

## ARTICLE

Received 4 Apr 2015 | Accepted 28 Aug 2015 | Published 7 Oct 2015

DOI: 10.1038/ncomms9502

OPEN

# Identifying causal gateways and mediators in complex spatio-temporal systems

Jakob Runge<sup>1,2</sup>, Vladimir Petoukhov<sup>1</sup>, Jonathan F. Donges<sup>1,3</sup>, Jaroslav Hlinka<sup>4</sup>, Nikola Jajcay<sup>4,5</sup>, Martin Vejmelka<sup>4</sup>, David Hartman<sup>4,6</sup>, Norbert Marwan<sup>1</sup>, Milan Paluš<sup>4</sup> & Jürgen Kurths<sup>1,2,7,8,9</sup>

Identifying regions important for spreading and mediating perturbations is crucial to assess the susceptibilities of spatio-temporal complex systems such as the Earth's climate to volcanic eruptions, extreme events or geoengineering. Here a data-driven approach is introduced based on a dimension reduction, causal reconstruction, and novel network measures based on causal effect theory that go beyond standard complex network tools by distinguishing direct from indirect pathways. Applied to a data set of atmospheric dynamics, the method identifies several strongly uplifting regions acting as major gateways of perturbations spreading in the atmosphere. Additionally, the method provides a stricter statistical approach to pathways of atmospheric teleconnections, yielding insights into the Pacific-Indian Ocean interaction relevant for monsoonal dynamics. Also for neuroscience or power grids, the novel causal interaction perspective provides a complementary approach to simulations or experiments for understanding the functioning of complex spatio-temporal systems with potential applications in increasing their resilience to shocks or extreme events.

<sup>1</sup>Potsdam Institute for Climate Impact Research, PO Box 60 12 03, Potsdam 14412, Germany. <sup>2</sup>Department of Physics, Humboldt University, Newtonstrasse 15, Berlin 12489, Germany. <sup>3</sup>Stockholm Resilience Centre, Stockholm University, Kräftriket 2B, Stockholm 11419, Sweden. <sup>4</sup>Department of Nonlinear Dynamics and Complex Systems, Institute of Computer Science, Academy of Sciences of the Czech Republic, Pod vodárenskou věží 2, Prague 8 18207, Czech Republic. <sup>5</sup>Department of Atmospheric Physics, Charles University, V Holešovičkách 2, Prague 8 18000, Czech Republic. <sup>6</sup>Computer Science Institute, Charles University, Malostranské náměstí, 2/25, Prague 1, 18000, Czech Republic. <sup>7</sup>Institute for Complex Systems and Mathematical Biology, University of Aberdeen, Aberdeen AB24 3UE, UK. <sup>8</sup>Department of Control Theory, Nizhny Novgorod State University, Gagarin Avenue 23, Nizhny Novgorod 606950, Russia. <sup>9</sup>Institute of Applied Physics of the Russian Academy of Sciences, Ul Ulyanova 46, 603950, Nizhny Novgorod, Russia. Correspondence and requests for materials should be addressed to J.R. (email: jakobrunge@posteo.de).

A complex system's susceptibility to perturbations may crucially depend on where such a perturbation enters the system and how it is propagated. In the climate system, perturbations such as volcanic eruptions, extreme events<sup>1,2</sup> or anthropogenic manipulations such as air pollution and geoengineering<sup>3,4</sup> may have very different global effects if the region they occur in is strongly connected globally. The huge volcanic eruption of Mt. Pinatubo in June 1991 had a large impact on global climate<sup>5</sup> also because it is located in a climatologically sensitive region tied to atmospheric teleconnections, the tropical western Pacific<sup>6</sup>. Similarly, epileptic seizures in the brain<sup>7</sup>, blackouts in power grids<sup>8,9</sup>, epidemic spreading<sup>10,11</sup> or the failure of certain banks in the financial system<sup>12,13</sup> are key examples where subprocesses have a high cumulative effect on the whole complex system when perturbed, making them gateways of external influences spreading in the system. How can such subprocesses or regions be identified? Through which subprocesses are perturbations mainly mediated? These questions are key for understanding the dynamics and functioning of these systems, predicting their behaviour under perturbations and could help to make them more resilient. One way to address this problem is via active experiments, for example invasive brain stimulations in neuroscience (bearing ethical concerns) or by computer simulations, for example, in epidemic spreading models<sup>11</sup> or tracer experiments in climate<sup>14</sup>. Such simulations are, however, only possible if the underlying physical equations are known and even then the corresponding calculations, for example in climate research, are computationally expensive and may not adequately represent important processes<sup>15</sup>. Here we follow the complementary approach of using the data alone to retrieve information about the interaction dynamics of the complex system (exploiting passive or natural experiments). Data-driven analysis within the framework of complex networks<sup>16</sup> is a very active field of current research and has, among others, been applied to study the structure and function of complex systems in neuroscience<sup>17–20</sup> and more recently also in climate research<sup>21–26</sup>.

To identify how important individual subprocesses are in spreading and mediating perturbations in such spatio-temporal complex systems with time series typically given on a spatial grid, the subprocesses or nodes first need to be reconstructed since the gridded time series are often not the variables of interest. Secondly, the analysis should be based on a network that more faithfully than pairwise statistical associations represents possible pathways of perturbation propagation, requiring a causal definition of network links able to distinguish direct from indirect interactions. Last, even if all links in a network were established to be 'statistically causal', the toolbox of classical network measures is not rich enough for quantifying gateways and mediators of perturbations. Essentially, these measures—with many originating from the social sciences<sup>27</sup>—are based on a different definition of links, for example, two persons knowing each other, as opposed to dynamical interactions in a complex system. Hence, what is needed are quantitative measures that take into account the relative importance of causal pathways on which perturbations propagate in a complex system's interaction network.

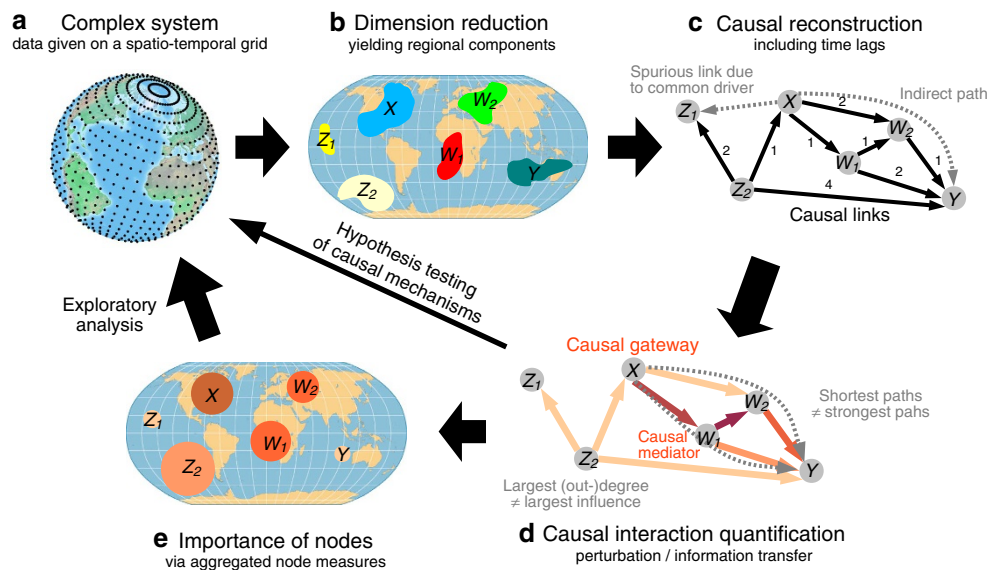
Here, we present such an approach based on three steps: First, a dimension reduction of a gridded data set using the Varimax approach<sup>28,29</sup> to a set of components representing relevant subprocesses defining the network's nodes. Secondly, a (multivariate) causal reconstruction of the network's links based on a causal discovery algorithm<sup>30–32</sup> and, thirdly, a causal interaction quantification utilizing Pearl's causal effect theory<sup>33–35</sup> to construct a causally weighted directed network on which we define network measures that are better suited for

quantifying key regions of causal perturbation spread and mediation compared to classical network measures such as the node degree and betweenness centrality<sup>36</sup>. The extent to which such a data-driven analysis allows for a causal interpretation depends on the included variables, time resolution of the data and assumptions such as stationarity. We demonstrate the potentials of our method on a global data set of surface pressure as a representative characteristic of atmospheric variability. Applied to test specific hypotheses, we find that within this pressure system the climatic interaction mechanism between the East Pacific limb of the El Niño Southern Oscillation (ENSO)<sup>37,38</sup>, and the Arabian Sea region, relevant for the Indian Monsoon system, is mainly mediated via the Indonesian archipelago. This application of our method also incorporates more rigorously the concept of atmospheric teleconnections, which were previously defined based on pairwise correlations<sup>39,40</sup>. As an exploratory tool, the method identifies several strongly uplifting regions of major convergence of low-level air masses and high-level air uplifts above the tropical oceans. These subprocesses integrate incoming perturbations at the surface and transport them vertically into the higher troposphere from which they again influence other surface processes via atmospheric downdrafts. This mechanism explains the key importance of these regions as gateways of perturbation spread along causal pathways in the atmosphere. Our approach is of substantial value for several applications. In climate research it may allow to more efficiently allocate resources to understand, monitor and forecast these important subprocesses. For other applications, like epileptic seizure prevention, it may help to more reliably identify which brain regions are seizure foci to concentrate counter measures on. In summary, the novel causal interaction perspective provides a complementary first-step approach towards model simulations and experiments to better understand the dynamics and functioning of complex spatio-temporal systems and may help to inform design and engineering processes aiming at increasing their resilience.

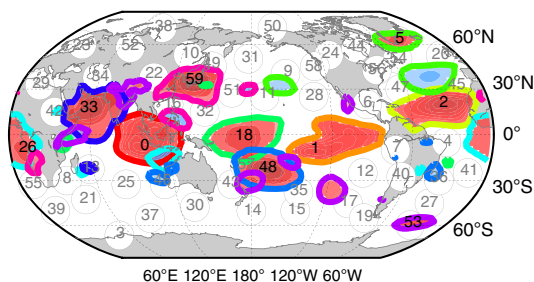
## Results

**Dimension reduction and causal reconstruction.** In the following, we explain and illustrate each of the three steps (see Fig. 1) with a climate example. More technical details are given in Methods.

In climate research, spatio-temporal data sets are typically given on a regular grid. Here we consider a reanalysis data set of surface pressures<sup>41</sup> for the period 1948–2012. At a resolution of 2.5° in latitude and longitude, the data set consists of 10,512 grid points with 3,339 samples for each time series on a weekly timescale. But towards an interpretation of perturbation propagation or information transfer<sup>42</sup>, such individual grid points are not the entity of interest, because they do not represent distinct climatological processes. For example, processes like ENSO require a special decomposition of the data fields for an efficient description of their spatio-temporal structure<sup>43</sup>. Also from a statistical perspective, a large number of variables with comparatively few observations presents an estimation problem<sup>44</sup>. The first step of our approach, therefore, is aimed at obtaining a small set of components that represent relevant subprocesses of the complex system. We choose Varimax-rotated principal components<sup>28,45</sup> here, combined with a subsequent significance test<sup>29,46</sup> to exclude components merely representing noise. For the atmospheric pressure data set, this dimension reduction algorithm yields a set of 60 components (cf. Methods and ref. 29). As shown for selected components in Fig. 2 (all components shown in Supplementary Figs 1, 2), the corresponding regions well represent several important climatological subprocesses. As further discussed in Methods,



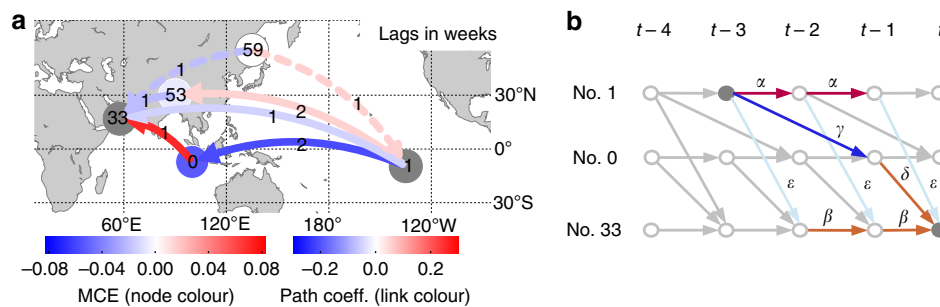
**Figure 1 | Schematic representation of our proposed approach illustrated for the complex system Earth.** (a) Climatological variables such as sea-level pressure or surface temperature are typically provided as time series at locations on a regular grid. (b) In the first step of our approach, a Varimax-dimension reduction yields a small set of regional components (here denoted by  $X$ ,  $Y$ , ...) representing climatological subprocesses with corresponding time series. (c) In the second step of our approach, this smaller set of variables allows for a reconstruction of the causal network (black links, labels denote time lags). An important pitfall in non-causal networks (as constructed from pairwise association measures such as cross-correlation) is that links can be spurious due to a common driving by another process or due to transitivity effects leading to indirect paths (grey dashed arrows in c). Perturbations cannot propagate along common-driver links (for example, between  $X$  and  $Z_1$ ) affecting network measures like the degree. Further, indirect paths such as from  $X$  to  $Y$  affect shortest path lengths in non-causal networks. (d) In the third step of our approach, the aim is to directly quantify the causal effect between pairs of components based on the causal network (Tigramite approach) and detect through which components and how much the causal effect is mediated. In the linear framework studied here this can be achieved by causal effect measures based on suitable link weights, where the weight of a link, for example  $X \rightarrow W_1$ , indicates the causal effect of a 1 s.d. perturbation in  $X$  on  $W_1$  (see Fig. 3b for a formula relating link weights to causal effects). Binary causal networks do not properly account for different link strengths which affects classical network measures (grey highlights in d). This analysis can be used to test specific hypotheses, but also to construct aggregate node measures (e) to identify components with high cumulative causal effect either as sources (causal gateways) or as intermediate nodes on causal pathways (causal mediators).



**Figure 2 | Dimension reduction of an atmospheric data set.** Map of the 60 Varimax components, enumerated from 0 to 59 (grey nodes). The location of the nodes is determined by the largest absolute spatial loading. For selected components (all components shown in Supplementary Figs 1, 2), we show the core 98% area of loadings. The colour of the surrounding line identifies the parts belonging to one component, sometimes several regions belong to one component (examples include dipole patterns). Some components can be associated with well-known climatic processes like the El Niño Southern Oscillation<sup>37</sup> (ENSO, No. 0 representing the western uplift and No. 1 the eastern downdraft limbs), or the North Atlantic Oscillation (NAO, No. 5 with a pronounced dipole structure over the Icelandic Low and the Azores High). Others are related to global monsoon systems: No. 33 in the Arabian Sea high-surface-pressure sector of the Indian Monsoon region<sup>53</sup> and No. 26 in the tropical Atlantic West African Monsoon system<sup>43</sup>. Not all components are regionally confined, for example, No. 53 with loadings in the South Atlantic as well as in the Himalayas.

all components are anomaly (seasonal cycle removed from the mean and variance) and standardized. Here we study intraseasonal interactions at a weekly time resolution.

Given that an external perturbation occurs in one of these components representing a certain subprocess: On which paths can it propagate and which other subprocesses can it possibly reach? Suppose a perturbation enters in subprocess  $X$  in the example in Fig. 1c, then it can only reach nodes further 'downstream' on causal paths like  $W_1$ ,  $W_2$  and  $Y$ , but not  $Z_1$ , even though they are statistically associated. Such spurious links can be unveiled by causal discovery algorithms<sup>30,47</sup> which iteratively test whether an association can be explained by another process in the network. Note that this notion of causality is only to be interpreted with respect to the included variables and unobserved drivers can still cause spurious links. The causal reconstruction steps are detailed in Methods (see also Supplementary Fig. 3). In essence, for the second step of our approach we employ a causal discovery algorithm adapted to time series<sup>31,32</sup> and a subsequent thresholding step to study the robustness of all further analyses at different link densities. This approach yields the causal time series graph<sup>31,32,48</sup> which is a special type of a graphical model<sup>49</sup> and encodes the conditional dependencies of the components at different time lags. For the example of the ENSO—Indian Ocean teleconnection studied next, Fig. 3 depicts two different representations: Fig. 3b shows the time series graph on which causal paths and the measures of causal effect (the third step of our approach) are based, while the aggregated causal network shown in Fig. 3a can be better visualized.



**Figure 3 | Results for the teleconnection between ENSO and the Indian Ocean within the causal network of pressure components.** (a) Exemplary selection of causal paths (solid lines) for the effect of No. 1, the eastern limb of ENSO, on No. 33 in the Arabian Sea at a lag of 3 weeks via components No. 0 above the Indonesian archipelago and No. 53 over East Asia. The edge colour corresponds to the path coefficient (labels denote the lags in weeks) and the node colour to the MCE. The causal measures exclude confounding effects due to common drivers of both processes, such as No. 59 (dashed links), but also from component No. 1 at lags further in the past as the more detailed time series graph representation reveals. (b) Time series graph depicting only links belonging to the most relevant causal paths (line color indicating the path coefficient) between Nos. 1 and 33 (grey nodes) at lag  $\tau = 3$  weeks via No. 0. As further discussed in Methods, one can estimate the total causal effect by summing over the products of path coefficients (link labels) along each path:  $I_{1 \rightarrow 33}^{CE}(\tau = 3) = \alpha^2 \epsilon + \alpha \epsilon \beta + \epsilon \beta^2 + \gamma \delta + \text{negligible other paths} = -0.08 \pm 0.01$ . The mediated effect via component No. 0 is given by the only path through this node:  $I_{1 \rightarrow 33|0}^{MCE}(\tau = 3) = \gamma \delta = -0.053 \pm 0.006$ . The s.e. are computed from a residual bootstrap (see Methods). Intermediate components can also counteract the total causal effect if MCE is of different sign than CE.

In our climate application, we consider time lags up to  $\tau_{\max} = 4$  weeks, since we are interested in atmospheric interactions where dependencies typically decay within a month<sup>50</sup>, but our results are robust also for longer time lags. Contemporaneous associations (possibly because of unobserved common drivers or faster interactions) can be represented as undirected links in the time series graph<sup>31,32</sup>, but these are not taken into account here since they are not regarded as causal. In the following, we discuss results for a link density of 20% in the causal network consisting of  $N = 60$  components, but our main results are also robust for link densities between 10 and 50% and other parameters of the method.

**Quantifying causal effect.** The causal time series graph allows to qualitatively determine which causal paths a perturbation can possibly take. Now we employ measures to quantify the causal effect of hypothetical perturbations and their mediation along causal paths, exemplified on the teleconnection mechanism by which component No. 1 in the East Pacific ENSO region influences component No. 33 describing surface pressures in the Arabian Sea with relevance also for the Indian Monsoon rainfall<sup>37,38,43</sup> (see component regions in Fig. 2).

We approach this problem by using measures of causal effect in the framework of structural equation modelling<sup>33,35</sup>. With the reconstructed time series graph as a causal hypothesis, we fit a linear regression model to the multivariate component time series  $\mathbf{X}$  with non-zero coefficients for every link in the time series graph. The standardized regression coefficient for a direct causal link  $X_{t-\tau}^i \rightarrow X_t^j$  between two components  $i, j \in \{0, \dots, N-1\}$  at lag  $\tau$  (in weeks) is then called the path coefficient<sup>33,51</sup>. This makes the time series graph a causally weighted directed network. The matrices of path coefficients between all components are shown in Supplementary Fig. 4.

Rather than studying only causal effects between adjacent nodes in the causal time series graph, here we are interested in the total causal effect (CE) also along indirect causal paths. Under certain assumptions (see Methods), the CE between two components  $i$  and  $j$  at lag  $\tau$ , denoted  $I_{i \rightarrow j}^{CE}(\tau)$ , can be evaluated by summing over the products of the path coefficients along each causal path<sup>35</sup> and carries the causal interpretation as the expected change in  $X_t^j$  (in units of its s.d. and relative to the unperturbed regime) at time  $t$  if  $X_t^i$  was perturbed at time  $t - \tau$  by a one s.d. delta

peak. The matrices of CEs between all components are shown in Supplementary Fig. 5. Similarly, the mediated causal effect (MCE)  $I_{i \rightarrow j|k}^{MCE}(\tau)$  via another component  $k$  can be measured by summing only over those paths that pass through component  $k$  (at any lag). These measures are now illustrated for the teleconnection mechanism between components Nos. 1 and 33.

In the atmospheric pressure system we find 31 indirect causal paths between the Eastern ENSO component No. 1 and the Arabian Sea component No. 33 at a lag of 3 weeks (only a selection via components No. 0 above the Indonesian archipelago and No. 53 over East Asia shown in Fig. 3a). The total CE sums up to  $-0.08 \pm 0.01$  here, implying that a perturbation of 1 s.d. in the East Pacific yields a decrease in No. 33 of  $\sim 8\%$  in units of its s.d. (note that component No. 1 has deviations of several s.d. from the seasonal mean). Further, we find that component No. 0 above the Indonesian archipelago mediates  $-0.053 \pm 0.006$  of that effect, here resulting from the causal chain No. 1  $\xrightarrow[-0.21 \pm 0.01]{2 \text{ weeks}}$  No. 0  $\xrightarrow[0.25 \pm 0.03]{1 \text{ week}}$  No. 33. This MCE explains more than 60% of the

total CE while other paths, for example via East Asia (No. 53), contribute less than 10%. For comparison, counting just the fraction of causal paths passing through a given node, in analogy to betweenness centrality<sup>36</sup>, here yields 1 from a total of 31 paths for No. 0 and, for example, 8 for component No. 35, even though the latter's mediated effect is much weaker. We estimated the CE for a link density of 20% here, in Supplementary Table 1 we show that our results are largely robust to this choice. On the other hand, we find that the interaction was much weaker in the first half of the data set (1948–1980).

Climatologically, our present analysis implies that of the many possible climatic mechanisms linking sea-level pressure anomalies in the ENSO region to pressure variability west of India, only the mechanism via No. 0 is relevant, at least within the intraseasonal timescale of the atmospheric surface pressure system and integrated over all seasons. Note that conclusions about an effect of ENSO on the Indian Monsoon are also complicated by the apparent non-stationarity of this relationship<sup>52</sup>. More detailed analyses taking into account additional climatological variables such as temperature, only certain seasons (for example, during El Niño phases), and filtering out non-relevant time scales (such as from oceanic drivers) can provide more accurate estimates of CEs for more specific climatological hypotheses.

**Causal gateways and mediators.** The foregoing case study introduced the CE measures and is an example of causal modelling for testing specific hypotheses about interaction mechanisms. Now we study aggregated node measures based on the causally weighted directed network in a more exploratory analysis to identify the importance of components as gateways for spreading and mediating hypothetical perturbations in the network.

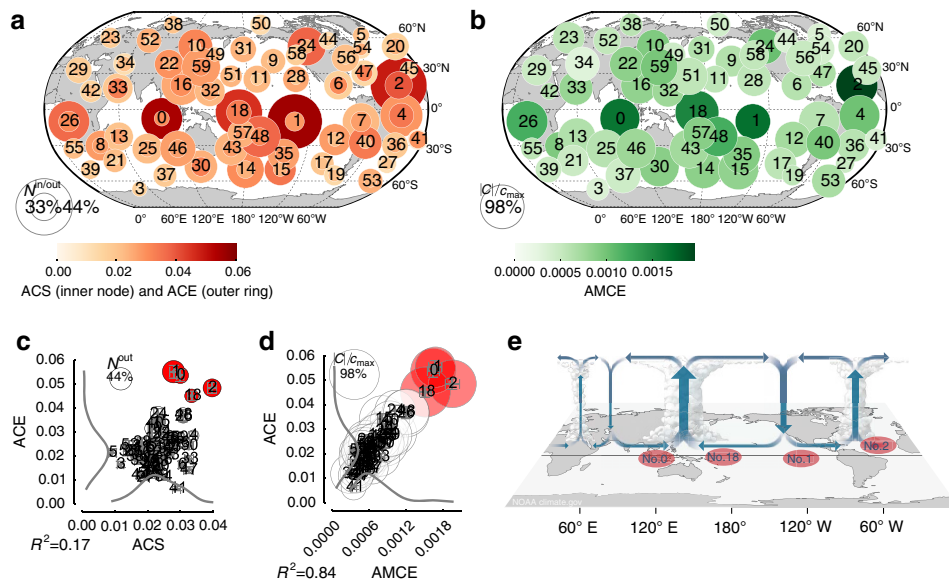
As aggregated first-order measures of CE, we consider the matrix of CEs between all pairs of components (Supplementary Fig. 5) taken at the lag with maximum absolute effect  $I_{i \rightarrow j}^{CE, \max} = \max_{\tau} |I_{i \rightarrow j}^{CE}(\tau)|$ . Then we define the mean along each column as the average causal effect (ACE) that a component has on the rest of the system and the row-mean as the average causal susceptibility (ACS) as a measure of how sensitive a component is to perturbations entering in other parts of the system. To measure how strong a subprocess mediates CEs spreading throughout the complex system, we propose the average mediated causal effect (AMCE) of a component  $k$  by averaging the previously defined MCE over all interaction pairs with causal paths through  $k$  (more details in Fig. 4 and Methods). As opposed to a path-based network measure like betweenness centrality<sup>36</sup>, AMCE depends not so much on the number of paths through a given component, but more on how strong the CE along these paths is.

In Fig. 4 we depict ACE, ACS and AMCE for the atmospheric pressure system. Although the distribution of susceptibilities is quite broad (Fig. 4c), few components have a very strong ACE (red nodes Nos. 0, 1, 2, 18) and are also rather susceptible. These components reach a large fraction of processes (node size in

Fig. 4a,c) and correspond to processes in the tropical oceans (No. 0 over Indonesia and the East Indian Ocean, No. 2 in the Atlantic, Nos. 1 and 18 in the East and West Pacific). A one s.d. perturbation entering these processes has a large effect of up to 0.3 on other processes and each of them drives more than 10 other processes with a CE of at least 0.1 (Supplementary Fig. 6). These components, thus, act as major gateways of perturbation spread and also belong to the most susceptible processes being causally driven by  $\sim 20$ –30% of the other components with an ACS above 0.05.

Figures 4b,d demonstrate that there is not much correlation between the fraction of interactions with a path through a certain component (node size) and its AMCE ( $R^2 = 0.36$ ). Components Nos. 0, 1, 2, 18 (but also Nos. 26 and 48) are the most dominant causal mediators being involved in more than 80% of all interactions with an AMCE between 0.0015 and 0.002. Note that the average non-zero CE between any pair is only about 0.02. In Supplementary Figs. 7, 8 we show that these results are robust for different link densities and other parameters of the method, in particular the maximum lag  $\tau_{\max}$  and the significance level in the algorithm. The results are also robust if only the first (1948–1980) or second half (1981–2012) of the data set is used.

Climatologically, components Nos. 0, 1, 2, 18 correspond to major convergence regions with ascending motion of air masses (cf. Fig. 4e)<sup>43,53</sup>. In particular, component No. 0 (and to a smaller degree No. 18) represents the uplifting western limb of the Walker circulation over Indonesia. This region is one of the strongest atmospheric convergence zones where moist air masses rise up and affect global tropical and extratropical climate via



**Figure 4 | Aggregated measures of causal effect in the causal network of pressure components.** (a) For each node/component, the fraction of components  $N_j^{\text{in}}$  affecting component  $j$  with  $I_{i \rightarrow j}^{CE, \max} = \max_{\tau} |I_{i \rightarrow j}^{CE}(\tau)| > 0.05$  (that is, a 5% effect taken at the lag with maximum absolute effect) scales with the inner node diameter and the fraction of processes  $N_i^{\text{out}}$  reached by component  $i$  scales with the width of the outer ring. The colours denote the ACS  $I_j^{\text{ACS}} = \frac{1}{N-1} \sum_{i \neq j} I_{i \rightarrow j}^{CE, \max}$  (inner node) and average causal effect  $I_j^{\text{ACE}} = \frac{1}{N-1} \sum_{i \neq j} I_{i \rightarrow j}^{CE, \max}$  (outer ring). (b) For each node/component  $k$ , the fraction of pairs  $|C_k|/c_{\max}$  (with  $c_{\max} = (N-1)(N-2)$ ) where this node is a mediator on a causal path (up to a lag of  $\tau_{\max} = 4$  weeks) scales with the node diameter. The colour shows the average mediated causal effect  $I_k^{\text{AMCE}} = \frac{1}{|C_k|} \sum_{(i,j) \in C_k} \max_{\tau} |I_{i \rightarrow j}^{\text{MCE}}(\tau)|$ . (c) and (d) depict ACE versus ACS and AMCE, respectively. The thick grey lines denote the density of the marginal distributions (arbitrary units). The error bars show the standard errors computed from a residual bootstrap (see Methods). (e) shows a schematic view of the global equatorial Walker circulation during normal conditions. Note that during La Niña events the uplift and downdraft regions are enhanced and during El Niño events they are reversed. Our analysis constitutes an average over all seasons (adapted with permission from NOAA Climate.gov drawing by Fiona Martin). The red components in (c,d) with high ACE, ACS and AMCE correspond to major convergence regions that integrate incoming perturbations at the surface and transport them vertically into the higher troposphere from which they influence other components via atmospheric downdrafts. In Supplementary Fig. 6 we depict more detailed results for these major gateways.

teleconnections in the upper troposphere. Component No. 2, located in the tropical Atlantic, also features strong uplifting deep convection<sup>43,53</sup>. The core ENSO region represented by component No. 1 plays a double role depending on the state of the ENSO system<sup>37,38</sup>. During normal ENSO conditions (depicted in Fig. 4e) it is a region of descending upper tropospheric air masses and, thus, not as much governed by surface pressures. During El Niño events, on the other hand, it is a region of strong uplifts. These effects are mixed in our analysis and more detailed studies can further distinguish seasons to obtain a more precise picture of seasonal climatic interactions. In summary, these strongly uplifting regions integrate incoming perturbations at the surface and transport them vertically into the higher troposphere from which they again influence other surface processes via atmospheric downdrafts. This mechanism explains the key importance of these regions as causal gateways and mediators of perturbations spreading in global climate via atmospheric teleconnections. Our analysis considered delta peak perturbations of 1 s.d. Since often perturbations reach extreme deviations of several s.d.<sup>1,2,54</sup>, and, even more importantly, multiple perturbations can accumulate, these findings reflect the large global influence of these regions.

## Discussion

We have introduced a three-step approach for the analysis of multivariate spatio-temporal data sets, consisting of a dimension reduction, causal reconstruction, and CE quantification to identify subprocesses in complex systems that are important gateways for spreading and mediating perturbations entering the system in one subprocess. While this approach lends itself also to other spatio-temporal complex systems such as the brain<sup>18</sup>, for applications to financial data or food webs, the causal reconstruction can already be applied to, for example, economic indices or species abundances, and in complex systems like power grids or transportation networks, the first two steps of our approach could be skipped since the network structure is naturally given.

The causal quantification approach takes classical analyses of functional brain networks<sup>18</sup> or climate networks<sup>21,23</sup>, which were previously mostly based on pairwise association measures, to a new level. Consequently, our proposed node measures can be seen as dynamical and causal alternatives to classical measures for functional networks such as the degree or betweenness centrality. In Supplementary Fig. 9 and Supplementary Note 1 we compare our results with these measures where we find that the latter have only weak predictive power ( $R^2 \approx 0.4$ ) for perturbation propagation and mediation. The pathway analysis goes substantially beyond pure causal network reconstructions<sup>24,32</sup> and also provides, for the first time, a stricter statistical approach for characterizing atmospheric teleconnections, which are of paramount importance for studies of climate change and in particular climate extremes<sup>55</sup>, and which were previously formulated more phenomenologically based on (lagged) correlations<sup>39,40,56</sup>.

Like any data-driven approach our method is limited by several assumptions: causal sufficiency<sup>30,33</sup> assumes that the common drivers of all variables are taken into account and the causal Markov condition assumes all ‘error terms’ of the nodes in the time series graph to be independent. In our climate analysis, for example, we only excluded common drivers from within the pressure system, but on larger monthly time scales the underlying sea surface temperature field certainly interacts with the faster atmospheric pressure field over the oceans<sup>57</sup> and can confound the assessment of CEs. It is, therefore, important to interpret CEs only relative to the variables that were taken into account.

A further complication are CEs that are faster than the weekly resolution considered here, which appear as contemporaneous in our analysis, but are not taken into account since they are not regarded as causal links. Our weekly time resolution reflects a balance between resolving causal directionality and a multiple testing problem if too many lags are considered (in our example 30 days). Also the interplay of different time scales<sup>58,59</sup> could be further addressed, for example by singular spectrum analysis<sup>45</sup>, and one could possibly also account for time-varying time delays of interactions<sup>60</sup>. To estimate the time series graph and CEs from the observed time series, we assume stationarity such that these properties do not change over time. More detailed research questions can take into account non-stationarity, for example, due to seasonality in climate (here we used the whole time series). While the linear approach can also be adapted from delta-peak perturbations to more general scenarios with multiple or different types of perturbations<sup>35</sup>, the perturbations must be small enough to conserve the dynamics and causal structure of the system such that the conditional distributions remain the same<sup>35</sup>. The effects of large unprecedented perturbations cannot be predicted from observed data alone. We introduced the method using simple linear measures here, but the framework can to some extent also be implemented with nonlinear quantifiers, for example using information-theoretic measures<sup>42,61</sup>.

We see the proposed method as a complementary first-step approach towards model simulations and experiments to help guide decision making in several ways: in climate, the knowledge of subprocesses or regions with large perturbative effect, either as gateways or mediators, can help to optimally design computationally expensive simulations such as tracer experiments<sup>14</sup>, geoengineering impact assessments<sup>3,4</sup>, or extreme event attribution studies<sup>62</sup>. Such experiments allow to conduct counterfactual analyses, for example, with and without anthropogenic influences, to conclude on necessary and sufficient CEs<sup>63</sup>. In neuroscience, it could help to optimize therapeutic interventions for preventing seizures by targeting selected brain regions with large CE or mediating CE. In power grids, nodes with strong mediating effect are the ones that one would best block to prevent a blackout perturbation from spreading throughout the network. Summarizing, the novel causal interaction perspective provides a general approach to better understand the possible influence of perturbations on complex spatio-temporal systems and may guide further research to inform design and engineering processes aiming at increasing their resilience against shocks or extreme events.

## Methods

**Data and software availability.** The climatological reanalysis data set<sup>41</sup> studied here can be downloaded from <http://www.esrl.noaa.gov/psd/data/gridded/data.ncep.reanalysis.html>. Code for the dimension reduction step is available from co-author M. Vejmelka at <https://github.com/vejmelkam/ndw-climate/blob/master/scripts>. A Python software script by J. Runge to estimate the causal network can be obtained from <http://tocsy.pik-potsdam.de/tigramite.php>.

**Dimension reduction.** Our dimension reduction approach is based on Varimax-rotated principal components<sup>28,45</sup> and a subsequent significance test to eliminate components merely representing noise<sup>29</sup>. As further discussed in ref. 29, the rotation of principal components maximizes the sum of the variances of the squared principal component weights (loadings) and better represents regionally confined processes than principal components. The data preprocessing steps to obtain the component weights are discussed in more detail in ref. 29,

here we give a brief summary: Monthly gridded time series are first anomalized to remove the annual cycle not only from the seasonal means but also from the seasonal variance. After a linear detrending, the covariance matrix is estimated on cosine-transformed data to account for the area a grid point represents (poles are excluded), and the eigenvectors are computed. These are then rotated using the Varimax criterion<sup>28</sup> and a limited number of components is selected based on a comparison of eigenvalues of original data (not components) to those from surrogate data which preserve the autocorrelation structure, but destroy dependencies between the grid point time series. Here this algorithm yields  $N=60$  significant components. Finally, the component weight matrix is multiplied with the daily original gridded time series (that have been preprocessed by anomalization in mean and variance, standardization, linear detrending and cosine transform as above), and the daily component time series are aggregated to a weekly resolution which reflects a balance between causal time resolution and the multiple testing problem in the causal reconstruction step. In contrast to principal components, where the diagonal entries corresponding to the eigenvalues can be interpreted as the explained variability, for rotated principal components, the off-diagonal entries are not zero anymore and one cannot simply attribute an ‘explained variance’ to each component. We enumerate the components by the entry on the diagonal starting with the largest value (component No. 0). Here monthly time series were used for the extraction of the components for computational reasons. Carrying out the decomposition directly on the daily or weekly time series might have provided a slightly different set of components, as the decomposition would also take into account higher frequency variability. In Supplementary Figs 1, 2 we show the loadings and time series of all components.

**Causal reconstruction.** To reconstruct the causal network from the component time series, we utilize a causal discovery algorithm<sup>31,32</sup> which is based on the PC algorithm (named after its inventors Peter Spirtes and Clark Glymour<sup>30</sup>). This algorithm can be used in an information-theoretic framework<sup>31</sup> as well as employing linear partial correlation<sup>24,32</sup>. Here we choose the linear approach as a first-order approximation. The significance level used in the causal algorithm is not a very reliable indicator for the final significance level of causal links because links are tested sequentially and, therefore, Bonferroni corrections cannot be easily applied. To overcome this problem, we use the causal algorithm only as a variable selection for a subsequent ‘causal regression’. The time series graph reconstruction, thus, consists of three steps as described below.

First, variable selection of the causal parents: The parents  $\mathcal{P}_j$  of each component  $j$  are selected with the causal algorithm described in ref. 32 and Supplementary Note 2. The algorithm’s parameters here are: maximum time lag  $\tau_{\max}=4$  weeks, (two-tailed) significance level  $\alpha=0.001$  (Student’s  $t$ -test), initial number of conditions  $n_0=3$ . For the causal algorithm to consistently converge to the true parents, one needs to assume causal sufficiency and the causal Markov condition, that is, the independence of error terms driving each subprocess, and faithfulness<sup>30</sup> which guarantees that the graph entails all conditional independence relations true for the underlying process and can be violated in certain pathological cases<sup>30</sup>. Since we estimate partial correlations from time series data here, we also assume stationarity. Ref. 64 discusses the computational complexity of the algorithm. In Supplementary Fig. 3b we show the distribution of the number of parents for every component (black dashed line) and in Supplementary Tables 2, 3 we list the parents for all components (median number of parents is 8).

Secondly, estimating the causal regression matrix: the lagged causal regression matrix  $\mathbf{C}(\tau)$  of shape  $(N, N, \tau_{\max})$  is estimated using the above selected parents by

$$C_{i \rightarrow j}(\tau) = r_{ji, \mathcal{P}(j)}(\tau) \quad \text{for } \tau = 1, \dots, \tau_{\max} \quad \text{and } i, j = 0, \dots, N-1, \quad (1)$$

where  $r$  denotes the standardized regression coefficient of component  $X_{t-\tau}^i$  in the multiple regression model of  $X_t^j$  on  $\{X_{t-\tau}^i, \mathcal{P}(X_t^j)\}$  using ordinary least squares regression. For infinite sample sizes, these ‘causal’ regression coefficients would be non-zero only for the parents of each component as estimated with the algorithm. In Supplementary Fig. 3a we show the sorted coefficients for every component  $j$ . The largest coefficients are typically associated with the past lag of a component. After a sharp decay, most of the coefficients have absolute values below 0.1.

Thirdly, constructing the time series graph: The causal time series graph is constructed from thresholding the causal regression matrix with cross-links ( $i \neq j$ ) and autolinks ( $i = j$ ) defined by

$$X_{t-\tau}^i \rightarrow X_t^j \Leftrightarrow |C_{i \rightarrow j}(\tau)| \geq \theta, \quad (2)$$

with the threshold  $\theta$  chosen to obtain a given link density in the corresponding aggregated causal network (Supplementary Fig. 9b), that is,  $\rho_{\text{dens}} = \frac{\# \text{ cross-links in causal network}}{N(N-1)}$ , where autolinks are not counted and multiple links between two components are only counted once. For the link density  $\rho_{\text{dens}}=0.2$  analysed here, there are 708 such links, while the time series graph, thresholded at  $\theta=0.0585$ , has a link density of 0.062. Because of the assumed stationarity, the subscript  $t$  in equation (2) can be dropped. For the network analysed in the main article at 20% link density, the median of parents in the time series graph is 14 which determines the non-zero coefficients in the CE-estimation model (4) described below. Supplementary Fig. 3b shows the distribution of parents for different link densities.

**Causal effect estimation.** There are different ways to use the reconstructed causal network to further quantify causal interactions between subprocesses. We call the general idea to use the time series graph for quantifying general causal interactions the Tigramite approach (time series graph-based measures of information transfer), which is also the abbreviation of the accompanying software package. Here we consider a measure  $I$  to quantify the linear CE of perturbations for its reliable estimation and interpretability, generalizations in an information-theoretic framework are discussed in refs 42,61.

The approach is based on the CE estimator for multivariate time series proposed in ref. 35 in a linear application of Pearl’s causal framework<sup>33</sup>, considering delta-peak perturbations (called atomic interventions in ref. 35). Within this framework, the CE of a perturbation of setting  $X_{t-\tau}^i$  to  $x^*$  on  $X_t^j$  is given by

$$\mathbb{E}[X_t^j \mid do(X_{t-\tau}^i) = x^*], \quad (3)$$

where  $\mathbb{E}[\dots]$  denotes the expectation. It is important to note that the  $do$ -operator does not pertain to a conditional expectation, but refers to the experiment of intervening in the system and forcing the variable to a certain value. From observational data, this effect can only be estimated (or identified) under certain assumptions<sup>33,35</sup>. Here we assume a linear model based on the reconstructed time series graph with all relevant variables (or confounders) included (thus, satisfying the back-door criterion<sup>33,35</sup>):

$$\mathbf{X}_t = \sum_{\tau=1}^{\tau_{\max}} \Phi(\tau) \mathbf{X}_{t-\tau} + \varepsilon_t \quad \text{with } \Phi_{ji}(\tau) \neq 0 \text{ only for } X_{t-\tau}^i \rightarrow X_t^j. \quad (4)$$

Note that we do not fit a full autoregressive model here, but a sparse one where we estimate only those coefficients corresponding to causal links in the time series graph (including cross- and autolinks). A standardized coefficient  $\Phi_{ji}(\tau)$  is called path coefficient<sup>34,51</sup> and stands for the change in the expectation of  $X_t^j$  (in units of its s.d.) induced by raising  $X_{t-\tau}^i$  by 1 s.d., while keeping all other parents of  $X_t^j$  constant. The matrices of path coefficients between all components are shown in Supplementary Fig. 4. Then the CE of a perturbation  $x^* = 1$ , that is, a 1 s.d. for the standardized component time series, is given by<sup>35</sup>

$$I_{i \rightarrow j}^{\text{CE}}(\tau) = \Psi_{ji}(\tau), \quad (5)$$

where  $\Psi(\tau)$  can be iteratively computed from matrix products of the estimated coefficient matrices  $\Phi(\tau)$  by

$$\Psi(\tau) = \sum_{s=1}^{\tau} \Phi(s) \Psi(\tau-s), \quad (6)$$

for example,

$$\begin{aligned} \Psi(0) &= \mathbb{I}, \\ \Psi(1) &= \Phi(1), \\ \Psi(2) &= \Phi^2(1) + \Phi(2), \\ \Psi(3) &= \Phi^3(1) + \Phi(1)\Phi(2) + \Phi(2)\Phi(1) + \Phi(3), \end{aligned}$$

where  $\mathbb{I}$  is the identity matrix. An entry  $\Psi_{ji}(\tau)$  here yields the sum over the products of path coefficients along all causal paths as explained for the climate example in Fig. 3. The framework also allows to encompass more complex types of perturbations such as the effect of multiple perturbations<sup>35</sup>. In Supplementary Fig. 5 we depict the matrices of CEs between all pairs of components for all considered link densities.

The MCE through a component  $k$  is defined as the sum over the products of path coefficients only along causal paths through  $k$ . From the matrices  $\Psi$  it can be derived as

$$I_{i \rightarrow j|k}^{\text{MCE}}(\tau) = \Psi_{ji}(\tau) - \Psi_{ji}^{(k)}(\tau), \quad (8)$$

where  $\Psi^{(k)}(\tau)$  is computed from equation (6) with modified path coefficient matrices  $\Phi^{(k)}(\tau)$  where all links towards component  $k$  are set to zero,

$$\Phi_{ki}^{(k)}(\tau) = \begin{cases} 0, & \text{for all links } X_{t-\tau}^i \rightarrow X_t^k \\ \Phi_{ki}(\tau), & \text{otherwise} \end{cases} \quad (9)$$

which blocks all paths through component  $k$  at any lag. The causal interpretation is that an indirect effect via the component  $X^k$  measures the change we would see in  $X_t^j$  while holding  $X_{t-\tau}^i$  constant and setting component  $X^k$  to whatever value it would have obtained under a unit change in  $X_{t-\tau}^i$ <sup>33,65</sup>.

**Aggregated measures.** For the aggregated causal effect measures ACE and ACS we are interested not so much in the lag at which the interaction occurs and, therefore, base these measures on the lag with maximum effect:

$$I_{i \rightarrow j}^{\text{CE,max}} = \max_{0 < \tau \leq \tau_{\text{max}}} |I_{i \rightarrow j}^{\text{CE}}(\tau)| \quad (10)$$

$$I_i^{\text{ACE}} = \frac{1}{N-1} \sum_{j \neq i} I_{i \rightarrow j}^{\text{CE,max}} \quad (11)$$

$$I_j^{\text{ACS}} = \frac{1}{N-1} \sum_{i \neq j} I_{i \rightarrow j}^{\text{CE,max}}. \quad (12)$$

$I_i^{\text{ACE}}$  quantifies by how much (in units of its s.d.) any of the  $N-1$  remaining components is changed on average by a one unit increase in component  $i$  (at the lag with maximum absolute

effect). This serves as a quantitative measure of how much a component is a gateway of perturbations.  $I_j^{\text{ACS}}$ , on the other hand, measures by how much a component  $j$  is changed on average by a one unit perturbation in any of the  $N-1$  remaining components. Further, we denote the fraction of components that  $i$  is influencing with  $|I_{i \rightarrow j}^{\text{CE,max}}| > 0.05$  by  $\mathcal{N}_i^{\text{out}}$  and as  $\mathcal{N}_i^{\text{in}}$  the fraction of components that  $j$  is affected by, at any lag within  $0 < \tau \leq \tau_{\text{max}}$ . Normalizing ACE and ACS by these quantities results in a measure that is not as robust as desired because it depends on the chosen threshold. In Supplementary Fig. 7 we show ACE versus ACS for different network link densities and reconstruction parameters (maximum lag  $\tau_{\text{max}}$ , significance level  $\alpha$ ), and also for different segments of the data set.

The AMCE is based on causal paths passing through a given node:

$$I_k^{\text{AMCE}} = \frac{1}{|\mathcal{C}_k|} \sum_{(i,j) \in \mathcal{C}_k} \max_{0 < \tau \leq \tau_{\text{max}}} |I_{i \rightarrow j|k}^{\text{MCE}}(\tau)|, \quad (13)$$

where  $\mathcal{C}_k$  is the set of interactions between all non-identical pairs  $i, j \neq k$  at all lags  $0 < \tau \leq \tau_{\text{max}}$  where  $k$  is an intermediate component (at any lags) and  $|\mathcal{C}_k|$  denotes its cardinality. Here we take the absolute value  $|I_{i \rightarrow j|k}^{\text{MCE}}|$ , but one could further distinguish between enhancing (where the sign of MCE equals that of CE) and counteracting (opposite signs) effects. In general, there can be maximally  $c_{\text{max}} = (N-1)(N-2) = 3,422$  interacting non-identical pairs and in Fig. 4b,d we depict the fraction of interaction pairs  $|\mathcal{C}_k|/c_{\text{max}}$  where a component  $k$  is an intermediate node as the size of the nodes. In Supplementary Fig. 8 we show ACE versus AMCE as in Fig. 4d for different setups.

**Uncertainty quantification.** To estimate the standard errors for all causal effect measures considered above, we employ a residual-based bootstrap procedure<sup>66</sup>. Each bootstrap surrogate  $\mathbf{X}_t^*$  is constructed from running model (4) with a joint random sample  $\varepsilon_t^*$  (with replacement) of the original multivariate residual time series  $\varepsilon_t = \mathbf{X}_t - \sum_{\tau=1}^{\tau_{\text{max}}} \Phi(\tau) \mathbf{X}_{t-\tau}$  and with the original coefficient matrices  $\Phi(\tau)$ ,

$$\mathbf{X}_t^* = \sum_{\tau=1}^{\tau_{\text{max}}} \Phi(\tau) \mathbf{X}_{t-\tau}^* + \varepsilon_t^*. \quad (14)$$

From this bootstrap surrogate time series,  $\Phi^*(\tau)$  is estimated from which the other quantities are derived. We use 200 bootstrap surrogates here to estimate the standard errors of all quantities defined above (given as  $\pm$  in the main text as well as Supplementary Table 1 and as error bars in scatter plots in Fig. 4 and in the Supplementary Figs 7, 8).

## References

1. Rahmstorf, S. & Coumou, D. Increase of extreme events in a warming world. *Proc. Natl Acad. Sci. USA* **108**, 17905–17909 (2011).
2. Ghil, M. *et al.* Extreme events: dynamics, statistics and prediction. *Nonlin. Process. Geophys.* **18**, 295–350 (2011).
3. Orgis, T. *et al.* Influence of interactive stratospheric chemistry on large-scale air mass exchange in a global circulation model. *Eur. Phys. J. Spec. Top.* **174**, 257–269 (2009).
4. Vaughan, N. E. & Lenton, T. M. A Review of Geoengineering Proposals. *Climatic Change* **109**, 745–790 (2011).
5. Bassett, G. W. & Lin, Z. Breaking global temperature records after Mt. Pinatubo. *Climatic Change* **25**, 179–184 (1993).
6. Trenberth, K. E., Fasullo, J. T., Branstator, G. & Phillips, A. S. Seasonal aspects of the recent pause in surface warming. *Nat. Clim. Change* **4**, 911–916 (2014).
7. Zubler, F. *et al.* Detecting functional hubs of ictogenic networks. *Brain Topogr.* **28**, 305–317 (2014).
8. Albert, R., Albert, I. & Nakarado, G. L. Structural vulnerability of the North American power grid. *Phys. Rev. E* **69**, 025103 (2004).
9. Menck, P. J., Heitzig, J., Kurths, J. & Schellnhuber, H. J. How dead ends undermine power grid stability. *Nat. Commun.* **5**, 3969 (2014).

10. Newman, M. E. J. Spread of epidemic diseases on networks. *Phys. Rev. E* **66**, 016128 (2002).
11. Klemm, K., Serrano, M. A., Egulluz, V. M. & Miguel, M. S. A measure of individual role in collective dynamics. *Sci. Rep.* **2**, 292 (2012).
12. Lenzu, S. & Tedeschi, G. Systemic risk on different interbank network topologies. *Physica A* **391**, 4331–4341 (2012).
13. Haldane, A. G. & May, R. M. Systemic risk in banking ecosystems. *Nature* **469**, 351–355 (2011).
14. Brovkin, V. *et al.* Geoengineering climate by stratospheric sulfur injections: Earth system vulnerability to technological failure. *Climatic Change* **92**, 243–259 (2009).
15. Stocker, T. & Qin, D. *Climate Change 2013: The Physical Science Basis* (Cambridge University Press, 2013).
16. Newman, M. E. J. *Networks: An Introduction* (Oxford University Press, 2010).
17. Friston, K. J. Functional and effective connectivity in neuroimaging: a synthesis. *Hum. Brain Mapp.* **2**, 56–78 (1994).
18. Bullmore, E. & Sporns, O. Complex brain networks: graph theoretical analysis of structural and functional systems. *Nat. Rev. Neurosci.* **10**, 186–198 (2009).
19. Schinkel, S., Zamora-Lopez, G., Dimingen, O., Sommer, W. & Kurths, J. Functional network analysis reveals differences in the semantic priming task. *J. Neurosci. Methods* **197**, 333–339 (2011).
20. Simpson, S. L., Bowman, F. D. & Laurienti, P. J. Analyzing complex functional brain networks: fusing statistics and network science to understand the brain. *Stat. Surv.* **7**, 1–36 (2013).
21. Tsonis, A. A., Swanson, K. L. & Wang, G. On the role of atmospheric teleconnections in climate. *J. Climate* **21**, 2990–3001 (2008).
22. Yamasaki, K., Gozolchiani, A. & Havlin, S. Climate networks around the globe are significantly affected by El Niño. *Phys. Rev. Lett.* **100**, 228501 (2008).
23. Donges, J. F., Zou, Y., Marwan, N. & Kurths, J. The backbone of the climate network. *Europhys. Lett.* **87**, 48007 (2009).
24. Ebert-Uphoff, I. & Deng, Y. Causal discovery for climate research using graphical models. *J. Climate* **25**, 5648–5665 (2012).
25. Deng, Y. & Ebert-Uphoff, I. Weakening of atmospheric information flow in a warming climate in the Community Climate System Model. *Geophys. Res. Lett.* **41**, 193–200 (2014).
26. Boers, N., Bookhagen, B., Barbosa, H., Marwan, N. & Kurths, J. Prediction of extreme floods in the eastern Central Andes based on a complex networks approach. *Nat. Commun.* **5**, 5199 (2014).
27. Freeman, L. C. A set of measures of centrality based on betweenness. *Sociometry* **40**, 35–41 (1977).
28. Kaiser, H. F. The varimax criterion for analytical rotation in factor analysis. *Psychometrika* **23**, 187–200 (1958).
29. Vejmelka, M. *et al.* Non-random correlation structures and dimensionality reduction in multivariate climate data. *Climate Dyn.* **44**, 2663–2682 (2015).
30. Spirtes, P., Glymour, C. & Scheines, R. *Causation, Prediction, and Search* (The MIT Press, 2000).
31. Runge, J., Heitzig, J., Petoukhov, V. & Kurths, J. Escaping the curse of dimensionality in estimating multivariate transfer entropy. *Phys. Rev. Lett.* **108**, 258701 (2012).
32. Runge, J., Petoukhov, V. & Kurths, J. Quantifying the strength and delay of climatic interactions: The ambiguities of cross correlation and a novel measure based on graphical models. *J. Climate* **27**, 720–739 (2014).
33. Pearl, J. *Causality: Models, Reasoning, and Inference* (Cambridge University Press, 2000).
34. Pearl, J. Linear models: a useful ‘microscope’ for causal analysis. *J. Causal Inference* **1**, 155–170 (2013).
35. Eichler, M. & Didelez, V. On Granger-causality and the effect of interventions in time series. *Lifetime Data Anal.* **16**, 3–32 (2010).
36. Freeman, L. C. Centrality in social networks conceptual clarification. *Soc. Networks* **1**, 215–239 (1978).
37. Philander, S. G. H. *El-Niño and the Southern Oscillation* (Academic press, 1990).
38. Cane, M. A. The evolution of El Niño, past and future. *Earth Planet. Sci. Lett.* **230**, 227–240 (2005).
39. Wallace, J. M. & Gutzler, D. S. Teleconnections in the geopotential height field during the Northern Hemisphere winter. *Mon. Weather Rev.* **109**, 784–812 (1981).
40. Ghil, M. & Mo, K. Intraseasonal Oscillations in the Global Atmosphere. Part I: Northern Hemisphere and Tropics. *J. Atmos. Sci.* **48**, 752–779 (1991).
41. Kalnay, E. *et al.* The NCEP/NCAR 40-year reanalysis project. *Bull. Am. Meteorol. Soc.* **77**, 437–471 (1996).
42. Runge, J. Quantifying information transfer and mediation along causal pathways in complex systems, Preprint at [http://arxiv.org/abs/1508.03808\[stat.ME\]](http://arxiv.org/abs/1508.03808[stat.ME]) (2015).
43. Webster, P. J. *et al.* Monsoons: processes, predictability, and the prospects for prediction. *J. Geophys. Res. Oceans* **103**, 14451–14510 (1998).
44. Hastie, T., Tibshirani, R. & Friedman, J. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction* 2nd edn (Springer, 2009).
45. Vautard, R. & Ghil, M. Singular spectrum analysis in nonlinear dynamics, with applications to paleoclimatic time series. *Physica D* **35**, 395–424 (1989).
46. Hlinka, J. *et al.* Reliability of inference of directed climate networks using conditional mutual information. *Entropy* **15**, 2023–2045 (2013).
47. Spirtes, P. & Glymour, C. An algorithm for fast recovery of sparse causal graphs. *Soc. Sci. Comput. Rev.* **9**, 62–72 (1991).
48. Eichler, M. Graphical modelling of multivariate time series. *Probab. Theory Relat. Fields* **153**, 233–268 (2012).
49. Lauritzen, S. L. *Graphical Models* (Oxford University Press, 1996).
50. Storch, H. V. & Zwiers, F. W. *Statistical Analysis in Climate Research* (Cambridge University Press, 1999).
51. Wright, S. The method of path coefficients. *Ann. Math. Stat.* **5**, 161–215 (1934).
52. Kumar, K. K., Rajagopalan, B. & Cane, M. A. On the weakening relationship between the Indian Monsoon and ENSO. *Science* **284**, 2156–2159 (1999).
53. Hosking, J. S., Russo, M. R., Braesicke, P. & Pyle, J. A. Tropical convective transport and the Walker circulation. *Atmos. Chem. Phys.* **12**, 9791–9797 (2012).
54. Petoukhov, V., Rahmstorf, S., Petri, S. & Schellnhuber, H. J. Quasiresonant amplification of planetary waves and recent Northern Hemisphere weather extremes. *Proc. Natl Acad. Sci. USA* **110**, 5336–5341 (2013).
55. Lau, W. & Kim, K. The 2010 Pakistan flood and Russian heat wave: teleconnection of hydrometeoro-logical extremes. *J. Hydrometeor.* **13**, 392–403 (2012).
56. Ghil, M. & Robertson, A. W. ‘Waves’ versus ‘particles’ in the atmosphere’s phase space: a pathway to long-range forecasting? *Proc. Natl Acad. Sci. USA* **99**, 2493–2500 (2002).
57. Lau, K. & Yang, S. in *Encyclopedia of Atmospheric Sciences*. (ed. Holton, J. R.) 2505–2510 (Academic Press, 2003).
58. Paluš, M. Multiscale atmospheric dynamics: cross-frequency phase-amplitude coupling in the air temperature. *Phys. Rev. Lett.* **112**, 078702 (2014).
59. Moron, V., Robertson, A. W., Qian, J.-H. & Ghil, M. Weather types across the Maritime Continent: from the diurnal cycle to interannual variations. *Front. Environ. Sci.* **2**, 65 (2015).
60. Coluzzi, B., Ghil, M., Hallegatte, S. & Weisbuch, G. Boolean delay equations on networks in economics and the geosciences. *Int. J. Bifurcat. Chaos* **21**, 3511–3548 (2011).
61. Runge, J., Heitzig, J., Marwan, N. & Kurths, J. Quantifying causal coupling strength: a lag-specific measure for multivariate time series related to transfer entropy. *Phys. Rev. E* **86**, 061121 (2012).
62. Pall, P. *et al.* Anthropogenic greenhouse gas contribution to flood risk in England and Wales in autumn 2000. *Nature* **470**, 382–385 (2011).
63. Hannart, A., Pearl, J., Otto, F., Naveau, P. & Ghil, M. Causal counterfactual theory for the attribution of weather and climate-related events. *Bull. Am. Meteorol. Soc.* Early online release at <http://dx.doi.org/10.1175/BAMS-D-14-00034.1> (2015).
64. Runge, J., Donner, R. & Kurths, J. Optimal model-free prediction from multivariate time series. *Phys. Rev. E* **91**, 052909 (2015).
65. VanderWeele, T. *Explanation in causal inference: methods for mediation and interaction* (Oxford University Press, 2015).
66. Hardle, W., Horowitz, J. & Kreiss, J.-P. Bootstrap methods for time series. *Int. Stat. Rev.* **71**, 435–459 (2003).

## Acknowledgements

J.R. received support by the German National Academic Foundation (Studienstiftung), a Humboldt University Postdoctoral Fellowship, and the German Federal Ministry of Science and Education (Young Investigators Group CoSy-CC<sup>2</sup>, grant no. 01LN1306A). J.F.D. thanks the Stordalen Foundation and BMBF (project GLUES) for financial support. D.H. has been funded by grant ERC-CZ CORES LL-1201 of the Czech Ministry of Education. M.P. and N.J. received funding from the Czech Science Foundation project No. P303-14-02634S and from the Czech Ministry of Education, Youth and Sports, project No. DAAD-15-30. J.H. was supported by the Czech Science Foundation project GA13-23940S and Czech Health Research Council project NV15-29835A. We thank Mary Lindsey from the National Oceanic and Atmospheric Administration for her kind help with Fig. 4e. NCEP Reanalysis data provided by NOAA/OAR/ESRL PSD, Boulder, Colorado, USA, from their web site at <http://www.esrl.noaa.gov/psd/>.

## Author contributions

J.R. designed the study, J.R., N.J. and D.H. prepared the data, J.R. carried out the analysis and prepared the manuscript. All authors discussed the results and contributed to editing the manuscript. N.M., M.P. and J.K. supervised the study.

## Additional information

**Supplementary Information** accompanies this paper at <http://www.nature.com/naturecommunications>

**Competing financial interests:** The authors declare no competing financial interests.

**Reprints and permission** information is available online at <http://npg.nature.com/reprintsandpermissions/>

**How to cite this article:** Runge, J. *et al.* Identifying causal gateways and mediators in complex spatio-temporal systems. *Nat. Commun.* 6:8502 doi: 10.1038/ncomms9502 (2015).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>