmosphere dynamics[43].

To investigate further, we plot the spatial extent of synchronization of drought onsets in the two most major ocean hubs observed in Fig. 1(b) – the IPWP and the Pacific Ocean. The results are shown in Supplementary Fig. 7a and b, respectively. We observe that the droughts in IPWP (approximate spatial extent considered marked in Supplementary Fig. 7a) have synchronization with South Asia, South Africa, South America, and Australia, whereas the Pacific Ocean is connected to the Middle East, East Africa and Western North America in addition to various oceanic regions. El-Niño Southern Oscillation (ENSO) in the Pacific Ocean, which changes the precipitation pattern over the IPWP, is known to alter global precipitation patterns by atmospheric and oceanic teleconnections[65,66]. Hence, the connections between land regions might be an impression of lagged associations of land regions to a common driver in the ocean.

To test this, we show the connections of drought onsets on land regions using CNs with and without considering ocean regions in Supplementary Figure 8. Supplementary Fig. 8a-d shows connections of land drought hubs without considering ocean regions, and Supplementary Fig. 8e-h shows the connections of the same regions after considering ocean grids as nodes in CN analysis. Drought onsets in western North America and the Middle East are connected to each other in land-only analysis (Supplementary Fig. 8a and b); however when oceans are included (Supplementary Fig. 8e and f), we observe that both are synchronized with the central Pacific Ocean. This pattern is due to the dependence of droughts in the Middle East and West North America on ENSO[67,68]. Droughts in Middle east and North America are modulated by ENSO through the zonal shifts of subtropical jet stream and Rossby waves. There exists stronger convergence (divergence) during warm (cold) phases of ENSO along with moisture transport from Arabian sea towards the Middle East[69,70]. At the same time, ENSO modulates the strength and location of polar and subtropical Jet streams through Rossby wave breaking and controls the moisture transport to the parts of North America[71]. Similarly, drought onsets over South Africa and Australia are linked to each other and are also connected to the IPWP (Supplementary Fig. 8c and g), as both are driven by ENSO[72,73]. Sahel region (Supplementary Fig. 8d) shows some connections with drought onsets in south Asia and shows very limited synchronization with the IPWP region. These results prove that the existence and role of global land drought hubs are overestimated in literature due to the CN
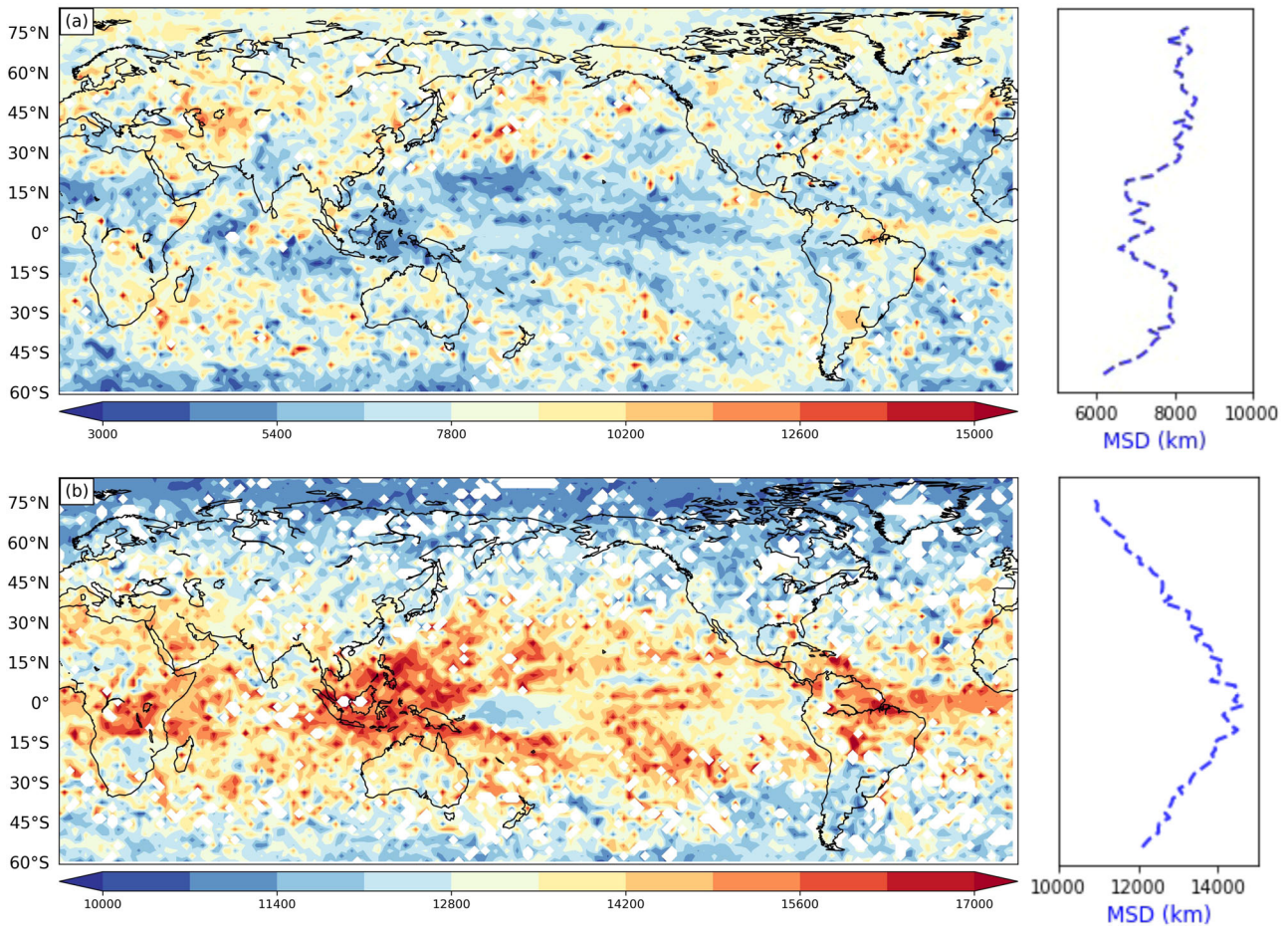
Fig. 2 The mean synchronization distance (MSD) of drought teleconnections. MSD (Eq. 9) of links generated using event synchronization after removing all links with distances less than 1000 km (**a**), and 10,000 km (**b**). The right panels show the variation of mean MSD along the latitude. MSD is computed only for those links which were found statistically significant at 90% confidence (methods). The Figure shows that the equatorial regions have the farthest drought teleconnections. The maritime continent is the biggest hub having the farthest teleconnections followed by the Atlantic Ocean, South America, the western Indian Ocean, and the equatorial region of Africa.

model artifact of not considering oceanic regions. The global drought hubs that drive global drought onsets and, potentially, modulate the land drought hubs are in ocean regions instead of land regions.

Revisiting the rich club phenomenon of drought hubs

To quantify the spatial extent of teleconnections of droughts hubs found in the ocean, we compute the mean synchronization distance (MSD, Eq. 9) of each node in CN. For every node, MSD measures the average spatial distance of all nodes which are connected to it. If the connectivity of a node or its spatial scale of synchronization is assumed to be its wealth in a network, the existence of a small fraction of nodes with a large degree of connectivity or MSD is analogous to a 'rich club' and has been argued to be a prominent topological characteristic of networks[42]. Figure 2 and Supplementary Figure 5 show MSD computed using links that were found statistically significant at 90% and 95% confidence, respectively. Since nearby grids are expected to be highly connected, we generate MSD considering links with distances more than 1000 km (Fig. 2a, Supplementary Fig. 5a) and 10,000 km (Fig. 2b, Supplementary Fig 5b).

We find that with a threshold of 1000 km (Fig. 2a, Supplementary Fig. 5a), extratropical regions (30N-60N and 30S-60S) have higher MSD than the equatorial regions with an exception in the equatorial Indian Ocean, northern South America and the Atlantic Ocean, which also have high MSDs. Among land regions, the

Middle East, northern South America, South Africa, western North America, and Australia are the regions with the farthest connectivity. IPWP and the central Pacific Ocean have low MSD values, reflecting a high number of local connections. All the grids in the Pacific Ocean interact with each other increasing the number of links at around a distance of 1000-5000 km, leading to a reduction in the mean value.

The Zonal Mean of MSD, shown on the right panel of Fig. 2a, decreases from the north to the south, with a dip at the equator, when oceanic regions are considered. On the contrary, Mondal et al.[42] from their analysis, excluding the oceanic regions, reported an increasing MSD of land regions from the northern to the southern hemisphere, concluding that drought hubs in the southern hemisphere constitute a 'rich club' in the CN. We performed the CN analysis with land-only grids without any threshold of distance to reproduce the results of Mondal et al.[42] in Supplementary Fig. 6a. The right panel (Supplementary Fig. 6a) shows the latitude variation of average MSD (blue) and percentage land area (red). It is clearly visible that the variation in MSD with latitude is opposite to the variation in the percentage of land with latitude. The southern hemisphere has fewer land areas, which are far from each other. Many of these land regions can also be connected to a common oceanic source, leading to overestimating MSD in the southern hemisphere. After considering the ocean regions in CN, the steep north-to-south increase in the average MSD of land regions vanishes, and we observe a slight

increase in the MSD of land regions (Supplementary Fig. 6b), showing only land regions from a CN which includes nodes on oceans as well) from north to south. Hence, we conclude that the tagging of the southern hemisphere as a "*rich club*" in the previous study[42] resulted from a statistical artifact due to the asymmetrical distribution of land in both hemispheres and the non-inclusion of oceans in the analysis.

Since MSD in Fig. 2a can be influenced by a high number of local connections, we increase the threshold to 10,000 km (Fig. 2b and Supplementary Fig. 6b) to reveal average spatial scales of only distant teleconnections. We observe a significant increase in the MSD of the equatorial region, with the IPWP emerging as the biggest hub with the farthest teleconnections. Apart from the IPWP, eastern portions of the Atlantic and the Indian oceans also have high MSD values. High DC with high BC indicates that these nodes form hotspots in the network which combined with the high MSD found in Fig. 2b, shows that they have the most distant teleconnections. Hence, the 'rich club phenomenon' in the CN of global drought onsets contains its most significant clusters in the equatorial region with the IPWP in the Maritime Continent being the most significant among them, having farthest teleconnections. Supplementary Fig. 7a and b show the distribution of MSD from each land drought hub along with regions in the Atlantic Ocean, maritime continent, and the Pacific Ocean (all regions shown in Fig. 4f) considering thresholds of 1000 km and 1000 km, respectively. Supplementary Fig. 7 shows that the fat-tailed nature of MSD becomes prominent after considering a threshold of 10,000 km, with the region in the Maritime continent (IPWP) containing the heaviest tail followed by the Pacific Ocean, Atlantic Ocean, and the Middle East. The fat-tailed distributions of ocean regions support the hypothesis that ocean regions are the drivers of most of the drought onsets on land regions.

Drought hubs in the ocean confound drought hubs on land

The above-mentioned results demand further investigation into the possibility of land drought hubs being simultaneously modulated by the ocean-atmosphere processes, like ENSO, acting as a confounder. Figure 3 shows the annual precipitation anomaly from ERA5 for major drought hubs along with the state of ENSO. El Niño and La Niña years are marked as red and blue circles at the end of each bar. The right panel shows the correlation of the annual rainfall anomaly with the Oceanic Niño Index (ONI). The Middle East and western North America have a statistically significant ($p < 0.05$) positive correlation with ONI and Australia, and South America has a statistically significant ($p < 0.05$) negative correlation with ONI. The Sahel and South Africa have a negative correlation; however, it is not statistically significant. These results indicate that the interannual variability in precipitation of many the drought hubs is significantly driven by ENSO, which can cause simultaneous or lagged droughts on multiple land regions based on the ENSO state and their time scales of interaction with ENSO. A few weak correlations in Fig. 3 could be because ENSO has a spatio-temporally varying impact on global precipitation. There are regions where interannual variability of precipitation has only a weak association with ENSO because either their relationship with ENSO has changed over many decades or there exists a strong seasonality in the strength of their teleconnections to the Pacific Ocean. Western Africa is known to have a strong association with the Pacific sea surface temperatures (SSTs) during the early months of the year. However, this relationship vanishes, and a stronger association with the Atlantic and Indian ocean gets established in later parts of the year for boreal summer monsoon and autumn rainfall, respectively[74]. Results from Janicot et al. 1996[75] demonstrate that the relationship between ENSO and Sahel rainfall is not constant and has a multidecadal scale of emergence. They show that the correlation between ENSO and Sahel rainfall has changed sign after 1970s which explains weak

correlation over the total time period of past 70 years in Fig. 3. To clarify the spatially and temporally varying global effect of ENSO on precipitation, we have plotted the correlation of ONI with average annual precipitation from GPCP (Supplementary Fig. 12) spatially showing only statistically significant grids at 95% confidence level. Weak correlation of ONI with rainfall in Sahel and South Africa is also visible. Supplementary Fig. 13 shows the precipitation anomaly across the globe during El Niño and La Niña years for the summer (June, July, August (JJA)), and winter seasons (October, November, December (OND)) of the northern hemisphere. During El Niño years, the composite of summer precipitation (Supplementary Fig. 13a) shows a simultaneous reduction of precipitation over IPWP, South Asia, East Africa, Some parts of Australia, and northern South America. Hence, drought onsets in these regions can potentially be synchronized. During La Niña years, JJA precipitation (Supplementary Fig. 13c) shows opposite behavior where the above-mentioned regions receive surplus rainfall with small deficits in southern South America and northern North America, which are synchronized mainly with the reduced precipitation of central Pacific. Similarly, for winter months, during El Niño years (Supplementary Fig. 13 b), the reduction in precipitation in IPWP is synchronized with South Africa, Australia, and South America. During La Niña years, OND months receive a simultaneous reduction in precipitation in East Africa, the Middle east, North America, and Southern South America, which are synchronized with a reduction in precipitation in the central Pacific. The above results show that ocean sources are the confounders of drought onsets on land. For example, in Fig. 1a, links originating from regions South Africa or South America can be because droughts in these regions are simultaneously modulated by teleconnections from the IPWP as shown in Supplementary Fig. 13b.

ES cannot separate a direct causal association from a confounding effect. Any two drought regions on land, if modulated by the same ocean source, will be delineated as synchronized unless the influence of oceanic confounder is removed. Hence, to test for the confounding behavior of oceans, we apply a causal network learning algorithm – PCMCI (methods) – to delineate causal graphs from the monthly time series of SPI on land and ocean regions. The results are shown in Fig. 4. Nodes are various land drought hubs (regions shown in Fig. 4f), and their links indicate a causal connection between monthly SPI at different nodes. A link is only shown if found statistically significant at a 95% confidence level. Link labels, if present, indicate the lag at which the connection was found on a monthly scale; else, the absence of link labels means that the link was found at zero lag. The node color represents autocorrelatoin whereas link color represents the strenght of directional link. Cross mark at the ends of links indicate that the directionality of that link couldn't be delineatad. Figure 4a contains connections between land regions when no ocean regions are considered. Figure 4b adds SPI from the Atlantic Ocean (AO) to the network, Fig. 4c adds the Maritime Continent (MC) in addition to the AO, which is followed by Fig. 4d, which adds a region from the central Pacific Ocean (PO). The incremental addition is done to add more ocean regions to the conditioning set one by one and observe the linkages between the land drought hubs, which are kept constant throughout the experiments.

When no ocean sources are present (Fig. 4a), we observe 5 links between land variables showing the Middle East (ME) as the most connected node with an incoming connection from western North America (WNA) and South America (SAM), and an outgoing link to Australia (AU) and Sahel (SH). In addition, we also get a link from AU to South Africa (SAF). As the Atlantic Ocean (AO) is added as a node to the network the connection between the AO and ME emerges. The Atlantic ocean sea surface temperatures (SSTs) modulate the strength of westerly jet stream, controlling the geopotential height in the upper troposphere over the Eurasia
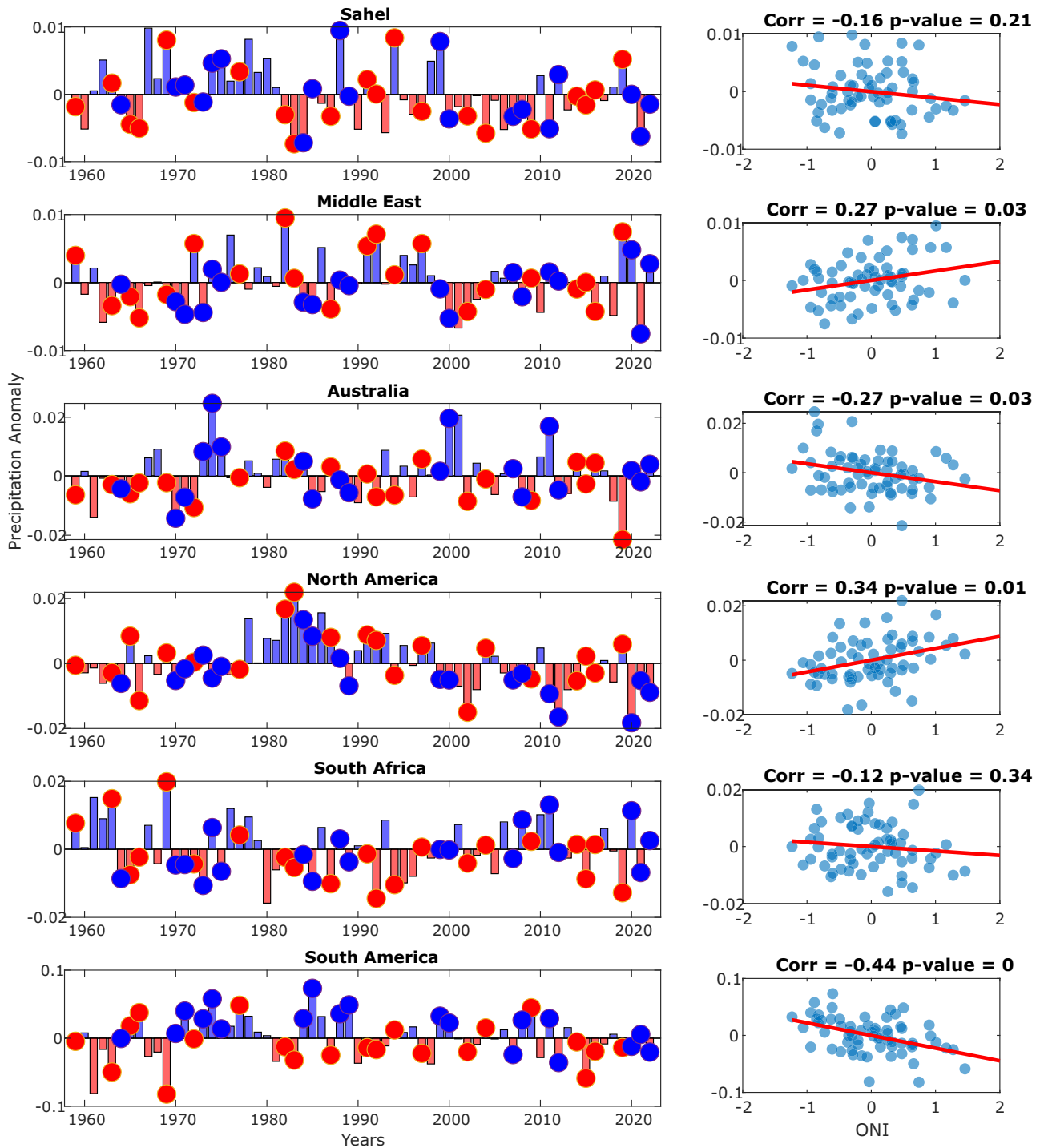
Fig. 3 Association of annual precipitation from land hubs with ENSO. Precipitation anomalies from ERA5 from various land drought hubs in Fig. 1a with the ENSO years. El Niño and La Niña years are marked as red and blue circles, respectively. The right panel shows the scatter plot and correlation between the precipitation anomaly and the ONI index.

ultimately altering moisture transport over the Middle east[76]. As the Maritime Continent (MC) and Pacific Ocean (PO) are added to the Network, links between land and ocean regions emerge. Some prominent links are the link from MC to AU, PO to WNA and PO to SAM. MC to AU links is known to exist via atmospheric moisture transport and varies with ENSO and the Indian Ocean Dipole (IOD)[56]. Link from PO to WNA represents the modulation of subtropical and polar jet streams by the PO and the link from PO to South America (SAM) represents the control of well-known Pacific-

South American (PSA) mode on south American rainfall at Inter-annual and quasi-biennial time scales[71,77]. A link between PO and MC emerges which represents the modulation of rain between west and central pacific through changes in walker circulation by ENSO. The most significant result from the causal analysis is the vanishing of WNA-ME link upon addition of oceanic sources (Fig. 4a and d). This shows that the link is not a true causal link (or a process-based teleconnection) but a synchronization which appears because of a confounding effect of PO on both these
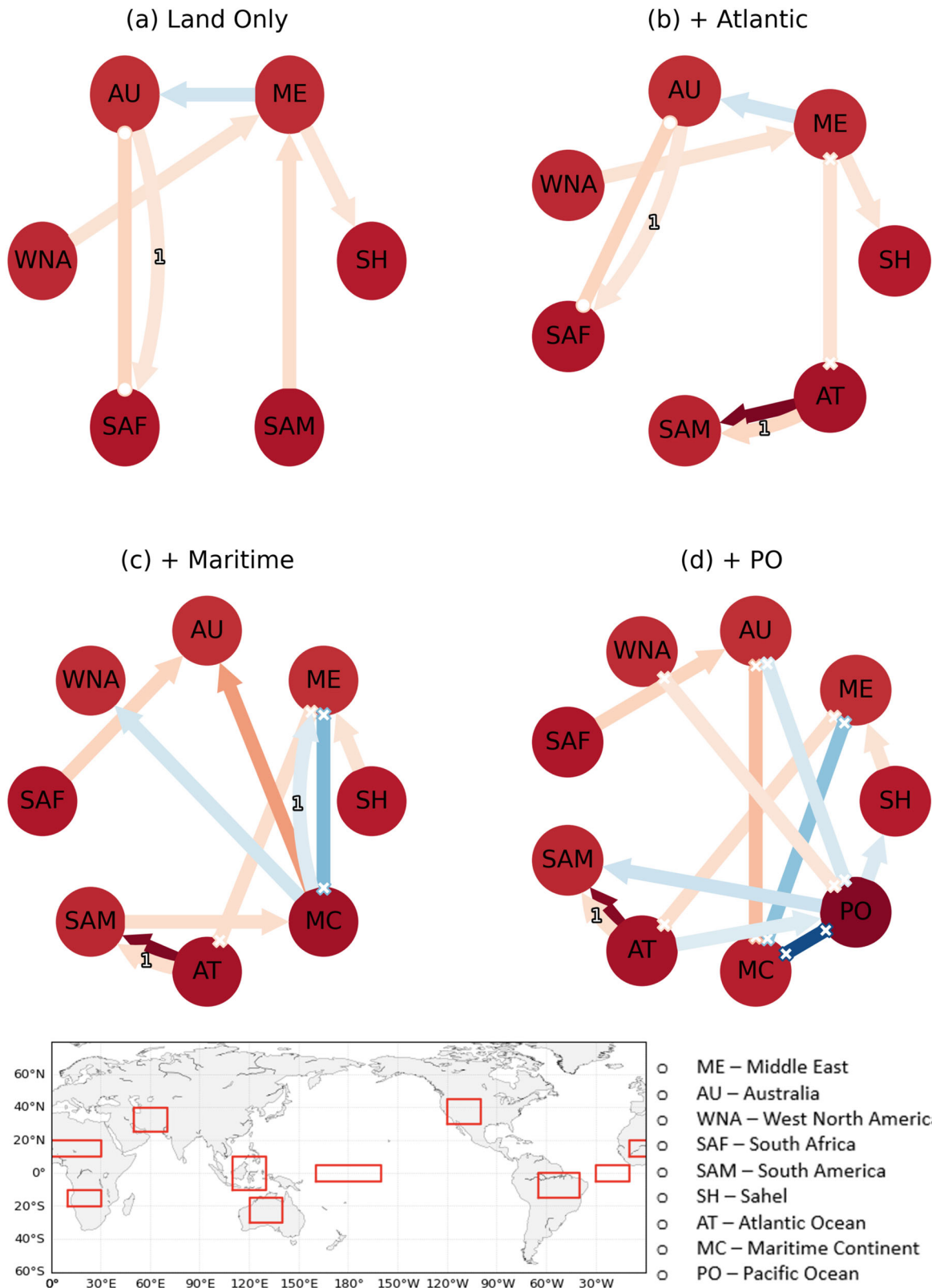
Fig. 4 Oceans confounding land drought hubs. Networks between SPI from different regions (f) generated using causal discovery algorithm, PCMCI. **a** Networks generated using only land drought hubs as nodes. **b–d** Consecutively add SPI from the Atlantic Ocean (AO), the IPWP in the Maritime continent (MC), followed by the Pacific Ocean (PC). A link is only shown if found statistically significant at a 95% confidence level. The node color represents autocorrelatoin whereas link color represents the strenght of directional link. Cross mark at the ends of links indicate that the directionality of that couldn't be delineatad. The lag at which the link was found significant is shown as link label, absence of which indicates that the link was found at zero lag. We observe an increase in the number of links as ocean sources are incrementally added to the network. The link between ME and WNA vanishes upon addition of MC as a node.

regions. This result indicates the role of ocean regions as confounders of global droughts leading to their synchronisations.

In this study, we employ event synchronization (ES) technique to compute a global complex network (CN) of drought onsets, and find regions of high degree centrality (DC), betweenness centrality (BC), and mean synchronisation distance (MSD) in oceans. Our results show that after including ocean regions in CN analysis, the high MSD of the southern hemisphere land regions, as reported in the literature[42], vanishes, and the equatorial regions, particularly the Indo-Pacific Warm Pool (IPWP) in the Maritime Continent, the Western Indian Ocean, and the Atlantic Ocean with some equatorial land regions in Africa and South America emerge as regions with farthest teleconnections. Among all oceanic regions, the Maritime continent has the farthest teleconnections, and we contend that it can be considered as the largest hub in global drought teleconnections. We argue that a CN made using just land grids constitutes a subset of a larger land-ocean network (including both land and oceans) exhibiting a rich club phenomenon with the most significant node clusters, which hold the capacity to modulate the entire network being in the ocean. Hence, the emergence of southern hemisphere rich club phenomenon is a statistical artifact arising due to missing nodes on ocean regions in a spatially embedded global network of drought onsets. Next, we use a causal network learning algorithm to examine the process connectivity of drought teleconnections showing that ocean regions confound global drought onsets and cause simultaneous droughts in multiple regions which leads to their synchronisation.

Our results significantly advance our understanding of global teleconnections of drought onsets and highlight the role of oceanic regions as global drought hubs. Since ocean regions can modulate multiple droughts simultaneously, identification of drought hubs in the ocean is critical for early warnings of global droughts. Hence, the utility of drought hubs in the Maritime Continent, western Indian Ocean, and Atlantic Ocean in improving the prediction of drought onsets on land regions is the potential future extension of this study. While quantifying synchronization of drought onsets, we make no distinction between teleconnected droughts and a spatially propagating drought. The same drought, if it spatially propagates from ocean to land will be identified as a synchronization link between land and ocean grids if its onset on land occurs within 6 months of its occurrence over the ocean. Similarly, a teleconnection of drought onset between ocean and land (for example, droughts modulated by ENSO on multiple land regions) will also appear synchronised if there is a temporal consistency of occurance within 6 months of lead/lag time. An analysis to distinguish between these two types of connectivity between drought onsets is possible using directed synchronization networks which we leave as a future scope of this study.

METHODS
Data
We used gridded monthly precipitation data from European Centre for Medium-Range Weather Forecast Reanalysis Version 5 (ECMWF, ERA-5)[62] from 1959 to 2022. We also used gridded monthly precipitation data and monthly sc-PDSI data from 1901 to 2021 obtained from Climate Research Unit (CRU). All gridded data were aggregated to 2°x 2° spatial resolution. Observed precipitation data is from Global Precipitation Climatology Project (GPCP)[61] at monthly scale from 1979 to 2022. We used the Oceanic Niño Index (ONI) provided by the Physical Sciences Laboratory (PSL), National Oceanic and Atmospheric Administration (NOAA).

Drought onset and characteristics
In this study, we consider the onset of moderate (or more severe) droughts similar to recent literature[42]. We classify drought onsets based on three well-known indices, Standardized Precipitation Index (SPI), Standardadized Precipitaion-Evapotranspiration Index (SPEI), and self-calibrating Palmer Drought Severity Index (sc-PDSI)[42,43,78–80]. SPI/SPEI have been generated using 12-month cumulative precipitation (and evaportranspiration) and hence they represent 12-month cumulative droughts (SPI12/SPEI12 hereafter called SPI/SPEI). Drought characteristics like onset, termination, duration, intensity, and severity are considered based on threshold approach on drought indices[81]. Here we consider moderate droughts using a threshold of -1 on SPI/SPEI and a threshold of -2 on sc-PDSI as shown in Supplementary Figure 1. Hence, a 'drought onset' is defined as the month when the value of cumulative 12-month SPI/SPEI goes below -1, and 'drought termination' is defined when the value recovers from the threshold as shown in Supplementary Figure 1. The time period between consecutive onset and termination of a drought event is called 'drought duration'. The intensity of drought can be defined as average value of SPI/SPEI/sc-PDSI between onset and termination. The severity of drought is defined as cumulative deficit measured using SPI/SPEI within the duration. To identify teleconnections of only drought onsets and to ensure enough sample size we do not use any thresholds of intensity or severity before classifying a drough onset as an event.

Event synchronisation (ES)
Droughts are driven by atmospheric processes, ocean teleconnections, and local feedbacks, which bring spatial variability in their onsets across the globe. Explaining this variability can shed light on the drought co-occurrence which can ultimately improve drought prediction. Hence, attempts have been made in the literature to understand the spatiotemporal variability of drought onsets using multiple statistical tools. Event Synchronisation (ES)[40,41], is one such tool that has gained popularity in earth science because of its computational efficiency and ability to delineate non-linear associations from data. ES is a non-parametric similarity measure that delineates temporal dependencies as well as delays between time series of various events (drought onsets in this case). In this study, ES is calculated for gridded data of drought onsets as follows:

We compute the varying time delay, $\tau^{i,j}_{l,m}$, between onsets at any two grid locations i, j in an m x n grid, having a total of s_i and s_j onsets, respectively, as

$$\tau^{i,j}_{l,m} = \min\left\{\frac{t^i_{l+1}-t^i_l, t^i_l - t^i_{l-1}, t^i_{m+1}-t^i_m, t^i_m - t^i_{m-1}}{2}\right\} \quad (1)$$

Where t^i_l (t^i_m) represents l^{th} (m^{th}) event occurring at grid location $i(j)$, with $l = 1, 2, 3, \ldots, s_i$ and $m = 1, 2, 3, \ldots, s_j$. $\tau^{i,j}_{l,m}$ provides the maximum permitted time delay between drought onsets to call them synchronised. In this study, we have added an upper limit of 6 months to $\tau^{i,j}_{l,m}$, hence, no delays exceeding 6 months shall be considered as synchronised.

Next, ES is estimated between each grid pair by counting the temporally coinciding events, under a condition that for each pair of grids, the absolute value of the temporal delay between any two synchronous events must not exceed $\tau^{i,j}_{l,m}$ (or 6 months, whichever is smaller).

$$ES_{i,j} = \frac{c(i|j) + c(j|i)}{\sqrt{s_i s_j}} \quad (2)$$

Where $c(i, |, j)$ measures number of times drought onset at i succeeds onset at j. Hence,

$$c(i|j) = \sum_{l=1}^{s_i} \sum_{m=1}^{s_j} J_{i,j} \tag{3}$$

And,

$$J_{i,j} = \begin{cases} 1 & \text{if } 0 < t_l^i - t_m^i < \tau_{l,m}^{i,j} \\ 0.5 & \text{if } t_l^i = t_m^i \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

In this study, we use gridded monthly SPI-12 from ERA-5 reanalysis product at a spatial resolution of 2° x 2° within latitudes 60°S and 85°N including both land and ocean regions. This gives us a total number of grids as 13140 (73 × 180). Hence, the final ES matrix generated is of size 13140 × 13140, which represents the strength of synchronisation between drought onsets across the globe. We also apply ES to drought onsets over only land regions classified using SPI-12 and sc-PDSI data from CRU, both at a spatial resolution of 2° x 2° within latitudes 60°S and 85°N. In this case, the total number of grids are 3846 (only land grids between latitudes 60°S and 85°N) and hence, the ES matrix generated is of size 3846 × 3846, which represents the strength of synchronisation between drought onsets among land regions of the globe.

Complex network (CN) analysis

The ES matrix can be converted to a CN adjacency matrix ($A_{i,j}$) where nodes i and j are the grid locations and the links between them represent the presence of pairwise synchronisation between the drought onsets on the grids. The ES matrix is symmetric because while estimating ES in Eq. (2), we count the events in both the direction (i to j, and j to i). Hence, it does not provide any information about the direction of synchronisation. In this study, we test each $ES_{i,j}$ for statistical significance using method of shuffled surrogates as suggested by Boers et al. 2019[41]. We randomly reshuffle the time series of events (onsets) and compute ES 10000 times to generate a null distribution of $ES_{i,j}$ and perform hypothesis testing at 90% and 95% confidence. Hence the Adjacency matrix thus generated contains only strongest and most reliable links in the global connections between drought onsets.

$$A_{i,j} = \begin{cases} 1 & \text{if } ES_{i,j} \text{ is statistically significant} \\ 0 & \text{otherwise} \end{cases} \tag{5}$$

From the network generated above, we analyse the importance of nodes and their spatial embedding using node importance measures, namely Degree Centrality (DC), Betweenness Centrality (BC), and Mean synchronisation Distance (MSD), as explained below.

Correction for errors due to projection system

Most projection systems used by CN literature in earth science have increasing number of grids per unit area from equator to either pole. Due to this, a greater number of links will be generated from higher latitudes than on the equator for the same physical area. This can lead to overestimation of network measures which include links from higher latitudes, which if not corrected, can cause misleading inferences. In this study, we address this problem by converting the binary adjacency matrix to a weighted matrix which uses cosine of the latitude of source node as a weight.

$$A_{ij)} = \begin{cases} \cos(\varphi) & \text{if } ES_{i,j} \text{ is statistically significant} \\ 0 & \text{otherwise} \end{cases} \tag{6}$$

Where, φ is the latitude of node j. Hence, while computing the measures of node importance in CNs, the nodes at higher latitudes get penalised. In this study, we compute DC and MSD using A_{ij}.

Measures of node importance in CNs

For a network having a total N number of nodes, Degree Centrality (DC) is the simplest measure of node importance which measures the number of connections of a node.

$$DC_j = \frac{\sum_{i=1}^{N-1} A_{i,j}}{N-1} \tag{7}$$

Hence, any grid having a high degree of connectivity means that drought onsets on the location (represented by the grid) are synchronised with a large number of droughts onsets across the globe.

Betweenness Centrality (BC) is a measure of node importance which quantifies the extent to which a node acts as a bridge or intermediary, facilitating the flow of information, resources, or interactions between other nodes in the network. It is calculated by determining the proportion of shortest paths in the network that pass through a particular node. BC is calculated as follows:

$$BC_j = \sum_{l \neq m \neq j \in V} \frac{n_{l,m}(j)}{N_{l,m}} \tag{8}$$

Where, $n_{l,m}(j)$ is the count of all possible shortest paths from node l to m that pass-through node j among all nodes (V). $N_{l,m}$ is the total number of shortest paths from node l to m. Hence, drought onsets on grids with higher BC scores can be considered to have greater control or influence within the network.

Spatial embedding of a network can be interpreted by analysing the actual geographical distances between the nodes in the network. We estimate the mean synchronisation distance (MSD) of a grid as an average of all the distances the grid is connected to weighted as per the strength of synchronisation (ES).

$$MSD = \frac{\sum_{i=1}^{N-1} ES_{ij} A_{ij} d_{ij}}{\sum_{i=1}^{N-1} ES_{ij} A_{ij}} \tag{9}$$

Where d_{ij} is the physical distance between points i and j.

PCMCI

Identifying causal relationships within complex systems is a fundamental challenge across various scientific disciplines and has also gained prominence in earth science in recent past. Peter and Clark Momentary Conditional Independence – PCMCI[82] – is a causal discovery algorithm which builds on the principles of conditional independence to delineate causal associations from observations while accounting for temporal delays. It is able to remove spurious links even in presence of high dimensional datasets[50,51,82] – a well-known problem in conditional independence based causal discovery approaches. PCMCI[51] achieves it by reducing the dimensions of the conditioning set prior to estimation of conditional independence based causal discovery. A brief description of PCMCI is as follows (for detailed algorithm please refer to Runge et al. 2019[82]).

The first step uses a modified PC algorithm (named after the inventors, Peter and Clark[83]) to generate the reduced conditioning set for each variable – called "Parents". From a set of variables $\overline{X}_t = X_t^1, X_t^2, \ldots, X_t^N$, for each variable $\overline{X}_t^j$, the PC stage starts with initialising preliminary parents $\overline{\mathcal{P}}(X_t^j) = (\overline{X}_{t-1}, \overline{X}_{t-2}, \ldots, \overline{X}_{t-\tau_{max}})$ and iteratively removes variables which are redundant and add no unique information when present in the conditioning set. The first iteration removes uncorrelated variables from $\overline{\mathcal{P}}(X_t^j)$, and the second iteration removes independent variables found after conditioning on the most correlated variables in first iteration.

Next, variables which are found independent after conditioning on two strongest drivers from the previous iteration are removed from $\overline{\mathcal{P}}(X_t^j)$. This is done until no variables are left to condition on. As recommended by Runge et al[50,51,82], we take an alpha level of 0.02 for hypothesis testing at this stage to not lose any true links from the $\overline{\mathcal{P}}(X_t^j)$.

The second stage, called the MCI stage uses the above-generated parents $\overline{\mathcal{P}}(X_t^j)$ for each variable X_t^j and tests the following null hypothesis at $\alpha = 0.05$ to find causal relationship between all variable pairs $X_{t-\tau}^i \rightarrow X_t^j$ at multiple lags $\tau = \{1, 2, \ldots, \tau_{\max}\}$:

$$\text{MCI} : X_{t-\tau}^i \perp\!\!\!\perp X_t^j | \overline{\mathcal{P}}(X_t^j) \setminus \{X_{t-\tau}^i\}, \overline{\mathcal{P}}(X_{t-\tau}^i) \forall X_{t-\tau}^i \in X_t^- \qquad (10)$$

where, $X_t^- = (X_{t-1}, X_{t-1}, \ldots, X_{t-\tau_{\max}})$. In this study, we have used a maximum lag $\tau_{\max}$ of 1 months and a partial correlation based conditional independence within PCMCI called 'ParCorr'. To avoid the penalty of high dimentionality and maintain high statistical power in conditional independence tests of PCMCI, we stick to a limited number of regions and test for causal discovery only at maximum lag 1.

We apply PCMCI on monthly SPI generated from ERA-5 reanalysis product to test for the presence of confounding behaviour of ocean regions. We first use PCMCI on monthly SPI of drought hubs on land and then perform three incremental addition of ocean variables. The first experiment contains SPI from land drought hubs only, the second experiment adds SPI from the Atlantic Ocean (AO), the third experiment adds the Maritime Continent (MC) in addition to the AO and the last experiment adds the Pacific Ocean (PO).

DATA AVAILABILITY

CODE AVAILABILITY

REFERENCES

1. Sheffield, J. & Wood, E. F. *Drought: past problems and future scenarios*. (Routledge, 2011).
2. Pokhrel, Y. et al. Global terrestrial water storage and drought severity under climate change. *Nat. Clim. Chang.* **11**, 226–233 (2021).
3. Satoh, Y. et al. The timing of unprecedented hydrological drought under climate change. *Nat. Commun.* **13**, 3287 (2022).
4. Zhang, Y., Keenan, T. F. & Zhou, S. Exacerbated drought impacts on global ecosystems due to structural overshoot. *Nat. Ecol. Evol.* **5**, 1490–1498 (2021).
5. McDowell, N. et al. Mechanisms of plant survival and mortality during drought: why do some plants survive while others succumb to drought? *New Phytol.* **178**, 719–739 (2008).
6. Ichoku, C. et al. Biomass burning, land-cover change, and the hydrological cycle in Northern sub-Saharan Africa. *Environ. Res. Lett.* **11**, 095005 (2016).
7. Stanke, C., Kerac, M., Prudhomme, C., Medlock, J. & Murray, V. Health Effects of Drought: a Systematic Review of the Evidence. *PLOS Cur.* **5**, ecurrents.dis.7a2-cee9e980f91ad7697b570bcc4b004 (2013).
8. Bryan, K. et al. The health and well-being effects of drought: assessing multi-stakeholder perspectives through narratives from the UK. *Clim. Chang.* **163**, 2073–2095 (2020).
9. Balbus, J. Understanding drought's impacts on human health. *Lancet Planet. Heal.* **1**, e12 (2017).
10. Apurv, T. & Cai, X. Impact of Droughts on Water Supply in U.S. Watersheds: The Role of Renewable Surface and Groundwater Resources. *Earth's Futur.* **8**, e2020EF001648 (2020).
11. Jiang, J. & Zhou, T. Agricultural drought over water-scarce Central Asia aggravated by internal climate variability. *Nat. Geosci.* **16**, 154–161 (2023).
12. Grayson, M. Agriculture and drought. *Nature* **501**, S1–S1 (2013).
13. Krishnamurthy, R. P. K., Fisher, J. B., Choularton, R. J. & Kareiva, P. M. Anticipating drought-related food security changes. *Nat. Sustain.* **5**, 956–964 (2022).
14. Zhu, P. et al. Warming reduces global agricultural production by decreasing cropping frequency and yields. *Nat. Clim. Chang.* **12**, 1016–1023 (2022).
15. Malhi, Y. et al. Climate change and ecosystems: threats, opportunities and solutions. *Philos. Trans. Royal Soc. B* **375**, 20190104 (2020).
16. Chang, H. & Bonnette, M. R. Climate change and water-related ecosystem services: impacts of drought in california, usa. *Ecosyst. Heal. Sustain.* **2**, e01254 (2016).
17. Li, P., Omani, N., Chaubey, I. & Wei, X. Evaluation of Drought Implications on Ecosystem Services: Freshwater Provisioning and Food Provisioning in the Upper Mississippi River Basin. *Int. J. Environ. Res. Public Heal.* **14**, 496 (2017).
18. Freire-González, J., Decker, C. & Hall, J. W. The Economic Impacts of Droughts: A Framework for Analysis. *Ecol. Econ.* **132**, 196–204 (2017).
19. Naumann, G., Cammalleri, C., Mentaschi, L. & Feyen, L. Increased economic drought impacts in Europe with anthropogenic warming. *Nat. Clim. Chan.* **11**, 485–491 (2021).
20. Yin, J. et al. Future socio-ecosystem productivity threatened by compound drought–heatwave events. *Nat. Sustain.* **6**, 259–272 (2023).
21. IPCC. Summary for Policymakers. *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* 3–32 (2021)
22. IPCC. Climate Change and Land: an IPCC Special Report on Climate Change, Desertification, Land Degradation, Sustainable Land Management, Food Security, and Greenhouse Gas Fluxes in Terrestrial Ecosystems. in 437–550 (Cambridge University Press, 2022).
23. Vicente-Serrano, S. M. et al. Global drought trends and future projections. *Philos. Transact. Royal Soc* **380**, 20210285 (2022).
24. WHO. Drought. https://www.who.int/health-topics/drought (2023).
25. UNCCD. *Drought in Numbers* https://www.unccd.int/sites/default/files/2022-06/Drought%20in%20Numbers%20%28English%29.pdf (2022).
26. Tijdeman, E., Blauhut, V., Stoelzle, M., Menzel, L. & Stahl, K. Different drought types and the spatial variability in their hazard, impact, and propagation characteristics. *Nat Hazard Earth Sys* **22**, 2099–2116 (2022).
27. Hao, Z., Singh, V. P. & Xia, Y. Seasonal Drought Prediction: Advances, Challenges, and Future Prospects. *Rev. Geophys.* **56**, 108–141 (2018).
28. AghaKouchak, A. et al. Status and prospects for drought forecasting: opportunities in artificial intelligence and hybrid physicalstatistical forecasting. *Philos. Transac. Royal Soc* **380**, 20210288 (2022).
29. Loon, A. F. V. et al. Drought in the Anthropocene. *Nat. Geosci.* **9**, 89–91 (2016).
30. Zhang, X. et al. Drought propagation under global warming: Characteristics, approaches, processes, and controlling factors. *Sci. Total Environ.* **838**, 156021 (2022).
31. Schumacher, D. L., Keune, J., Dirmeyer, P. & Miralles, D. G. Drought self-propagation in drylands due to land–atmosphere feedbacks. *Nat. Geosci.* **15**, 262–268 (2022).
32. Scheffer, M. et al. Anticipating Critical Transitions. *Science* **338**, 344–348 (2012).
33. Fair, K. M., Zachreson, C. & Prokopenko, M. Creating a surrogate commuter network from Australian Bureau of Statistics census data. *Sci. Data* **6**, 150 (2019).
34. Ruddell, B. L. & Kumar, P. Ecohydrologic process networks: 1. Identification. *Water Resources Res.* **45**, 3419 (2009).
35. Fang, X., Yang, Q. & Yan, W. Modeling and analysis of cascading failure in directed complex networks. *Safety Sci.* **65**, 1–9 (2014).
36. Fan, J. et al. Network-based Approach and Climate Change Benefits for Forecasting the Amount of Indian Monsoon Rainfall. (2020).
37. Ash, J. & Newth, D. Optimizing complex networks for resilience against cascading failure. *Phys. A: Stat. Mech. Appl.* **380**, 673–683 (2007).
38. Tsonis, A. A., Swanson, K. L. & Roebber, P. J. What do networks have to do with climate? *Bulletin of the American Meteorol. Soc.* **87**, 585–595 (2006).
39. Malik, N., Bookhagen, B., Marwan, N. & Kurths, J. Analysis of spatial and temporal extreme monsoonal rainfall over South Asia using complex networks. *Clim. Dyn.* **39**, 971–987 (2012).
40. Boers, N., Bookhagen, B., Marwan, N., Kurths, J. & Marengo, J. Complex networks identify spatial patterns of extreme rainfall events of the South American Monsoon System. *Geophys. Res. Lett.* **40**, 4386–4392 (2013).

41. Boers, N. et al. Complex networks reveal global pattern of extreme-rainfall tele-connections. *Nature* **566**, 373–377 (2019).

42. Mondal, S., Mishra, A. K., Leung, R. & Cook, B. Global droughts connected by linkages between drought hubs. *Nat. Commun.* **14**, 144 (2023).

43. Hassan, W. U. & Nayak, M. A. Global teleconnections in droughts caused by oceanic and atmospheric circulation patterns. *Environ. Res. Lett.* **16**, 014007 (2021).

44. Smith, M. ~D. et al. Drought-Net: A global network merging observations, experiments, and modeling to forecast terrestrial ecosystem sensitivity to drought. **2016**, B11J-06 (2016).

45. Mondal, S., Mishra, A. K. & Leung, L. R. Spatiotemporal Characteristics and Propagation of Summer Extreme Precipitation Events Over United States: A Complex Network Analysis. *Geophys. Res. Lett.* **47**, e2020GL088185 (2020).

46. Mukherjee, S., Mishra, A. K., Zscheischler, J. & Entekhabi, D. Interaction between dry and hot extremes at a global scale using a cascade modeling framework. *Nat Commun* **14**, 277 (2023).

47. Zeng, J., Gensheimer, M. F., Rubin, D. L., Athey, S. & Shachter, R. D. Uncovering interpretable potential confounders in electronic medical records. *Nat. Commun.* **13**, 1014 (2022).

48. Jager, K. J., Zoccali, C., MacLeod, A. & Dekker, F. W. Confounding: What it is and how to deal with it. *Kidney Int* **73**, 256–260 (2008).

49. Runge, J., Gerhardus, A., Varando, G., Eyring, V. & Camps-Valls, G. Causal inference for time series. *Nat. Rev. Earth Environ.* 1–19 (2023)

50. Nowack, P., Runge, J., Eyring, V. & Haigh, J. D. Causal networks for climate model evaluation and constrained projections. *Nat. Commun.* **11**, 1–11 (2020).

51. Runge, J. Causal network reconstruction from time series: From theoretical assumptions to practical estimation. *Chaos* **28**, 075310 (2018).

52. Guimerà, R. & Sales-Pardo, M. Missing and spurious interactions and the reconstruction of complex networks. *Proc. Natl. Acad. Sci.* **106**, 22073–22078 (2009).

53. Smirnov, D. A. Spurious causalities with transfer entropy. *Phys. Rev. E - Stat. Nonlinear Soft Matter Phys.* **87**, 1–12 (2013).

54. Jana, S., Rajagopalan, B., Alexander, M. A. & Ray, A. J. Understanding the Dominant Sources and Tracks of Moisture for Summer Rainfall in the Southwest United. *States. J. Geophys. Res.: Atmos.* **123**, 4850–4870 (2018).

55. Cheng, T. F., Lu, M. & Dai, L. Moisture channels and pre-existing weather systems for East Asian rain belts. *npj Clim. Atmos. Sci.* **4**, 32 (2021).

56. Reid, K. J., King, A. D., Lane, T. P. & Hudson, D. Tropical, Subtropical, and Extratropical Atmospheric Rivers in the Australian Region. *J. Clim.* **35**, 2697–2708 (2022).

57. Timmermann, A. et al. El Niño–Southern Oscillation complexity. *Nature* **559**, 535–545 (2018).

58. Guo, L. et al. Moisture sources for East Asian precipitation: mean seasonal cycle and interannual variability Moisture sources for East Asian precipitation: mean seasonal cycle and interannual variability. *J. Hydrometeorol.* **20**, 657–672 (2019).

59. Dey, D. & Döös, K. Tracing the Origin of the South Asian Summer Monsoon Precipitation and Its Variability Using a Novel Lagrangian Framework. *J. Clim.* **34**, 8655–8668 (2021).

60. Jiang, R., Xie, J., He, H., Luo, J. & Zhu, J. Use of four drought indices for evaluating drought characteristics under climate change in Shaanxi, China: 1951–2012. *Nat. Hazards* **75**, 2885–2903 (2015).

61. Adler, R. F. et al. The Version-2 Global Precipitation Climatology Project (GPCP) Monthly Precipitation Analysis (1979–Present). *J. Hydrometeorol.* **4**, 1147–1167 (2003).

62. Hersbach, H. et al. The ERA5 global reanalysis. *Quart. J. Royal Meteorol. Soc.* **146**, 1999–2049 (2020).

63. Singh, J. et al. Enhanced risk of concurrent regional droughts with increased ENSO variability and warming. *Nat. Clim. Chang.* **12**, 163–170 (2022).

64. Miralles, D. G., Gentine, P., Seneviratne, S. I. & Teuling, A. J. Land–atmospheric feedbacks during droughts and heatwaves: state of the science and current challenges. *Ann. N. York Acad. Sci.* **1436**, 19–35 (2019).

65. Deckker, P. D. The Indo-Pacific Warm Pool: critical to world oceanography and world climate. *Geoscience Lett* **3**, 20 (2016).

66. Jian, Z. et al. Warm pool ocean heat content regulates ocean–continent moisture transport. *Nature* **612**, 92–99 (2022).

67. Barlow, M. et al. A Review of Drought in the Middle East and Southwest Asia. *J. Clim.* **29**, 8547–8574 (2015).

68. Barlow, M., Cullen, H. & Lyon, B. Drought in Central and Southwest Asia: La Niña, the Warm Pool, and Indian Ocean Precipitation. *J. Clim.* **15**, 697–700 (2002).

69. Sandeep, S. & Ajayamohan, R. S. Modulation of Winter Precipitation Dynamics Over the Arabian Gulf by ENSO. *J. Geophys. Res.: Atmos.* **123**, 198–210 (2018).

70. Kumar, K. N. & Ouarda, T. B. M. J. Precipitation variability over UAE and global SST teleconnections. *J. Geophys. Res.: Atmos.* **119**, 10,313–10,322 (2014).

71. Ryoo, J.-M. et al. Impact of Rossby Wave Breaking on U.S. West Coast Winter Precipitation during ENSO Events. *J. Clim.* **26**, 130614135346002 (2013).

72. Zhang, W. et al. Tropical Indo-Pacific Compounding Thermal Conditions Drive the 2019 Australian Extreme Drought. *Geophys. Res. Lett.* **48**, e2020GL090323 (2021).

73. Ummenhofer, C. C., Kulüke, M. & Tierney, J. E. Extremes in East African hydro-climate and links to Indo-Pacific variability on interannual to decadal timescales. *Clim. Dyn.* **50**, 2971–2991 (2018).

74. Balas, N., Nicholson, S. E. & Klotter, D. The relationship of rainfall variability in West Central Africa to sea-surface temperature fluctuations. *Int. J. Clim.* **27**, 1335–1349 (2007).

75. Janicot, S., Moron, V. & Fontaine, B. Sahel droughts and Enso dynamics. *Geophys. Res. Lett.* **23**, 515–518 (1996).

76. Ehsan, M. A. et al. Atlantic Ocean influence on Middle East summer surface air temperature. *npj Clim. Atmos. Sci.* **3**, 5 (2020).

77. Mo, K. C. & Paegle, J. N. The Pacific–South American modes and their downstream effects. *Int. J. Clim.* **21**, 1211–1229 (2001).

78. Beguería, S., Vicente-Serrano, S. M., Reig, F. & Latorre, B. Standardized precipitation evapotranspiration index (SPEI) revisited: parameter fitting, evapotranspiration models, tools, datasets and drought monitoring. *Int. J. Clim.* **34**, 3001–3023 (2014).

79. Kumar, K. N., Rajeevan, M., Pai, D. S., Srivastava, A. K. & Preethi, B. On the observed variability of monsoon droughts over India. *Weather Clim. Extre.* **1**, 42–50 (2013).

80. Gupta, V. & Jain, M. K. Investigation of multi-model spatiotemporal mesoscale drought projections over India under climate change scenario. *J. Hydrol.* **567**, 489–509 (2018).

81. Mishra, A. K. & Singh, V. P. A review of drought concepts. *J. Hydrol.* **391**, 202–216 (2010).

82. Runge, J., Nowack, P., Kretschmer, M., Flaxman, S. & Sejdinovic, D. Detecting causal associations in large nonlinear time series datasets. *Sci. Adv.* **5**, eaau4996 (2017).

83. Spirtes, P. & Glymour, C. An Algorithm for Fast Recovery of Sparse Causal Graphs. *Social Sc. Comp. Rev.* (1991).

ACKNOWLEDGEMENTS

AUTHOR CONTRIBUTIONS

T.C. and S.G. conceived the idea and designed the problem. T.C. performed the analysis with inputs from V.C. All authors analysed the results. T.C. wrote the manuscript, V.C. and S.G. reviewed the manuscript.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41612-023-00558-1.

Correspondence and requests for materials should be addressed to Subimal Ghosh.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.