scientific reports

Check for updates

OPEN

Controllability and heterogeneity of river networks using spectral graph theory approach

Shiblu Sarker[1], Arvind Singh[1✉], Alexander Veremyev[2], Vladimir Boginski[2] & Scott Peckham[3]

River networks are important landscape features that have been extensively studied over many years. While seminal works have focused on characterizing the topological properties of river networks, the quantification of their spectral properties has received limited attention. In this study, through a graph-theoretic formulation of river network topology, we investigate the eigenvalue spectra of its connectivity matrix (i.e., adjacency matrix). First, we explain the observed range of zero eigenvalues on the spectra using the notion of multiplicity (i.e., algebraic and geometric multiplicity) for both undirected and directed river networks. Next, we investigate the physical meaning of the multiplicity of zero eigenvalues on the dynamics of the river network. We show that multiplicity of zero eigenvalues is sufficient to determine the minimum set of driver nodes on the river network. The ratio of the number of driver nodes vs total number of nodes is a measurement of controllability of the river network, which is essential for a comprehensive understanding of the system's dynamics under external forcing. Using both synthetic and natural river networks, we show that with increasing heterogeneity, quantified via Tokunaga c-value, the number of zero eigenvalues increases indicating that basins in humid climate require more number of driver nodes to control their network dynamics. Finally, we show that driver nodes tend to avoid critical nodes identified via pairwise connectivity. Our results indicate that the multiplicity of zero eigenvalues in the eigenvalue spectrum can serve as a valuable tool for understanding and quantifying the physical and dynamical properties of river networks, such as controllability and heterogeneity. Furthermore, our findings establish a clear connection between controllability metrics and the vulnerability of river networks.

Branching patterns of river networks exhibit complex topology and depend on climatic, geologic, biologic and ecologic conditions[1–4]. They have been studied extensively over the past several decades. These dendritic networks serve as essential pathways for transport of water, sediment, nutrients and interact with surrounding ecologic and biotic activities[3,5–10]. Quantifying their structure and dynamics is crucial for understanding their evolution under changing external forcings and for predicting the environmental fluxes that operate within them[1,3,5–8,11–17].

Disparate literature exists on quantifying the connection between river network (RN) topology and its hydrologic response. For example, Rodriguez-Iturbe et al.,[3] explored geomorphic and topologic properties of RNs under different climate and their interaction with biodiversity, vegetation, human populations, and water-borne diseases. More recently, Abed-Elmdoust et al.,[11] showed that precipitation pattern has significant effects on the structure and evolution of a RN. Their study adopted a graph-theoretical approach to explore topologic properties of synthetic RNs for a range of precipitation patterns.

In a graph-theoretical approach, a network is represented as a set of nodes and links, which can be expressed using a network connectivity matrix (e.g., adjacency or Laplacian). Several complex network metrics have been extensively used to characterize networks across diverse fields, including social networks, transportation, communication, bioinformatics, molecular chemistry, mathematics, and geoscience[18–23]. Quantification of connectivity, vulnerability, influence of nodes and complexity of the networks, among others, offer deeper insight into network structure and dynamics and has been a focus of research in past few decades[24]. For example, the topologic and dynamic complexity of delta channel networks has been investigated through a graph-theoretic approach[25,26]. Sarker et al.[27] developed a framework to understand the vulnerability and resilience of RNs under external disruptions using a graph-theoretic approach.

[1]Department of Civil, Environmental and Construction Engineering, University of Central Florida, Orlando, Florida 32816, USA. [2]Department of Industrial Engineering and Management Systems, University of Central Florida, Orlando, Florida 32816, USA. [3]Institute for Arctic and Alpine Research (INSTAAR), University of Colorado Boulder, Boulder, Colorado, USA. ✉email: Arvind.Singh@ucf.edu

nature portfolio

1

In a recent study, Abed-Elmdoust et al.[28] adopted spectral graph theory framework to investigate the eigenvalue spectrum of the graph adjacency matrix. They found that the eigenvalue spectrum exhibits a forbidden range of zero eigenvalues (spectral gap) which is dictated by the branching pattern of the RN. They also suggested spectral gap and nullity (number of zero eigenvalues) are independent of size and shape of the basin and the number of outlets on the landscape. However, studies that specifically relate physical and spectral properties of a RN are limited. In addition, to the best of our knowledge, the multiplicity of eigenvalue on the eigenvalue spectrum of the matrix associated with RN and their relationship to climate and geomorphology has never been explored.

In this study, we characterize physical connections and arrangements of channels and junctions in a RN using a tree network topology, where nodes represent stream junctions and links represent stream segments. The RNs were obtained through numerical simulations using the optimal channel network (OCN) approach, which ignores processes such as ground infiltration and evaporation, as well as from natural basins across the United States. The simulated and natural RN topologies were analyzed to understand: (1) the spectral properties of the connectivity (adjacency) matrix and their response to changing climatic and physical processes, and (2) the range of zero eigenvalues in the eigenvalue spectrum, using the concept of eigenvalue multiplicity and its implications in the context of RNs. Additionally, we investigate how the multiplicity of zero eigenvalues (nullity) is related to network controllability and heterogeneity in both synthetic and natural RNs under varying physical processes and climate conditions (i.e., humid and dry environments). Furthermore, we propose a systematic method for identifying number of critical nodes[27] whose removal could significantly impact overall network controllability, offering a valuable tool for water resource management and conservation planning. While existing methods form the foundation of our analysis, the key advancement lies in their novel application and interpretation within the context of RN topology, leading to insights with direct implications for environmental sustainability and hydrological resilience.

Results and discussion

Eigenvalue and eigenvalue spectrum of synthetic RNs

For a given network (here RN), based on the linear transformation, eigenvalues and eigenvectors can be obtained as:

$$Ax = \lambda x \tag{1}$$

where A is a square matrix, x are vectors on the vector space (i.e., eigenvectors) and λ are the corresponding scale factors (i.e., eigenvalues). In general, eigenvectors are the directions in which the matrix acts solely as a scaling transformation, and eigenvalues are the corresponding scale factors[29–31].

Based on spectral graph theory, properties of the networks can be analyzed using the eigenvalue spectrum of the network connectivity matrix to determine the network's complex and heterogeneous behavior[32,33]. Previous studies found that the eigenvalue spectrum of adjacency matrix exhibits striking features such as spectral gap and nullity which are closely related to the branching patterns of the network[28]. To investigate the range and multiplicity of zero eigenvalues in their eigenvalue spectrum, in this study we generate synthetic RNs using the optimal channel network (OCN) approach. These networks were simulated on an arbitrary basin shape with an initial grid size of 50×50 nodes, with varying energy decay exponent γ ranging from 0.1 to 0.9 (see Fig. 1, and also Methods for details). Figure 1 illustrates different branching patterns of RNs that emerge with varying γ. This variation in branching patterns affects the eigenvalue spectrum. For example, as depicted in Fig. 1 (bottom right panel), the number of zero eigenvalues (λ_0) decreases with increasing γ and exhibits a non-linear behavior. Note that the eigenvalue spectrum were computed for networks that achieved steady state with minimum energy (optimized), shown as subplot in each panel Fig. 1(a-e), as a function of γ values. While previous research[28] has introduced an empirical relationship for the λ_0 in terms of the γ of the RNs, encompassing their probability density function with specific skewness (indicating varying branching topology), the physical significance of λ_0 and their spatial distribution (location) within the RN has not been thoroughly examined. In the following section, we explore the notion of multiplicity in the eigenvalue spectrum for a small hypothetical stream network.

Algebraic and geometric multiplicity of eigenvalues on the eigenvalue spectrum

Figure 2 shows a small undirected (top left panel) and directed (bottom left panel) exemplary stream networks with 4 nodes and 3 links. The corresponding real eigenvalues for both undirected and directed networks are computed using Eq. (1) and presented in Fig. 2. Based on spectral graph theory, the number of times an eigenvalue appears in an eigenvalue spectrum is referred to as algebraic multiplicity of that eigenvalue (denoted by $M_A(\lambda)$)[29,30,34]. Figure 2 depicts that the algebraic multiplicity of zero eigenvalues is $M_A(\lambda_0) = 2$ for the undirected network, whereas $M_A(\lambda_0) = 4$ for the directed network. For undirected binary tree networks, a general expression for $M_A(\lambda_0)$ can be derived (see details in Appendix 1). Besides algebraic multiplicity, the geometric multiplicity of an eigenvalue ($M_G(\lambda)$), i.e., the number of linearly independent eigenvectors associated with it, can also be characterized and may provide useful information regarding branching patterns. Mathematically, geometric multiplicity of an eigenvalue $M_G(\lambda)$ is the dimension of the null space of that eigenvalue.

Figure 2 depicts that the geometric multiplicity of zero eigenvalues $M_G(\lambda_0)$ is 2 for both undirected and directed network since there are two independent rows that exist on the column canonical form of zero eigenvalues of the adjacency matrix (i.e., $\lambda = \lambda_0$). Detail calculation steps for $M_A(\lambda)$ and $M_G(\lambda)$ are presented in Fig. 3. Based on Figs. 2 and 3, one can observe that the algebraic multiplicity $M_A(\lambda)$ and geometric multiplicity $M_G(\lambda)$ of an eigenvalue can differ. However, the geometric multiplicity can never exceed the algebraic multiplicity (i.e., $M_A(\lambda) \geq M_G(\lambda)$)[29,35]. In addition, geometric multiplicity can be also obtained from the rank of the adjacency matrix, which is related to the arrangement of nodes in the network topology. In
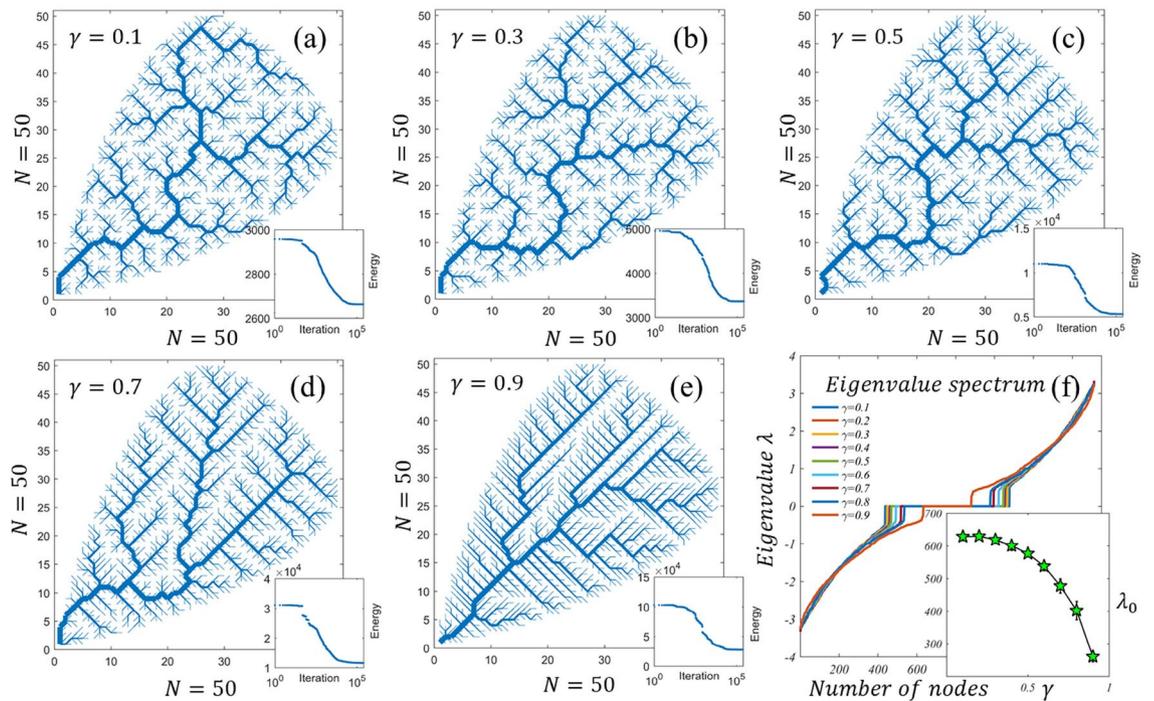
Fig. 1. Synthetic RNs obtained from OCN model for different energy exponent γ (**a**-**e**) and their corresponding eigenvalue spectrum (**f**). Here only five networks are shown (i.e. γ = 0.1, 0.3, 0.5, 0.7 and 0.9) for brevity. The bottom right subplots in panels (**a**-**e**) show the total energy expenditure as a function of number of iteration and subplot in panel (**f**) show the number of zero eigenvalue λ_0 as a function of γ.
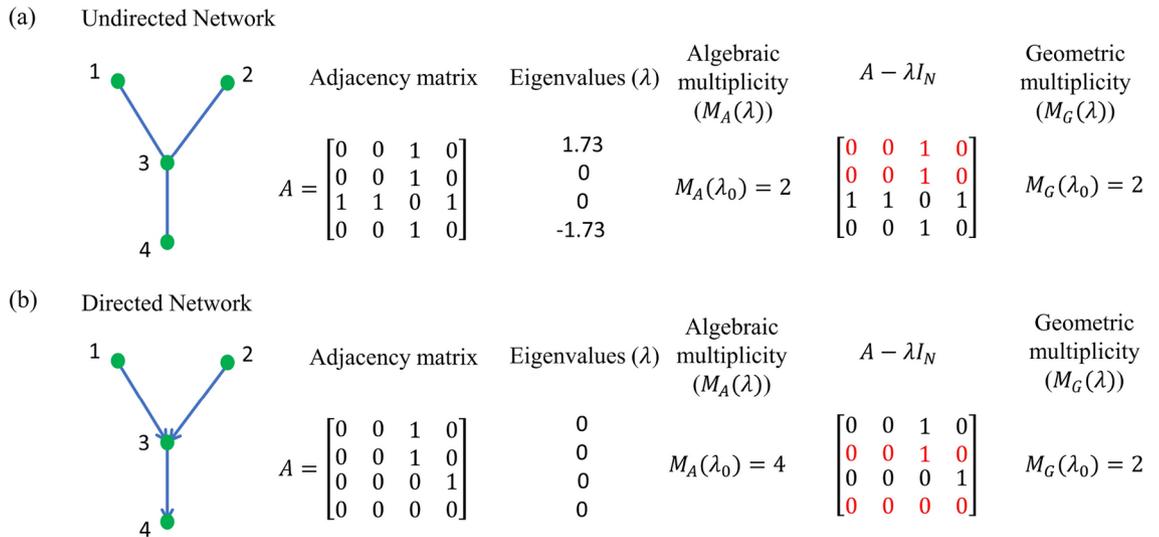

Fig. 2. Examples of algebraic and geometric multiplicities of adjacency matrix for undirected (**a**) and directed (**b**) hypothetical networks. I_N represents the identity matrix. The rows that are linearly dependent on other rows in the $(A - \lambda I_N)$ are marked by red.

the following sub-section we explore the physical implications of the multiplicity of zero eigenvalues and how the arrangement of nodes is related to the $M_G(\lambda_0)$.

Physical meaning of multiplicity of zero eigenvalues

To investigate the physical meaning of the multiplicity of zero eigenvalues and their dependence on branching pattern, we use a relatively larger synthetic network (synthetic RN) (compared to 4 nodes shown in Fig. 2) with 16 nodes but different link/channel arrangements (Fig. 4) and employ the concept of matching[36]. For an undirected graph, matching is defined as an independent set of links that do not share common nodes (see

Adjacency matrix	Eigenvalues	Algebraic multiplicity	Geometric multiplicity

$$A = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

Eigenvalues:
1.73
0
0
-1.73

$M_A(\lambda)$
= Number of repetition

$M_G(\lambda)$
= $\max[1, N - rank(A - \lambda I_N)]$

$$(A - \lambda I_N) \text{ for } 1.73, = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} - 1.73 * \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} -1.7 & 0 & 1 & 0 \\ 0 & -1.7 & 1 & 0 \\ 1 & 1 & -1.7 & 1 \\ 0 & 0 & 1 & -1.7 \end{bmatrix} \rightarrow \textit{there are NO zero rows} \rightarrow \text{rank is } 4 \rightarrow \max(1,4-4) = 1$$

$$(A - \lambda I_N) \text{ for } -1.73, = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} + 1.73 * \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1.7 & 0 & 1 & 0 \\ 0 & 1.7 & 1 & 0 \\ 1 & 1 & 1.7 & 1 \\ 0 & 0 & 1 & 1.7 \end{bmatrix} \rightarrow \textit{there are NO zero rows} \rightarrow \text{rank is } 4 \rightarrow \max(1,4-4) = 1$$

$$(A - \lambda I_N) \text{ for } 0, = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} - 0 * \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \rightarrow \textit{there are 2 zero rows} \rightarrow \text{rank is } 2 \rightarrow \max(1,4-2) = \max(1,2) = 2$$

Eigenvalues(λ)	-1.73	0	1.73
$M_A(\lambda)$	1	2	1
$M_G(\lambda)$	1	2	1

Fig. 3. Computation of algebraic and geometric multiplicity of a hypothetical network using the adjacency matrix shown in Fig. 2(a).

(a)
Pure Branching
7 Branching nodes
8 source nodes

(b)
4 Branching nodes
3 side branching nodes
8 source nodes

(c)
3 Branching nodes
4 side branching nodes
8 source nodes

6 eigenvalues are zero
6 unmatched nodes

4 eigenvalues are zero
4 unmatched nodes
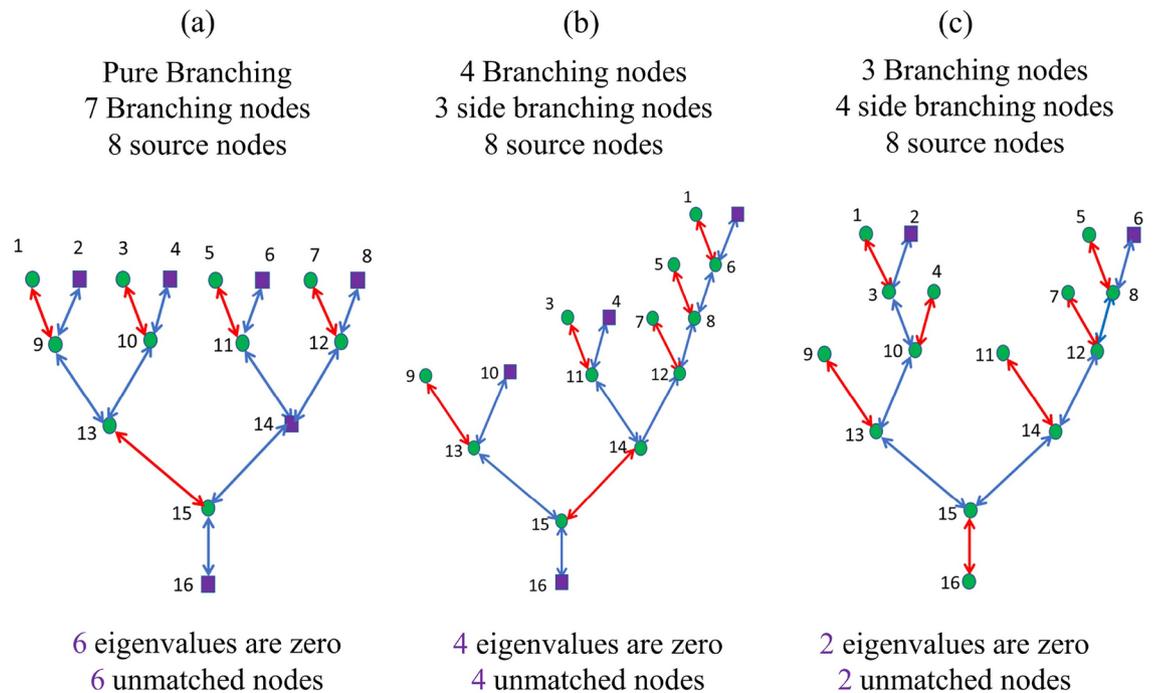
2 eigenvalues are zero
2 unmatched nodes

Fig. 4. Schematic of three undirected hypothetical networks with same number of nodes but different channel arrangement: pure branching (**a**), higher number of branching nodes than side-branching nodes (**b**), and higher number of side-branching nodes than branching nodes (**c**). Corresponding number of zero eigenvalue and unmatched nodes are also shown. Violet solid squares depict the unmatched nodes (also termed as driver nodes for an undirected network) and the red links are the set of links representing maximum matching.

Fig. 4) and a maximum matching is a matching that contains the largest possible number of links[36]. Whereas for a directed network, matching is defined as a set of links that do not share common in-nodes or out-nodes, where an in-node for one link can be an out-node for another. Therefore, for a directed network, a node is considered matched when it is an endpoint of one of the links in the matching, while the rest of the nodes are called

unmatched nodes (see Fig. 4). For both undirected and directed cases, the minimum number of unmatched nodes can be achieved by a maximum matching.

It is important to note that in the context of RNs, both directed and undirected network models may provide valuable insights. Directed network models are especially crucial for characterizing processes like streamflow[37], which predominantly occur in one direction (e.g., upstream to downstream). This makes such models indispensable for hydrological applications, including flood forecasting, watershed management, and modeling pollutant transport[11]. On the other hand, undirected networks focus on the connectivity between nodes without considering flow direction. These models are particularly useful for studying the role of connectivity (e.g., topological properties) in ecological[9,10,15] and geomorphological[28] contexts, where factors like habitat connectivity and species migration-such as fish movement-are key considerations.

Figure 4 shows that algebraic multiplicity of zero eigenvalues $M_A(\lambda_0)$ is equal to the minimum number of unmatched nodes determined using maximum matching. In general, the fact that the number of unmatched nodes is equal to the $M_A(\lambda_0)$ is always true for any undirected tree (connected graph with N nodes and $N-1$ links) has been shown in prior research[38,39]. According to Liu et al.[40], unmatched nodes are sufficient to understand the controllability of a dynamic network system. Controllability is a metric commonly used in network control theory which measures the ability to control the dynamics of a network under external influence (see Methods, for details). Among all the nodes from a network topology, for controlling a network, one needs to control a set of minimum nodes to guide the dynamics of the network. This set of minimum nodes is called the driver nodes (N_D) which are the unmatched nodes identified based on maximum matching. In that sense, for an undirected network, algebraic multiplicity of zero eigenvalue can be used as a physical measure of network controllability.

Figure 4a-c show schematics of three undirected hypothetical networks each with 8 source nodes and 16 total nodes. Network 1 (Fig. 4a) is pure branching with 7 branching nodes; Network 2 (Fig. 4b) has 4 branching nodes and 3 side-branching nodes; whereas, Network 3 (Fig. 4c) has 3 branching nodes and 4 side-branching nodes. However, they all have different number of driver nodes i.e., 6, 4 and 2, respectively, suggesting that algebraic multiplicity of zero eigenvalues can vary within a fixed network size but different channel arrangement. More specifically, it depends on the branching pattern of the RN. Note that branching nodes occur when two channels of the same Horton-Strahler order meet, forming a higher-order channel, while side branching nodes form when channels of different Horton-Strahler orders merge[41–43] (see Methods, for details). It can also be seen that, for undirected RNs with side-branching, driver nodes avoid junction nodes, however, for pure branching (e.g., Fig. 4a), driver nodes position can be at junction nodes. In addition, in the case of pure branching, the number of driver nodes can be determined using a general solution applicable to undirected binary tree networks (see details in Appendix 1). Hence, one can conclude that the branching pattern of a RN can influence the location of driver nodes, and thus the controllability of a network. Although, as discussed above, $M_A(\lambda_0)$ provides a quantitative intuition to detect driver nodes on a network topology and explain the controllability of undirected network, however, $M_A(\lambda_0)$ has limitations to characterize the controllability for a directed network.

For directed RNs, the number of unmatched nodes is equal to the number of source nodes[40]. Figure 5a shows the hypothetical network presented in Fig. 4c as a directed network along with the locations of the driver nodes based on the maximum matching (Figs. 5b-c). As can be seen from Fig. 5c, for a directed network, the number of driver nodes is equal to the number of source nodes (i.e., $N_D = N_S$), suggesting that the controllability for a directed network depends on the number of source nodes. Yuan et al.[35] indicated that not all driver nodes are located at the source nodes and geometric multiplicity of zero eigenvalue may provide both number and location of the driver nodes. To explore the actual position of the driver nodes using the formulation by Yuan et al.[35], we use a larger size network such as a synthetic RN generated by OCN. Figure 6a shows the location of driver nodes on the synthetic RN. Figure 6b visually confirms that not all the driver nodes are at the source nodes.

In the rest of the manuscript, we aim to address how the multiplicity of zero eigenvalue and driver nodes are related to the geomorphic and climatic properties of the synthetic and natural RNs.

Driver nodes on synthetic RNs

To further investigate the role of branching patterns on network dynamics, i.e., controllability, systematically as a function of physical processes and climate, we compute the number of driver nodes and identify their locations using the Yuan et al. framework[35]. We generate several synthetic RNs based on OCN model with a fixed number of nodes ($N = 1513$), however with varying energy exponent $\gamma = 0.1$ to 0.9. Figure 6a shows, as an example, a generated synthetic RN. The locations of the superimposed driver nodes on the network can be seen in Figs. 6a and 6b. Note that for each γ, 15 independent networks were generated, and their ensemble averaged metrics were computed in order to minimize the effect of random network initialization.

Figure 7a illustrates a comparative analysis of N_D between undirected and directed versions of synthetic RNs. In particular, Figure 7a shows the algebraic multiplicity of zero eigenvalues $M_A(\lambda_0)$ in the eigenvalue spectra of the connectivity matrix, which is equal to the number of unmatched nodes (driver nodes) for undirected networks, as a function of γ. As the percentage of zero eigenvalue represents controllability, C ($C = N_D/N$, representing the ratio of number zero eigenvalues to the total number of eigenvalues) of a network, from Fig. 7a, it can be observed that controllability decreases with increasing γ where the drainage pattern changes drastically from an intertwisted RN to a relatively straightened pattern[28]. Note that controllability C for an undirected RN has been observed to be independent of the RN size and basin shape[28]. Likewise, for the case of directed RN, controllability also decreases with increasing γ with a similar trend but is higher than for the undirected RN (Fig. 7a). As discussed above, for a directed network, the number of driver nodes (N_D) computed based on geometric multiplicity of zero eigenvalues $M_G(\lambda_0)$ is equal to the number of source nodes (N_S). However, a significant number of driver nodes are not located at the source nodes (Fig. 7b). We observed that $\sim 4.5\% - 7\%$ driver nodes are located at branching nodes and $\sim 12.5\% - 16\%$ driver nodes are located at side-branching
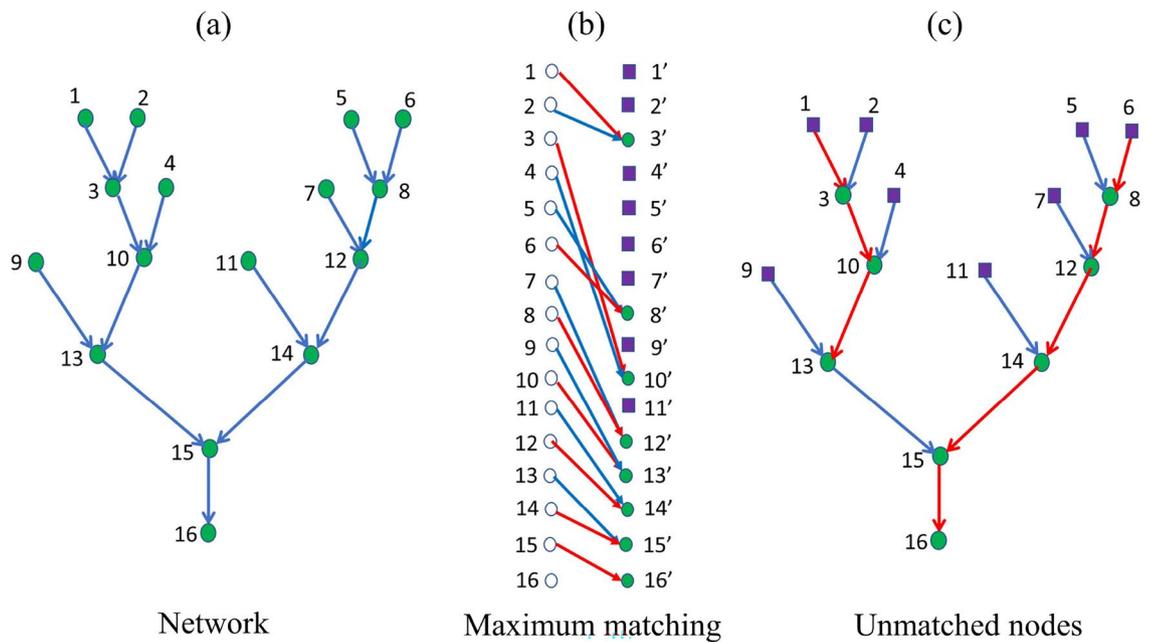
Fig. 5. Maximum matching concept shown for a hypothetical directed network for the case of Fig. 4c. Violet squares nodes depict unmatched nodes (driver nodes) to achieve full control of network. Note that for a directed network, the number of driver nodes is equal to the number of source nodes but not all source nodes are driver nodes (see Fig. 6). Red links show the set of links exhibiting maximum matching for the directed network.
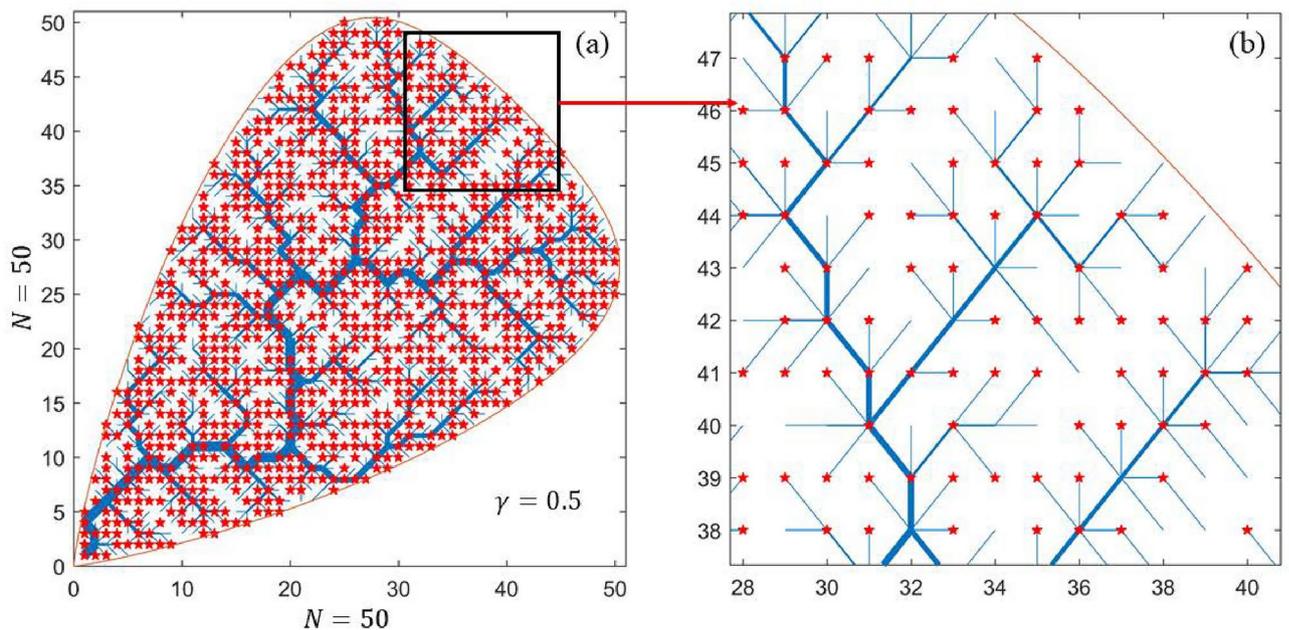

Fig. 6. Exact location of driver nodes for a directed synthetic RNs (shown in Fig. 1(c)) generated by OCN. Driver nodes (star) superimposed on synthetic RN (**a**), and zoomed-in image showing location of the driver nodes (**b**).

nodes. For detailed information on branching and side-branching nodes, refer to the Methods section and Fig. 8(a), which provides a schematic representation of a RN with branching and side-branching junctions for illustration. It is also observed that as the γ increases, the percentage of driver nodes located on side-branching nodes increases slightly, despite the fact that the total percentage of driver nodes decreases. This essentially indicates that as the drainage pattern changes from an intertwisted channel network to a relatively straightened
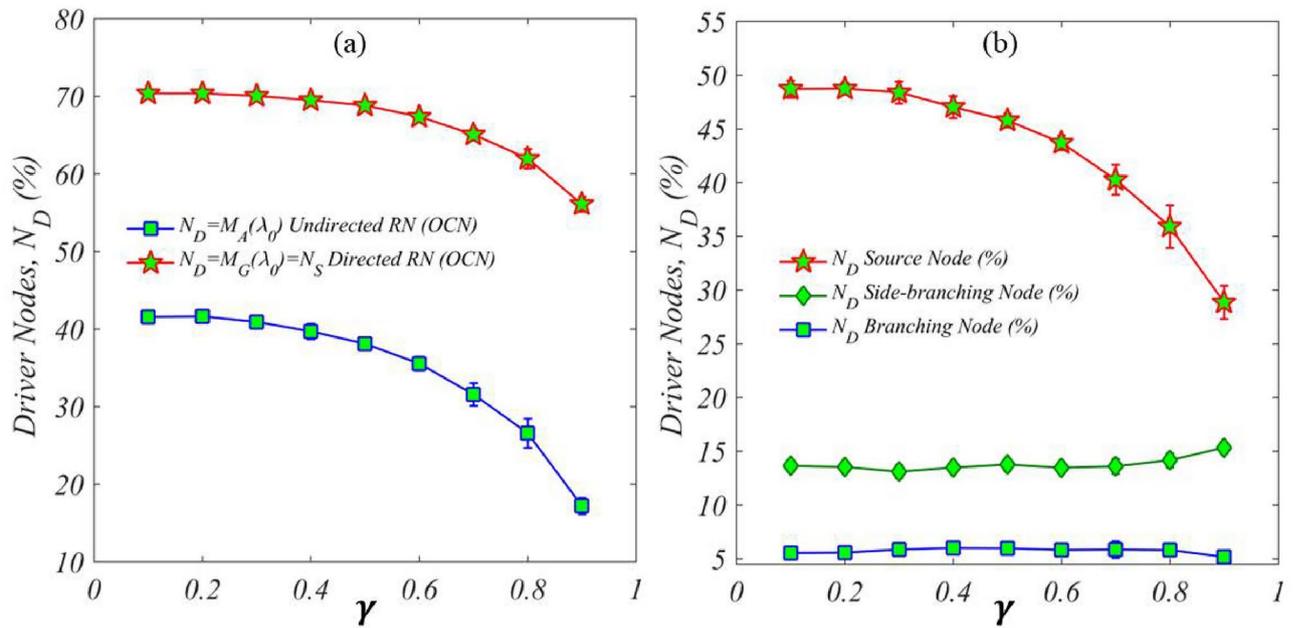
Fig. 7. Percentage of driver node as a function of γ for synthetic RNs (**a**). For undirected network driver nodes, N_D (%); Zero eigenvalue λ_0 (%) as a function of γ (**a**). For directed network number of driver node ($N_D = M_G(\lambda_0)$); number of source node N_S as a function of γ (**a**), driver node (%) located on source node, branching and side-branching node (**b**). Networks were generated using OCN for $\gamma = 0.1$ to $\gamma = 0.9$.
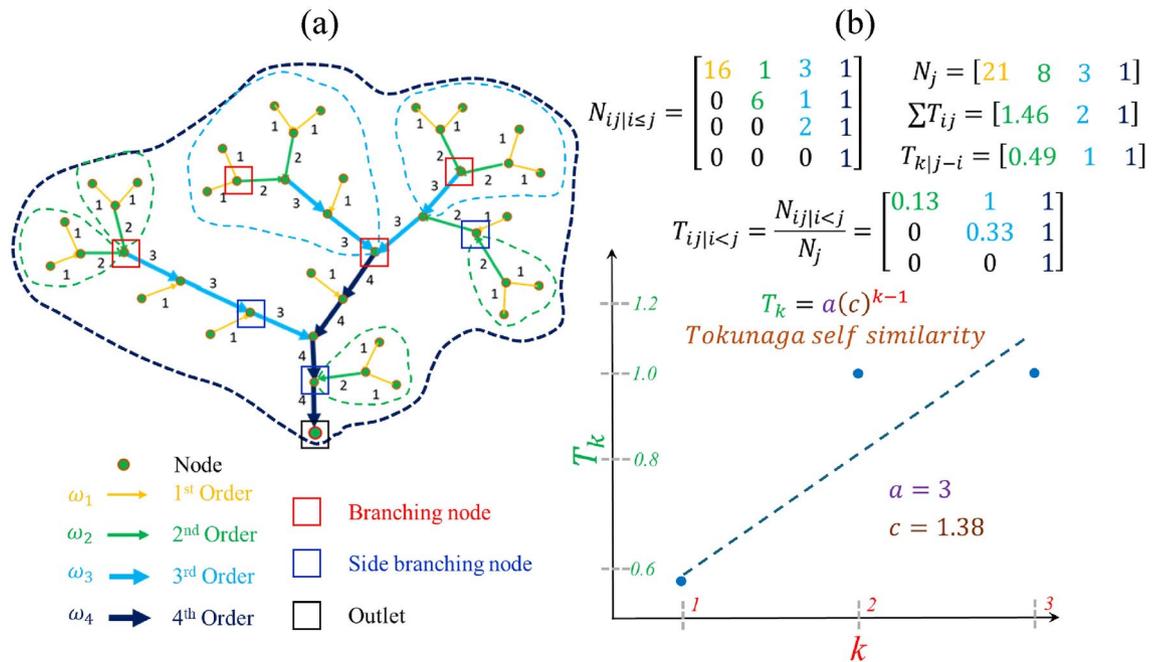

Fig. 8. An example of a hypothetical 4th-order RN with branching and side-branching nodes (**a**) and corresponding Tokunaga self-similarity model along with the computation of parameters a and c (**b**); different colors indicate different order (see[41]).

pattern (as observed in steeper topographies), besides the source node, the controllability is mostly governed by the side branching nodes and vice versa.

Heterogeneity of synthetic RNs

Network heterogeneity is a metric that describes the diversity of characteristics within a network, often related to the uniformity of its organization (i.e., the arrangement of links and nodes) and is critical for understanding the

emergence and evolution of RNs[44,45]. In this study, we utilize the Tokunaga self-similarity model (see, for details, Methods and Fig. 8(b), which illustrates the computation of Tokunaga model parameters for the RN example shown in Fig. 8a) to quantify the heterogeneity of the network. This model is frequently used to characterize the branching patterns of RNs[11,46,47]. It offers advantages over the Horton-Strahler[41–43] indexing system by accounting for side-branching channels (see Methods and Fig. 8 for details), which frequently occur in natural RNs[43]. The Tokunaga parameter c-value, describes the varying degree of side-branching and has been suggested to provide deeper insights into the structure and function of networks[47], as structural changes can indicate functional changes in a RN[7,48]. Figure 9 illustrates the relationship between the c-value and controllability obtained from synthetic RNs simulated using OCN approach for various γ values. The observed increasing trend from this figure suggests that larger heterogeneity in RNs is associated with increased controllability. In the next section, we explore the relationship between controllability and heterogeneity in natural RNs.

Controllability and heterogeneity of natural RNs under varying climate

In this section, we investigate the controllability computed based on the multiplicity of zero eigenvalues on the eigenvalue spectrum of natural RNs. For this, we extracted RNs from the digital elevation model (DEM) of 56 natural basins across the United States, based on the availability of LiDAR data. Although the flow of water in RNs is directional, due to biodiversity and ecological considerations (as discussed above), the eigenvalue spectrum was computed on an undirected adjacency matrix considering only the notion of the network connectivity. This connectivity has been shown to play a major role in biological and ecological communities in riverine ecosystems and their processes[9,15]. Note that, in this study, the networks were extracted using a curvature-based method from high-resolution (1m) topographic data[49].

In order to understand the distinct climatic signature on the eigenvalue spectrum, the long-term climate was considered in the form of the climate aridity index (*CAI*), which is commonly defined as the ratio of mean annual potential evaporation (E_p) to precipitation (P)[50–56]. According to Budyko, regions where the aridity index is higher than 1, are generally classified as dry, and regions with aridity index less than 1 are classified as humid regions. In addition, *CAI* has also been associated with a broader range of climatic regimes, such as arid $12 > CAI \geq 5$, semi-arid $5 > CAI \geq 2$, subhumid $2 > CAI \geq 0.75$, and humid $0.75 > CAI \geq 0.375$ [50,53,55].

The eigenvalue spectrum computed from the adjacency matrix of natural RNs for the different climatic regions exhibits a distinct range of zero eigenvalue. Similarly, to the synthetic RN, the eigenvalue range for natural RN's can be explained by the algebraic multiplicity $M_A(\lambda_0)$ and the geometric multiplicity $M_G(\lambda_M)$.

Figure 10(a-b) show two sample examples of natural basins with superimposed RN for humid (a) and dry (b) climates. Figure 10c show controllability as a function of *CAI* for all the extracted natural RNs considered here. As can be seen, controllability shows a decreasing trend with *CAI*, indicating that controllability is higher for RNs in humid climates as opposed to dry climates.

The dependence of controllability on climate underscores the significant influence of climate on the branching patterns of RNs. Zanardo et al.[57] explored the relationship between the Tokunaga c-value and climatic variables such as rainfall, storm frequency, and duration and argued that the c-value increases with increasing precipitation. More recently, Ranjbar et al.[47] demonstrated a positive correlation between the c-value and the complexity of RNs. Figure 10d shows the c-value as a function of *CAI* for the basins analyzed here, suggesting
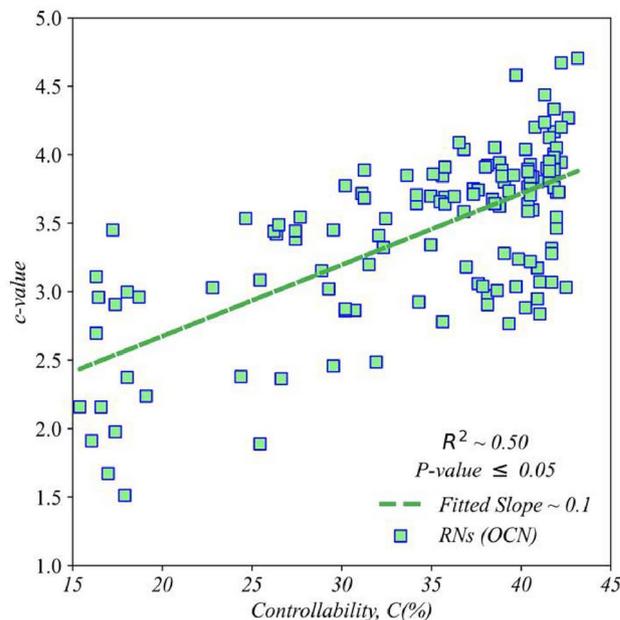

Fig. 9. Relation between heterogeneity based on Tokunaga analysis (c-value) and controllability, C (%) for undirected networks generated using OCN for a range of γ values.
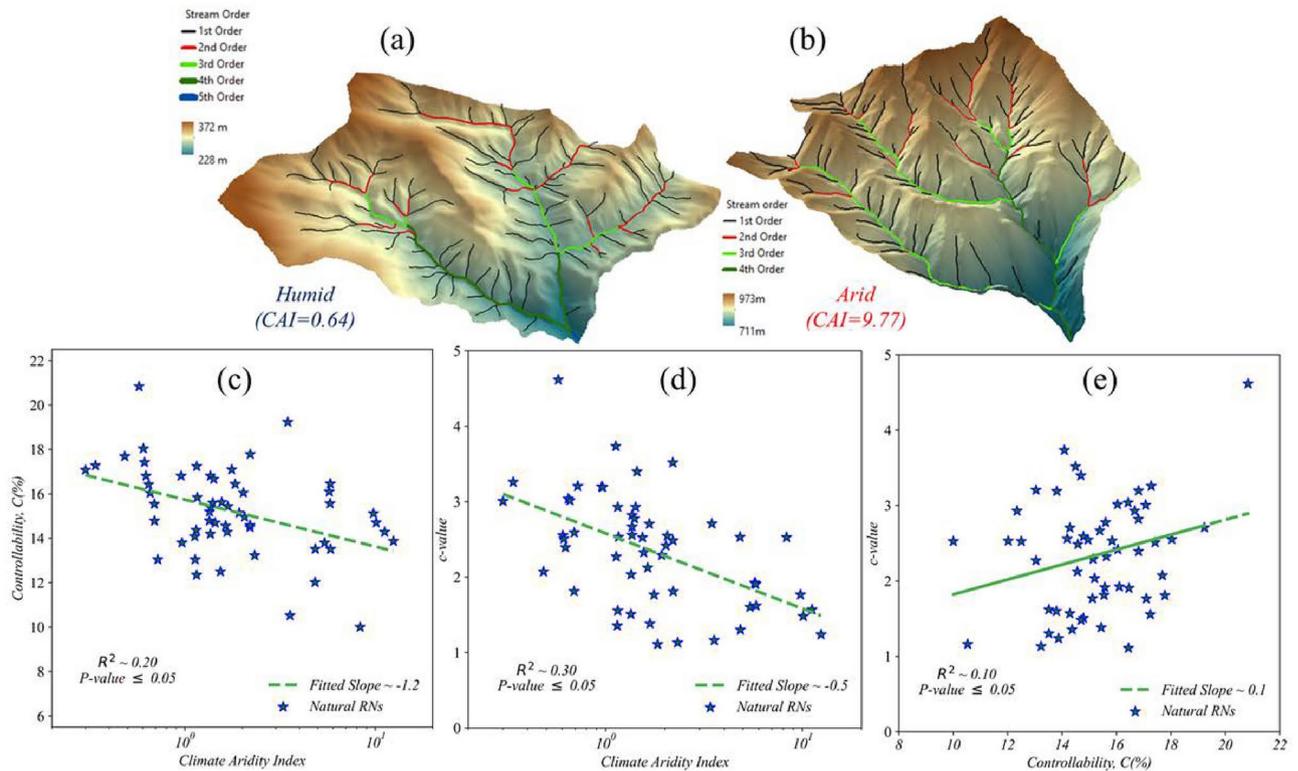
Fig. 10. Example of DEMs and superimposed RNs for basins in humid (**a**) and arid (**b**) climate; the channel order based on the Horton-Strahler ([41]) ordering scheme is shown with different colors. Controllability C (%) (**c**) and heterogeneity represented by c-value (**d**) for natural basins as a function of climate aridity index, c-value plotted against Controllability C (%) (**e**) for natural basins. Note that the significance of these relationships was tested using t test which showed a $p - value < 0.05$, within a 95% confidence interval.

that side-branching decreases with increasing *CAI*. This observation highlights the role of side-branching in the network, indicating that both controllability and heterogeneity increase as *CAI* decreases.

To further explore whether the controllability metric depends on the branching pattern, we plot the c-value as a function of controllability in Figure 10e. We observe a general increasing trend, where higher c-values correspond to greater controllability, indicating that networks with more side-branching tend to exhibit enhanced controllability. Although the correlation is not as high as observed in synthetic RNs, the computed $p - value$ suggests the relationship is significant. The lower correlation in the case of natural RNs (as compared to synthetic RNs, see Fig. 9) can be attributed to varying basin size, hydrology, geology and soil heterogeneity, among others. Our results suggest that intricate, inhomogeneous branching structures with more side-branching, commonly found in humid environments[57], may produce RNs less prone to change under increasing precipitation. In contrast, arid climate basins, with more homogeneous branching structures, may exhibit more dynamic behavior under similar precipitation conditions. These quantitative insights offer valuable tools for a comparative and comprehensive understanding of humid versus dry basins.

Relation with the critical nodes and network vulnerability

Critical nodes are the set of nodes whose removal result in the maximum fragmentation of the network[27,58–60]. A network structure or connection which is more vulnerable may create more fragmentation under the extreme external influence (see Methods). Therefore, we compute the vulnerability (defined by the power-law exponent of pairwise connectivity; see methods for details) of RNs for varying climate (see Fig. 11(a)). For example, in Fig. 11a we show pairwise connectivity as a function of number of critical nodes (CN) removed and the computed power-law slopes for two natural basins (shown in Fig. 10) from humid and arid environments (see also Fig. 11b). As can be seen, the computed power-law slopes, indicating network vulnerability, differ for different climatic conditions.

Figure 11(b) shows that vulnerability increases with increasing *CAI* indicating dry basins are more vulnerable than humid basins. In addition, Fig. 11(b) also illustrates the distribution of vulnerability across different basin types classified according to the *CAI*: Humid, Subhumid, Semiarid, and Arid. The black solid star in each climatic regime indicates the mean vulnerability value of each category. In particular, vulnerability tends to increase from the humid to arid basins, suggesting a potential relationship between aridity and vulnerability. The observed higher variability, particularly within the Subhumid category, indicates that the vulnerability may be influenced by additional underlying factors specific to each basin type. This quantification of vulnerability has important implications in the context of ecological and biological considerations. Specifically, riverine ecosystems often
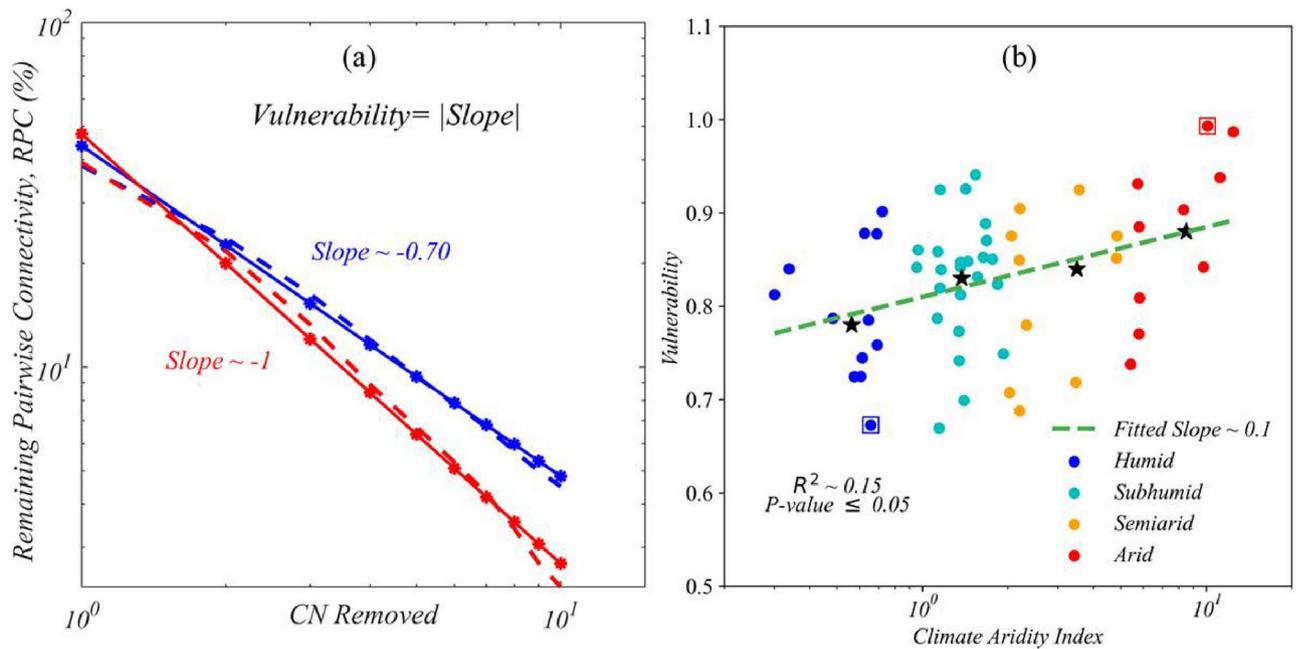
Fig. 11. Remaining pairwise connectivity as a function of the number of critical nodes (CN) removed, with numerical values indicating the power-law fitted slope (**a**), where red and blue colors correspond to basins represented by red and blue squares in panel (**b**). Vulnerability as a function of the climate aridity index (**b**); here different colors indicate a broader range of climatic regimes[50,53], categorized as humid ($0.375 \leq CAI < 0.75$), subhumid ($0.75 \leq CAI < 2$), semiarid ($2 \leq CAI < 5$), and arid ($5 \leq CAI < 12$). The black solid star in (**b**) indicates the mean value within each climate category.

occur in spatially structured habitats where fragmentation directly plays a key role in their processes such as diversity, productivity and even resilience of the metapopulation[9,10,15].

As discussed above, the controllability framework allows us to identify the driver nodes which correspond to the set of nodes that controls the dynamic response of a system. In the context of RNs, it is worth pointing out that most of those nodes correspond to the headwater locations (source nodes). A question of particular interest is to explore whether those junction driver nodes coincide with critical nodes as critical nodes result in maximum fragmentation of the network.

To investigate the characteristics of driver nodes and their relationship with the critical nodes, we compute the number of driver nodes that are located on the junction nodes (N_{DJ}). In addition to that, we also identify critical nodes which coincide with the driver junction nodes ($N_{CN \cap N_{DJ}}$) for specific conditions where the number of critical nodes (k) i) is equal to the number of junction driver nodes, i.e., $k = N_{DJ}$, ii) is equal to the total number of driver nodes in that network, i.e. $k = N_D$, and iii) for two different fixed values of k, i.e. $k = 5$ and $k = 10$. It is observed that when $k = N_{DJ}$, no driver nodes and critical nodes coincide for any RN, indicating that driver nodes tend to avoid critical nodes. For $k = N_D$, we observe a few common nodes. For example, Table 1 shows that out of 56 natural basins, 40 basins exhibit a very small amount ($\sim$ 1-3.5 %) of junction driver nodes (N_{DJ}). 37 out of those 40 basins show common critical and driver nodes (i.e., $N_{CN \cap N_{DJ}}$). In addition, based on fixed values of k, we observed percentage of common nodes increases as k increases. Figure 12(a-b) shows an example of a natural basin along with superimposed critical nodes and driver nodes. Out of 238 nodes, 34 nodes are driver nodes (N_D), while only 3 nodes are located on the junction. In other words, controllability is mainly governed by the headwater locations (source nodes). In addition, for $k = N_D$, although, the 3 junction driver nodes coincide with critical nodes found from a set of 34 critical nodes group, Fig. 12(c-d) (percentage of N_{DJ} and $N_{CN \cap N_{DJ}}$ as a function of CAI) indicates that as the climate changes from humid to arid, the amount of $N_{CN \cap N_{DJ}}$ increases. It is also worth pointing out that the relation between N_{DJ} and $N_{CN \cap N_{DJ}}$ were obtained from 37 basins out of 56 basins as the 19 basins did not have common critical and driver junction nodes.

Summary and Concluding remarks

In this paper, we investigated the eigenvalue spectrum of the adjacency matrix associated with both synthetic and natural river networks (RNs). Synthetic RNs were generated using the optimal channel network (OCN) approach, while natural RNs were extracted from a 1-meter resolution digital elevation model. We explored the spectral properties of these networks in the context of controllability and heterogeneity, with particular focus on the range of zero eigenvalues by analyzing their multiplicity. Our findings suggest that the multiplicity of zero eigenvalues has the potential to quantify the dynamical properties (e.g., controllability) of RNs. For synthetic RNs, we observed a significant correlation between the c-value (characterizing RNs heterogeneity) and controllability,

Basin's CAI	$C = N_D/N$ (%)	Vulnerability	c-value	N_{DJ}/N (%)	$N_{CN \cap N_{DJ}}$ (%) $k=5$	$N_{CN \cap N_{DJ}}$ (%) $k=10$	$N_{CN \cap N_{DJ}}$ (%) $k=N_D$
0.30	17.07	0.81	3.01	0.61	0	0	0.61
0.34	17.28	0.84	3.26	0	0	0	0
0.48	17.69	0.79	2.07	0.68	0.34	0.34	0.68
0.57	20.83	0.72	4.62	0	0	0	0
0.61	18.03	0.72	2.55	0.41	0	0	0
0.61	17.42	0.74	2.51	0.43	0	0.43	0.43
0.62	16.81	0.88	2.39	0.57	0	0.28	0.57
0.64	16.42	0.79	3.04	0.86	0.43	0.43	0.86
0.66	16.04	0.67	3.02	0	0	0	0
0.69	15.54	0.88	1.81	1.01	0	0	1.01
0.69	14.78	0.76	2.59	0.82	0	0	0.41
0.72	13.04	0.90	3.21	1.52	0	0.43	1.09
0.95	16.81	0.84	3.20	0.85	0	0	0.28
0.96	13.80	0.86	3.19	0.61	0	0	0.61
1.13	14.08	0.79	3.73	0.14	0	0	0.14
1.13	13.04	0.86	2.27	0	0	0	0
1.15	14.37	0.67	1.36	1.45	0.15	0.15	1.45
1.15	12.35	0.82	2.93	0	0	0	0
1.15	17.24	0.92	1.56	1.72	1.72	1.72	1.72
1.16	15.84	0.84	2.53	0.78	0.16	0.31	0.62
1.34	14.79	0.77	1.51	0.97	0	0.19	0.78
1.35	15.20	0.74	2.03	0.49	0.25	0.25	0.25
1.36	16.81	0.85	2.82	0	0	0	0
1.36	14.19	0.81	2.56	0.83	0	0	0.67
1.36	15.38	0.84	2.67	0	0	0	0
1.40	15.60	0.70	2.78	0.46	0	0	0.46
1.42	16.67	0.93	2.93	2.27	0.76	0.76	1.52
1.44	14.69	0.85	3.40	0	0	0.00	0
1.54	12.50	0.94	2.53	2.08	1.04	2.08	2.08
1.56	15.63	0.83	2.33	0.78	0	0.78	0.78
1.64	14.56	0.85	2.12	0	0	0	0
1.67	14.29	0.89	2.70	1.07	0	0.36	1.07
1.68	15.44	0.87	1.38	0.67	0	0	0.67
1.76	17.09	0.85	1.76	0.43	0	0	0
1.84	16.44	0.82	1.11	2.05	0.68	1.37	0.68
1.93	15.14	0.75	2.29	0.40	0	0	0.40
2.03	16.05	0.71	2.42	1.85	0.62	1.23	1.23
2.06	14.96	0.88	2.55	0.88	0	0.59	0.88
2.20	14.49	0.85	3.52	0.72	0	0	0.72
2.20	14.58	0.69	2.49	1.04	0	1.04	1.04
2.20	17.78	0.90	1.81	1.11	0	1.11	0
2.32	13.23	0.78	1.13	0.53	0	0.53	0.53
3.47	19.23	0.72	2.71	0	0	0	0
3.57	10.53	0.92	1.16	0.67	1.32	1.32	0.67
4.82	12.02	0.85	2.53	0	0	0	0
4.84	13.51	0.88	1.30	0	0	0	0
5.42	13.79	0.74	1.60	0.34	0	0.34	0.34
5.73	16.10	0.93	1.92	0	0	0	0
5.79	15.57	0.77	1.91	0	0	0	0
5.80	16.46	0.89	1.90	0.42	0	0	0.42
5.82	13.51	0.81	1.62	0	0	0	0
Continued							

Basin's CAI	$C = N_D/N$ (%)	Vulnerability	c-value	N_{DJ}/N (%)	$N_{CN \cap N_{DJ}}$ (%) $k = 5$	$N_{CN \cap N_{DJ}}$ (%) $k = 10$	$N_{CN \cap N_{DJ}}$ (%) $k = N_D$
8.31	10.00	0.90	2.53	3.33	3.33	3.33	3.33
9.77	15.12	0.84	1.77	0.58	0	0.58	0.58
10.11	14.70	0.99	1.48	0	0	0	0
11.19	14.29	0.94	1.57	1.26	0	0	1.26
12.45	13.87	0.99	1.24	1.54	0.77	1.54	1.54

Table 1. Vulnerability, critical nodes (*CN*) and driver nodes (N_D) characteristics of natural RNs. *C* denotes controllability, N_{DJ} is the number of driver nodes on junctions of a network, *N* is the total number of nodes in a network, $N_{CN \cap N_{DJ}}$ is the number of common driver and critical nodes, and *k* is the total number of identified critical nodes in a network.
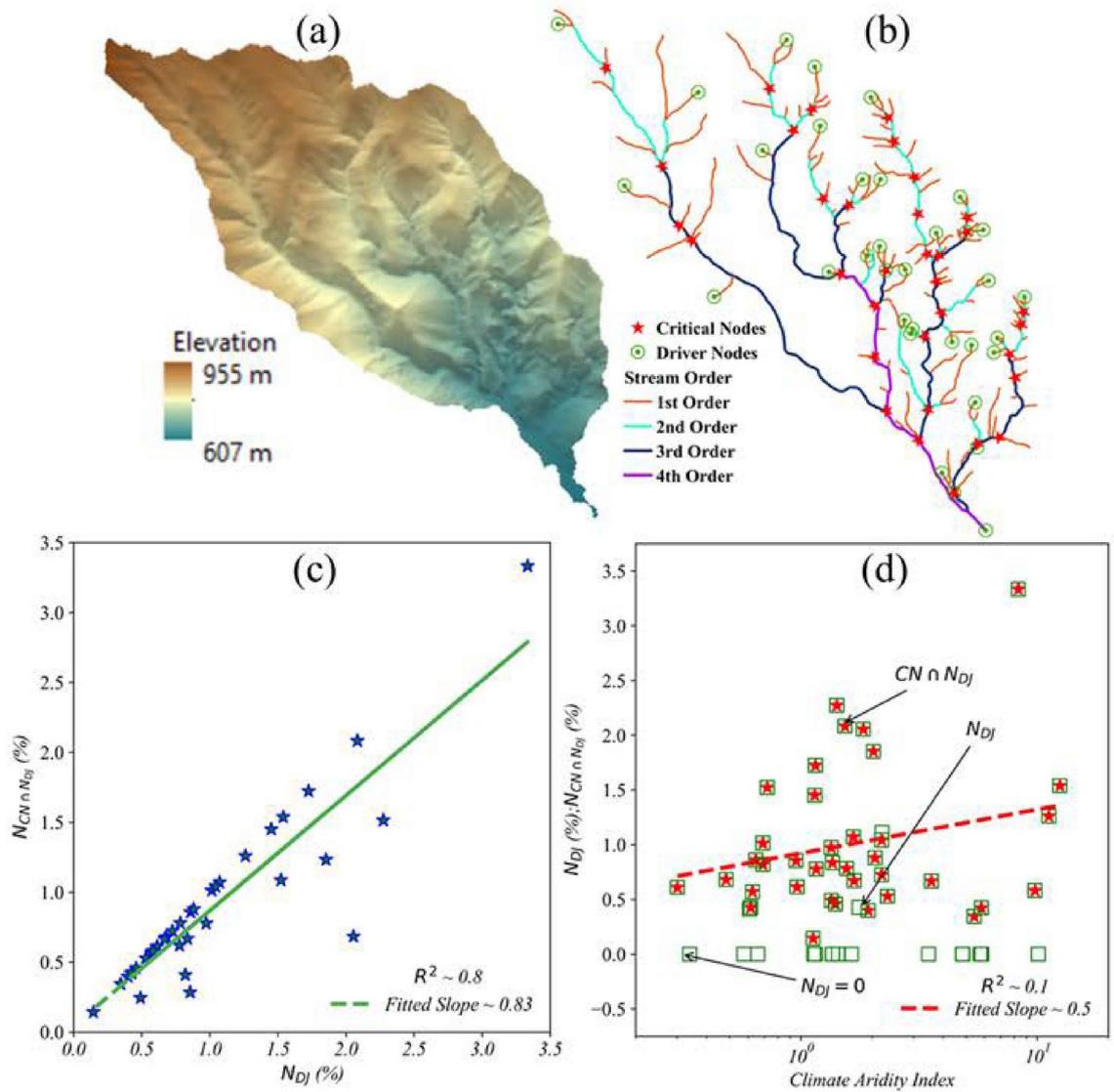

Fig. 12. Example of a natural DEM (1-m resolution) (**a**) and corresponding extracted RN with superimposed critical nodes and driver nodes (**b**), the correlation between N_{DJ} and $N_{CN \cap N_{DJ}}$ (**c**), and percentage of N_{DJ} and $N_{CN \cap N_{DJ}}$ as a function of climate aridity index (**d**).

indicating that networks with higher heterogeneity exhibit greater controllability, i.e., a higher capacity to regulate dynamics under external influences, compared to those with lower heterogeneity. Similar observations were made in natural RNs, where a positive correlation between controllability and the c-value further suggested that humid basins are more heterogeneous and exhibit higher controllability than dry basins. Our comparative analysis between driver and critical nodes indicates that the vulnerability of an RN increases with increasing climate aridity index. In other words, RNs in arid climates tend to be more vulnerable (dynamically unstable) than those in humid climates. Additionally, our results suggest that driver nodes tend to avoid critical nodes and are primarily located at source nodes. The primary contribution of our research lies in the novel integration of network theory principles with hydrologically relevant connectivity, providing new insights into the structural robustness and functional dynamics of river systems. Unlike conventional hydrological approaches that mainly focus on physical and empirical modeling, our study introduces a spectral analysis framework to quantify network heterogeneity and assess the influence of critical nodes on system vulnerability. Overall, our findings highlight the relevance of the proposed controllability and heterogeneity metrics for understanding natural RNs and their response to climate variability.

Methods

Synthetic network generation

In this study, we generate synthetic RNs using optimal channel network (OCN) approach, which has been extensively explored from various hydrologic and geomorphic perspectives[11,14,61–64]. Previous studies found that OCNs can mimic commonly observed topologic and geometric properties of real RNs[42]. OCNs can also be used to generate DEMs which have been argued to represent properties of natural landscapes[11]. In general, OCN modeling relies on the local minimization of the topologic energy defined as $E = \sum_{i=1}^{N-1} L_i Q_i^\gamma$, where $L_i Q_i^\gamma$ represents the energy dissipated in the ith link of the network. Here, L_i and Q_i are the length and discharge of the link, respectively[42]. The energy exponent γ, which varies between 0 and 1, characterizes the mechanics of erosional processes and defines the branching pattern of the channel network[42]. The simulations were conducted on an arbitrarily shaped area designed to mimic real basins, using an initial grid of size 50×50 nodes (Fig. 1). We used a lobe-like shape that mimics a natural basin boundary. While the square grid consisted of 2,500 nodes, the lobe-like shape contained 1,513 nodes. For simplification, the topologic energy was computed assuming a unit distance between adjacent nodes and uniform precipitation over the entire simulated basin[11,27,28]. We also assumed that the network drains the entire amount of precipitated unit rainfall from every node within the network[28,61,62]. Here, we generate networks for nine different γ values, ranging from 0.1 to 0.9, while keeping the initial random tree network the same for each simulation (Fig. 1). The generated networks reproduce several topological properties, which are analyzed using a graph-theoretical framework (see also, Abed-Elmdoust et al.[11], for details).

Matrices adopted to calculate spectral properties

Previous studies have suggested that the flow path in a RN can be defined through a directed graph which can be denoted by a $N \times N$ adjacency matrix A, where:

$$A_{ij} = \left\{ \begin{array}{ll} 1: & i \text{ flows to } j; \\ 0: & \text{otherwise} \end{array} \right. \tag{2}$$

Adjacency matrix (also referred to as connectivity matrix[31]) defines the connection between nodes and the flow direction in RNs. To compute the real eigenvalue spectrum of drainage networks, undirected adjacency matrix can be considered by ignoring the flow direction[28] which can be represented as:

$$B_{ij} = \left\{ \begin{array}{ll} 1: & i \text{ and } j \text{ adjacent}; \\ 0: & \text{otherwise.} \end{array} \right. \tag{3}$$

In addition, the relation between matrices A and B can be expressed as $B = A + A^T$, where A^T is the transpose of A.

We also employ the degree matrix, which is a diagonal matrix of the network and contains information about the degree of each node, i.e., the number of links connected to each node. The degree matrix D can be expressed as:

$$D_{ij} = \left\{ \begin{array}{ll} deg(i): & \text{if } i = j; \\ 0: & \text{otherwise} \end{array} \right. \tag{4}$$

where $deg(N_i)$ of a node N_i represents the number of links terminates at that node. For a directed graph, the term degree of a node defines either in-degree (number of incoming links) or out-degree (number of outgoing links).

Identification of Driver nodes: network controllability

In this study, we identify the driver nodes (N_D) on a network to understand the controllability of RNs and investigate the dependency of controllability on RN topological structure, specifically the branching pattern. Controllability is a measure of the ability of a network to guide a dynamic system and the knowledge of the driver nodes offer the full control over the entire network[40]. In particular, the control of these driver nodes is enough to achieve full control of the network system dynamics. Liu et al.[40] proposed a framework to identify a set of driver nodes for a directed network using the maximum matching in the network. However, Yuan et al.[35]

arguably provides an accurate framework to identify the driver nodes for both directed and undirected networks based on the spectral graph theory. Their framework was based on the multiplicity of eigenvalues and suggested that controllability is simply a measure of ratio between minimum number of driver nodes and the total number of nodes. Additionally, Yuan et al. depicted how geometric multiplicity of eigenvalue, $M_G(\lambda)$, relates to the position of the driver nodes. More specifically, geometric multiplicity for a network can be calculated as:

$$M_G(\lambda) = max\left[1, N - rank(A - \lambda I_N)\right] \tag{5}$$

where N is the dimension (i.e., number of nodes in a network) of square matrix A (adjacency matrix), λ is any eigenvalue from the eigenvalue spectrum, I_N is the N dimensional identity matrix and $A - \lambda I_N$ matrix is the column canonical form of λ eigenvalue of the adjacency matrix. We use Yuan et al. (2013) framework[35] to detect driver nodes for both synthetic and natural RNs and map them based on the multiplicity of the eigenvalue spectrum. We explore the use of geometric multiplicity of eigenvalue as a spectral metric to locate the driver nodes on the RN adjacency matrix and eigenvalue spectra. We further quantify the controllability of RNs and investigate the dependency of controllability on branching structure of the RN. The controllability C of a network can be expressed as:

$$C = \frac{N_D}{N} = \frac{M_G(\lambda_M)}{N} = \frac{max\left[1, \ N - rank(A - \lambda I_N)\right]}{N} \tag{6}$$

where $N =$ is the total number of nodes on a RN and, $N_D =$ is the number of driver nodes, Here, $\lambda = \lambda_M$ the eigenvalues that appear most frequently in an RN's eigenvalue spectrum, (in this case $\lambda = 0$) (see Figs. 2-3).

Equation (6) has a significant advantage over the maximum matching framework proposed by Liu et al.[40] to detect driver nodes as it provides the exact number and position of driver nodes on a network topology based on the geometric multiplicity and can be applied to both directed and undirected networks. In that sense, for any RN, geometric multiplicity M_G is sufficient to understand dynamics and quantify controllability under external influence.

Network heterogeneity

As discussed above, network heterogeneity is a metric that describes the variability in the arrangement of links and nodes and is important for understanding emergence and evolution of a RN[44,45]. We used Horton-Strahler stream ordering system to quantitatively describe the branching structure of a RN by classifying streams based on their hierarchical connectivity[41]. In this system, first-order streams, which have no tributaries, serve as the fundamental units. When two streams of the same order merge, they form a stream of the next higher order (see Fig. 8(a), as an example). However, when streams of different orders merge, the resulting stream retains the higher order. Higher-order streams generally exhibit greater discharge and drainage area. Additionally, branching nodes occur when two streams of the same order meet to form a higher-order channel, whereas side-branching nodes form when streams of different orders merge[11,41,65]. In this study, we employ the Tokunaga self-similarity model[46] to quantify network heterogeneity and characterize the branching patterns of RNs. The Tokunaga relation (see Fig. 8(b)) can be expressed as

$$T_k = ac^{k-1} \tag{7}$$

where Tokunaga indices $T_{ij} = T_{i(i+k)} = T_k$ and $k = j - i$. The matrix T_{ij} can be computed as $T_{ij} = \frac{N_{ij}}{N_j}$

, where N_{ij} denotes the average number of streams of order i connected to streams of order j, and $i < j$ (see[43,47,66]). This framework assumes that the mean of branches of order i connecting to randomly selected branch of order j, T_{ij}, is independent of the branch orders and only depends on the difference $k = j - i$ and follow exponential relationship with k (Eq. 7). a is a constant and c describes the degree of side-branching and can be seen as a measure of heterogeneity since it indicates the connectivity of low-order channels with respect to the high-order channels of the RN[47,57]. The parameter c is known as the Tokunaga parameter and is referred to as the c-value in the current manuscript (see Fig. 8(b)).

Critical nodes and vulnerability

Critical nodes are the set of nodes whose deletion maximizes the network fragmentation[27,58–60]. This can be achieved by minimizing the size of the largest remaining connected components or pairwise connectivity, i.e., the total number of node pairs connected by a path. This disruption metric can be used for understanding network vulnerability (i.e., optimal response of a network to an external attack) and protection (i.e., network defense) purposes. For this purpose, we have adopted the recently proposed critical node identification (CNI) framework[27] on a suite of natural RNs obtained from digital elevation model (DEM) data to understand its vulnerability.

In order to identify critical nodes, a RN can be represented by a simple undirected graph $G = (V, E)$ with a set of nodes V and links E. The links connecting node $i \in V$ and $j \in V$ are represented by a pair $(i, j) \in E$. Let $\mathbb{N}(i) = \{j : (i, j) \in E\}$ denote the neighborhood of node i. We assume that up to k nodes in this graph are deleted as critical nodes. For any node $i \in V$, we define the indicator variable v_i as

$$v_i = \begin{cases} 1, & \text{if node } i \text{ is deleted as a CN,} \\ 0, & \text{otherwise.} \end{cases} \tag{8}$$

Then, for each pair of nodes $i, j \in V$ ($i \neq j$), we define the indicator variable u_{ij} as

$$u_{ij} = \begin{cases} 1, \text{if nodes } i, j \in V \text{ are connected by a path in the remaining graph,} \\ 0, \text{otherwise.} \end{cases} \tag{9}$$

The objective function, which quantifies the number of connected node pairs in the remaining graph, and the limit on the number of removed nodes can be expressed as $\sum_{i,j \in V} u_{ij}$ and $\sum_{i \in V} v_i \leq K$, respectively. Critical nodes can then be determined via minimizing the objective function using linear integer programming. For more details on method to identify critical nodes see[27,58–60,67].

In this study, the term vulnerability (i.e., the ability to hold integrity under external influence) is used to explain the dis-connectivity of the RN. We employ the fitted power-law exponent of the remaining pairwise connectivity vs removed critical node as a measure of vulnerability[27].

Extraction of natural River Networks
We use RNs obtained from 56 natural basins located across the United States corresponding to different hydrologic and geomorphic conditions. These RNs were extracted from 1 m resolution digital elevation models (DEMs) using a curvature-based method[49]. The DEMs were obtained from National Elevation dataset website (https://apps.nationalmap.gov/viewer/). ArcGIS 10.4 and Python 3.1 were used to analyze the data and generate the figures (see Figs. 10 and 12).

Data availability
We utilized USGS LiDAR data to extract natural river networks. The processed DEM datasets of the natural basins are available in the Zenodo repository: https://doi.org/10.5281/zenodo.14799822.

Appendix: $M_A(\lambda_0)$ for perfect binary tree networks
A perfect binary tree network (PBTN) is one in which every node at the same height has 2 branches (degree 3), except for the leaf nodes at the maximum height (h). The total number of nodes for these tree graphs is given by 2^h for trees with a root node (e.g., RNs) and $2^h - 1$ otherwise. By utilizing Eq. (1) in conjunction with the characteristic polynomial attributes of matrix A, one can represent $z_h + e_h = 2^h$, where $z_h = M_A(\lambda_0)$ and e_h indicates the number of nonzero eigenvalues of A. In this case, h can be considered the highest order of RN. Therefore, we obtain the following recurrence relation for z_h based on the observation $z_{h+1} = e_h$.

$$z_{h+1} + z_h = 2^h \tag{10}$$

The general solution to this recurrence is

$$z_h = \frac{1}{3}\left(2^h - (1 + 3C_1)(-1)^h\right) \tag{11}$$

Therefore, for a rooted PBTN (e.g., RN), we have the values of z_h where $C_1 = -1$. Similarly, for $h = 2, 3, 4, 5, \ldots$, the corresponding values of z_h are $2, 2, 6, 10, \ldots$.

For an unrooted, PBTN a slightly different recurrence holds

$$z_{h+1} + z_h = 2^{h+1} \tag{12}$$

Hence, the general solution can be expressed as

$$z_h = \frac{1}{3}\left(2^{h+1} + (-1)^h\right) \tag{13}$$

As a result, for $h = 2, 3, 4, 5, \ldots$, the corresponding values of z_h are $3, 5, 11, 21, \ldots$. Fiorini et al.[68] have shown that for any tree graph with N nodes and maximum degree of D, the maximum z_h is

$$\eta(T) = N - 2\left\lceil \frac{(N-1)}{D} \right\rceil \tag{14}$$

For binary trees, with $D = 3$, this implies that

$$\eta(T) \leq N - 2\left\lceil \frac{(N-1)}{3} \right\rceil \tag{15}$$

Applying this inequality to our rooted PBTN, we have

$$z_h \leq \frac{1}{3}\left(2^h + 2\right) \tag{16}$$

This is consistent with Eq. (11), and, in fact, this upper bound is equal to Eq. (11) when h is an even number. Rojo and Robbiano[69] derived an explicit formula for the eigenvalues (and their multiplicities) of A for any Bethe tree.

The unrooted PBTN is a special case with $D = 2$. In their notation, κ represents the number of levels in the tree, with the tree height given by $h = \kappa - 1$. They demonstrated that all eigenvalues are expressed as:

$$\lambda(q, s) = 2\sqrt{2} \cos\left(\frac{\pi q}{s + 1}\right) \tag{17}$$

where s ranges from 1 to $h + 1$ and q ranges from 1 up to s. The largest eigenvalue (or spectral radius) occurs when the argument to the cosine is smallest, or when $q = 1$ and $s = h + 1$. Note that the limiting value of the spectral radius as h approaches infinity is therefore $2\sqrt{2}$; also found by Liu[70]. The smallest eigenvalue occurs for an (q, s) pair such that $\frac{q}{(s+1)}$ is less than, but as close as possible, to 1/2. Rojo and Robbiano[69] state that the multiplicity of an eigenvalue given by Eq. (17) is given by 2^{h-s}. Note that this is independent of q. Equation (17) give the same eigenvalue for different (q, s) pairs. In particular, the eigenvalue of 0 is attained whenever $\frac{q}{(s+1)} = \frac{1}{2}$, or $q = \frac{(s+1)}{2}$. So to obtain the multiplicity of the eigenvalue 0, it is necessary to compute the sum

$$z_h = \sum_{s=1}^{h} N(s) 2^{h-s} \tag{18}$$

where $N(s)$ is the number of integer q-values for a given s such that $q = \frac{(s+1)}{2}$. We therefore find that $N(s) = 1$ if s is odd, and $N(s) = 0$ if s is even. For example, when $h = 3$, we have $N(1) = 1$, $N(2) = 0$ and $N(3) = 1$. Equation (18) gives the correct result when h is odd, however, requires the addition of 1 when h is even[69].

References

1. Tucker, G. E. & Slingerland, R. Drainage basin responses to climate change. *Water Resources Research* **33**, 2031–2047 (1997).
2. Tucker, G. E. Drainage basin sensitivity to tectonic and climatic forcing: Implications of a stochastic model for the role of entrainment and erosion thresholds. *Earth Surface Processes and Landforms* **29**, 185–205 (2004).
3. Rodriguez-Iturbe, I., Muneepeerakul, R., Bertuzzo, E., Levin, S. A. & Rinaldo, A. River networks as ecological corridors: A complex systems perspective for integrating hydrologic, geomorphologic, and ecologic dynamics. *Water Resources Research* **45** (2009).
4. Hooshyar, M., Singh, A. & Wang, D. Hydrologic controls on junction angle of river networks. *Water Resources Research* **53**, 4073–4083 (2017).
5. Bertuzzo, E. *et al.* On the space-time evolution of a cholera epidemic. *Water Resources Research* **44** (2008).
6. Czuba, J. A. & Foufoula-Georgiou, E. Dynamic connectivity in a fluvial network for identifying hotspots of geomorphic change. *Water Resources Research* **51**, 1401–1421 (2015).
7. Zaliapin, I., Foufoula-Georgiou, E. & Ghil, M. Transport on river networks: A dynamic tree approach. *Journal of Geophysical Research: Earth Surface* **115** (2010).
8. Hansen, A. & Singh, A. High-frequency sensor data reveal across-scale nitrate dynamics in response to hydrology and biogeochemistry in intensively managed agricultural basins. *Journal of Geophysical Research: Biogeosciences* **123**, 2168–2182 (2018).
9. Carrara, F., Altermatt, F., Rodriguez-Iturbe, I. & Rinaldo, A. Dendritic connectivity controls biodiversity patterns in experimental metacommunities. *Proceedings of the National Academy of Sciences* **109**, 5761–5766 (2012).
10. Terui, A. et al. Metapopulation stability in branching river networks. *Proceedings of the National Academy of Sciences* **115**, E5963–E5969 (2018).
11. Abed-Elmdoust, A., Miri, M.-A. & Singh, A. Reorganization of river networks under changing spatiotemporal precipitation patterns: An optimal channel network approach. *Water Resources Research* **52**, 8845–8860 (2016).
12. Ranjbar, S., Hooshyar, M., Singh, A. & Wang, D. Quantifying climatic controls on river network branching structure across scales. *Water Resources Research* **54**, 7347–7360 (2018).
13. Smith, T. R. & Bretherton, F. P. Stability and the conservation of mass in drainage basin evolution. *Water Resources Research* **8**, 1506–1529 (1972).
14. Molnár, P. & Ramírez, J. A. Energy dissipation theories and optimal channel characteristics of river networks. *Water Resources Research* **34**, 1809–1818 (1998).
15. Benda, L. et al. The network dynamics hypothesis: how channel networks structure riverine habitats. *BioScience* **54**, 413–427 (2004).
16. Rinaldo, A., Rigon, R., Banavar, J. R., Maritan, A. & Rodriguez-Iturbe, I. Evolution and selection of river networks: Statics, dynamics, and complexity. *Proceedings of the National Academy of Sciences* **111**, 2417–2424 (2014).
17. Hooshyar, M., Singh, A., Wang, D. & Foufoula-Georgiou, E. Climatic controls on landscape dissection and network structure in the absence of vegetation. *Geophysical Research Letters* **46**, 3216–3224 (2019).
18. Boccaletti, S., Latora, V., Moreno, Y., Chavez, M. & Hwang, D.-U. Complex networks: Structure and dynamics. *Physics reports* **424**, 175–308 (2006).
19. Chung, F., Lu, L. & Vu, V. The spectra of random graphs with given expected degrees. *Internet Mathematics* **1**, 257–275 (2004).
20. Cvetkovic, D. M. & Rowlinson, P. Spectral Graph Theory. *Topics in algebraic graph theory, eds. L. W. Beinke, R. J. Wilson, Cambridge University Press* 88–112 (2004).
21. Gutman, I. Chemical graph theory–the mathematical connection. *Advances in Quantum Chemistry* **51**, 125–138 (2006).
22. Van Mieghem, P. *Performance analysis of communications networks and systems* (Cambridge University Press, 2009).
23. Mohar, B. & Poljak, S. Eigenvalues in combinatorial optimization. In *Combinatorial and graph-theoretical problems in linear algebra*, 107–151 (Springer, 1993).
24. Bonetti, S., Bragg, A. D. & Porporato, A. On the theory of drainage area for regular and non-regular points. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* **474**, 20170693 (2018).
25. Tejedor, A., Longjas, A., Zaliapin, I. & Foufoula-Georgiou, E. Delta channel networks: 1. a graph-theoretic approach for studying connectivity and steady state transport on deltaic surfaces. *Water Resources Research* **51**, 3998–4018 (2015).
26. Tejedor, A., Longjas, A., Zaliapin, I. & Foufoula-Georgiou, E. Delta channel networks: 2. metrics of topologic and dynamic complexity for delta comparison, physical inference, and vulnerability assessment. *Water Resources Research* **51**, 4019–4045 (2015).
27. Sarker, S., Veremyev, A., Boginski, V. & Singh, A. critical nodes in river networks. *Scientific reports* **9**, 1–11 (2019).

28. Abed-Elmdoust, A., Singh, A. & Yang, Z.-L. Emergent spectral properties of river network topology: An optimal channel network approach. *Scientific reports* **7**, 11486 (2017).
29. Golub, G. & Van Loan, C. Matrix computations, vol. 3 baltimore. *MD: JHU Press.[Google Scholar]* (2012).
30. Nering, E. D. *Linear algebra and matrix theory* (University of Michigan, Tech. Rep., 1970).
31. Biggs, N., Biggs, N. L. & Norman, B. *Algebraic graph theory*. 67 (Cambridge university press, 1993).
32. Chun, F. Spectral graph theory. cbms regional conference series in mathematics. *American Mathematical Society* (1997).
33. Rai, A. et al. Understanding cancer complexome using networks, spectral graph theory and multilayer framework. *Scientific reports* **7**, 1–16 (2017).
34. Fraleigh, J. B. *A first course in abstract algebra* (Pearson Education India, 2003).
35. Yuan, Z., Zhao, C., Di, Z., Wang, W.-X. & Lai, Y.-C. Exact controllability of complex networks. *Nature communications* **4**, 2447 (2013).
36. Lovász, L. & Plummer, M. D. *Matching theory*, vol. 367 (American Mathematical Soc., 2009).
37. Istalkar, P. & Biswal, B. Streamflow prediction in ungauged basins: How dissimilar are drainage basins?. *Journal of Hydrology* **637**, 131357 (2024).
38. Cvetković, D. M. & Gutman, I. M. The algebraic multiplicity of the number zero in the spectrum of a bipartite graph. *Matematički vesnik* **9**, 141–150 (1972).
39. Gutman, I. & Borovicanin, B. Nullity of graphs: an updated survey. *Zbornik Radova* **14**, 137–154 (2011).
40. Liu, Y.-Y., Slotine, J.-J. & Barabási, A.-L. *Controllability of complex networks. nature* **473**, 167 (2011).
41. Horton, R. E. Erosional development of streams and their drainage basins; hydrophysical approach to quantitative morphology. *Geological society of America bulletin* **56**, 275–370 (1945).
42. Rodriguez-Iturbe, I. & Rinaldo, A. *Fractal river basins: chance and self-organization* (Cambridge University Press, 2001).
43. Tarboton, D. G. A new method for the determination of flow directions and upslope areas in grid digital elevation models. *Water resources research* **33**, 309–319 (1997).
44. Snijders, T. A. The degree variance: an index of graph heterogeneity. *Social networks* **3**, 163–174 (1981).
45. Bell, F. K. A note on the irregularity of graphs. *Linear Algebra and its Applications* **161**, 45–54 (1992).
46. Tokunaga, E. Consideration on the composition of drainage networks and their evolution. *Geogr. Rep. Tokyo Metrop. Univ.* **13**, 1–27 (1978).
47. Ranjbar, S., Singh, A. & Wang, D. Controls of the topological connectivity on the structural and functional complexity of river networks. *Geophysical Research Letters* **47**, e2020GL087737 (2020).
48. Roy, J., Tejedor, A. & Singh, A. Dynamic clusters to infer topologic controls on environmental transport of river networks. *Geophysical Research Letters* **47**, e2021GL096957 (2022).
49. Hooshyar, M., Wang, D., Kim, S., Medeiros, S. C. & Hagen, S. C. Valley and channel networks extraction based on local topographic curvature and k-means clustering of contours. *Water Resources Research* **52**, 8081–8102 (2016).
50. Arora, V. K. The use of the aridity index to assess climate change effect on annual runoff. *Journal of hydrology* **265**, 164–177 (2002).
51. Budyko, M. I., Miller, D. H. & Miller, D. H. *Climate and life*, vol. 508 (Academic press New York, 1974).
52. Henning, D. & Flohn, H. *Climate Aridity Index (Budyko-Ratio)* (UN Environmental Project, 1977).
53. Ponce, V. M., Pandey, R. P. & Ercan, S. Characterization of drought across climatic spectrum. *Journal of Hydrologic Engineering* **5**, 222–224 (2000).
54. Chen, X., Alimohammadi, N. & Wang, D. Modeling interannual variability of seasonal evaporation and storage change based on the extended budyko framework. *Water Resources Research* **49**, 6067–6078 (2013).
55. Chen, X. & Sivapalan, M. Hydrological basis of the budyko curve: Data-guided exploration of the mediating role of soil moisture. *Water Resources Research* **56**, e2020WR028221 (2020).
56. Biswal, B. Dynamic hydrologic modeling using the zero-parameter budyko model with instantaneous dryness index. *Geophysical Research Letters* **43**, 9696–9703 (2016).
57. Zanardo, S., Zaliapin, I. & Foufoula-Georgiou, E. Are american rivers tokunaga self-similar? new results on fluvial network topology and its climatic dependence. *Journal of Geophysical Research: Earth Surface* **118**, 166–183 (2013).
58. Veremyev, A., Prokopyev, O. A. & Pasiliao, E. L. An integer programming framework for critical elements detection in graphs. *Journal of Combinatorial Optimization* **28**, 233–273 (2014).
59. Veremyev, A., Prokopyev, O. A. & Pasiliao, E. L. Finding groups with maximum betweenness centrality. *Optimization Methods and Software* **32**, 369–399 (2017).
60. Arulselvan, A., Commander, C. W., Elefteriadou, L. & Pardalos, P. M. Detecting critical nodes in sparse graphs. *Computers and Operations Research* **36**, 2193–2200 (2009).
61. Rigon, R., Rinaldo, A., Rodriguez-Iturbe, I., Bras, R. L. & Ijjasz-Vasquez, E. Optimal channel networks: a framework for the study of river basin morphology. *Water Resources Research* **29**, 1635–1646 (1993).
62. Rodríguez-Iturbe, I. et al. Energy dissipation, runoff production, and the three-dimensional structure of river basins. *Water Resources Research* **28**, 1095–1103 (1992).
63. Paik, K. & Kumar, P. Emergence of self-similar tree network organization. *Complexity* **13**, 30–37 (2008).
64. Rinaldo, A., Rodriguez-Iturbe, I., Rigon, R., Ijjasz-Vasquez, E. & Bras, R. Self-organized fractal river networks. *Physical review letters* **70**, 822 (1993).
65. Peckham, S. D. New results for self-similar trees with applications to river networks. *Water Resources Research* **31**, 1023–1029 (1995).
66. Wang, K. et al. Side tributary distribution of quasi-uniform iterative binary tree networks for river networks. *Frontiers in Environmental Science* **9**, 792289 (2022).
67. Pavlikov, K. Improved formulations for minimum connectivity network interdiction problems. *Computers & Operations Research* **97**, 48–57 (2018).
68. Fiorini, S., Gutman, I. & Sciriha, I. Trees with maximum nullity. *Linear algebra and its applications* **397**, 245–251 (2005).
69. Rojo, O. & Robbiano, M. An explicit formula for eigenvalues of bethe trees and upper bounds on the largest eigenvalue of any tree. *Linear algebra and its applications* **427**, 138–150 (2007).
70. Liu, L. The spectral radius of the adjacency matrix of a complete binary tree as the number of nodes approaches infinity. In *2nd International Conference on Applied Mathematics, Modelling, and Intelligent Computing (CAMMIC 2022)*, vol. 12259, 322–328 (SPIE, 2022).

Acknowledgements

Author contributions

S.S. conducted the simulations and wrote the first draft. S.S. and A.S. analyzed the results. A.V., V.B., S.P. and A.S.

provided feedback in interpreting results. All authors contributed in preparing the manuscript.

Declarations

Competing interests
The authors declare no competing interests.

Additional information
Correspondence and requests for materials should be addressed to A.S.

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

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