Investigating Topologic and Geometric Properties of Synthetic and Natural River Networks under Changing Climate

Shiblu Sarker

University of Central Florida

Part of the Civil Engineering Commons

Find similar works at: https://stars.library.ucf.edu/etd2020

University of Central Florida Libraries http://library.ucf.edu

This Doctoral Dissertation (Open Access) is brought to you for free and open access by STARS. It has been accepted for inclusion in Electronic Theses and Dissertations, 2020- by an authorized administrator of STARS. For more information, please contact STARS@ucf.edu.

STARS Citation


https://stars.library.ucf.edu/etd2020/965
INVESTIGATING TOPOLOGIC AND GEOMETRIC PROPERTIES OF SYNTHETIC AND NATURAL RIVER NETWORKS UNDER CHANGING CLIMATE

by

SHIBLU SARKER
B.Sc. Bangladesh University of Engineering and Technology, 2013

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctoral of Philosophy, in the Department of Civil, Environmental and Construction Engineering in the College of Engineering and Computer Science at the University of Central Florida Orlando, Florida

Spring Term
2021

Major Professor: Arvind Singh
© 2021 Shiblu Sarker
ABSTRACT

River networks are important landscape features that collect and transport water, sediment and nutrients from regions of higher elevation to lower elevations. These networks have been studied for several decades from a range of geomorphological and hydrological perspectives. Investigating the geometric and topological properties of river networks is important for developing predictive models describing the network dynamics under changing climate as well as for quantifying the physical processes operating upon them. Although these networks have been characterized for a wide range of geomorphic properties, topological properties, and in particular, spectral properties of river networks received limited attention. In this dissertation, we propose a framework to identify critical nodes on river networks in the context of vulnerability under external disruptions. In addition, through a graph-theoretic formulation of river network topology, we investigate the observed range of zero eigenvalues on the spectra using the notion of multiplicity, that can be related to controllability of the river network for a comprehensive understanding of the dynamics of a system under external forcing. Furthermore, we use topological and geometrical signatures of the river networks and their organizational complexity to study advection and diffusion of fluxes on the network. The findings of this research reveal great potential to understand external forcing, e.g. climatic, control on river networks’ geometric and topological properties.
ACKNOWLEDGMENTS

I would like to express my thanks to my advisor Dr. Arvind Singh, for his guidance throughout my research. I would also like to say a special thank you to Dr. Alexander Veremyev and all of my committee members Dr. Dingbao Wang, Dr. Talea Mayo, and Dr. Vladimir Boginski for their insightful comments and help.

I would like to thank my past and present colleagues at the CHAMPS lab for always sharing their opinions and insights regarding my research in the past four years. Working in the CHAMPS lab was an amazing and inspiring experience for me. Most importantly, I am thankful to my Grandmother Parul Rani Dey, my cousin Sumon Dey, and all other family members for their constant encouragement and my wife Tanni Sarker who supported me during this journey. Further, without the inspiration of my childhood mentor and teacher Parimol Ball, I would not be able to come this far and accomplish my achievements. I am also thankful to Grant Sanderson and his YouTube channel 3Blue1Brown that helps me throughout the study.
TABLE OF CONTENTS

LIST OF FIGURES ........................................................................................................................................ viii
LIST OF TABLES ............................................................................................................................................... xiv
CHAPTER 1: INTRODUCTION ........................................................................................................................ 1
CHAPTER 2: SYNTHETIC AND NATURAL RIVER NETWORKS ................................................................. 5
   2.1 Simulated River Networks .................................................................................................................. 5
   2.2 Natural River Networks ................................................................................................................... 6
CHAPTER 3: CRITICAL NODE DETECTION FRAMEWORK ....................................................................... 8
   3.1 Centrality and Node Importance ...................................................................................................... 10
   3.2 Critical Node Identification ............................................................................................................. 13
   3.3 Central Nodes vs Critical Nodes ...................................................................................................... 18
   3.4 Critical Nodes on Natural River Networks ....................................................................................... 20
   3.5 Critical Node Implications Towards Ecological and Biological Contexts .................................... 23
CHAPTER 4: CONTROLLABILITY AND HETEROGENEITY OF A RIVER NETWORK .................................. 25
   4.1 Matrices adopted to calculate spectral properties ........................................................................... 25
   4.2 Identification of Driver nodes: network controllability .................................................................. 26
   4.3 Network heterogeneity ...................................................................................................................... 29
   4.4 Eigenvalue and eigenvalue spectrum of synthetic river networks ................................................. 31
   4.5 Algebraic and geometric multiplicity of eigenvalues on the eigenvalue spectrum ....................... 33
   4.6 Physical meaning of multiplicity of zero eigenvalues ....................................................................... 34
   4.7 Driver nodes on synthetic river networks ....................................................................................... 38
   4.8 Heterogeneity of synthetic river network ...................................................................................... 41
4.9 Controllability and heterogeneity of natural river network for varying climate ..........42
4.10 Relation with the Critical Nodes ..............................................................................45

CHAPTER 5: A UNIFIED GRAPH-THEORETIC FRAMEWORK TO UNDERSTAND THE PHYSICAL INFEERENCE OF ENVIRONMENTAL FLUX DYNAMICS ON RIVER NETWORKS ..................................................................................................................................................51

5.1 Advection diffusion equation on a river network ............................................................52
5.2 Gradient operator of flux concentration on a river network ..........................................53
5.3 Divergence operator of flux concentration on a river network .......................................54
5.4 Laplacian operator and diffusion on a river network ......................................................55
5.5 Matrix related to advection process on a river network ................................................56
5.6 Spectral properties of the Laplacian matrix .................................................................57
5.7 Connection between topology and dynamics on river network ......................................58
5.8 Generation of width function ........................................................................................59
5.9 Hydrograph generation from width function ...............................................................63
5.10 The dynamics of hierarchical aggregation ....................................................................64
5.11 Physical inference of $\mu_2$ and $\mu_N$ from spectral point of view ...............................66

CHAPTER 6: TOPOLOGIC AND GEOMETRIC COMPLEXITY OF RIVER NETWORK ....70

6.1 Betweenness centrality as Combinatorial Organization of River Network .................71
6.2 Approximate Entropy .....................................................................................................75
6.3 Energy expenditure on the landscape ............................................................................77
6.4 Entropy and Energy from spectral point of view ..........................................................77
6.5 Geomorphological descriptors .....................................................................................79
6.6 Network Metric to Capture Complexity .........................................................................80
<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.7</td>
<td>Types of Complexity and their comparison</td>
<td>82</td>
</tr>
<tr>
<td>6.8</td>
<td>Topologic vs Geometric Complexity</td>
<td>83</td>
</tr>
<tr>
<td>6.9</td>
<td>Response of Hydro-climatic Variable on the River Network Complexity</td>
<td>85</td>
</tr>
<tr>
<td>6.10</td>
<td>Entropy and Energy from Spectral Point of View</td>
<td>87</td>
</tr>
<tr>
<td><strong>CHAPTER 7: GEOMETRIC ROUGHNESS OF THE NATURAL RIVER NETWORK UNDER CHANGING CLIMATE</strong></td>
<td></td>
<td>89</td>
</tr>
<tr>
<td>7.1</td>
<td>Tokunaga self-similar river networks</td>
<td>89</td>
</tr>
<tr>
<td>7.2</td>
<td>Power spectral density of a signal</td>
<td>90</td>
</tr>
<tr>
<td>7.3</td>
<td>Hydro-climatic forcings</td>
<td>91</td>
</tr>
<tr>
<td>7.4</td>
<td>Metrics in context of dynamics and vulnerability</td>
<td>92</td>
</tr>
<tr>
<td>7.5</td>
<td>Inhomogeneity of river network</td>
<td>92</td>
</tr>
<tr>
<td>7.6</td>
<td>Roughness of river network</td>
<td>94</td>
</tr>
<tr>
<td>7.7</td>
<td>Dynamics and vulnerability of river network</td>
<td>96</td>
</tr>
<tr>
<td><strong>CHAPTER 8: SUMMARY AND CONCLUSIONS</strong></td>
<td></td>
<td>99</td>
</tr>
<tr>
<td>REFERENCES</td>
<td></td>
<td>102</td>
</tr>
</tbody>
</table>
Figure 2.1 Synthetic river networks generated using OCN model within a defined basin boundary and uniform rainfall for (a) γ = 0.3, (b) γ = 0.5, and (c) γ = 0.7. The size of the square grid used was $50 \times 50$. The line thickness represents the channel order. The bottom right subplots in each panel show the total energy expenditure as a function of number of iterations.

Figure 3.1 Schematic example of a tree network to illustrate the concepts of central and critical nodes. Here, size of the group is considered to be equal to 1 ($S = 1$) and $a$, $b$, and $c$ are the remaining connected components (fragmentation) after the removal of node 7. Therefore, in this case, local centrality measure and the group measure are the same. However, in the case of group size $S > 1$, the group BC measure is simply fraction of the shortest paths captured by at least one of the nodes in that group.

Figure 3.2 Betweenness Centrality (undirected) for (a) γ = 0.3, (b) γ = 0.5, and (c) γ = 0.7. The dashed horizontal line in (a), (b) and (c) shows the threshold level above which 10 most central nodes were selected. Plots (d), (e) and (f) show locations of the 10 most central nodes superimposed on the river network for γ = 0.3, 0.5, and 0.7, respectively.

Figure 3.3 Locations of critical nodes on the synthetic networks for (a) γ = 0.3, (b) γ = 0.5, and (c) γ = 0.7. The associated sub-basins emerged as a result of the removal of the critical nodes can be seen in figures (d)-(f) for γ = 0.3, 0.5, and 0.7, respectively.

Figure 3.4 Comparison of critical node locations on a synthetic $OCN$ corresponding to $γ = 0.5$ for two different group sizes (i.e. $S = 5$ and $S = 10$). The critical nodes ID is also presented for visualization purposes. Among 5 critical nodes (shown as blue circles) in $S = 5$ (bottom left
panel), 3 are common in \( S = 10 \) (bottom right panel), however, 2 critical nodes (shown as red boxes) change their locations in the group size represented by \( S = 10 \). 

**Figure 3.5** Fragmentation of network resulting in sub-basins due to removal of the (a) most central nodes and (b) the most critical nodes. Notice that the removal of critical nodes results in approximately linear slopes between the number of sub-basins and number of nodes removed as a function of \( \gamma \). 

**Figure 3.6** Pairwise connectivity of the network for different \( \gamma \) values due to removal of the most central nodes (a) and the most critical nodes (b). Inset in (b) shows the pairwise connectivity for critical nodes on a log-log plot, suggesting power-law behavior.

**Figure 3.7** DEMs and superimposed river networks for six natural basins used in this study. Identified critical nodes are also superimposed on the network for visualization purposes. The resolution of collected DEMs was 10 m. The channel order based on the Horton-Strahler (Horton, 1945) ordering scheme is shown with different colors.

**Figure 3.8** (a) Number of sub-basins emerged and (b) pairwise connectivity (%) as a function of number of critical nodes removed for different natural basins. The inset in (b) shows pairwise connectivity on a log-log scale. The numerical values represent the average slope fitted to linear regression lines in (a) and to the power-law relationships shown in inset in (b).

**Figure 4.1** Examples of algebraic and geometric multiplicities of adjacency matrix for undirected (a) and directed (b) hypothetical networks. \( IN \) represents the identity matrix. The rows that are linearly dependent on other rows in the \( A - \lambda IN \) are marked by red.

**Figure 4.2** Computation of algebraic and geometric multiplicity of a hypothetical network using the adjacency matrix shown in figure 4.1(a).
Figure 4.3 Tokunaga self-similarity model and computation of $a$ and $c$ for a hypothetical and a 4th order natural river network; different colors indicates different order (Horton, 1945).

Figure 4.4 Example of linear transformation on a 10 x10 grid obtained from an exemplary basin (a) Original vector space (b) and vector space after transformation (c). Corresponding exemplary basin on the grid is also shown.

Figure 4.5 Schematic of three undirected hypothetical networks with same number of nodes but different channel arrangement: pure branching (a), higher branching nodes than the side-branching nodes (b), and higher side branching nodes than the 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.

Figure 4.6 Maximum matching concept shown for a hypothetical directed network for the case of figure 4.5c. 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 figure 4.7). Red links show the set of links under maximum matching for the directed network.

Figure 4.7 Exact location of driver nodes for a directed synthetic network (shown in Fig. 2(c)) generated by OCN. Driver nodes (star) superimposed on synthetic river network (a), and zoomed-in image showing location of the driver nodes (b).

Figure 4.8 Percentage of driver node as a function of $\gamma$ for synthetic network (a). For undirected network driver nodes, $N_D$ (%); Zero eigenvalue $\lambda_0$ (%) as a function of $\gamma$ (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). Network generated by OCN for γ=0.1 to γ=0.9..........................................................39

**Figure 4.9** Relation between heterogeneity obtain from Tokunaga analysis (c-value) and node degree variance (a), c-value and controllability, C (%) (b), and node degree variance and controllability C (%) for undirected networks (c)........................................................................41

**Figure 4.10** Example of DEMs and superimposed river networks for basin in humid (a) and arid (b) climate; the channel order based on the Horton-Strahler (Horton, 1945) ordering scheme is shown with different colors. Controllability C (%) (c), heterogeneity represented by node degree variance (d) and c-value (e) for natural basins as a function of climate aridity index. ........................................................................................................................................43

**Figure 4.11** Vulnerability as a function of climate aridity index. The inset shows pairwise connectivity as a function of the number of CN removal and numerical values represent the power-law fitted slope........................................................................................................................................47

**Figure 4.12** Example of a natural DEM (1-m resolution) and corresponding extracted RN with superimposed critical nodes and driver nodes (a), and percentage of $N_{DJ}$ and $N_{CN \cap N_{DJ}}$ as a function of climate aridity index (b). The inset in (b) shows the correlation between $N_{DJ}$ and $N_{CN \cap N_{DJ}}$. ........................................................................................................................................50

**Figure 5.1** Example of matrixes associated with a hypothetical river network: The hypothetical river basin along with river network (a), Adjacency matrix (b), Degree matrix (c), Incidence matrix (d), transpose of Incidence matrix (e), Laplacian matrix (f). .........................52

**Figure 5.2** Mathematical operations on adjacency matrix of an exemplary stream network (presented in figure 5.1) to deduce width function. ........................................................................................................................................59
Figure 5.3 Width function extraction example based on proposed framework of flux dynamics.

Figure 5.4 Space time discretization on a natural river network extracted from DEM (a), Computed incremental flux concentration at outlet as a function of time (b), Width function of extracted natural river network (c), Unit hydrograph extraction from width function due to a rectangular pulse (d).

Figure 5.5 Computed cluster dynamics on a natural river network (presented in figure 5.4) based on proposed framework (a), Phase transition example in cluster dynamics (b).

Figure 5.6 Synchronizability example on two hypothetical river networks: maximum branching (a), minimum branching (b).

Figure 5.7 Response of hydro climatic variable on synchronizability of river network: climate aridity index (a), mean annual rainfall (b) and Response of c-value (c).

Figure 6.1 BC computation for a hypothetical network. The combinatorial description is presented along with the modified BC. Different color indicates the different sub-tree network for the different nodes along the main channel from the outlet. The source node, internal nodes, and outlet are also presented with a distinct color.

Figure 6.2 Proposed nodewise topological, geometrical and the stream order metric (i.e. BCn, BCL and BCO score) presented as an example of a natural river network. Different color indicates different channel order (Horton, 1945).

Figure 6.3 Details of algorithm to compute Approximate entropy on a BC series for river network.

Figure 6.4 Proposed modified BC score for a natural DEM (a), extracted river network (b), topological vs geometric BC signal (c), and topological vs stream order BC signal (d).
**Figure 6.5** Topological and geometric signal observed for a natural river network shown in figure 6.4........................................................................................................................................................................81

**Figure 6.6** Correlation test between the stream order and the topological complexity (a), stream order complexity and c-value (b), proposed topological and geometric complexity (c) and topological vs geometric complexity computed by using geomorphic signal on the watershed (d). ........................................................................................................................................................................83

**Figure 6.7** Influence of hydro-climatic variable, climate aridity index (top panel) and mean annual rainfall (bottom panel) on proposed topological and geometric complexity. ..........................86

**Figure 6.8** Role of the hydro-climatic variable on river network organizational complexity and their corresponding energy dissipation on the landscape. .................................................................86

**Figure 6.9** Role of the hydro-climatic variable on spectral energy expenditure (graph energy) and entropy (von Neumann entropy) of the river network .................................................................88

**Figure 7.1** River network inhomogeneity (c-value) to the response of climate aridity index (a), Mean annual precipitation (b), and NDVI (c) .........................................................................................................93

**Figure 7.2** Geometric time series signal generation on river network extracted from the natural DEM (a); different color indicates different channel order (Horton, 1945), analogy of computed time series with the width function (b), and roughness computation using $PSD$ (c)...95

**Figure 7.3** Response of geometric roughness under the influence of c-value (a), climate aridity index (b), mean annual precipitation (c), and NDVI (d).........................................................................................96

**Figure 7.4** Correlation test between geometric roughness to the energy expenditure on landscape (a), controllability (b), and vulnerability of river networks (c) .........................................................98
LIST OF TABLES

Table 3.1 Climatic and geomorphic properties of natural basins. ...............................23

Table 4.1 Vulnerability, critical nodes (CN) and driver nodes (N_D) characteristics of natural RNs. C denotes controllability, $N_D$ is the number of driver nodes on junctions of a network, $N$ is the total number of nodes in a network, $N_{CN \cap N_D}$ is the number of common driver and critical nodes, and $k$ is the total number of identified critical nodes in a network.................................48
CHAPTER 1: INTRODUCTION

The structure and the evolution of river networks are controlled by the external forcings such as climate and tectonics (Abed-Elmdoust et al., 2016; Bonnet & Crave, 2003; Gabet et al., 2004; Gasparini et al., 2008; Godard et al., 2013; Goren et al., 2014; Ranjbar et al., 2018; Rinaldo et al., 1995; A Singh et al., 2015; Tucker, 2004; Tucker & Slingerland, 1997; Whittaker, 2012). These river networks exhibit dendritic patterns and serve as primary pathways for transport of sediment, water, and other environmental fluxes. They also provide necessary ecosystems to a variety of ecologic and biotic activities (Bertuzzo et al., 2008; Czuba & Foufoula-Georgiou, 2015; Rodriguez-Iturbe et al., 2009; Zaliapin et al., 2010). However, these networks are facing significant threats under changing climate and anthropogenic activities. Therefore, the knowledge of the topologic and geometric structure and their dynamics of river networks is crucial for better understanding of their emergence and evolution under change as well for prediction and management of environmental fluxes operating upon them (Bertuzzo et al., 2008; Czuba & Foufoula-Georgiou, 2015; Ghotbi, Wang, Singh, Blöschl, et al., 2020; Ghotbi, Wang, Singh, Mayo, et al., 2020; Rodriguez-Iturbe et al., 2009; Tucker & Slingerland, 1997; Zaliapin et al., 2010). Several studies have focused on quantifying river network structure and their response to changing external forcings (Abed-Elmdoust et al., 2016; Coulthard et al., 2000; Godard et al., 2013; Meade, 2005; Rinaldo et al., 1995; Tucker & Slingerland, 1997). Rinaldo et al., (1995) investigated the effects of cyclic climatic forcing on drainage density using a numerical model, and suggested that under no active uplift, drainage density during successive wet periods tends to decline progressively. They also showed that the influence of climatic oscillations on landscape evolution and argued that the period of input precipitation signal imprints a distinct response on
sediment flux (Godard et al., 2013). More recently, Abed-Elmdoust et al., (2016) examined the effect of non-uniform spatio-temporal rainfall on simulated river networks and their branching patterns and reported significant reorganization of these patterns with different geomorphic and hydrologic signatures.

Branching patterns of river networks have been argued to exhibit complex topology and share similar properties such as scaling and self-similarity with other complex systems (Rodríguez-Iturbe & Rinaldo, 2001). A common way to quantify network complexity is via graph-theoretical approaches that have been used extensively in characterizing networks from diverse fields including but not limited to social networks, transportation networks, communication networks, and networks from computer science and mathematics (Albert & Barabasi, 2002; Barabasi & Albert, 1999; Boccaletti et al., 2006; I Gutman, 2006; Mohar & Poljak, 1993). In chapter 2, we discuss in detail the generation of synthetic river networks and extracted natural river networks used throughout the study. Although there are many different approaches for quantifying node importance (Borgatti, 2006), in chapter 3, we focus on the one which is based on the effect of node removal to the network structure. The corresponding problem is known as the critical node detection problem (Lalou et al., 2018; Sarker et al., 2019). Specifically, we consider the number of connected pairs of nodes in the remaining network as a measure of vulnerability and fragmentation of river networks and find the set of nodes whose removal minimizes this measure. As a solution approach for finding exact locations of critical nodes, we use the linear integer programming problem formulation developed in Veremyev et al., (2014), due to its simplicity of implementation and effectiveness on the considered networks. In addition, since all river networks investigated here are trees, we demonstrate that the considered critical node identification problem is equivalent to the problem of finding a group of nodes with the highest group betweenness
centrality (Veremyev et al., 2017), which can be interpreted as a quantitative measure of the role of a group of nodes as intermediaries in the process monitoring and control of flow in a network. The implications of these results are discussed from a network protection perspective under potential natural or man-made disruptions (Sarker et al., 2019).

In chapter 4, we characterize physical connection and arrangement of the simulated and natural river network using a tree network topology to understand the Spectral properties associated with adjacency matrix and their behavior as a function of $\gamma$, the mathematical and physical meaning of eigenvalue and eigenvector in terms of linear transformation, mathematical explanation of range of zero eigenvalue in the spectrum using the notion of multiplicity of eigenvalue, and how multiplicity of zero eigenvalues is related to the network controllability (Liu et al., 2011; Yuan et al., 2013) and heterogeneity (Snijders, 1981) for both synthetic and natural river network. Natural river networks extracted using 1m digital elevation model (DEM) to show network controllability and heterogeneity for distinct climatic region (humid and dry) across the United States.

In chapter 5, we have proposed a unified graph-theoretic framework to understand the physical inference of environmental flux dynamics using the notion of advection diffusion equations defined on a static river network topology. In addition to that we show how proposed framework can reproduce the width function as well as the hierarchical aggregation of environmental fluxes which is well known geometric representation of watershed. Proposed framework also provides synchronizability metric that can be a comprehensive understanding of the environmental flux delivery system on watershed.

In chapter 6, the network of extracted natural rivers was further evaluated to understand the applicability of network centrality concept to capture river network organization, what
modification of betweenness centrality ($BC$) could be useful to show natural sequence in river network topology and geometry. In addition to that, we investigate the significance of approximate entropy ($ApEn$) concept on the network centrality sequence along the longest channel from the outlet to quantify the organizational complexity of the river network. Furthermore, the comparison of topologic and geometric complexity was interpreted in the context of energy expenditure.

In chapter 7, we proposed an abstract geometric roughness metric computed on the basis of the time series of flux concentration at the outlet as a proxy for the width function, which can be interpreted as a self-similar property. The proposed metric is able to reproduce the notion of fractal nature of geometry of the river network. Furthermore, the proposed metric is also tested for the following three hypotheses: (a) information encoded in geometric roughness may be correlated to the Tokunaga self-similar property of river network which is commonly used inhomogeneity metric of river network, (b) geometric roughness metric encodes the information on the influence of climate and presence of vegetation on the watershed, and (c) there is an association between the dynamics and vulnerability with the proposed geometric roughness.

In chapter 8, we include a summary of each chapter and the concluding remarks as well as some future research directions.
CHAPTER 2: SYNTHEtic AND NATURAL RIVER NETWORKS

2.1 Simulated River Networks

In this dissertation, we generate river networks using optimal channel network (OCN) approach. The OCNs recreate several topologic and geometric properties commonly observed in real river networks. They have been explored extensively in the past and recent studies from a range of hydrologic and geomorphic perspectives (Abed-Elmdoust et al., 2016; Molnar & Ramírez, 1998; Paik & Kumar, 2008; Rigon et al., 1993; Rinaldo et al., 1993; Rodríguez-Iturbe et al., 1992). In general, OCN modeling is based 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 \( i^{th} \) link of the network, and \( L_i \) and \( Q_i \) are its length and discharge, respectively (Rodríguez-Iturbe & Rinaldo, 2001). The energy exponent \( \gamma \), varies between 0 and 1, characterizes the mechanics of erosional processes and defines branching pattern of the channel network (Rodríguez-Iturbe & Rinaldo, 2001). For simplification, the topologic energy was computed assuming the unit distance between two adjacent nodes and a uniform precipitation over the entire simulated basin was applied (Abed-Elmdoust et al., 2016, 2017). Here, we generate channel networks for different values of \( \gamma \) (from 0.1 to 0.9; representing varying geomorphic processes) keeping the initial random tree network the same for each simulation.

The OCN basins were simulated on an arbitrary shape mimicking real basins using an initial square grid of size \( N \times N \) nodes, where \( N = 50 \) (see Abed-Elmdoust et al., 2016). Figure 2.1 shows simulated OCNs within a defined basin boundary obtained under uniform rainfall for \( \gamma = 0.3 \) (Fig. 2.1a), \( \gamma = 0.5 \) (Fig. 2.1b), and \( \gamma = 0.7 \) (Fig. 2.1c). Note that the networks shown here corresponded to the iteration with minimum energy (\( \sim 2 \times 10^5 \)).
Figure 2.1 Synthetic river networks generated using OCN model within a defined basin boundary and uniform rainfall for (a) $\gamma = 0.3$, (b) $\gamma = 0.5$, and (c) $\gamma = 0.7$. The size of the square grid used was $50 \times 50$. The line thickness represents the channel order. The bottom right subplots in each panel show the total energy expenditure as a function of number of iterations.

2.2 Natural River Networks

In addition to the synthetic river network, natural river network also extracted from natural basins from 10 m resolution digital elevation models (DEMs) from six different regions across the United States (Tarboton, 1997) (https://viewer.nationalmap.gov/basic/) to compare and visualize the critical (vulnerable) nodes with the synthetic river network.

In addition, to investigate the other notable properties such as controllability, synchronizability and complexity under different climatic conditions (arid vs humid climate), basins across the United States are selected, based on the availability of LiDAR data. The LiDAR
data were obtained from (https://lta.cr.usgs.gov/lidar_digitalelevation). The long-term climate (arid vs humid) can be quantified based on the aridity index (CAI) defined as the ratio of mean annual potential evaporation ($E_p$) to precipitation ($P$) (Budyko et al., 1974; Henning & Flohn, 1977). To extract a more accurate network from the LiDAR data (DEM) we have adopted a curvature-based method (Hooshyar et al., 2016, 2017, 2019; Keylock et al., 2021; A Singh et al., 2015).
CHAPTER 3: CRITICAL NODE DETECTION FRAMEWORK

In this chapter, based on the graph-theoretic approach, we use linear integer programming formulation to detect exact locations of critical nodes on the river network topology, i.e., the nodes, whose removal minimizes the number of connected node pairs in the remaining graph. Specifically, we consider a slight variation of compact formulation techniques described in Veremyev, Boginski, et al., (2014). The presented framework allows to find exact optimal solutions on sparse graphs using the existing state-of-the art optimization solvers and standard computational power (desktop computer or laptop) within a reasonable time. We use Python 2.7 with Gurobi 7.5.2 (Gurobi Optimization, Inc., 2016) as an optimization solver to implement and solve the corresponding linear integer programs. Below we formally present general idea behind this formulation and the corresponding integer program in more details.

A river network is assumed to be represented by a simple undirected graph \( G = (N,E) \) with a set of nodes \( N \) and edges \( E \). The edge connecting node \( i \in N \) and \( j \in N \) is represented by a pair \( (i,j) \in E \). Let \( N(i) = \{j: (i,j) \in E\} \) denote the neighbourhood of node \( i \). Assume that up to \( K (\leq |S|) \) nodes in this graph are deleted as critical nodes. For any node \( i \in N \), we define the indicator variable \( v_i \) as

\[
v_i = \begin{cases} 1, & \text{if node } i \text{ is deleted as a critical node} \\ 0, & \text{otherwise} \end{cases}
\]  

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

\[
u_{ij} = \begin{cases} 1, & \text{if nodes } i,j \in V \text{ are connected by a path in the remaining graph} \\ 0, & \text{otherwise} \end{cases}
\]
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 \( \sum_{i,j \in N} u_{ij} \) and \( \sum_{i \in N} v_i \leq K \), respectively.

Then, the problem formulation can be written in the following simple form:

\[
\begin{align*}
\text{minimize} & \quad \sum_{i,j \in V} u_{ij} \\
\text{subject to} & \quad u_{ij} \geq 1 - v_i - v_j, \quad \forall (i,j) \in E \\
& \quad u_{ij} \geq u_{kj} - v_i, \quad \forall (i,j) \notin E, k \in N(i) \\
& \quad \sum_{i \in V} v_i \leq K \\
& \quad v_i, u_{ij} \in \{0,1\}, \quad \forall i,j \in N
\end{align*}
\]  

(3) \hspace{2cm} (4) \hspace{2cm} (5) \hspace{2cm} (6) \hspace{2cm} (7)

The key idea behind this formulation is to recursively model connectivity variables using constraints (4) and (5). Note that this formulation is slightly different from the one developed in Veremyev, Boginski, et al., (2014) and omits some constraints. Those constraints are not required due to the minimization nature of the problem and such formulation allows to relax variables \( u_{ij} \) to be nonnegative, as shown by Pavlikov, (2018). In addition, for larger networks, we use more compact formulation which does not consider nodes with degree 1 (see Veremyev, Boginski, et al., (2014) and Pavlikov, (2018) for more details).
As a remark, we note that since all river networks analyzed here are trees and the considered critical node detection problem has been shown to be polynomial-time solvable on trees Di Summa et al., (2011), it can be solved using the polynomial-time algorithm, in contrast to the general case of graphs where this problem is NP-hard. However, we use the approach that is applicable to all types of graphs (rather than the specialized polynomial-time algorithm designed for trees) due to its simplicity of implementation and a very short computational time for all considered network instances. In addition, we point out that there are some other studies in the critical node detection area employing integer programming or other techniques which can be used to identify critical nodes in river networks (Pavlikov, 2018; Santos et al., 2018; Shen et al., 2012; Di Summa et al., 2011; Walteros et al., 2018). However, they are either more complex or require some preprocessing procedures. For more information on critical node detection problems and solution techniques, we refer the reader to the recent survey by Lalou et al., (2018).

3.1 Centrality and Node Importance
Centrality is one of the most fundamental concepts in network analysis which is used to quantify the `importance' or `influence' of a node to the network structure. Although various centrality measures have been proposed and investigated in many different contexts (Borgatti & Everett, 2006; M. Newman, 2018) including network vulnerability analysis (Iyer et al., 2013; Lü et al., 2016), in this paper we focus on one particular centrality measure, betweenness centrality, due to its natural interpretation in the context of nodes ``controlling" the environmental fluxes (e.g. water, sediments, nutrients) through river networks, as well as due to its direct relation to the critical node detection problem which we will state formally later. Specifically, we identify the locations of the most `central' nodes according to the betweenness centrality measure and compare them with the locations of the most `critical' nodes.
Betweenness centrality (BC) is a commonly used topological measure for identifying location and evaluating influence of a node on the overall network (Freeman, 1977). It is a local measure (assigns a score for each node) and is based on the number of shortest paths captured by each node traversing through it from the entire network. For a given network with nodes $N$, if $n_{st}(u)$ is the number of shortest paths from node $s$ to node $t$ that pass-through node $u$ and $n_{st}$ is the total number of shortest paths from node $s$ to node $t$ from the entire network, then BC (or influence) score (not scaled) of node $u$ can be mathematically expressed as:

$$C(u) = \sum_{s \neq u} \frac{n_{st}(u)}{n_{st}} \quad (8)$$

**Figure 3.1** Schematic example of a tree network to illustrate the concepts of central and critical nodes. Here, size of the group is considered to be equal to 1 ($|S| = 1$) and $a$, $b$, and $c$ are the remaining connected components (fragmentation) after the removal of node 7. Therefore, in this case, local centrality measure and the group measure are the same. However, in the case of group size $|S| > 1$, the group BC measure is simply fraction of the shortest paths captured by at least one of the nodes in that group.
Figure 3.1 shows a hypothetical network with 12 nodes and illustrates the concept of BC. For example, out of $12 \times (12 - 1) = 132$ shortest paths among all pairs of nodes (note that there is only one shortest path between any pair of nodes in this network), the number of shortest paths captured by node $u$, here $u = 7$, (i.e., shortest paths containing node $u$) is 100, hence, $C(u) = 100$.

Notice, in this example, $BC$ score (scaled) of node 7 is $100/132$; this is an individual or local $BC$ measure for node 7. Most central nodes can be identified on any network based on the individual $BC$ score.

![Graphs showing Betweenness Centrality](image)

**Figure 3.2** Betweenness Centrality (undirected) for (a) $\gamma = 0.3$, (b) $\gamma = 0.5$, and (c) $\gamma = 0.7$. The dashed horizontal line in (a), (b) and (c) shows the threshold level above which 10 most central nodes were selected. Plots (d), (e) and (f) show locations of the 10 most central nodes superimposed on the river network for $\gamma = 0.3$, 0.5, and 0.7, respectively.
To identify the locations of the most central nodes and their influence (strength) on the network, we computed $BC$ for each node in the entire network for varying energy exponent $\gamma$ (note that the number of shortest paths between any pair of nodes $s, t$ is $n_{st} = 1$ if a graph is a tree which is the case in our setup). Figure 3.2 shows, as an example, the computed node-wise $BC$ (in particular, undirected $BC$) using equation (8) for $\gamma = 0.3$ (Fig. 3.2a), $\gamma = 0.5$ (Fig. 3.2b), and $\gamma = 0.7$ (Fig. 3.2c). Assuming that we are interested in identifying 10 most central nodes on the network, we selected 10 nodes for each $\gamma$ above a threshold value represented by dashed lines (Fig. 3.2, top panels). Figures 3.2 (d), (e) and (f) show locations of the 10 most central nodes superimposed on the OCN for $\gamma = 0.3$, 0.5, and 0.7, respectively. Also note that although we have computed networks and their characteristics for $\gamma = 0.1 - 0.9$, we only show results for $\gamma = 0.3$, 0.5, and 0.7 due to space considerations.

3.2 Critical Node Identification

Critical node identification (CNI) problem deals with the optimal deletion of nodes from a network to maximize the network fragmentation. This can be achieved by minimizing the size of the largest remaining connected components or pair-wise connectivity, i.e., the total number of node pairs connected by a path (Arulselvan et al., 2009; Veremyev, Boginski, et al., 2014; Veremyev, Prokopyev, et al., 2014). The pair-wise connectivity is a network disruption metric which 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. Therefore, critical nodes represent a group (subset) of nodes that are crucial for maintaining the integrity of a network (Lalou et al., 2018).

In this chapter we use the graph-theoretic framework and integer programming formulations to identify critical nodes on the river network topology (see Methods for details).
This framework can be used to detect critical nodes on a tree network. Since river networks exhibit tree like structures which can be delineated numerically using stream ordering schemes (Horton, 1932, 1945; R L Shreve, 1966, 1967; Ronald L Shreve, 1974), of interest would be to explore locations of critical nodes on these networks.

Moreover, since the considered river networks are trees, any pair of nodes is connected by exactly one shortest path. Hence, a pair of nodes becomes disconnected if a node belonging to the corresponding shortest path is removed. Thus, the number of pairs of nodes that can be disconnected by removal of one node \( u \) is equal to the \( BC \) score \( C(u) \) of a corresponding node.

Therefore, the detection of critical nodes in our network topology by minimizing the pairwise connectivity is equivalent to identifying most influential (central) group of nodes with the highest ‘group betweenness centrality’ (Veremyev et al., 2017) with same predefined size as number of critical nodes. In particular, the pairwise connectivity is the inverse of network fragmentation and the set of critical nodes is equivalent to a group of nodes with maximum group betweenness centrality which is a group centrality measure as opposed to individual node measure (i.e. \( BC \)).

For a group of nodes \( S \subseteq N \) in a given undirected graph \( G \), the group betweenness centrality of \( S \) in \( G \) can be defined as

\[
C(S) = \sum_{s,t \in N} \frac{n_{st}(S)}{n_{st}}
\]  

(9)

This group betweenness centrality measure is based on the number of the shortest paths \( n_{st}(S) \) between any pair of nodes \((s, t)\) in a graph \( G \) that pass through at least one of the nodes in a group \( S \), thus relating pairwise connectivity with group betweenness centrality. Therefore, the
critical nodes act together as a ‘group capturing the highest number of shortest paths’ between all pairs of nodes in the network.

In other words, in the group betweenness centrality maximization formulation, the group $S$ is the set of critical nodes whose removal fragments the network into a number of connected components that can be inferred as sub-basins of the river network $G$. In the example shown in Fig. 3.1, $a$, $b$ and $c$ are the remaining connected components due to the deletion of node 7. For a group size of 1, node 7 is the only node whose removal provides minimum pairwise connectivity (32=6+6+20 connected node pairs remain within connected components $a$, $b$ and $c$, respectively) or maximum fragmentation of the network. For this group size, critical node is also the most central node with the highest BC score, as the number of shortest paths traversing through node 7 plus the number of remaining connected node pairs when node 7 is removed is 132, which is the total number of shortest paths among all pairs of nodes. Since, in the case of a tree network, any pair of nodes is connected by one and exactly one shortest path as there is no loop in the tree network, in such a case, critical nodes are equivalent to group of nodes with highest $BC$ score.
Figure 3.3 Locations of critical nodes on the synthetic networks for (a) $\gamma = 0.3$, (b) $\gamma = 0.5$, and (c) $\gamma = 0.7$. The associated sub-basins emerged as a result of the removal of the critical nodes can be seen in figures (d)-(f) for $\gamma = 0.3$, 0.5, and 0.7, respectively.

Figure 3.3 (a-c) depicts the locations of critical nodes on the synthetic networks generated by OCN approach for $\gamma = 0.3$, $\gamma = 0.5$, and $\gamma = 0.7$ (Figs. 3.3a-c, respectively) and the associated sub-basins emerged as a result of the critical nodes deletion (Figs. 3.3d-f). The definition of equation (9) also implies that, for $|S| = 1$, group BC becomes individual node $BC$. In other words, if one looks for a ‘single’ `most important” node, then the most central node in a river network is also the most critical node. However, the location of the critical node is dependent on the group size $|S|$. The example shown in Fig. 3.4 illustrates the locations of the critical nodes based on the group sizes of $|S| = 5$ (circle) and $|S| = 10$ (stars) on the same synthetic network. It can be seen that the critical nodes for $|S| = 5$ are not necessarily a subset of critical nodes for $|S| = 10$. There are 3 critical nodes out of total 5 that are common in the group size $|S| = 5$ and
$|S| = 10$, indicating that the location of the critical nodes is dependent on the group size. In this chapter, we use a user defined group size to minimize pairwise connection of the network under node removal. In short, critical nodes for a given group size (e.g. $|S| = 5$) is not necessarily a subset of critical nodes for larger group size (e.g. $|S| = 10$). However, note that, one can constrain an optimization problem where a group of identified critical nodes can be fixed in case new critical nodes need to be identified on the network (e.g. for a larger group size).

| Group size $|S|$ | Critical node ID | Pairwise connectivity (%) |
|----------------|-----------------|--------------------------|
| 1              | 795             | 39                       |
| 2              | 525 795         | 21.5                     |
| 3              | 525 795 1156   | 17.7                     |
| 4              | 525 582 887 1156| 12.6                     |
| 5              | 92 525 582 887 1156| 10.2                     |
| 6              | 92 457 525 582 887 1156| 8.6   |
| 7              | 92 457 525 582 887 1041 1256| 7.4   |
| 8              | 92 457 525 582 887 1016 1041 1256| 6.4   |
| 9              | 92 296 457 525 582 887 1016 1041 1256| 5.6   |
| 10             | 92 296 457 508 653 887 905 1016 1041 1256| 5.1   |

**Figure 3.4** Comparison of critical node locations on a synthetic OCN corresponding to $\gamma = 0.5$ for two different group sizes (i.e. $|S| = 5$ and $|S| = 10$). The critical nodes ID is also presented for visualization purposes. Among 5 critical nodes (shown as blue circles) in $|S| = 5$ (bottom left panel), 3 are common in $|S| = 10$ (bottom right panel), however, 2 critical nodes (shown as red boxes) change their locations in the group size represented by $|S| = 10$.

To investigate the characteristics of sub-basins formed by the removal of critical nodes from the network as a function of $\gamma$, we analyze the number of fragments (sub-basins) from the
modeled channel networks obtained using OCN approach for different γ values. These results are shown in Fig. 3.5 and discussed in the following sections.

### 3.3 Central Nodes vs Critical Nodes

Comparing locations of a set of most central nodes with most critical nodes on a network (Figs. 3.2 and 3.3), it can be seen that the central nodes tend to be more clustered (localized) than critical nodes. To further understand and compare the characteristics of sub-basins induced by the central nodes and critical nodes, we analyze the number of sub-basins versus number of nodes removed. Figure 3.5 shows the fragmentation of the network resulting in sub-basins due to the removal of the most central nodes (Fig. 3.5a) and the most critical nodes (Fig. 3.5b). Here, 10 most central and most critical nodes are considered for comparison. Unlike in Fig. 3.5a, number of sub-basins follow a linear trend when plotted against the number of critical nodes removed for varying γ with an average slope $\sim 2.43 \pm 0.26$. Moreover, the number of sub-basins (fragmentations) induced by critical node deletion is much larger in the case of critical nodes than central nodes. For example, excluding $\gamma = 0.8$ and $\gamma = 0.9$ (Fig. 3.5a), on average the number of sub-basins generated in the case of central nodes is roughly half of that as in the case of critical nodes, as a result of a closer node localization in the case of most central nodes.
Figure 3.5 Fragmentation of network resulting in sub-basins due to removal of the (a) most central nodes and (b) the most critical nodes. Notice that the removal of critical nodes results in approximately linear slopes between the number of sub-basins and number of nodes removed as a function of $\gamma$.

Figure 3.6 Pairwise connectivity of the network for different $\gamma$ values due to removal of the most central nodes (a) and the most critical nodes (b). Inset in (b) shows the pairwise connectivity for critical nodes on a log-log plot, suggesting power-law behavior.
Figure 3.6 shows the pairwise connectivity (i.e. number of connected node pairs in the remaining network) of the network for different $\gamma$ values due to removal of the most central nodes (Fig. 3.6a) and the most critical nodes (Fig. 3.6b). A few observations can be made from Fig. 3.6. i) The pairwise connectivity (PC) shows higher variability as a function of $\gamma$ for central nodes, whereas it almost remains same for critical nodes. ii) Overall, PC is higher for central nodes for larger number of nodes removed compared to critical nodes. iii) While the PC decreases with increasing $\gamma$ for both most central and critical nodes, a clear power-law decay is observed in the case of critical nodes with slope parameter approximately $\sim -0.86$ (inset in Fig. 3.6b). Figures 3.5 and 3.6 also suggest that the removal of subsequent central nodes on the list of most central nodes captured a lot of the same paths that were already served by the removal of previous nodes on the list. On the other hand, the group measure of BC has the ability to resolve this issue.

Based on above observations and defining vulnerability as the maximum fragmentation of a network, we argue that critical nodes, as opposed to central nodes, provide a more consistent measure to quantify the vulnerability of a river network under node disruptions.

### 3.4 Critical Nodes on Natural River Networks

Natural river networks were extracted from 10 m resolution DEMs from six different regions across the United States (see Fig. 3.7 and Table 1). The networks were extracted based on same threshold flow accumulation value for channel initiation (Tarboton, 1997). Similar to simulated networks, 10 critical nodes were identified using CNI algorithm discussed above and in Methods and were superimposed on the network for visualization purposes (Fig. 3.7). As can be seen from Fig. 3.7, the identified critical nodes are scattered around the network on the basin. Note that for visual perception purposes, only five critical nodes are shown on each network in the figure;
however, both pairwise connectivity and number of sub basins were computed based on 10 most
critical nodes.

Figure 3.8a shows the number of sub-basins generated as a function of number of critical
nodes removal for different natural basins, exhibiting approximately linear relationship with
removed critical nodes. However, a significant variability is observed between basins, in particular
for a larger set of critical nodes removed. Also, note a higher fragmentation in case of river network
from South Dakota; this may be due to higher drainage density (ratio of total channel length to
total basin area (Hooshyar et al., 2017; Ranjbar et al., 2018; Rodríguez-Iturbe & Rinaldo, 2001;
Tucker & Slingerland, 1997) of the South Dakota basin (see Table 1). Figure 3.8b shows the
relation between the percentage of PC and number of nodes removed. As seen from Fig. 3.8b, the
fraction of the remaining connected pairs of nodes when critical nodes are removed from real river
network also exhibits a power-law behavior for all six analyzed basins. The average power-law
exponent ($slope = -1.02$ ; obtained from log-log plots from individual basins by averaging)
for the real basins is slightly steeper than observed from synthetic basins. This observation of
power-law slope of $\sim -1$ and its origin for both synthetic and natural basins require further
investigation and will be the focus of a future study.
Figure 3.7 DEMs and superimposed river networks for six natural basins used in this study. Identified critical nodes are also superimposed on the network for visualization purposes. The resolution of collected DEMs was 10 m. The channel order based on the Horton-Strahler (Horton, 1945) ordering scheme is shown with different colors.

Figure 3.8 (a) Number of sub-basins emerged and (b) pairwise connectivity (%) as a function of number of critical nodes removed for different natural basins. The inset in (b) shows pairwise connectivity on a log-log scale. The numerical values represent the average slope fitted to linear regression lines in (a) and to the power-law relationships shown in inset in (b).
Table 3.1 Climatic and geomorphic properties of natural basins.

<table>
<thead>
<tr>
<th>Basins</th>
<th>Drainage area ((km^2))</th>
<th>Total channel length ((km))</th>
<th>Drainage density ((1/km))</th>
<th>Maximum channel order</th>
<th>Average temperature (^\circ(F))</th>
<th>Average annual precipitation (rainfall-inch)</th>
<th>Total number of nodes</th>
</tr>
</thead>
<tbody>
<tr>
<td>California</td>
<td>1931.05</td>
<td>387.78</td>
<td>0.201</td>
<td>4</td>
<td>64.1</td>
<td>12.83</td>
<td>76</td>
</tr>
<tr>
<td>Florida</td>
<td>4166.05</td>
<td>967.82</td>
<td>0.232</td>
<td>4</td>
<td>68.3</td>
<td>56.52</td>
<td>166</td>
</tr>
<tr>
<td>Texas</td>
<td>3365.24</td>
<td>761.14</td>
<td>0.226</td>
<td>4</td>
<td>64.3</td>
<td>40.97</td>
<td>128</td>
</tr>
<tr>
<td>Virginia</td>
<td>2215.18</td>
<td>585.17</td>
<td>0.264</td>
<td>4</td>
<td>56</td>
<td>43.11</td>
<td>92</td>
</tr>
<tr>
<td>South Dakota</td>
<td>4571.53</td>
<td>1350.91</td>
<td>0.295</td>
<td>5</td>
<td>43</td>
<td>21.76</td>
<td>188</td>
</tr>
<tr>
<td>Washington</td>
<td>1439.84</td>
<td>364.205</td>
<td>0.252</td>
<td>3</td>
<td>52.65</td>
<td>37.13</td>
<td>72</td>
</tr>
</tbody>
</table>

3.5 Critical Node Implications Towards Ecological and Biological Contexts

In this chapter, we identified critical nodes using the notion of connectivity. The reduction in the connectivity of a network is a measure of the integrity of the network. Biological and ecological communities in riverine ecosystems often occur in spatially structured habitats where connectivity directly plays a key role in their processes. Some previous works suggested that the hierarchical and branching structure of river networks can explain the observed traits of the riverine biodiversity (Carrara et al., 2012) and has been argued to potentially affect the biological diversity and productivity in riverine ecosystems (Benda et al., 2004). For example, Carrara et al., (2012) suggested that constrained species dispersal on the dendritic network face higher extinction risks and heterogeneous habitats sustain higher levels of among-community biodiversity. Based on theoretical observations and long-term dataset, Terui et al., (2018) found that the branching complexity of riverine structure dampens the temporal variability of metapopulations and a loss of such complexity may affect resilience of the metapopulation (Terui et al., 2018). They also
suggested that scale-invariant characteristics of fractal river networks emerged as important stabilizing agent for riverine metapopulation (Terui et al., 2018). Although the flow of water in river networks is directional, due to such biodiversity and ecological considerations, undirected group betweenness centrality was considered to detect critical nodes on the river network topology. In that sense, critical nodes identification and their locations have significant potential to understand the stability and persistence of biodiversity of meta-communities to stochastic watershed disturbances (e.g. floods, fires, droughts, and storms) and maintain ecological integrity on the landscapes as well as for quantifying relationships between patchy and heterogeneous habitat to network fragmentation across multiple spatial and temporal scales.
CHAPTER 4: CONTROLLABILITY AND HETEROGENEITY OF A RIVER NETWORK

In this chapter, we characterize connections and arrangements of channels and junctions in a river network topology, where nodes represent stream junctions and links represent stream segments. The simulated and natural river network topology were analyzed to understand: (1) the spectral properties associated with the connectivity (adjacency) matrix and their behavior with changing climatic and physical processes, (2) the physical and mathematical meaning of eigenvalues and eigenvectors of the connectivity matrix in terms of linear transformation, and (3) the range of zero eigenvalue in the eigenvalue spectrum using the notion of multiplicity of eigenvalue. In addition, we investigate how the multiplicity of zero eigenvalues (nullity) is related to the network controllability and heterogeneity for both synthetic and natural river networks for varying physical processes and climate (i.e. humid and dry).

4.1 Matrices adopted to calculate spectral properties

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

$$A_{ij} = \begin{cases} 1: i \text{ flows to } j; \\ 0: \text{ otherwise} \end{cases}$$ (10)

Adjacency matrix (also referred to as connectivity matrix (Biggs, 1993)) defines the connection between nodes and the flow direction in river networks. Following properties of the adjacency matrix are associated with river networks (Abed-Elmdoust et al., 2017). (i) If there is a connection between upstream node i and downstream node j then the element of A corresponding to $i^{th}$ row and $j^{th}$ column will be 1; thus, all elements of the row associated to the outlet of the
network are zero because there is no downstream node for an outlet. (ii) The number of 1’s in every column of A represents how many numbers of nodes directly contribute from the immediate upstream to the corresponding node. And (iii), all the diagonal elements of A are zero.

To compute the real eigenvalue spectrum of drainage networks, undirected adjacency matrix can be considered by ignoring the flow direction (Abed-Elmdoust et al., 2017) which can be represented as:

\[
B_{ij} = \begin{cases} 
1: & i \text{ and } j \text{ adjacent;} \\
0: & \text{otherwise}
\end{cases}
\] (11)

In addition, the relation between matrix A and matrix B can be expressed as \( B = A + A^T \), however, \( A \) cannot be obtained from \( B \).

To further explore spectral properties of river networks, we also employ the degree matrix. The degree matrix of a network is a diagonal matrix which contains information about the degree of each node i.e., the number of links connected to each node. For any graph \( G = (N, E) \), the degree matrix \( D \) for the graph \( G \) is an \( N \times N \) diagonal matrix. Thus:

\[
D_{ij} = \begin{cases} 
\text{deg}(N_i): & \text{if } i = j; \\
0: & \text{otherwise}
\end{cases}
\] (12)

where the degree \( \text{deg}(N_i) \) of a node counts the number of times a link terminates at that node. For a directed graph, the term degree of a node defines either indegree (number of incoming edges) or outdegree (number of outgoing edges).

**4.2 Identification of Driver nodes: network controllability**

In this chapter, we identify the driver nodes \( (N_D) \) on a network to understand the controllability of river networks and investigate the dependency of controllability on river network 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 (Liu et al., 2011). In particular, the control of these driver nodes is enough to achieve full control of the network system dynamics. Liu et al. (2011) 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. (2013) 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:

$$M_G(\lambda) = \max[1, N - \text{rank}(A - \lambda I_N)] \quad (13)$$

where $N$ is the dimension (i.e. number of nodes in the network) of square matrix $A$ (adjacency matrix), $\lambda$ is any eigenvalue from the eigenvalue spectrum, $I_N$ is the $N$ dimensional identity matrix and $A - \lambda I_N$ matrix is called the column canonical form of $\lambda$ eigenvalue of adjacency matrix. We use Yuan et al. (2013) framework to detect driver nodes for both synthetic and natural river networks 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 river network based on column canonical form of adjacency matrix and eigenvalue spectra. We further quantify the controllability of river networks and investigate the dependency of controllability on branching structure of the river network. The controllability $C$ of a network can be expressed as:
\[ C = \frac{N_D}{N} = \frac{M_G(\lambda_M)}{N} = \max[1, N - \text{rank}(A - \lambda I_N)] \]  

where \( N = \) is the total number of nodes on a river network and, \( N_D = \) is the number of driver nodes. Here, \( \lambda = \lambda_M \) due to the fact that zero appears maximum times in the eigenvalue spectrum of river network (see figs 4.1-4.2).

Equation (14) has a significant advantage over the maximum matching framework proposed by Liu et al. (2011) 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 river network, geometric multiplicity \( M_G \) is sufficient to understand dynamics and quantify controllability under external influence.

**Figure 4.1** 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.
Network heterogeneity

Network heterogeneity is a metric that describes the diversity of characteristics often related to the uniformity of the network organization (i.e., arrangements of links and nodes) and is important for understanding emergence and evolution of a river network. In this section, we quantify network heterogeneity based on degree distribution of the network. Snijders, (1981) and Bell, (1992) proposed variance of node degree as a measure of network heterogeneity which can be expressed as:

$$D_{var} = \frac{1}{N} \sum_{i=1}^{N} (k_i - \langle k_i \rangle)^2$$

(15)
where, \( \langle k_i \rangle \) represents the average node degree in the network. Note that, in the case of a natural river network, the degree matrix was built, incorporating the link length in addition to the connectivity.

We also use the Tokunaga self-similarity model to quantify the network heterogeneity, which is an alternative way to characterize the branching pattern of river networks (Tokunaga, 1978). It has advantages over the Horton-Strahler indexing system which suffers from limitations due to high order side branching frequently occurring in natural river networks (Tarboton, 1997). The Tokunaga relation can be expressed as

\[
T_k = ac^{k-1}
\]

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 figure 4.3). This framework assumes that the mean 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 \( k \) (equation (16)). \( a \) 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 river network (Zanardo et al., 2013). The parameter \( c \) is known as the Tokunaga parameter (hereafter referred to as c-value).
Figure 4.3 Tokunaga self-similarity model and computation of \( a \) and \( c \) for a hypothetical and a 4th order natural river network; different colors indicates different order (Horton, 1945).

4.4 Eigenvalue and eigenvalue spectrum of synthetic river networks

The basic concept of eigenvalues can be explained by any square matrix that represents a transformation on a vector space. Based on a linear transformation, eigenvalues and eigenvectors can be obtained as:

\[
AX = \lambda X
\]  

(17)
where $A$ is any square matrix that represents a transformation of a vector space, $X$ are some vectors on the vector space (i.e. eigenvectors) and $\lambda$ are the corresponding scale factors (i.e. eigenvalues). In other words, eigenvectors are the directions in which the matrix acts solely as a scaling transformation, and eigenvalues are the corresponding scale factors. A schematic description of linear transformation and the scaling is shown in figure 1, which depicts an example of linear transformation on a $10 \times 10$ square grid (vector space). Figures 1a and 1b represent original and transformed vector spaces, respectively. From Figure 1 it can be seen that the green vectors change their direction after transformation, while red and blue preserve their direction after the linear transformation. This transformation is said to be linear because the underlying assumption is an original line vector remains a line after the transformation. For example, in figure
1, red and blue vectors are called eigenvectors due to the fact that they preserve their directions. In addition, the length of the red vectors remains the same after the transformation, however, the blue vectors increase their length with a certain scale factor. These scale factors are the corresponding eigenvalues of blue eigenvectors. On the other hand, eigenvalues for red eigenvectors are 1 because they do not change their length after the transformation.

In spectral graph theory, properties of the networks are analyzed based on the eigenvalue spectrum of the network connectivity matrix to determine the network’s complex and heterogeneous behavior (Chung, 2011; Rai et al., 2017). Previous studies found that the eigenvalue spectrum of adjacency matrix exhibits striking features such as spectral gap & nullity which are closely related to the branching pattern of the network (Abed-Elmdoust et al., 2017). In this chapter, we generate synthetic river networks using OCN approach with varying energy decay exponent $\gamma = 0.1$ to 0.9 (Figs 2; see also methods) to investigate the range of zero eigenvalues in their eigenvalue spectrum using multiplicity of eigenvalue. The following section presents the notion of multiplicity in the eigenvalue spectrum for a small exemplary stream network.

4.5 Algebraic and geometric multiplicity of eigenvalues on the eigenvalue spectrum

Figure 4.1 depicts a small undirected (top left panel) and directed (bottom left panel) exemplary stream networks with 4 nodes and 3 links. Corresponding real eigenvalues for both undirected and directed networks are computed using equation (17) and presented in figure 4.1. Based on spectral graph theory, the number of times an eigenvalue appears in an eigenvalue spectrum is known as algebraic multiplicity of that eigenvalue (denoted by $M_A(\lambda)$) (Fraleigh, 2003; Golub & Van Loan, 2013; Nering, 1970). Figure 4.1 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. Besides algebraic multiplicity, the geometric multiplicity of an eigenvalue \( M_G(\lambda) \), i.e. the number of linearly independent eigenvectors associated with it, is also characterized. Mathematically, geometric multiplicity of an eigenvalue \( M_G(\lambda) \) is the dimension of the null space of that eigenvalue.

Figure 4.1 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 figure 4.2. Based on figure 4.1 and figure 4.2 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., \( M_A(\lambda) \geq M_G(\lambda) \)) (Golub & Van Loan, 2013; Yuan et al., 2013). 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 the following sub-section we explore the physical explanation of the multiplicity of zero eigenvalues and how the arrangement of nodes is related to the \( M_G(\lambda_0) \).

### 4.6 Physical meaning of multiplicity of zero eigenvalues

In order to investigate the physical meaning of the multiplicity of zero eigenvalues, we adopt the concept of matching. For an undirected graph, matching is defined as the independent set of links that do not share common vertices (see figure 4.5), whereas for a directed network, set of links do not share common in-nodes or out-nodes (here in-node for any link can be out-node for another link). Furthermore, a node is said to be matched when it is an endpoint of one of the links in the matching, and the rest of the nodes in that graph are called unmatched nodes (see figure 4.5). The minimum possible number of unmatched nodes can be achieved by maximum matching. The
maximum matching is a special condition of matching, where if any link is added to the matching set, it will no longer be a matching of that graph. The number of members in that set is also referred to as the cardinality of the maximum matching.

**Figure 4.5** Schematic of three undirected hypothetical networks with same number of nodes but different channel arrangement: pure branching (a), higher branching nodes than the side-branching nodes (b), and higher side branching nodes than the 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.

Figure 4.5 shows that algebraic multiplicity of zero eigenvalues $M_A(\lambda_0)$ is exactly equal to the minimum number of unmatched nodes which is determined based on the concept of maximum matching. According to Liu et al. (2011), unmatched nodes are sufficient to understand the controllability of a dynamic network system. Controllability is a physical quantity commonly used
in network control theory which measures the ability to control the dynamics of a network under external influence (formally addressed in the method section). Among all the nodes from 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. Note that, as discussed above, although $M_A(\lambda_0)$ 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.

Figures 4.5a-c show schematics of three undirected hypothetical networks with total number of nodes 16. Network 1 (figure 4.5a) contains 8 source nodes and is pure branching with 7 branching nodes, Network 2 (figure 4.5b) has 4 branching nodes and 3 side-branching nodes, whereas, Network 3 (figure 4.5c) 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 river network. It can also be seen that, for undirected river network with side-branching, driver nodes avoid junction nodes, however, for pure branching, driver nodes position can be at junction nodes. Hence, one can conclude that the branching pattern of a river network can influence the location of driver nodes, and thus the controllability of a network.

For a directed river network, the number of unmatched nodes is equal to the number of source nodes. Figures 4.6a shows the hypothetical network presented in Figure 4.5c as a directed network along with the locations of the driver nodes based on the maximum matching (Figs. 4.6b-
c). As can be seen from Fig. 4.6c, 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. (2013) 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 exact formulation by Yuan et al. (2013), we use larger size network such as synthetic network generated by OCN. Figure 4.7a shows the location of driver nodes based on the \( M_G(\lambda_M) \).

In addition, figure 8b visually confirms that not all the driver nodes are at the source nodes.

**Figure 4.6** Maximum matching concept shown for a hypothetical directed network for the case of figure 4.5c. 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 figure 4.7). Red links show the set of links under maximum matching for the directed network.
In the next sections, we aim to address how the multiplicity of zero eigenvalue and driver nodes are related to the geomorphic and climatic properties of the synthetic river networks.

4.7 Driver nodes on synthetic river networks

![Driver nodes on synthetic river networks](image)

**Figure 4.7** Exact location of driver nodes for a directed synthetic network (shown in Fig. 2(c)) generated by OCN. Driver nodes (star) superimposed on synthetic river network (a), and zoomed-in image showing location of the driver nodes (b).

In order to further investigate the role of branching patterns on network dynamics, i.e. controllability, we compute the number of driver nodes and identify their locations using Yuan et al. (2013). We generated several synthetic river networks based on OCN approach with a fixed number of nodes ($N = 1513$), however with varying energy exponent $\gamma = 0.1$ to 0.9. Figure 4.7a shows, as an example, a generated synthetic river network. The locations of the superimposed driver nodes on the network can be seen in Figs. 4.7a and b. Note that for each $\gamma$, 15 independent
networks were generated, and their ensemble averaged metrics were computed in order to minimize the effect of random network initialization.

Figure 4.8 Percentage of driver node as a function of $\gamma$ for synthetic network (a). For undirected network driver nodes, $N_D$ (%); Zero eigenvalue $\lambda_0$ (%) as a function of $\gamma$ (a). For directed network number of driver node ($N_D = M_G(\lambda_0)$); number of source node $N_S$ as a function of $\gamma$ (a), Driver node (%) located on Source node, branching and side branching node (b).

Network generated by OCN for $\gamma=0.1$ to $\gamma=0.9$.

Figure 4.8 shows that 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 $\gamma$. Thus percentage of zero eigenvalue represents controllability, $C$ for an undirected network. From Fig. 4.8a, it can be observed that controllability decreases with increasing $\gamma$ where the drainage pattern changes drastically from an intertwined river network to an entirely straightened pattern (Abed-Elmdoust et al., 2017). Note
that, controllability $C = \frac{N_D}{N}$ for an undirected river network represents the nullity (the ratio of number zero eigenvalues to the total number of eigenvalues) (Abed-Elmdoust et al., 2017). Furthermore, they show that the nullity is independent of basin size and shape indicating that controllability for undirected river network is independent of shape and size of a river network. Similarly, for the case of directed river network, controllability also decreases with increasing $\gamma$ with similar trend but higher than the undirected river network (figure 4.8a). Figure 4.8a shows the comparative trend as a function of $\gamma$ for both undirected and directed synthetic river network.

It is also observed that, for directed network number of driver nodes ($N_D$) computed based on geometric multiplicity of zero eigenvalues $M_G(\lambda_0)$ is exactly equal to the number of source nodes ($N_S$), which further confirm the previous argument (figure 4.8a). However, significant amount of driver nodes is not located at the source node (figure 4.8b). In addition, we also observed that, (~4.5%-7%) driver nodes are located at the branching node and (~12.5%-16%) driver nodes are located at the side-branching nodes. It also observed that as $\gamma$ increases percentage of driver nodes located on the side-branching nodes increases even though total percentage of driver nodes decreases. However, driver nodes located on the branching nodes shows higher variability and does not follow any notable trend.
4.8 Heterogeneity of synthetic river network

**Figure 4.9** Relation between heterogeneity obtain from Tokunaga analysis (c-value) and node degree variance (a), c-value and controllability, C (%)(b), and node degree variance and controllability C (%) for undirected networks (c).

Figure 4.9a shows a significant correlation between the heterogeneity obtained from the Tokunaga analysis (c-value; formally defined in the method section) and variance of node degree. In network science, variance of a node degree is a classical measure of degree heterogeneity (i.e. diversity of the network) (Bell, 1992; Snijders, 1981). According to (Bell, 1992; Snijders, 1981), two different aspects of a complex network can be quantified through a heterogeneity measure: (i) the diversity in node degrees, and (ii) the diversity of the structure on the network (Bell, 1992; Jacob et al., 2017; Snijders, 1981). Since, variance of node degree may offer deeper insight into the understanding of the structure and functions of the network. The changes in structure can reflect changes in functions of river networks. The structure can be more heterogeneous to serve heterogeneous to serve natural purposes (Anand et al., 2011; Anand & Bianconi, 2009). To further explore the response of heterogeneity in the structure and functions of river network, we tested the
correlation between the c-value (Tokunaga, 1978; Zanardo et al., 2013) and controllability. Similar correlation also observed in case of c-value and the controllability (see figure 4.9b), that suggest higher heterogeneous river network also exhibit high controllability. Figure 4.9c shows that the heterogeneity obtained from the node degree variance is perfectly correlated with the controllability of the synthetic river networks and both increases as c-value increases. Furthermore, these observed correlations suggesting that with increasing c-value river network should exhibit more heterogeneous behavior as well as higher controllability. We have further tested this hypothesis based on natural river networks under varying climate. In the next section we explore controllability and heterogeneity of the natural river networks.

4.9 Controllability and heterogeneity of natural river network for varying climate

In this section, we investigate the spectral metrics computed based on the eigenvalue spectrum of natural river networks. For this, we extracted river networks from digital elevation models of 60 natural basins across the United States, based on the availability of LiDAR data. Although the flow of water in river networks is directional, due to biodiversity and ecological considerations, 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 (Benda et al., 2004; Carrara et al., 2012). Furthermore, other spectral metrics obtained from the synthetic river networks such as network controllability, and heterogeneity were also computed for natural river networks under different climatic conditions. Note that, in this chapter, the networks were extracted using a curvature-based method from high-resolution (1 m) topographic data (Hooshyar et al., 2016).
Figure 4.10 Example of DEMs and superimposed river networks for basin in humid (a) and arid (b) climate; the channel order based on the Horton-Strahler (Horton, 1945) ordering scheme is shown with different colors. Controllability $C$ (%) (c), heterogeneity represented by node degree variance (d) and c-value (e) for natural basins as a function of climate aridity index.

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 ($P$) (Arora, 2002; Budyko et al., 1974; Henning & Flohn, 1977; Ponce et al., 2000). 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 related to broader range climatic regimes, such as arid $12 > AI \geq 5$, semiarid $5 > AI \geq 2$, subhumid $2 > AI \geq 0.75$, and humid $0.75 > AI \geq 0.375$ (Arora, 2002; Ponce et al., 2000).
The eigenvalue spectrum computed from the natural river networks for the different climatic regions that exhibit a distinct range of zero eigenvalue. Similarly, to the synthetic river network, the eigenvalue range for natural river network’s can be explained by the algebraic multiplicity $M_A(\lambda_0)$ and the geometric multiplicity $M_G(\lambda_M)$. Furthermore, our available data suggested that the algebraic multiplicity of zero eigenvalues $M_A(\lambda_0)$ is equal to the percentage of driver node for the undirected river network, that further strengthens the hypothesis that the range of zero eigenvalues can be used as a quantitative measurement of river network controllability. On the other hand, geometric multiplicity $M_G(\lambda_M)$ suggested that the number of driver nodes on the directed river network is always equal to the number of the source node and for the natural river network, which is exactly half of the total number of nodes on the network. In addition to the controllability, we also investigated the implication of eigenvalue multiplicity on heterogeneity of natural river network, heterogeneity computed on the metric obtained from the weighted adjacency matrix, where weight is considered the link length of river network topology. Note that, in the case of a natural river network, the degree matrix was built, incorporating the link length in addition to the connectivity.

Figure 4.10(a-b) shows two sample examples of natural basins with superimposed river network for humid (a) and dry (b) climate. Figures 4.10c and 4.10d show controllability and heterogeneity as a function of CAI for the all the extracted natural river networks considered here. As can be seen, although the slope of controllability with CAI is higher ($\sim 2.32$) as compared to slope of curve for heterogeneity ($\sim 0.38$), both controllability and heterogeneity show a decreasing trend with CAI, indicating that controllability and heterogeneity are higher for river networks in humid climates as opposed to dry climates.
To further explore the decreasing trend of controllability and heterogeneity with CAI and whether these metrics depend on the branching structure or not, we perform the Tokunaga analysis on the extracted natural river network and compute the parameter c-value. Figure 4.10e shows the c-value as a function of CAI suggesting that with increasing CAI side-branching decreases. This observation further indicates the role of side-branching, and the controllability and heterogeneity of a river network increases with decreasing CAI. In particular, our results suggest that, indeed, intricate branching structure specifically inhomogeneous branching structure with more side-branching (Zanardo et al., 2013) might result in heterogeneous nature of river network that is less dynamically active and usually found in a humid environment. In contrast, arid climate basins are comparatively more dynamically active that might be due to a more homogenous branching structure compared to the humid basins. These quantitative notions provide us a useful hypothesis for a comparative and comprehensive understanding of humid vs dry basins. In other words, the value of controllability is relatively higher for humid basins than for arid basins, which leads us to infer the dynamic resilience from the spectral property of the river networks. Furthermore, it is important to note that a fair logarithmic negative correlation \((slope \approx -2.32)\) was observed from our available data with \(R^2\) value of \(\sim 0.19\), however, \(t\) test results show a significant correlation with a 95% confidence interval. Similarly, for the case of heterogeneity of the river network, a power law relation was found with a slope of \(\sim -0.38\) from our available data with \(R^2\) value of \(\sim 0.14\), however, \(t\) test results confirmed the significance with a 95% confidence interval.

### 4.10 Relation with the Critical Nodes

The controllability framework further allows us to identify the driver nodes which correspond to the set of nodes that controls the dynamic response of a system. It is worth pointing
out that most of those nodes correspond to the headwater locations (source nodes). In this regard, we explore the percentage of the driver nodes that are located on the junction nodes using natural RNs as a function of climate. A question of particular interest is to explore whether those junction driver nodes coincide with critical nodes (critical nodes result in maximum fragmentation of network) obtained under the constraint of $k = \text{number of driver nodes}$, where $k$ is the size of the group of critical nodes that capture the highest number of shortest paths (see details in chapter 3). This common set of nodes may help us to understand and quantify the effect of junction driver nodes on the vulnerability of river network under varying climates.

As discussed above, a network structure or connection which is more vulnerable may create more fragmentation under the extreme external influence (see details in chapter 3). Therefore, we compute the vulnerability (defined by the power-law exponent of pairwise connectivity) of RNs as a function of $CAI$. Figure 4.11 shows that vulnerability increases with increasing $CAI$ indicating dry basins are more vulnerable than humid basins. This quantification of vulnerability has important implications in the context of ecological and biological considerations. Specifically, riverine ecosystems often 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 (Benda et al., 2004; Carrara et al., 2012; Terui et al., 2018).
Figure 4.11 Vulnerability as a function of climate aridity index. The inset shows pairwise connectivity as a function of the number of CN removal and numerical values represent the power-law fitted slope.

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 ($N_{CN\cap NDJ}$) 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 river network, indicating that driver nodes tend to avoid critical nodes. For $k = N_D$, we observe a few common nodes. For example, Table 4.1 shows that out of 60 natural basins, 42 basins exhibit a very small
amount (~1-3.5 %) of junction driver nodes \(N_{DJ}\). 39 out of those 42 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 4.12a 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. 4.12b (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 39 basins out of 60 basins as the 21 basins did not have common critical and driver junction nodes.

Table 4.1 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.

<table>
<thead>
<tr>
<th>Basin’s CAI</th>
<th>C = (N_D/N) (%)</th>
<th>Vulnerability</th>
<th>c-value</th>
<th>(N_{DJ}/N) (%)</th>
<th>(N_{CN \cap N_{DJ}}(%)) (k = 5)</th>
<th>(N_{CN \cap N_{DJ}}(%)) (k = 10)</th>
<th>(N_{CN \cap N_{DJ}}(%)) (k = N_D)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.30</td>
<td>17.07</td>
<td>0.81</td>
<td>3.01</td>
<td>0.61</td>
<td>0</td>
<td>0</td>
<td>0.61</td>
</tr>
<tr>
<td>0.31</td>
<td>19.15</td>
<td>0.68</td>
<td>6.45</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0.34</td>
<td>17.28</td>
<td>0.84</td>
<td>3.26</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0.48</td>
<td>17.69</td>
<td>0.79</td>
<td>2.07</td>
<td>0.68</td>
<td>0.34</td>
<td>0.34</td>
<td>0.68</td>
</tr>
<tr>
<td>0.57</td>
<td>20.83</td>
<td>0.72</td>
<td>4.62</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0.61</td>
<td>18.03</td>
<td>0.72</td>
<td>2.55</td>
<td>0.41</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0.61</td>
<td>17.42</td>
<td>0.74</td>
<td>2.51</td>
<td>0.43</td>
<td>0</td>
<td>0.43</td>
<td>0.43</td>
</tr>
<tr>
<td>0.62</td>
<td>16.81</td>
<td>0.88</td>
<td>2.39</td>
<td>0.57</td>
<td>0</td>
<td>0.28</td>
<td>0.57</td>
</tr>
<tr>
<td>0.64</td>
<td>16.42</td>
<td>0.79</td>
<td>3.04</td>
<td>0.86</td>
<td>0.43</td>
<td>0.43</td>
<td>0.86</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td></td>
</tr>
<tr>
<td>0.66</td>
<td>16.04</td>
<td>0.67</td>
<td>3.02</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>0.69</td>
<td>15.54</td>
<td>0.88</td>
<td>1.81</td>
<td>1.01</td>
<td>0</td>
<td>0</td>
<td>1.01</td>
</tr>
<tr>
<td>0.69</td>
<td>14.78</td>
<td>0.76</td>
<td>2.59</td>
<td>0.82</td>
<td>0</td>
<td>0</td>
<td>0.41</td>
</tr>
<tr>
<td>0.72</td>
<td>13.04</td>
<td>0.90</td>
<td>3.21</td>
<td>1.52</td>
<td>0</td>
<td>0.43</td>
<td>1.09</td>
</tr>
<tr>
<td>0.95</td>
<td>16.81</td>
<td>0.84</td>
<td>3.20</td>
<td>0.85</td>
<td>0</td>
<td>0</td>
<td>0.28</td>
</tr>
<tr>
<td>0.96</td>
<td>13.80</td>
<td>0.86</td>
<td>3.19</td>
<td>0.61</td>
<td>0</td>
<td>0</td>
<td>0.61</td>
</tr>
<tr>
<td>1.00</td>
<td>20.00</td>
<td>0.91</td>
<td>2.42</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1.00</td>
<td>13.18</td>
<td>0.75</td>
<td>3.76</td>
<td>0.96</td>
<td>0</td>
<td>0.32</td>
<td>0.96</td>
</tr>
<tr>
<td>1.13</td>
<td>14.08</td>
<td>0.79</td>
<td>3.73</td>
<td>0.14</td>
<td>0</td>
<td>0</td>
<td>0.14</td>
</tr>
<tr>
<td>1.13</td>
<td>13.04</td>
<td>0.86</td>
<td>2.27</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1.15</td>
<td>14.37</td>
<td>0.67</td>
<td>1.36</td>
<td>1.45</td>
<td>0.15</td>
<td>0.15</td>
<td>1.45</td>
</tr>
<tr>
<td>1.15</td>
<td>12.35</td>
<td>0.82</td>
<td>2.93</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1.15</td>
<td>17.24</td>
<td>0.92</td>
<td>1.56</td>
<td>1.72</td>
<td>1.72</td>
<td>1.72</td>
<td>1.72</td>
</tr>
<tr>
<td>1.16</td>
<td>15.84</td>
<td>0.84</td>
<td>2.53</td>
<td>0.78</td>
<td>0.16</td>
<td>0.31</td>
<td>0.62</td>
</tr>
<tr>
<td>1.34</td>
<td>12.95</td>
<td>0.89</td>
<td>3.61</td>
<td>0.30</td>
<td>0</td>
<td>0</td>
<td>0.30</td>
</tr>
<tr>
<td>1.34</td>
<td>14.79</td>
<td>0.77</td>
<td>1.51</td>
<td>0.97</td>
<td>0</td>
<td>0.19</td>
<td>0.78</td>
</tr>
<tr>
<td>1.35</td>
<td>15.20</td>
<td>0.74</td>
<td>2.03</td>
<td>0.49</td>
<td>0.25</td>
<td>0.25</td>
<td>0.25</td>
</tr>
<tr>
<td>1.36</td>
<td>16.81</td>
<td>0.85</td>
<td>2.82</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1.36</td>
<td>14.19</td>
<td>0.81</td>
<td>2.56</td>
<td>0.83</td>
<td>0</td>
<td>0</td>
<td>0.67</td>
</tr>
<tr>
<td>1.36</td>
<td>15.38</td>
<td>0.84</td>
<td>2.67</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1.40</td>
<td>15.60</td>
<td>0.70</td>
<td>2.78</td>
<td>0.46</td>
<td>0</td>
<td>0</td>
<td>0.46</td>
</tr>
<tr>
<td>1.42</td>
<td>16.67</td>
<td>0.93</td>
<td>2.93</td>
<td>2.27</td>
<td>0.76</td>
<td>0.76</td>
<td>1.52</td>
</tr>
<tr>
<td>1.44</td>
<td>14.69</td>
<td>0.85</td>
<td>3.40</td>
<td>0</td>
<td>0</td>
<td>0.00</td>
<td>0</td>
</tr>
<tr>
<td>1.54</td>
<td>14.56</td>
<td>0.85</td>
<td>2.12</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1.54</td>
<td>12.50</td>
<td>0.94</td>
<td>2.53</td>
<td>2.08</td>
<td>1.04</td>
<td>2.08</td>
<td>2.08</td>
</tr>
<tr>
<td>1.56</td>
<td>15.63</td>
<td>0.83</td>
<td>2.33</td>
<td>0.78</td>
<td>0</td>
<td>0.78</td>
<td>0.78</td>
</tr>
<tr>
<td>1.64</td>
<td>14.56</td>
<td>0.85</td>
<td>2.12</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1.67</td>
<td>14.29</td>
<td>0.89</td>
<td>2.70</td>
<td>1.07</td>
<td>0</td>
<td>0.36</td>
<td>1.07</td>
</tr>
<tr>
<td>1.68</td>
<td>15.44</td>
<td>0.87</td>
<td>1.38</td>
<td>0.67</td>
<td>0</td>
<td>0</td>
<td>0.67</td>
</tr>
<tr>
<td>1.76</td>
<td>17.09</td>
<td>0.85</td>
<td>1.76</td>
<td>0.43</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1.84</td>
<td>16.44</td>
<td>0.82</td>
<td>1.11</td>
<td>2.05</td>
<td>0.68</td>
<td>1.37</td>
<td>0.68</td>
</tr>
<tr>
<td>1.93</td>
<td>15.14</td>
<td>0.75</td>
<td>2.29</td>
<td>0.40</td>
<td>0</td>
<td>0</td>
<td>0.40</td>
</tr>
<tr>
<td>2.03</td>
<td>16.05</td>
<td>0.71</td>
<td>2.42</td>
<td>1.85</td>
<td>0.62</td>
<td>1.23</td>
<td>1.23</td>
</tr>
<tr>
<td>2.06</td>
<td>14.96</td>
<td>0.88</td>
<td>2.55</td>
<td>0.88</td>
<td>0</td>
<td>0.59</td>
<td>0.88</td>
</tr>
<tr>
<td>2.20</td>
<td>14.49</td>
<td>0.85</td>
<td>3.52</td>
<td>0.72</td>
<td>0</td>
<td>0</td>
<td>0.72</td>
</tr>
<tr>
<td>2.20</td>
<td>14.58</td>
<td>0.69</td>
<td>2.49</td>
<td>1.04</td>
<td>0</td>
<td>1.04</td>
<td>1.04</td>
</tr>
<tr>
<td>2.20</td>
<td>17.78</td>
<td>0.90</td>
<td>1.81</td>
<td>1.11</td>
<td>0</td>
<td>1.11</td>
<td>0</td>
</tr>
<tr>
<td>2.32</td>
<td>13.23</td>
<td>0.78</td>
<td>1.13</td>
<td>0.53</td>
<td>0</td>
<td>0.53</td>
<td>0.53</td>
</tr>
<tr>
<td>3.47</td>
<td>19.23</td>
<td>0.72</td>
<td>2.71</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3.57</td>
<td>10.53</td>
<td>0.92</td>
<td>1.16</td>
<td>0.67</td>
<td>1.32</td>
<td>1.32</td>
<td>0.67</td>
</tr>
<tr>
<td>4.82</td>
<td>12.02</td>
<td>0.85</td>
<td>2.53</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>4.84</td>
<td>13.51</td>
<td>0.88</td>
<td>1.30</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>5.42</td>
<td>13.79</td>
<td>0.74</td>
<td>1.60</td>
<td>0.34</td>
<td>0</td>
<td>0.34</td>
<td>0.34</td>
</tr>
<tr>
<td>5.73</td>
<td>16.10</td>
<td>0.93</td>
<td>1.92</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
Figure 4.12 Example of a natural DEM (1-m resolution) and corresponding extracted RN with superimposed critical nodes and driver nodes (a), and percentage of $N_{DJ}$ and $N_{CN\cap DJ}$ as a function of climate aridity index (b). The inset in (b) shows the correlation between $N_{DJ}$ and $N_{CN\cap DJ}$. 
CHAPTER 5: A UNIFIED GRAPH-THEORETIC FRAMEWORK TO UNDERSTAND THE PHYSICAL INference OF ENVIRONMENTAL FLUX DYNAMICS ON RIVER NETWORKS

A river network can be expressed as a tree network with nodes and links. The links represent channels, and the nodes correspond to the locations where one channel splits into new channels or two or more channels merge into a single channel. This splitting of channels is called bifurcation whereas merging of channels is called confluence. These bifurcation and confluence of channels create a tree network topology which can be expressed as a connected finite graph with no cycles, whereas, a finite graph on which any two nodes connected by one and only one unique simple path. In this chapter, a river network is considered as a tree network topology on a plane represented asymmetric (for directed) and symmetric (for undirected) adjacency matrix (Abed-Elmoust et al., 2016, 2017; Sarker et al., 2019; Tejedor et al., 2015a, 2015b). A symmetric adjacency matrix is useful to characterize the flow paths in a river network, while an asymmetric adjacency matrix is very useful for understanding network structure ignoring the flow direction. Here, we adopted both directed and undirected adjacency matrices for understanding and quantifying topological structure and dynamics by using the simple notion of linear algebra. A schematic representation of adjacency matrix of a hypothetical river network is shown in figure 5.1a; figure 5.1b represents directed and undirected river network, respectively. In this chapter, we also adopted different graph-theoretic matrices (e.g. Laplacian and Incidence matrices) to investigate environmental flux dynamics on river networks using the notion of advection diffusion processes. The following section presents the concept of advection diffusion processes on a river network topology.
Figure 5.1 Example of matrixes associated with a hypothetical river network: The hypothetical river basin along with river network (a), Adjacency matrix (b), Degree matrix (c), Incidence matrix (d), transpose of Incidence matrix (e), Laplacian matrix (f).

5.1 Advection diffusion equation on a river network

Here, we considered a directed network $G(N,E)$ with a set of nodes $N = \{1, 2, 3, \ldots, N\}$, and a set of links $E = \{1, 2, 3, \ldots, N - 1\}$. The dynamics of environmental flux in a one-dimensional river can be expressed as (Jin et al., 2019):

$$\frac{dC}{dt} = -D_f \frac{d^2C}{dx^2} + v \frac{dC}{dx} \quad (18)$$

where $C$ is the environmental flux density at location $x$ and time $t$, $D_f$ is the diffusion coefficient, and $v$ is the flow velocity. Equation (19) can be written as
\[
\frac{dC}{dt} = -Df \frac{d}{dx} \frac{dC}{dx} + v \frac{dC}{dx}
\]  
(19)

\[
\frac{dC}{dt} = -Df \nabla \cdot \nabla C + v \nabla C
\]  
(20)

Here, \(\nabla C\) represents the gradient of \(C\) and \(\nabla \cdot \nabla C\) represents the divergence of gradient of \(C\). In other words, \(\nabla C\) represents the gradient of environmental flux concentration, \(\nabla \cdot \nabla C\) is the divergence of that concentration gradient. We adopted matrix operators from the discrete calculus on graph \(G(N, E)\) that are useful to explain the advection diffusion equation of environmental flux on river networks. In this regard, gradient operator can be considered as Incidence matrix \((I)\) and divergence operator can be considered as transpose of Incidence matrix \((I)\) (Chapman, 2015; Rak, 2017), and divergence of gradient can be considered as Laplacian matrix \((L)\). Now, equation (20) can be written as

\[
\frac{dC}{dt} = -I^T (Df IC + vC)
\]  
(21)

The following section provides the details of Incidence and Laplacian matrix in the context of advection and diffusion of environmental flux on a river network.

5.2 Gradient operator of flux concentration on a river network

In this chapter, we consider a concentration function on the graph \(G(N, E)\), where, \(C : N \rightarrow \mathbb{R}^N\). In order to calculate the gradient of \(C\) to obtain information about the change in \(C\) with respect to the links, we can consider the \((n - 1) \times n\) matrix, where \((n - 1)\) is the number of links and \(n\)
is the number of nodes. This gradient operator can be mapped by Incidence matrix of same
dimensions (Chapman, 2015; Rak, 2017). Incidence matrix can be expressed as:

\[ I_{ij} = \begin{cases} 
-1: & \text{if } j \text{ node is upstream of } i \text{ link} \\
1: & \text{if } j \text{ node is downstream of } i \text{ link} \\
0: & \text{otherwise} 
\end{cases} \]  \quad (22)

Figure 5.1d shows that representation of Incidence matrix for an exemplary river network.
From the equation (22), following properties of incidence matrix can be listed: (i) \( i^{th} \) row of a \( j^{th} \)
specific column of an incidence matrix represents the connection of \( i^{th} \) link with \( j^{th} \) node. (ii) If
the sign of that element of an incidence matrix is negative, then \( i^{th} \) link is outgoing from the \( j^{th} \)
node (i.e. \( i^{th} \) link convey the flux from \( j^{th} \) node) and if the sign of that element is positive then
the \( i^{th} \) link is incoming from the \( j^{th} \) node (i.e. \( i^{th} \) link contribute the flux to \( j^{th} \) node). In other
words, incidence matrix is an arrangement of flux advection on junctions in a river network
topology. (iii) row wise algebraic summation of every element is zero due to the fact that, river
network is a tree and every pair of nodes are connected by one and only one unique path.

**5.3 Divergence operator of flux concentration on a river network**

The divergence operator is the transpose of the gradient operator. Intuitively, here we are
interested to capture the divergence each link, the sum of the flux on the links connecting to that
node. We consider divergence operator as \( n \ast (n - 1) \) matrix in order to map from the links to the
nodes. This divergence operator is transpose of gradient operator, i.e. \( I^T \) can be considered as the
divergence operator (Chapman, 2015; Rak, 2017).
Figure 5.1e depicts the transpose of incidence matrix for a hypothetical tree network, which is helpful to map divergence of environmental flux on a river network topology.

### 5.4 Laplacian operator and diffusion on a river network

Laplacian operator is defined as the divergence of the gradient of a flux concentration function, which is simply the second derivative and can be expressed as:

\[
\nabla^2 C = \nabla . \nabla C = \frac{d^2 C}{dx^2}
\]

(23)

In network science, the second derivative of any function can be expressed by using Laplacian matrix (\(L\)). The Laplacian matrix plays a key role in studying diffusion on a network (Chapman, 2015; Rak, 2017; Tejedor et al., 2015a). Mathematically, Laplacian matrix is related to the incidence matrix using following equation:

\[
L = I^T * I
\]

(24)

Hence, this Laplacian operator can be seen as the divergence of the gradient of the concentration of environmental fluxes at nodes on the network. In network science, Laplacian matrix is defined as:

\[
L_{ij} = \begin{cases} 
\text{deg}(N_i): & \text{if } i = j \\
-1: & \text{if } i \neq j \text{ and } N_i \text{ is adjacent to } N_j \\
0: & \text{otherwise}
\end{cases}
\]

(25)
Figure 5.1f shows the representation of Laplacian matrix for an exemplary river network. This Laplacian matrix plays an important role in understanding and quantifying environmental flux dynamics on the river network due to its direct connection to diffusion processes.

5.5 Matrix related to advection process on a river network

In this section, we adopt a discrete notion of advection from which the continuous formulation can be derived. The original continuous formulation of advection equation can be expressed as:

\[
\frac{dC}{dt} = \nabla (vC) \quad (26)
\]

Where, \(\nabla\) is the divergence operator, \(C\) is the scalar concentration quantity moving through \(v\) (Chapman, 2015; Rak, 2017). This formulation sets the change in flux concentration at a node equal to the flux through the node. Chapman, (2015) presented advection as a modified version of consensus dynamics, which can be described by (Rak, 2017):

\[
L_{adv} = D^{in} - A \quad (27)
\]

This \(L_{adv}\) can be referred to as an advective Laplacian matrix. As a result, the final advection diffusion equation can be written as below:

\[
\frac{dC}{dt} = -D_f LC + L_{adv}C \quad (28)
\]

Equation (29) provides us a quantitative space time notion of the transport of the environmental flux on a river network topology. As a result, numerically we can discretize the above equation as below:
\[
\frac{C_{i+1} - C_i}{\Delta t} = -D_f L C_i + L_{adv} C_i
\]  

(29)

where \( C_i \) and \( C_{i+1} \) is the flux concentration at every node at \( t \) and \( t + \Delta t \) time step, respectively. \( L \) and \( L_{adv} \) can be constructed based on space discretization on along the channel (i.e. one-dimensional).

5.6 Spectral properties of the Laplacian matrix

This Laplacian matrix \( (L) \) has some interesting, useful properties. In this section, we are interested to investigate dynamical properties on a network by using the eigenvalue spectrum of \( L \). In terms of linear algebra, eigenvalues and eigenvectors can be obtained from the following equation:

\[
LX = \mu X
\]

(30)

where \( L \) is Laplacian matrix which represents a transformation of a \( N \times N \) vector space, \( X \) are commonly known as eigenvectors due to the fact that they do not change the direction after the linear transformation of the vector space and \( \mu \) are the corresponding scale factors of vectors \( X \) after linear transformation, commonly known as eigenvalues (see details in chapter 4).

\[
\mu_i = \mu_1, \mu_2, \mu_3 \ldots \ldots \mu_N
\]

(31)

For an \( N \times N \) dimensional \( L \) matrix there are \( N \) number of eigenvalues. From the eigenvalue spectrum of a \( L \) matrix following properties can be observed: first, all the eigenvalues are positive, however, the first eigenvalue is 0. If the network is connected there will be one and only one 0 eigenvalue in the spectrum. In other words, at least one of its eigenvalues is zero for
the connected network. For disconnected networks, the number of zero eigenvalues in the spectrum indicates the number of fragmentations in the network. The largest eigenvalue ($\mu_N$) of the spectrum is called the dominant eigenvalue and the eigenvector corresponding to dominant eigenvalue is called the dominant eigenvector or homogeneity vector. The smallest non-zero eigenvalue ($\mu_2$) of the spectrum is called the algebraic connectivity of the network. Eigenvector corresponding to this algebraic connectivity is known as the Fiedler Vector of the network.

In this chapter, we investigate the physical inference of the spectral properties, specifically $\mu_2$ and $\mu_N$ in the context of environmental flux dynamics on the river network topology.

### 5.7 Connection between topology and dynamics on river network

The following relation helps us to understand how environmental flux dynamics on the river network is related to the concept of connectivity.

\[
\frac{dC}{dt} = -D_f LC + L_{adv} C
\]  \hfill (32)

Equation (32) can be expressed based on the adjacency matrix, which is useful to explain topological properties of river networks (Abed-Elmdoust et al., 2016, 2017).

\[
\frac{dC}{dt} = -D_f \left( (D^{in} + D^{out}) - B \right) C + (D^{in} - A)C
\]  \hfill (33)

where, $A$ and $B$ are the directed and the undirected adjacency matrix which is formally addressed in the chapter 4. $D$ is the is a diagonal matrix which contains information about the degree of each node i.e. the number of links connected to each node. In network science, this $D$ matrix is commonly referred as degree matrix, which can be further classified and expressed as sum of in degree ($D^{in}$) and out degree matrix ($D^{out}$) (see details in chapter 4).
For an undirected graph, the degree $\text{deg}(N_i)$ of a node counts the number of times a link terminates at that node, whereas, for a directed graph, the term degree of a node defines either indegree (number of incoming links) or outdegree (number of outgoing links). Therefore, one can argue that water, solute or sediment concentration flux can be modeled only by using a static river network connectivity matrix, which further helps us to understand how the static topology of the river network affects the dynamical process operating up on the network.

### 5.8 Generation of width function

Width function is a one-dimensional signal of two-dimensional river network topology, which represents the number of nodes in the network as a function of distance from the outlet along the flow. It is especially important in case of river networks as it provides valuable details on the river basin structure (Lashermes et al., 2007; Ranjbar et al., 2018, 2020). In addition, the width function is an important geometric descriptor of a river network which can be used to gauge different properties of a river basin (Rodríguez-Iturbe & Rinaldo, 2001). Width function can be extracted from the adjacency matrix.

![Diagram of a river network and the corresponding adjacency matrix operations](image)

**Figure 5.2** Mathematical operations on adjacency matrix of an exemplary stream network (presented in figure 5.1) to deduce width function.
Figure 5.2 depicted a hypothetical river basin with 8 nodes and 7 links, which can be represented by the directed adjacency matrix \( A \) which formally stated earlier. \( A_k = A^k \) representing the \( k^{th} \) power of \( A \) shows the number of paths, along the network links, of length \( k \). The number of non-zero element of \( 8^{th} \) column tells us the number of nodes in the network from the outlet at a distance \( k \). For example, calculating \( A_3 = A^3 \), one finds that the only nonzero elements at \( 8^{th} \) column are \((A_3)_{(3,8)}\) and \((A_3)_{(4,8)}\). This is a general property of the asymmetric adjacency matrix of a river network. In fact, an interesting outcome of the path length property is that by calculating the higher powers of \( A \), the number of nonzero elements decrease until eventually after certain power all elements become zero. Moreover, if \( K \) is the length of the longest path in a network, for any \( k > K \) one finds that \( A_k = 0 \). Hence, this property can be utilized to directly calculate the width function of a river network from the directed adjacency matrix \( A \) and its integer powers. Mathematically we can write as below:

\[
W(k) = \sum_{i=1}^{N} (A^k)_{i0}
\]  

(34)

which is simply at each \( k \), a sum over the \( O^{th} \) column of \( A^k \), where, \( O \) is the associated with outlet.
The width function can also be recreated using the proposed advection diffusion framework of flux dynamics on the river network, as discussed earlier (equation 29). Figure 5.3 shows the proposed framework of advection diffusion equation for a hypothetical river network with 8 nodes and 7 links, where every link length is considered unit length and every node can be regarded as unit channelized pixel area. Figure 5.3 also demonstrates space time discretization of the advection diffusion equation for unit time and unit flux concentration. Where, \( C \) is the column matrix can be regarded as flux concentration on each node on the river networks, which updates at every time step from 0 to a finite time \( t \). This finite time follow \( t \leq K \), where, \( K \) is the length of the longest path in a network. In other words, time \( t \) can be regarded as distance from the outlet for the unit velocity over the entire links of the river network. If unit flux concentration is applied to each node then the incremental concentration of two consecutive time at \( O^{th} \) row of \( C \) is related to the number of nodes in the river network at that time step. In other words, concentration difference between

**Figure 5.3** Width function extraction example based on proposed framework of flux dynamics.
time step 3 and 2 at 8\textsuperscript{th} row in C matrix is $5 - 3 = 2$, which is essentially tells us 2 nodes in the river network at 3 unit distance from the outlet of the basin. Therefore, width function can be reproduced from the proposed advection diffusion framework of flux concentration on river network under the condition of unit flux concentration and the unit space time discretization.

**Figure 5.4** Space time discretization on a natural river network extracted from DEM (a), Computed incremental flux concentration at outlet as a function of time (b), Width function of extracted natural river network (c), Unit hydrograph extraction from width function due to a rectangular pulse (d).

Figure 5.4a provides an example of a natural river network extracted from 1m resolution DEM along with the space discretization on it. Under the initial condition of the unit flux concentration at every node the incremental concentration obtained at the outlet as function of time is shown in figure 5.4b. This incremental concentration function obtained is quantitatively equal
to the width function extracted from the same river network. This also confirms our argument that
the proposed framework might be useful to deduce the width function (figure 5.4c) of the natural
river networks, as each channel transmits flows at constant speed through the entire landscape.

5.9 Hydrograph generation from width function

A hydrograph normally displays the discharge as a function of time at a certain point in the
river network, usually at the outlet, which is a very useful curve as it has information on the
watershed response under applied precipitation on landscapes. Early studies have reported the
simple connection between hydrological response (hydrograph) through the width function (Abed-
Elmdoust et al., 2016; Snell & Sivapalan, 1994). Under a uniform precipitation of certain duration,
the unit hydrograph can be expressed in terms of the width function, where the flow velocity is
constant over the entire network. Mathematically, the unit hydrograph can be written in terms of
the width function in following way (Abed-Elmdoust et al., 2016):

\[ Q(t, t_r) = \sum_j W(x_j) \prod \left( \frac{t - t_j}{t_r} \right) \]  \hspace{1cm} (35)

and,

\[ \prod \left( \frac{t - t_j}{t_r} \right) = \begin{cases} 1 & \text{if } t_j \leq t < t_j + t_r \\ 0 & \text{otherwise} \end{cases} \]  \hspace{1cm} (36)

where, \( Q(t, t_r) \) represents the amount of discharge at the outlet at a time \( t \), \( W(x_j) \) is the total
number of nodes that are contributing to the outlet at a distance of \( x_j \), \( t_j \) is the traveling time, \( t_r \) is
the uniform precipitation of duration.
Figure 5.4d depicted an example of a hydrograph generation based on the equation (35) and (36), which is based on a simple connection between the hydrograph and the width function (figure 5.4c) under a rectangular pulse of rainfall duration (inset figure 5.4d).

5.10 The dynamics of hierarchical aggregation

Besides reproduction of hydrological and geomorphological descriptors our proposed framework is able shed some light to answer the question on how environmental fluxes (water, solutes, and sediments) operate over the network. To understand the dynamics of fluxes on a river network, a static river network topology can be regarded as a dynamic tree (Zaliapin et al., 2010). The dynamic hierarchy can be built on the concept that, when two streams are connected, they both influence the downstream dynamics. In other words, two first order streams of different length do not automatically provide a second order stream (Zaliapin et al., 2010). So, we keep track of time to merge multiple individual particles coming from source nodes with each other to form clusters along the downstream direction to the outlet. Here, distance can be regarded as time due to the consideration of unit velocity through the entire landscape. In this chapter, we show that our proposed concept of advection diffusion framework can reproduce cluster dynamics to environmental transport on river networks. In addition, the phase transition phenomena (Zaliapin et al., 2010) in the cluster dynamics of river networks can also be explained by using the advection diffusion formulation of the flux dynamics.

The consecutive merging of river streams is commonly known as hierarchical aggregation. Here, we consider a process that starts at time $t = 0$ from the source nodes with unit concentration, which is a part of our initial condition $C_0$ of our advection diffusion equation. The number of non-zero element of the matrix $C$ can be considered as number of clusters which start to merge with
one another as time evolves from $t = 1$ to $k$. Again, $t$ cannot exceed $k$ due to the consideration of unit velocity over the entire landscape, therefore, time can be considered as distance ($i.e., t = \frac{d}{v}$).

When two clusters can merge at the same time the number of clusters decreases by one and this process continues until all particles have been merged into a single cluster of magnitude is equal to the number of source node. The magnitude of every element in matrix $C$ will be the magnitude of the cluster. The number of cluster and the maximum magnitude of the cluster has important information about the downstream transport along a river network (Zaliapin et al., 2010). Example of the hierarchical aggregation on a natural river network extracted from DEM which is presented earlier section is depicted in the figure 5.5. The number of cluster and the maximum magnitude of the cluster is plotted as a function of time. Figure 5.5a-b exhibits phase transition of a natural river network where the number of clusters of connected nodes decreases and the size of the largest cluster increases. Furthermore, the maximum magnitude of matrix $C$ changes from an exponential to a power-law function as the critical time ($t_c$) approaches.

![Figure 5.5](image-url)

**Figure 5.5** Computed cluster dynamics on a natural river network (presented in figure 5.4) based on proposed framework (a), Phase transition example in cluster dynamics (b).
Zaliapin et al., (2010) synthesized the statistical properties on hierarchical dynamics of river networks. Furthermore, they show phase transition phenomena based on cluster dynamics. In other words, at the small time steps the maximum magnitude of clusters increases exponentially with time, however, at some critical time \( t_c \) the maximal magnitude undergoes a drastic qualitative change and follows power law behavior. This phenomenon is commonly known as phase transition in network percolation theory. In simple words, if the probability of node connection pairs or the probability of pair of nodes getting connected through transportation of flux can be specified by \( p \), then the network will only be connected when \( p > p_c \), and the corresponding cluster is called the dominant cluster. On the other hand, if \( p < p_c \), then the network will be considered as an isolated cluster in dynamic point of view. In summary, river networks exhibit properties similar to network percolation which can be quantified based on the notion of advection diffusion of flux over the network.

5.11 Physical inference of \( \mu_2 \) and \( \mu_N \) from spectral point of view

In network science, \( \mu_2 \) is commonly known as Fiedler value that reflects how well connected the overall network is. This Fiedler value is dependent on the number of nodes, as well as the way in which nodes are connected. In other words, Fiedler value decreases with the number of nodes increases. In many literatures pertaining to network science, \( \mu_2 \) is occasionally entitled as spectral gap, however, in this section we are interested to explain \( \mu_2 \) in context of environmental flux dynamics. As we discussed before, the actual coefficient matrix of the diffusion equation is \(-D_f L\), and \( L \)'s eigenvalues are \( \mu_i \). According to the properties discussed in the method section, the coefficient matrix of the diffusion equation has one 0 eigenvalue and all the other eigenvalues are negative. Therefore, the smallest non-zero eigenvalue of \( L \), corresponds to the largest non-zero
eigenvalue of $-D_fL$, which shows the slowest exponential decay over time. Therefore, if $\mu_2$ is small then this decay takes a very long time, resulting in slow diffusion of environmental flux. On the other hand, larger $\mu_2$ the diffusion occurs quickly on the river network topology. Therefore, larger nodal connectivity exhibits higher diffusive nature of flow on the network from a dynamical point of view.

![Figure 5.6 Synchronizability example on two hypothetical river networks: maximum branching (a), minimum branching (b).](image)

Another inherent dynamical behavior of individual node can also be observed from the spectrum of the Laplacian matrix ($L$). Here, we want to show how the internal nodes of a river network exhibits synchronization properties from the dynamic point of view. Conceptually, synchronization is how nodewise environmental flux grows or shrinks over time. The metric of this synchronization properties is known as synchronizability. In this chapter we propose a metric to quantify the synchronization of internal nodes based on the spectrum of Laplacian matrix ($L$). We have found that the time normalized ratio $\mu_N/\mu_2$ can provide us with a useful comprehensive
interpretation of the synchronization of internal nodes from a dynamic point of view. It relies on the assumption that the properties of nodes of the river network are homogeneous and that they are linearly connected.

Figure 5.6a-b depicts two hypothetical networks with the same number of nodes but different in channel arrangement. Each network has 4 source nodes, 3 internal nodes and one outlet. Figure 5.6 shows the response of internal nodes under unit concentration applied to the source nodes. In other words, environmental flux grows or shrinks over time at the internal nodes. Visually we can see that network 1 exhibits more temporal similarities than network 2. The proposed metric shows network 1 has higher synchronizability than network 2, which provides us a comprehensive intuition about the role of branching structure on the environmental flux dynamics on river networks.

**Figure 5.7** Response of hydro climatic variable on synchronizability of river network: climate aridity index (a), mean annual rainfall (b) and Response of c-value (c).

In this section, we tested the relationship between synchronizability with hydro-climatic descriptors and branching structure of river networks. To understand the response of hydro-climatic variable on the synchronizability, we have investigated them as a function of climate
aridity index \( (E_p/P) \) and the mean annual rainfall \( (R) \) on the landscape. This correlation is tested based on the 67 river networks extracted from the natural basin across the different climatic regions of the United States. Based on the figure 5.7a-b we can argue that the river network from the humid climate has less synchronizability than the relatively arid climate. It is important to point out that a significant power law relationship \( (slope \approx 0.35) \) was observed from our available data with \( R^2 \) value of \( \approx 0.2 \), however, \( t \) test results show a significant correlation with a 95% confidence interval. Similar correlation was also observed with the mean annual rainfall, which again strengthens the argument that basin with higher rainfall exhibits less synchronizability than the lower rainfall.

Furthermore, we have also investigated the role of the branching structure on the synchronizability of the river network. The figure 5.7c exhibit the consistent relation between the synchronizability of river network with the c-value. As c-value increases the synchronizability of river network exponentially decreases with slope of \( \approx -0.17 \). Zanardo et al., (2013) showed a statistical dependency between hydro-climatic variable and the c-value, that helps us to provide a physical basis to argue that a large number of low-order channels with respect to the number of high-order channels river networks from the humid climate exhibits less synchronizability than the arid climate river network, which has a small number of low-order channels with respect to the number of high-order channels.
CHAPTER 6: TOPOLOGIC AND GEOMETRIC COMPLEXITY OF RIVER NETWORK

Complexity measurements are central to the research of networks in many areas (Albert & Barabasi, 2002; Boccaletti et al., 2006). Seminal approach to the quantification of network complexity using the concept of Shannon’s information theory has been extensively used (Ranjbar et al., 2018, 2020; Shannon & Weaver, 1949). According to Shannon’s theory, entropy describes probabilistic information about the network’s nodal structure or connection. The complexity of the network can be measured quantitatively using the notion of entropy on network metric which is able to describe the organization of a network (Bonchev & Buck, 2005). We introduced the principle of approximate entropy in the sense of network organization and extended it to the network metric which can capture the topological and the geometric structure (West et al., 2012). While approximate entropy is typically used in the time series, some literature proposed that it can also be used in the naturally ordered network metric (West et al., 2012). Usually, in the case of complex network concepts, the degree sequence is most popular. However, the question is how best to assign an entropy to a degree sequence for a river network? A simple application of entropy to a degree sequence, however, may not provide useful information due to the natural ordering system (self-organization etc.) in the case of a river network which could be overcome by a more sophisticated approach by assigning some combinatorial description and modification of the degree sequence. For this purpose, we adopt and calculate the betweenness centrality due to the fact that it is a path-based connectivity measure of the network, which is useful to describe the combinatorial organization of the network (Sarker et al., 2019). Betweenness centrality ($BC$) is a path based global measure of network flow passing through the individual node within the
network. It essentially tells us how the flow on the network is disrupted by the nodes. In the case of river network, $BC$ can be computed for the nodes along the longest channel to capture the influence of the nodes on connectivity along the longest channel due to the fact that the river flows from the upstream to the downstream and longest channel may account for most of the dynamical behavior of a basin. In other words, $BC$ along the longest channel has some information about the integrity of the network. Conventional $BC$ is useful to understand the disruption of the network but in this chapter, we have modified $BC$ in such a way that it can capture the topology and geometry of the river network at the same time. In addition to computing number of paths, we can also consider the distance of the paths passes through the nodes; if every link length is considered unit length, then, the $BC$ can be considered as a topologic organizational metric of the river network. On the other hand, when unit link length is replaced by Euclidean link length then $BC$ can be considered as a geometric organizational metric of river network. The $BC$ score is normalized by the total score of the nodes along the longest channel to justify a natural ordering system of the river network.

6.1 Betweenness centrality as Combinatorial Organization of River Network

Commonly, $BC$ is the number of shortest paths captured by nodes. This number of shortest paths captured by nodes ($BC$) along the main channel from the outlet can be considered as a combinatorial description and it has a natural order. In the figure 6.1, $BC$ of nodes 20, 19, 17, 15, 11, 7, 5, 2 can be considered as a natural ordering system for this river network topology. This $BC$ can be simply computed based on a directed river network. For example, node 11 has $u/s$ node-set $u$ and $d/s$ node-set $d$, where $u = \{1,2,3,4,5,6,7,8,9,10\}$ and $d = \{15,17,19,20\}$. So, the $BC$ (number of shortest paths captured) of node 11: $n_u * n_d = 10 \times 4 = 40$. This 40 is the
number of paths that passes through the node 11, however, these paths could be different in lengths. To incorporate the lengths of these paths, in this chapter we have modified the BC score based on the total path length captured by the nodes along the longest channel from the outlet. Following that, we can calculate the modified BC (length of the shortest path captured) of node 11:

\[ \text{modified BC} = \sum d(u,v) = \sum d(1,15), d(2,15), d(3,15), d(4,15), d(5,15), d(6,15), d(7,15), d(8,15), d(9,15), d(10,15), d(1,17), d(2,17), d(3,17), d(4,17), d(5,17), d(6,17), d(7,17), d(8,17), d(9,17), d(10,17), d(1,19), d(2,19), d(3,19), d(4,19), d(5,19), d(6,19), d(7,19), d(8,19), d(9,19), d(10,19), d(1,20), d(2,20), d(3,20), d(4,20), d(5,20), d(6,20), d(7,20), d(8,20), d(9,20), d(10,20) = 188. \]

![Figure 6.1](image_url) *Figure 6.1 BC computation for a hypothetical network. The combinatorial description is presented along with the modified BC. Different color indicates the different sub-tree network for the different nodes along the main channel from the outlet. The source node, internal nodes, and outlet are also presented with a distinct color.*
If every link is different in terms of length, the total number of path lengths captured by nodes along the longest channel will have the information of geometry. Here, we are interested in applying the notion of approximate entropy on this proposed modified $BC$ score to measure the network complexity.

Besides the hypothetical example, we can see from the figure 6.2 that, there are 16 nodes along the longest channel of a natural river network, where different link color indicates different order (Horton, 1945). Conventional $BC$ for each node is presented in the figure along with the modified $BC$ which entitled as topological measure of organization (i.e. $BC_n$). On the other hand, geometric organizational metric is entitling as $BC_L$ based on the natural link length. To obtain geometrical information, we compute the $\Delta_L = BC_L - BC_n$. Another combinatorial approach for topological metric could be using the link order, where the total number of link order captured by nodes along the longest channel (i.e. $BC_O$) can be seen as a topological stream order metric. In addition, we can show $\Delta_O = BC_O - BC_n$ as stream order organizational metric to calculate stream order complexity.

While Freeman (Freeman, 1977) proposed $BC$ centrality metric to evaluating influence of any node on the overall network, we have modified $BC$ according to the equations (37)-(40) to capture nodal combinatorial information (topological, stream order, and geometric) on the overall network.

$$BC(v) = \sum_{s \neq v \neq t \in V} n_{st}(v)$$ (37)

$$BC_n(v) = \sum_{s \neq v \neq t \in V} nl_{st}(v)$$ (38)
\[ BC_L(v) = \sum_{s \neq v \neq t \in V} \{ L_{st}(v) - n_{st}(v) \} \] (39)

\[ BC_O(v) = \sum_{s \neq v \neq t \in V} \{ SO_{st}(v) - n_{st}(v) \} \] (40)

Where, \( v \) is the desired node along the longest channel for which we are going to compute the betweenness centrality. \( n_{st}(v) \) = number of paths captured by the node \( v \), \( n_{st}(v) \) = number of links on the paths captured by the node \( v \), \( L_{st}(v) \) = length (m) captured by the node \( v \) and \( SO_{st}(v) \) = sum of the link order captured by the node \( v \) on the paths captured by the node \( v \).

Figure 6.2 Proposed nodewise topological, geometrical and the stream order metric (i.e. \( BC_n, BC_L \) and \( BC_O \) score) presented as an example of a natural river network. Different color indicates different channel order (Horton, 1945).
6.2 Approximate Entropy

Approximate entropy is a statistical technique used to quantify the unpredictability of fluctuations over a data series. When moment statistics such as mean, variance are not useful to distinguish between the series then the concept of entropy can provide useful information. While the calculation of entropy involves a large amount of data, S. M. Pincus, (1991) developed approximate entropy to deal with these limitations by modifying an exact regularity statistic. Although it was initially developed for the study of medical data, its applications later expanded to other fields (S. Pincus & Kalman, 2004; S. M. Pincus et al., 1991). Approximate Entropy and Sample Entropy are two algorithms to determine the regularity of data series based upon the existence of patterns (Delgado-Bonal & Marshak, 2019; Ranjbar et al., 2018). For example, we are interested to compute approximate entropy ($A_pEn$) of a data series $S$ containing $N$ data values, $S = \{x(1), x(2), x(3), ..., x(N)\}$. From this data, a series of vectors can be constructed as:

![Diagram](image)

**Figure 6.3** Details of algorithm to compute Approximate entropy on a $BC$ series for river network.
\[ X(1) = \{x(1), x(2), \ldots, x(m)\} \]  \hspace{1cm} (41)

\[ X(2) = \{x(2), x(3), \ldots, x(m+1)\} \] \hspace{1cm} (42)

\[ \ldots \ldots \ldots \ldots \ldots \] \hspace{1cm} (43)

\[ X(N-m+1) = \{x(N-m+1), x(N-m+2), \ldots, x(N)\} \] \hspace{1cm} (44)

The distance between two vectors \(X(i)\) and \(X(j)\) can be defined as the maximum difference in their respective corresponding elements.

\[ d(X(i), X(j)) = \max_{k=1,2,\ldots,m}(|X(i+k-1), X(j+k-1)|) \] \hspace{1cm} (45)

where, \(i = 1,2,\ldots,N-m+1\) and \(j = 1,2,\ldots,N-m+1\) and \(N\) is the number of data points in the series. For each vector \(X(i)\), a measure that describes the similarity between the vector \(X(i)\) and all other vectors \(X(j)\) \(j = 1,2,\ldots,N-m+1, j \neq i\) can be constructed as:

\[ C_i^m(r) = \frac{1}{(N-(m-1))} \sum_{j \neq i} \theta[r - d[X(i), X(j)]] \] \hspace{1cm} (46)

where,

\[ \theta[x] = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases} \] \hspace{1cm} (47)

The symbol \(r\) specifies a filtering level and related to the standard deviation of the series.

Finally, \(ApEn\) can be calculated by the following equation:

\[ ApEn(m, r) = \varnothing^m(r) - \varnothing^{m+1}(r) \] \hspace{1cm} (48)

where,
\[ \varphi^m(r) = \frac{1}{(N - (m - 1))} \sum ln[C^m_i(r)] \] (49)

The application of approximate entropy (ApEn) on BC data can be shown as the algorithm shown in figure 6.4.

6.3 Energy expenditure on the landscape

In this section, we have investigated how the entropy of the river network organization is related to the energy expenditure on the landscape. To answer this question, we have adopted an energy minimization principle to calculate the energy expenditure on the landscape (Abed-Elmdoust et al., 2016; Rodríguez-Iturbe & Rinaldo, 2001; Sarker et al., 2019). The energy dissipated in each link can be defined as \( LQ^\gamma \), where \( L \) and \( Q \) are its length and discharge, respectively and \( \gamma \) is known as the energy exponent, which characterizes the mechanics of erosional processes varies between 0 and 1 (Rodríguez-Iturbe & Rinaldo, 2001). We have computed \( \gamma \) from the slope area curve extracted from the natural high-resolution DEM (LiDAR) where channel network extracted based on the curvature-based method (Hooshyar et al., 2016, 2019; Keylock et al., 2021; A Singh et al., 2015). To compute the energy expenditure the discharge in each link is generated by using mean annual rainfall (\( mm \)) at each node and actual link length is considered. The overall energy expenditure is calculated using \( E = \sum_{i=1}^{N-1} L_iQ_i^\gamma \). Computed energy expenditure is normalized by using the rainfall amount and the drainage area.

6.4 Entropy and Energy from spectral point of view

In this section, the entropy of the network is computed based on the concept of betweenness centrality (BC). Besides the network centrality concept, the entropy can also be measured based
on the spectral properties of the network. To explore entropy with the help of spectral properties, we have adopted the von Neumann entropy of the network, which can be also used as a measure of network complexity. The von Neumann entropy of a network was introduced by Braunstein et al., (2006) and then further analyzed in many works (Anand et al., 2011; Anand & Bianconi, 2009; Passerini & Severini, 2008, 2009). Among them, Passerini and Severini investigated the use of the density matrix (normalized Laplacian) to compute the entropy of the network spectrum. We are interested to compute von Neumann entropy of a network system to understand how micro-states (nodes and links) are organized within that network system. Arrangement of the micro-states can be expressed by the Laplacian matrix ($L = D - A$). Normalized Laplacian matrix is commonly known as density matrix ($\rho = L/\text{tr}(L)$). Density matrices play a pivotal role in micro-states (nodes and links) and can be described by a density matrix $\rho$ and its von Neumann entropy is defined as:

$$S(\rho) = -\text{tr}(\rho \ln \rho)$$

(50)

where, $\text{tr}$ denotes the trace operator and $\ln$ denotes the matrix logarithm. The von Neumann entropy of $\rho$ can also be computed as the Shannon entropy of the spectrum of $\rho$, i.e.,

$$S(\rho) = -\sum_{i=1}^{N} \nu_i \ln \nu_i$$

(51)

where, $\nu_i = \nu_1, \nu_2, \ldots, \nu_N$ are the eigenvalue spectrum of the density matrix. Corresponding graph energy ($GE$) also can be obtained from the spectrum of the adjacency matrix ($B$), where, $GE = \sum_{i=1}^{N} |\lambda_i|$ and $\lambda_i = \lambda_1, \lambda_2, \ldots, \lambda_N$ are the spectrum of adjacency matrix (Ivan Gutman & Wagner, 2012). The graph energy is normalized by the number of nodes in the network to make it size invariant. To compute spectral energy and entropy we have ignored the flow
direction to obtain the real eigenvalue and due to the fact that spectrum only indicates the structure of the network.

6.5 Geomorphological descriptors

In order to validate our proposed complexity metrics, we have adopted commonly used geomorphological descriptors such as width functions. As discussed in chapter 5, a width function represents the number of nodes in the network which are positioned as a function of distance from the outlet. The number of incremental concentrations at an outlet as a function of time can be seen as a proxy of width function under the condition of unit rainfall on every pixel on a network with unit flow velocity. In addition, network geometry information is also encoded in this function as a distance. Furthermore, it can be seen as a topological signal under the assumption of unit link length for the entire river network. Approximate entropy is computed on the topologic and geometric signal to obtain topological and geometric complexity. Finally, we also adopted the Tokunaga self-similarity model to validate our proposed stream order complexity.
6.6 Network Metric to Capture Complexity

Figure 6.4 Proposed modified $BC$ score for a natural DEM (a), extracted river network (b), topological vs geometric $BC$ signal (c), and topological vs stream order $BC$ signal (d).

The figure 6.4 shows a natural DEM and its corresponding extracted river network along with the stream order. Extracted river networks can be seen as two-dimensional tree network topology which is the physical arrangement of nodes and links. The longest channel is also shown and there are 28 nodes on the longest channel where two or more channels meet to create confluence (or bifurcate at confluences). Nodes are numbering from the outlet to the channel head. We have proposed three series to describe topology, stream order arrangement and geometry of the two-dimensional river network depicted in the figure 6.4. Figure 6.4(c) depicts how nodewise topological and geometric information varies along the longest channel from the outlet. It is
obvious that there is some difference between the two series due to the fact that every link length is not the same. Similarly, figure 6.4(d) depicts how the information of topology and stream order varies on the longest channel. This information is the combinatorial information of topology and the geometry of the natural channel. In this section, we have try to distinguish the pattern of those series statistically, but not using the conventional moment statistics or rank order statistics. Here we use the entropy-based method to investigate the regularity, which indicates the complexity of the river network. Figure 6.5 shows computed topological and geometric signals of a natural river network under the assumption of unit rainfall over the entire network pixel and unit flow velocity.

![Figure 6.5 Topological and geometric signal observed for a natural river network shown in figure 6.4](image)

Figure 6.5 Topological and geometric signal observed for a natural river network shown in figure 6.4
6.7 Types of Complexity and their comparison

Figure 6.6 provides us a notion of relative comparison of topology, stream order, and geometric organizational complexity. Here, we investigate the topological complexity, stream order complexity and the geometric complexity based on the value obtained from the approximate entropy on the series computed for nodal $BC_n$, $BC_o$ and $BC_L$ score along the longest channel respectively (see methods section for details). Figure 6.6 (a-d) depicts a good correlation between topological complexity and the stream order complexity. This information helps us to infer that the combinatorial organization of links and combinatorial information of stream order carries similar types of information around the nodes and they both are part of river network topologic organization and their approximate entropy could be a metric of river network complexity. The correlation is tested based on the 67 river networks extracted from the natural basin across different climatic regions of the United States. A $t$ test was performed on the data, which shows significant at 95% confidence interval (i.e. $P$-value < 0.05). The correlation between the proposed stream order complexity computed based on $BC$ and the Tokunaga c-value further strengthen the implication to the complexity in branching patterns of the river network. Specifically, it suggests that higher c-value exhibits higher stream order complexity and vice versa. It is also depicted that the geometric complexity is slightly lower than the topologic and the stream order complexity, which is statistically confirmed by the $t$ test on the same networks obtained from the natural basins; significant at 95% confidence interval on linear correlation. The physical reason for lower geometric complexity than topologic can be explained using the notion of thermodynamics which is formally stated in the next section.
Figure 6.6 Correlation test between the stream order and the topological complexity (a), stream order complexity and c-value (b), proposed topological and geometric complexity (c) and topological vs geometric complexity computed by using geomorphic signal on the watershed (d).

6.8 Topologic vs Geometric Complexity

In order to explain why geometric complexity is lower than the topologic complexity, we need to understand the concept of entropy. To understand the concept of entropy from a physical point of view we need to recall the first law of thermodynamics, which essentially states that energy cannot be created or destroyed but it can be converted from one form to another form of energy. Energy transfers from higher energy to lower energy; this phenomenon is commonly known as energy dissipation (Kleidon et al., 2013). For example, a system can convert heat energy to mechanical energy, which is called heat energy dissipation. To know the amount of energy conversion from one form to another form we need to consider the second law of thermodynamics,
that indirectly tells us how much energy a system can convert. If a system has higher entropy, then it has more uncertainty to convert the energy. In other words, to describe that system more information is needed. The perfectly regular system which follows a specific arrangement has lower entropy, on the other hand, the irregular system has higher entropy because more information needed to describe the system. For example, a regular crystal has a lower entropy than the glass of water. Similarly, in information theory, if more data points are needed to predict the next term in any series of a system, then that system is said to have higher entropy i.e. higher complexity. The second law of thermodynamics also states that, as time proceeds the entropy increases that means the ability to convert energy is slowed down as a function of time. Physically, entropy is considered as the property of macro-state (system as a whole) not the micro-state (elements of system) of a system. It indicates how much disorganized micro variable/information inside a macro variable/information. In the context of a river network, nodes and links arrangement can be seen as micro information and the entire network can be seen as macro information of the river network system. In thermodynamics, the change in free energy is inversely related to the change of entropy.

From the evolution of a river network, we know as time proceeds, the river network is reorganized in such a way that it minimizes the energy dissipation on the landscape. The amount of energy dissipation is the change in the free energy of the system. In other words, the river network system itself wants to change its arrangement in such a way so that it can maximize the entropy of the network by minimizing the work done and minimizing the free energy of the network. More specifically, a river network wants to be more complex to achieve a steady state if there is no other external force. This phenomenon reminds us of the energy minimization principle or entropy maximization principle of the network system. From the energy minimization principle,
we know that total topologic energy dissipation is $E = LQ^r$. This energy is the change in the free energy of the river network system. In other words, a river network system is able to convert more energy from one form to another form. From the energy dissipation equation, we can see that, if the channel length is higher than energy conversion will be higher. That means one should expect lower entropy of the system. This argument strengthens our results and justifies the proposed hypothesis that geometric complexity is lower than the topologic complexity due to the fact that $BC_L$ convey the information about the actual link lengths or the path lengths around the nodes along the main channel. Further, we have investigated the influence of climate on our proposed river network complexity in the following section.

### 6.9 Response of Hydro-climatic Variable on the River Network Complexity

To understand the response of hydro-climatic variable on the river network complexity (topological and geometric) we have investigated them as a function of climate aridity index ($E_p/P$) and the mean annual rainfall ($R$) on the landscape. Based on the figure 6.7, we can argue that the river networks from the humid climate have more complexity (entropy) than the relatively arid climate. It is important to point out that a poor logarithmic negative correlation ($slope \approx -0.06$) was observed from our available data with $R^2$ value of $\sim 0.15 - 0.2$, however, $t$ test results show a significant correlation with a 95% confidence interval. Similarly, a positive correlation was also observed with the mean annual rainfall, which again strengthens the argument that a basin with higher rainfall exhibits more complex topology and geometry than the lower rainfall.
Figure 6.7 Influence of hydro-climatic variable, climate aridity index (top panel) and mean annual rainfall (bottom panel) on proposed topological and geometric complexity.

Figure 6.8 Role of the hydro-climatic variable on river network organizational complexity and their corresponding energy dissipation on the landscape.
Figure 6.8 exhibits a reverse relation of river network energy as a function of hydro-climatic variables (compared to complexity shown in Fig 6.7), indicating humid basin exhibits a more complex river network than the arid basin, however, the river network from the humid basin has dissipated less amount of energy compared to the river network from the arid climate. These results provide a physical basis to argue that rivers from a humid climate have less tendency to change their pattern than the rivers from arid climate, further indicating a quantitative insight on landscape dynamics under variable climate.

6.10 Entropy and Energy from Spectral Point of View

Besides the energy dissipation by the river on the landscape, network energy also can be calculated from the two-dimensional river network topology using spectral graph techniques. In network science, many studies have focussed on the entropy and energy of a network. For example, Anand et. al. investigated the relationship between Shannon entropy and the von Neumann entropy in the context of network science (Anand et al., 2011; Anand & Bianconi, 2009). In this section, we explore the behavior of energy and the von Neumann entropy as a function of a hydro-climatic variable. Although we have conducted our study based on the directed river network, the spectrum of the network is computed based on the undirected network, ignoring the flow direction. In addition, we have computed the real eigenvalue of the undirected network and found similar energy and entropy trends as a function of a hydro-climatic variable.
**Figure 6.9** Role of the hydro-climatic variable on spectral energy expenditure (graph energy) and entropy (von Neumann entropy) of the river network.
CHAPTER 7: GEOMETRIC ROUGHNESS OF THE NATURAL RIVER NETWORK UNDER CHANGING CLIMATE

In this chapter, we focus our attention on the geometric roughness of a river network. We define geometric roughness based on the slope of a power spectral density computed from the flux obtained at the outlet of a river network. To understand the roughness of river networks, we use the concept of fractals and their potential use to describe roughness. In addition to that, we also employ power spectral density to understand the roughness of river network geometry.

7.1 Tokunaga self-similar river networks

In this section, we concentrate on the fractal nature of trees due to the fact that river networks are a classical example of fractal characteristics (W. I. Newman et al., 1997; Tarboton, 1997; Tokunaga, 1978). Apart from seminal Horton-Strahler stream ordering system (Horton, 1932, 1945; Strahler, 1957), Tokunaga self-similarity is an alternative way to characterize the branching pattern of river networks (Keylock et al., 2020; Peckham & Gupta, 1999; Ranjbar et al., 2020; Tokunaga, 1978). It has advantages over the Horton-Strahler indexing system that suffers from limitations due to high order side branching which frequently occurs in natural river basins (Tarboton, 1997). In this consequence, the Tokunaga self-similarity is an extension of Horton-Strahler ordering system which takes into account the effect of side branching on river networks (W. I. Newman et al., 1997; Tarboton, 1997; Tokunaga, 1978). Tokunaga taxonomy details are presented in chapter 4. The parameter c-value indicates the connectivity of low-order channels with respect to the high-order channels and therefore it characterizes “featheredness” of river
networks (Zanardo et al., 2013). In this section, we discuss the use of c-value as a self-similar metric to describe river network structural organization as well as analogy to fractals.

Besides the Tokunaga self-similarity model there are many examples of self-similar series which can be extracted from landscapes and their processes (e.g. width function, drainage area etc). In this section, we use a width function series as a geometric descriptor of a basin assuming unit velocity of water flow over the entire river network. In addition, we perform power spectral density (using Fourier transform) and their behavior in the frequency domain.

7.2 Power spectral density of a signal

Power spectral density (PSD) is a frequency response measurement of the signal intensity or amplitude. In general, it provides a standard method to capture how the energy in a signal is distributed across different scales (Gardner & Robinson, 1989; Keylock et al., 2014; Ranjbar & Singh, 2020; Sandhu et al., 2016; Arvind Singh et al., 2011, 2012; Stoica et al., 2005; Stull, 2012). The PSD $S(\omega)$ of a discrete signal $g(t)$ can be computed as the average magnitude of the Fourier transform squared, over a time or space interval and expressed as follows:

$$S(\omega) = \frac{1}{2\pi} \sum_{t_1}^{t_2} g(t)e^{-i\omega t} = \frac{\hat{g}(\omega)\overline{\hat{g}}(\omega)}{2\pi}$$

(52)

where, $\hat{g}(\omega)$ is the discrete Fourier transform of $g(t)$ and $\overline{\hat{g}}(\omega)$ is its complex conjugate, and $\omega$ is the wave number. We analyzed this PSD in the log-log domain across the frequency $\omega$ in the following form:

$$S(\omega) \sim \frac{1}{\omega^\beta}$$

(53)
where, $\beta$ is the power-law exponent of the PSD; we referred to this as a measure of geometric roughness of river network and is computed using the slope of the linear regression fitted to the estimated PSD plotted on log-log scales (Pilgram & Kaplan, 1998).

7.3 Hydro-climatic forcings

In this section, we explore the relation between common hydro-climatic variables and the exponent $\beta$ i.e., geometric roughness of the natural river networks. In order to investigate the response of hydro-climatic forcings, we compute climate aridity index for selected basins across the United States (see details in chapter 2). Climate aridity index is defined as the ratio of mean annual potential evaporation ($E_p$) to precipitation ($P$) (Budyko et al., 1974; Henning & Flohn, 1977), reflects long-term climate of the watershed. Besides climate aridity index, we explore the response of mean annual rainfall and surface reflectance data on $\beta$. For surface reflectance data, we use corresponding images collected from United States Geological Survey (USGS) to calculate $NDVI$. The $NDVI$ is usually defined and computed as the normalized ratio of red ($R$) and near-infrared ($NIR$) reflectance of a sensor system that generally characterized the greenness of watershed vegetation (S. Tahsin et al., 2017; Subrina Tahsin et al., 2016, 2018, 2020, 2021). It is can be expressed as:

$$NDVI = \frac{NIR - R}{NIR + R}$$  \hspace{1cm} (54)

Here, surface reflectance products are atmospherically corrected and after image acquisition, all georeferenced images were clipped to the spatial extent of the corresponding
watershed area (see Subrina Tahsin et al., (2018) for details). Resampling and projection were done using local coordinates that were implemented utilizing ArcGIS.

7.4 Metrics in context of dynamics and vulnerability

In this section, we investigate the influence of dynamics and connectivity on the proposed geometric roughness of the river network. River networks are the backbone of landscapes; they are not static as they appear and reorganize by shifting and moving across landscapes over a larger period of time. In other words, river networks are constantly changing their structure to collect and transport water, sediment, organic matter, and nutrients from upland mountain regions to the oceans until they establish equilibrium between tectonic uplift and river erosion (Willett et al., 2014). To achieve this equilibrium, river networks minimize their energy expenditure on the landscape and creates dendritic and complex riverine networks and ecosystem (Abed-Elmdoust et al., 2016; Rinaldo et al., 1992; Rodríguez-Iturbe & Rinaldo, 2001). Although naturally emerging inhomogeneous networks are difficult to control, some studies suggested the metrics to capture controllability of networks under external influences (Liu et al., 2011; Yuan et al., 2013). We use controllability metric based on structural controllability theory that was suggested by Yuan et al., (2013). To infer the role of network fragmentation under external influence on the proposed geometric roughness, we also implement the critical node framework proposed by Sarker et al., (2019). In this context, we referred to the power-law exponent of the remaining pairwise connectivity vs removed critical node as vulnerability (formally addressed in the chapter 4).

7.5 Inhomogeneity of river network

Tokunaga self-similarity model is a commonly used model to characterize branching structure of river networks (Tokunaga, 1978). This model could be analogous to the notion of
fractal dimensions in a way that it could capture some sort of roughness of natural river networks. However, the parameter c-value only captures topologic information and limits in characterizing geometry (Zanardo et al., 2013), although it is significantly dependent on the magnitude, frequency and duration of precipitation on watersheds which is a key climatic parameter that dictates the shape of the watershed (Zanardo et al., 2013). Figure 7.1a-c suggests that c-value significantly decreases as climate aridity index increases. In other words, the humid basins have more inhomogeneous branching structure than the dry basins (see also figure 7.1b). It is also observed that higher value of NDVI exhibits higher c-value (figure 7.1c). It is important to point out that a significant logarithmic correlation ($slope = \sim 1.3 - 1.4$) was observed from our available data with $R^2$ value of $\sim 0.2 - 0.3$, however, $t$ test results show a significant correlation with a 95% confidence interval. Similarly, linear correlation ($slope = \sim 2.07$) was also observed in case of c-value vs NDVI plot with $R^2$ value of $\sim 0.2$, with significance $t$ test on our available data. Our results not only further strengthen Zanardo et al., (2013) argument, but also show the presence of vegetation has a significant impact on river network inhomogeneous branching structure.

**Figure 7.1** River network inhomogeneity (c-value) to the response of climate aridity index (a), Mean annual precipitation (b), and NDVI (c).
7.6 Roughness of river network

One of the commonly used metrics to capture the river network geometry is the width function, which basically is the one-dimensional representation of the two-dimensional geometry of the network (Abed-Elmdoust et al., 2016; Lashermes & Foufoula-Georgiou, 2007; Marani et al., 1994; Ranjbar et al., 2020). In this study, we use width function to achieve a comprehensive understanding of geometric roughness of natural river networks. As a general rule, the width function can be computed by the number of channelized pixels that have the same distance from the basin outlet. Although the concept of distances is usually measured along the longest flow path of the channel network, however, in this section we have computed time as a proxy of distance assuming the unit velocity over the entire river network (see details in the chapter 5). In addition, we have computed the corresponding flux concentration at outlet assuming unit rainfall over the entire watershed. Therefore, the resulting signal, which is a time series, mimics the width function and can be further used to investigate the fractal nature of the basin and hence the geometric roughness of the river network. We have analyzed the obtained time series to calculate the distribution of energy across frequencies or wavenumbers using power spectral density. Note, we refer to geometric roughness based on the exponent of power-law of the PSD across the frequency (Pilgram & Kaplan, 1998; Witt & Malamud, 2013) A higher PSD slope relates to more roughness in channel geometry.

Figure 7.2a shows an example of a natural DEM along with extracted river network. Here, unit space discretization is presented to show the computation of the incremental concentration at the outlet assuming unit uniform rainfall based on channelized pixel representing distance from the outlet as a function of time (see details in chapter 5). Figure 7.2b shows the computed time series of incremental concentration (a proxy of width function which captures the geometry of the
extracted river network). In addition, figure 7.2c presents the power PSD where slope indicates the strength of variation across frequency.

**Figure 7.2** Geometric time series signal generation on river network extracted from the natural DEM (a); different color indicates different channel order (Horton, 1945), analogy of computed time series with the width function (b), and roughness computation using PSD (c).

To investigate the response of magnitude of PSD slopes and controls on this metric we compare c-value with slope of PSDs. Figure 7.3a exhibits the linear positive correlation ($R^2 = \sim 0.15$) between c-value (inhomogeneity of branching structure) and PSD slopes (geometric roughness of river network). In other words, higher inhomogeneity shows higher geometric roughness and vice versa. This hypothesis is further confirmed by the $t$ test results that show significance with a 95% confidence interval on our available data.
In addition to the c-value, the proposed geometric roughness shows a consistent behavior with climate and vegetation. Based on these results, we can conclude that humid climate and vegetation have a significant impact on geometric roughness. Specifically, excessive rainfall, humidity and vegetation increase the roughness of river network geometry and vice versa.

Figure 7.3 Response of geometric roughness under the influence of c-value (a), climate aridity index (b), mean annual precipitation (c), and NDVI (d).

7.7 Dynamics and vulnerability of river network

One interesting question to ask is how likely a river network can change (reorganize) over time and how can we quantify this change. The idea of reorganization in a river network can be
inferred by the energy expenditure (Rodriguez-Iturbe & Rinaldo, 2001). The detailed description of this principle is described based on the optimal channel network (OCN) model (see chapter 2). This model is based on the minimization of the dissipated energy by river network on the landscape. This energy is a function of length, discharge carried by the channel and slope area relationship of basin. In addition, several studies have found that the dynamics and reorganization are related to the intensity of the rainfall (Abed-Elmdoust et al., 2016). These conclusions drive us to investigate whether we can correlate the inherent dynamics of the river network with the geometric roughness. To explore this, we have computed the energy dissipation per unit area of natural watershed and plot them against the geometric roughness ($\beta$) of the corresponding watershed. Strikingly, we observe that beta increases logarithmically with the energy expenditure per unit area (figure 7.4a). Significance of this observation is further tested using t test on our available data. Another way to understand the dynamics can be achieved by measuring controllability of the river network. To test our hypothesis, we compute controllability and plot it against $\beta$. The visual observation helps us to confirm our conclusion on the dynamic behavior of the river network. In other words, significant linear correlation based on our available data suggest that higher geometric roughness has higher controllability and vice versa (figure 7.4b). In addition, Figure 7.4c exhibits that vulnerability (quantified based on pairwise connectivity) linearly decreases ($slope = \sim - 0.62$) with the increase of geometric roughness and vice versa. Note, although our data show $R^2$ value of $\sim 0.13$, however, $t$ test results show a significant correlation with a 95% confidence interval.
Figure 7.4 Correlation test between geometric roughness to the energy expenditure on landscape (a), controllability (b), and vulnerability of river networks (c).
CHAPTER 8: SUMMARY AND CONCLUSIONS

Natural drainage networks emerge as an interplay between several external and internal factors such as climate, tectonics, and vegetation. Climate has been identified as one of the most significant controls on landscape evolution. In Chapter 1, we provide a detailed discussion on the need of studying river networks under changing external forcings (e.g. climate). In chapter 2, we have discussed the techniques to generate synthetic river networks as well as natural river network extraction.

In chapter 3, we identified critical nodes on synthetic (obtained from optimal channel network model) and natural river networks to understand the vulnerability of river networks under potential external disruptions. We compared the identified critical nodes with the most central nodes. We also investigated the characteristics of sub-basins induced by the fragmentation due to removal of central and the critical nodes from the network for both synthetic and natural river networks through their network disruption metrics.

In chapter 4, we discussed the spectral properties of these networks in the context of controllability and heterogeneity. To achieve this, we explore the range of zero eigenvalues in the eigenvalue spectrum using the notion of multiplicity of zero eigenvalues. In addition, we also explore algebraic multiplicity of zero eigenvalues based on the concept of maximum matching. We show that, for undirected network algebraic multiplicity of zero eigenvalues is equal to the number of driver nodes in the network which is a quantification of the ability to control the dynamics of a river network. It is seen that; networks with higher side branching are more controllable (i.e. have
higher ability to control the dynamics) than the lower side branching river networks and driver nodes avoid higher degree nodes on the directed network. A significant correlation was observed between the heterogeneity obtained from the node degree variance and c-value; these findings further suggested that the higher heterogeneity exhibits higher controllability and vice versa. In addition, we have observed higher controllability in case of river network obtained from the humid basin than the arid basin.

In chapter 5, a unified graph theoretic framework is proposed to understand the environmental flux dynamics on river networks. Proposed framework can be seen as an analog of an advection diffusion model that is able to reproduce the characteristics of the geometric structure of a drainage network and dynamics of hierarchical aggregation on river networks. Proposed framework suggested a metric (synchronizability) that may have potential for comprehensive understanding of delivery of environmental flux on the river network including under climatic fluctuations. Using natural river networks, we explore the response to hydro climatic variables on synchronizability; we observe lower synchronizability for humid climate, while comparatively higher synchronizability for arid climate.

In chapter 6, we further explore betweenness centrality (BC) in the context of topological and geometric organization of river networks. We use Approximate entropy (ApEn) on the modified BC (accounting channel length) array along the longest channel from the outlet to quantify the organizational complexity of the channel network. This analysis provides additional understanding on the complex nature of watershed. Furthermore, it is also demonstrated that the topologic and
geometric complexity comparison of natural watersheds can be related to the energy and entropy based on the well-established second law of thermodynamics.

Finally, in chapter 7, we hypothesized that the temporal signal of fluxes at an outlet can provide information about the network structure governed by the geometry. This framework is general enough to account for different types of natural river networks as well as their physical processes operated on the system. The study suggests that the notion of the fractal dimension can capture some sort of roughness of the shapes across scales in addition to the self-similarity properties. Tokunaga self-similarity model, characterizing river network inhomogeneity, can be explained in the context of stream order roughness. The notion of roughness can also be interpreted through the analysis of power spectral density of any signal, here width function. In addition, temporal signals of fluxes at outlets illustrate a prominent feature of roughness in their frequency domain that is referred to as geometric roughness of the river network. The proposed geometric roughness framework provides additional information to understanding the influence of climate and vegetation as well as dynamics and vulnerability on the geometry of the river network.

Interesting research directions for future work can be listed as below:

1. Explore the role of eigenvectors associated with the eigenvalue spectrum on synthetic and natural river networks.
2. The comparison between complexity of the river networks of natural versus urbanized areas can provide information on anthropogenic vs natural influences on riverine ecosystems.
REFERENCES


New York.


Research Letters, 46(6), 3216–3224.


