Noname manuscript No.
(will be inserted by the editor)

Random Self-similar Trees: Emergence of Scaling Laws

Yevgeniy Kovchegov · Ilya Zaliapin · Efi Foufoula-Georgiou

Received: date / Accepted: date

Abstract The hierarchical organization and emergence of scaling laws in complex systems – geophysical, biological, technological, and socio-economic – have been the topic of extensive research at the turn of the 20th century. Although significant progress has been achieved, the mathematical origin of and relation among scaling laws for different system attributes remain unsettled. Paradigmatic examples are the Gutenberg-Richter law of seismology and Horton's laws of geomorphology. We review the results that clarify the appearance, parameterization, and implications of scaling laws in hierarchical systems conceptualized by tree graphs. A recently formulated theory of random self-similar trees yields a suite of results on scaling laws for branch attributes, tree fractal dimension, power-law distributions of link attributes, and power-law relations between distinct attributes. Given the relevance of power laws to extreme events and hazards, our review informs related theoretical and modeling efforts and provides a framework for unified analysis in hierarchical complex systems.

Yevgeniy Kovchegov
Department of Mathematics, 368 Kidder Hall, Oregon State University, Corvallis, OR 97331
ORCID: 0000-0001-6332-9741
E-mail: kovchegy@math.oregonstate.edu

Ilya Zaliapin
Department of Mathematics and Statistics, 1664 N. Virginia st., University of Nevada, Reno, NV 89557
ORCID: 0000-0001-6257-0517
E-mail: zal@unr.edu

Efi Foufoula-Georgiou
Departments of Civil and Environmental Engineering and Earth System Science, 5400 Engineering Hall, University of California, Irvine, CA 92697
ORCID: 0000-0003-1078-231X
E-mail: efi@uci.edu

Keywords Geophysical hazards · Hierarchical system · Scaling · Self-similarity · Horton's laws · Tokunaga model · Hierarchical Branching Process

Article Highlights

- Theory of random self-similar trees provides a unifying framework for studying scaling laws in complex systems
- Hierarchical Branching Process explains power laws for system attributes, system fractal dimension, and other scalings
- A one-parameter critical Tokunaga model closely fits the key data and scalings of geomorphology

Declarations

Funding. This research was partially supported by National Science Foundation grants DMS-1412557 (YK), EAR-1723033 (IZ), EAR-2122191 (IZ), DMS-1839336 (EF-G), EAR-1811909 (EF-G), and ECCS-1839441(EF-G).

Conflicts of interest/Competing interests. The authors declare they have no conflict of interest and/or competing interests.

Availability of data and material. All data and material are available in the main text.

Code availability. The algorithms for the reviewed models are described in the text. The codes are available upon request.

Contents

1 Motivation

The emergence of extremes in complex natural systems – tectonic, hydrological, climatic, biological – has been a topic of extensive research, recognizing the catastrophic impact that hazards produced by these systems exert on population, economy and the environment. The complexity science, developed and proliferated at the turn of the 20th century, recognizes several fundamental traits shared by natural extremes and hazards of different origin: (i) They are generic in systems that have *hierarchical organization.* Notably, such a hierarchy *"need not be manifest in the object but may arise in the construction of a model"* (Badii and Politi, 1999). (ii) The hazard-generating systems exhibit *scalings*, often expressed as power-law distributions and/or power-law relations among system's attributes (Barenblatt, 1996). (iii) The scalings are often connected to *self-similarity* – a property of the system to retain its statistical properties after being zoomed in or out via a suitable transformation (Mandelbrot, 1982; Barenblatt, 1996). The results reviewed in this work originated in analysis of systems whose hierarchical organization is particularly evident, and is commonly represented by a *tree graph*, and whose scalings and self-similarity are well established empirically.

An example of such a system is the Earth lithosphere that generates *earthquakes.* The two staples of statistical seismology are the power-law distribution of earthquake moments, which is equivalent to the celebrated Gutenberg-Richter law of earthquake magnitudes (Gutenberg and Richter, 1954), and power-law temporal decay of event rate within aftershock series (Omori, 1894; Utsu et al., 1995). Multiple other physical and statistical scalings of earthquake attributes are summarized in (Ben-Zion, 2008, Table 2). Hierarchical representation of seismicity by branching processes is also well explored, starting from the pioneering works of Kagan (1973),Kagan and Knopoff (1976, 1981), and Vere-Jones (1976). A very popular Epidemic Type Aftershock Sequence (ETAS) model of earthquake dynamics (Ogata, 1998), is a Galton-Watson branching process with a power-law offspring distribution and space-time-magnitude marks (Saichev et al., 2005; Baró, 2020). A trajectory of this process is a tree graph, whose vertices represent individual earthquakes and edges – triggering processes. Alternative tree representations of seismicity are discussed in (Baiesi and Paczuski, 2004; Holliday et al., 2008; Zaliapin et al., 2008; Yoder et al., 2013; Zaliapin and Ben-Zion, 2013). More conceptually, the lithosphere can be thought as *"a hierarchy of blocks separated by boundary zones, with densely fractured nodes at junctions and intersections"* (Keilis-Borok, 2002). This hierarchy spans a wide range from the seven major tectonic plates of continental size to nearly 10^{25} grains of rocks. The earthquakes are produced by complex dynamics and interaction of these blocks (Burridge and Knopoff, 1967; Allegre et al., 1982; Gabrielov et al., 1990; Zaliapin et al., 2003; Soloviev and Ismail-Zadeh, 2003).

Another prime hierarchical system prone to natural hazards is geomorphological landscape evolution, which is associated with and in part driven by mass-movement processes like *sediment transport, rockfalls, debris flows*, and *landslides.* One of the system's fundamental outputs is the network of stream channels that spans the continental earth in the form of permanent river and delta networks and ephemeral drainage pathways that extend to the grain scales. This network is naturally linked to such hazards as *floods* (Gupta et al., 1994, 2007; Tessler et al., 2015) and *coastal and hillslope erosion* (Roering et al., 1999). Geomorphological networks are conventionally represented by trees (for converging river channels) or directed acyclic graphs (for diverging deltaic systems or braided rivers) (Sapozhnikov and Foufoula-Georgiou, 1996; Lashermes et al., 2007; Passalacqua et al., 2010; Tejedor et al., 2017, 2015a,b), and their scaling laws have been recognized since the groundbreaking work of Horton (Horton, 1945); see (Rodriguez-Iturbe and Rinaldo, 2001, and refs. therein).

Geomorphic hazards like *landslides, avalanches,* and *forest-fires* are characterized by scaling laws similar to those in seismicity (Malamud et al., 1998; Turcotte et al., 2002; Malamud et al., 2004a,b), and have been successfully examined within a hierarchical framework (Turcotte, 1999; Turcotte et al., 1999, 2002). Biological hazards, such as spread of *human, animal, and plant epidemics*, are naturally modeled by time-oriented trees where vertices represent infected subjects. The discussed phenomenology is relevant in other areas beyond hazard studies where hierarchical organization and related scalings have been reported. These areas include computer science (Flajolet et al., 1979; Drmota and Prodinger, 2006), statistical physics of fracture (Zaliapin et al., 2003; Davidsen et al., 2007; Herrmann and Roux, 2014), vascular analysis (Kassab, 2000), brain studies (Cassot et al., 2006), ecology (Grant et al., 2007), scaling of biomass in river streams (Barnes et al., 2007; Gangodagamage et al., 2007), fractal hydraulic conductivity (Neuman, 1990; Molz et al., 1997), and allometric scaling laws in biology (West et al., 1997; Turcotte et al., 1998).

The multitude of systems traditionally studied via a prism of tree representation and associated hierarchical dynamics calls for a unifying framework to address the following questions:

What is a self-similar tree?
How to model a self-similar tree?
How to test for self-similarity with limited data (in a single tree)?
What does self-similarity imply for the scalings of observed attributes?

This survey summarizes the currently available answers to these questions. We show that the key manifestation of self-similarity is *Horton's laws* that describe scaling of various tree attributes. A geomorphologic origin of the Horton's laws and the fact that tree representation of river networks is direct and intuitive affected our choice of examples and terminology. Our main results however have universal applicability and are formulated in generic graph-theoretic terms.

2 Introduction

In a pioneering study "*of streams and their drainage basins*," Robert E. Horton took the first steps toward exploring "*the problems of the development of land*

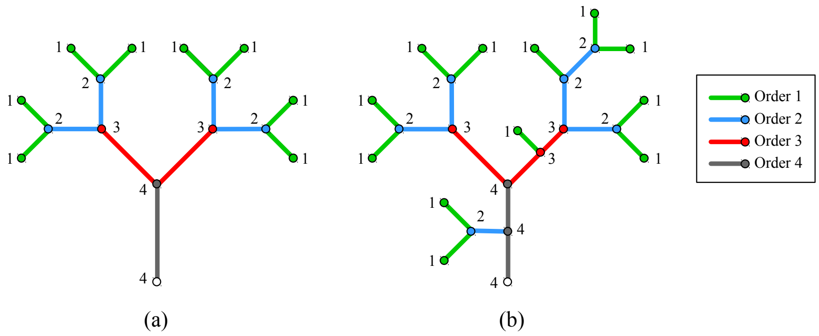

Fig. 1 Horton-Strahler orders in a binary tree. Different colors correspond to different orders of vertices and edges, as indicated in legend. (a) Perfect binary tree – orders are inversely proportional to vertex/edge depth. (b) General binary tree – orders represent vertex importance in the hierarchy, from leaves (smallest order) to the root (largest order).

forms, particularly drainage basins and their stream nets, along quantitative lines" (Horton, 1945). Starting from William Playfair's shrewd observation of "*a nice adjustment of [stream] declivities*" that produces "*system of valleys, communicating with one another*", Horton revealed deeper regularities in organization of river streams. He introduced the concept of stream order and formulated two fundamental laws of the composition of stream-drainage nets (Horton, 1945, p. 291). *The Law of Stream Numbers* postulates a geometric decay of the numbers N_K of streams of increasing order K, with the exponent R_B; see Sect. 3.3, Eq. (8). *The Law of Stream Lengths* postulates a geometric growth of the average length L_K of streams of increasing order K, with the exponent R_L. The initial Horton's ordering scheme has been later adjusted by Arthur Newell Strahler (1957) to its present form (which we call Horton-Strahler orders, Sect. 3.2), preserving the laws of stream numbers and lengths. Horton-Strahler orders are illustrated in Figs. 1,2.

During the 20th century, geometric dependence on the order has been documented for multiple physical and combinatorial stream attributes, including upstream area, magnitude (number of upstream sources), the total channel length, the longest stream length, link slope, mean annual discharge, and energy expenditure (Rodriguez-Iturbe and Rinaldo, 2001, and refs. therein). A geometric scaling of an arbitrary river stream attribute with order is called *Horton's law* and the respective geometric index is called *Horton exponent*; see Sect. 3.3, Eq. (9). The Horton's laws play an elemental role in studies of drainage networks. Being important in their own right, Horton's laws imply power law tails for the empirical frequencies of link attributes (Sect. 3.5) and power law relations between different attributes (Sect. 3.4). A celebrated example is the upstream contributing area $A_{(i)}$ of a link i and the length $\Lambda_{(i)}$ of the longest channel from the link i to the basin divide. Each of these attributes has a power-law empirical frequency,

$$\#\{i : A_{(i)} \geq a\} \propto a^{-\beta_A}, \quad \#\{i : \Lambda_{(i)} \geq l\} \propto l^{-\beta_\Lambda}, \tag{1}$$

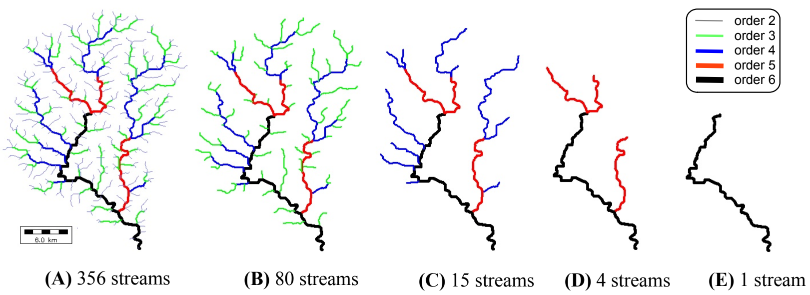

Fig. 2 Stream network of Beaver creek, Floyd County, KY. (A) Streams (branches) of orders $K = 2, \ldots, 6$ are shown by different colors (see legend on the right). Streams of order $K = 1$ (source streams) are not shown for visual convenience. Accordingly, this is the first Horton pruning of the network. (B)–(E) Consecutive Horton prunings of the river network; uses the same color code for branch orders as panel (A). The basin has order $K = 6$ since it is completely eliminated in 6 Horton prunings. The channel extraction is done using RiverTools software (`http://rivix.com`).

where i spans a large collection of links. The expression $a \propto b$ means that a is proportional to b, that is $a = \text{Const.} \times b$. Furthermore, the two quantities are related via a power-law $\Lambda_{(i)} \propto A_{(i)}^{\mathbf{h}}$, with $\mathbf{h} \approx 0.6$. This relation is known as the Hack's law (Hack, 1957); it is often reported for the area and the length of the longest stream in a basin (Rigon et al., 1996; Rodriguez-Iturbe and Rinaldo, 2001).

Intriguingly, the key parameters of these and other scaling laws can often be expressed via the Horton exponents R_B and R_L. For example, De Vries et al. (1994) and La Barbera and Rosso (1989) have shown, under some simplifying assumptions, that

$$\beta_A = 1 - \frac{\log R_L}{\log R_B} \quad \text{and} \quad \beta_\Lambda = \frac{\log R_B}{\log R_L} - 1, \tag{2}$$

and the fractal dimension $\mathbf{d}$ of a large tree is given by

$$\mathbf{d} = \frac{\log R_B}{\log R_L}. \tag{3}$$

This yields simple relations among the examined quantities:

$$\beta_A = 1 - \mathbf{h}, \quad \beta_\Lambda = \mathbf{d} - 1, \quad \text{and} \quad \mathbf{h} = \frac{1}{\mathbf{d}}. \tag{4}$$

Despite their recognized importance, the Horton's laws remain an empirical finding and their origin and apparent ubiquity remain unsettled. A first attempt at rigorous explanation of the Horton's laws and related scalings was made by Ronald L. Shreve (Shreve, 1966), who claimed that "*the statistical nature and remarkable generality of Horton's law of stream numbers suggest the speculation that the law of stream numbers arises from the statistics of a large number of randomly merging channels in somewhat the same fashion that the law of perfect gases arises from the statistics of a large number of randomly colliding* *gas molecules.*" To substantiate this claim, Shreve examined a "*topologically random population of channel networks, defined as a population within which all topologically distinct networks with given number of first-order streams are equally likely*". This model is equivalent to the critical binary Galton-Watson process with a given progeny (Burd et al., 2000; Pitman, 2006; Kovchegov and Zaliapin, 2020). Shreve's calculations imply that in this model the Horton's law of stream numbers hold with $R_B = 4$. Although not attempted by Shreve, it can be shown (Burd et al., 2000) that the law of stream lengths also holds here with $R_L = 2$ under assumption of constant or equally distributed edge lengths. This corresponds to

$$\beta_A = 1/2, \quad \beta_\Lambda = 1, \quad \mathbf{d} = 2, \quad \text{and} \quad \mathbf{h} = 1/2. \tag{5}$$

Albeit insightful and mathematically tractable, the random topology model deviates from the observations. This has been explicitly noted by De Vries et al. (1994) who examined the observed area scaling and have shown that $\beta_A \approx 0.45 \neq 0.5$, and by Peckham (1995) who has shown in a detailed analysis of river networks that $R_B \approx 4.5 \neq 4$. This called for developing alternative modeling approaches.

Versatile modeling efforts of the past decades have proven challenging to develop an approach that would be mathematically tractable and flexible enough to fit a range of observations. One end of the modeling spectrum is occupied by conceptual models, such as the Peano fractal basin (Rodriguez-Iturbe and Rinaldo, 2001, Sect. 2.4) that has already appeared in Horton's work under a different name (Horton, 1945, Fig. 25), or Scheidegger's lattice model (Scheidegger, 1967; Takayasu et al., 1988). These models provide an invaluable insight into the origin of the observed scalings; they however lack realistic dendritic patterns and values of scaling exponents because of a stiff geometry. On the other end are simulation approaches that are successful in generating visually appealing networks and closely fitting selected exponents, but can be analytically opaque. The Optimal Channel Network (OCN) model (Rinaldo et al., 1992; Rigon et al., 1993; Balister et al., 2018) is a particularly recognized simulation technique. Following the energy expenditure minimization principle, the model constructs random drainage basins on a planar lattice (or more general graph) and fits a variety of observed scaling laws. We refer to (Rodriguez-Iturbe and Rinaldo, 2001) for a comprehensive discussion of these and other models.

Despite the progress achieved by the modeling efforts of the 20th century, the following essential questions remain unanswered:

What are sufficient conditions for the Horton's laws?
What are the values of the Horton exponents? and
How are the Horton exponents for different stream attributes related to each other and to other basin parameters?

There is a consensus that Horton's laws are connected to the self-similar structure of a basin, which is generally understood as invariance of basin's statistical structure under changing the scale of analysis (zooming in or out).

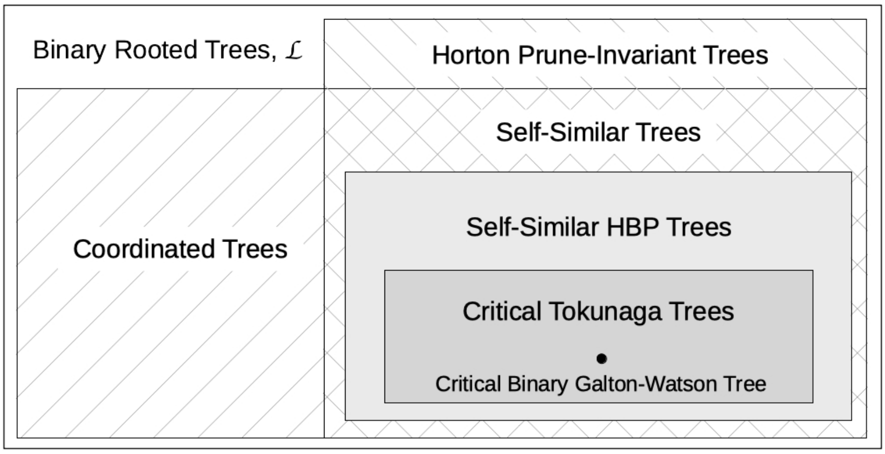

Fig. 3 Tree spaces examined in this survey: A Venn diagram. We work with binary rooted trees with edge lengths from the space $\mathcal{L}$ formally defined in Sect. 3.1. The **Self-Similar Trees** belong to the intersection of **Coordinated Trees** (Sect. 4.2) and **Horton Prune-Invariant Trees** (Sect. 4.3). The **Hierarchical Branching Process (HBP)** of Sect. 8 generates a particularly symmetric (infinite-dimensional) family of self-similar tree distributions, which we propose as a default model for applications. A one-parametric family of **Critical Tokunaga Trees** (Sect. 9) is a model proposed for river networks. It includes the celebrated **Critical Binary Galton-Watson Tree** with exponential edge lengths. The combinatorial part of this tree, being conditioned on the tree size, is equivalent to the Shreve's random topology model.

However, a consensus is still lacking about a suitable formal definition of tree self-similarity. Three alternative definitions have been studied: the Toeplitz property of the Tokunaga coefficients (Peckham, 1995; Newman et al., 1997); invariance of a distribution with respect to the Horton pruning (cutting source streams) in Galton-Watson trees (Burd et al., 2000); and statistical self-similarity of basin attributes (Gupta and Waymire, 1989; Peckham and Gupta, 1999). This triggers the questions:

How are the alternative definitions of tree self-similarity related? and
Is self-similarity (any version) sufficient for selected Horton's laws?

Answers to these questions require a rigorous toolbox that would go beyond the conventional heuristic approaches, which, albeit able to suggest quick routes to useful results, may lead to contradictions. For example, the classical works of De Vries et al. (1994), La Barbera and Rosso (1989) and many later studies adopted the assumption of *an ideal basin* that obeys the exact Horton's laws of stream numbers and lengths with $R_L < R_B$ (Rodriguez-Iturbe and Rinaldo, 2001, Sect. 2.5). The mean size (the number of links) of a basin of order K in this model is asymptotically given by (Rodriguez-Iturbe and Rinaldo, 2001, Eq. (2.91))

$$\frac{R_B^K - R_L^K}{R_B - R_L}. \tag{6}$$

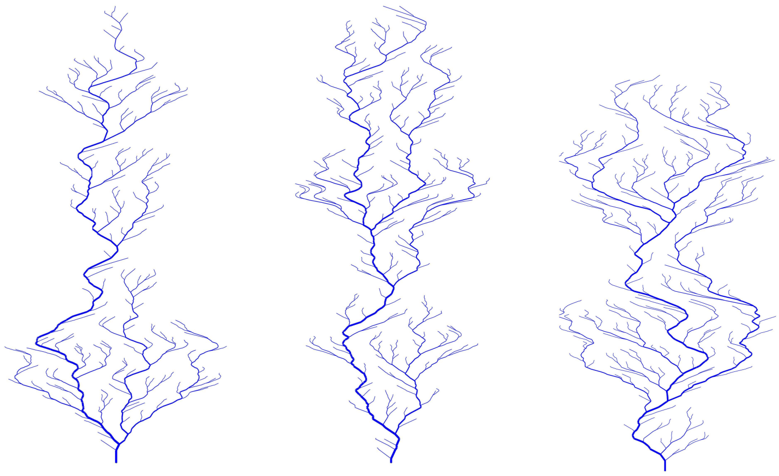

Fig. 4 Examples of HBP trees (Sect. 8). The trees are generated by the critical Tokunaga process with $c = 2.3$ and order $K = 5$ (Sect. 9). The line width is proportional to the contributing area approximated by $\sum \ell_i^2$, where the sum is taken over all upstream edges. The figure accurately represents the tree combinatorial structure; the edge lengths are scaled for a better planar embedding. We notice that the HBP generates trees with no planar embedding. The current figure uses an ad-hoc embedding; accordingly, the related purely geometric properties, such as junction angles or spacing between channels, are not a part of the model.

Observe that the mean basin size also equals twice the mean number N_1 of leaves, which is $2N_1 = 2R_B^{K-1}$, where the equality holds because of the exact Horton's law for stream numbers. Equating the two expressions for the mean basin size and considering the limit of large K one obtains $R_B = 2R_L$. If in addition one expects the fractal dimension of a basin to be $\mathbf{d} = 2$, then $R_B = R_L^2$, in accordance with (3). The two constraints lead to the unique solution $R_B = 4$, $R_L = 2$ that corresponds to the Shreve's random topology model (or, more precisely, to the critical binary Galton-Watson model, since the basin size is random). However, the initial assumption of the exact Horton's law does not hold in the Shreve's model. Moreover, the mean tree size in this model scales as $\frac{4}{3}R_B^{K-1}$; see Sect. 9, Eqs. (72),(73). In general, the ideal basin assumption is unrealistic in analysis of river networks as follows from the results of the present paper (see Appendix G). This prompts one to carefully validate the results yielded by heuristic approaches and calls for developing formal techniques. (We must notice that despite the mentioned contradiction, the key heuristic results obtained in the classical works are valid and are reproduced by formal techniques.)

We answer the questions posed above and develop a rigorous toolbox of working with branching structures capitalizing on a self-consistent mathematical theory of random self-similar trees recently developed by the authors (Kovchegov and Zaliapin, 2020; Kovchegov et al., 2021). The main goal of the current work is to present the theory in relation to the empirical and modeling constraints accumulated in the studies of river networks.

The theory builds on the self-similarity concepts developed by Horton (1945), Strahler (1957), Hack (1957), Shreve (1966, 1969), Tokunaga (1978), Mandelbrot (1982), Tarboton et al. (1988), La Barbera and Rosso (1989), Gupta and Waymire (1989), Tarboton et al. (1989), Leopold et al. (1992), Rinaldo et al. (1992), Rigon et al. (1993), Tarboton (1996), Maritan et al. (1996), Turcotte (1997), Gupta and Waymire (1998), and many others. Technically, its impetus is provided by the works of Peckham (1995), Newman et al. (1997), Turcotte et al. (1998), Peckham and Gupta (1999), Burd et al. (2000), Veitzer and Gupta (2000), McConnell and Gupta (2008).

2.1 How to Use This Survey

The survey has a three-fold goal: (i) to outline the key technical tools for examining scaling laws in trees; (ii) to present a number of scaling results for familiar tree attributes; and (iii) to propose the critical Tokunaga process as a model for river networks. To help a reader to promptly find the desired material, we briefly (and informally) summarize the main results of the work below in Sect. 2.2. Section 2.3 discusses the organization of material throughout the paper. Finally, Table 1 provides cross references to relevant equations and figures for each of the attribute and scaling exponent examined in the work.

2.2 The Main Results: A Brief Overview

Here we take a short stroll through the main staples of the theory of random self-similar trees and give an overview of the results presented in this work, before these are expanded in detail in the sections that follow.

We work with systems represented by binary tree graphs. The action takes place on the space $\mathcal{L}$ of all such trees, with a root and positive edge lengths. In essence, the survey examines a series of consecutively narrower subspaces of trees with consecutively stronger symmetries related to scaling laws – Fig. 3 illustrates the examined hierarchy.

The key element of the theory is the operation of *Horton pruning* $\mathcal{R}$ that removes (using the hydrological terminology) the source streams from a basin. The number of Horton prunings necessary to remove a link from a basin defines the link's Horton-Strahler order, hence the pruning name. An alternative (equivalent) counting approach to assigning the Horton-Strahler orders is described in Sect. 3.2. Figures 1, 2 illustrate Horton-Strahler orders and Horton pruning in simple binary trees and in the stream network of Beaver Creek, KY.

Self-similarity is defined as distributional invariance with respect to the Horton pruning in trees that satisfy the coordination property. *Coordination*

Table 1 Tree attributes and scaling exponents examined in the study, with corresponding equations and figures.

Attribute	Description	Eqns.	Figs.
$A_{(i)}$	Contributing area of vertex i	(1),(14),(99),(100)	15,16a
$\overline{A}_k$	Empirical average of the contributing areas of order-k branches	(17),(18)	
A_k	Mean contributing area of an order-k branch	(37),(38),(59),(76),(77)	5c,10c
$L_{[i]}$	Length of branch i	(105)	16b
$\overline{L}_k$	Empirical average of the lengths of order-k branches	(17),(18)	
L_k	Mean length of an order-k branch	(51),(60)	5d,9
$L_k^{\rm tot}$	Mean total channel length upstream of an order-k branch		5c
$\Lambda_{(i)}$	Length of the longest stream to the divide from vertex i (height)	(101),(102),(103)	15
$\overline{\Lambda}_k$	Empirical average of the heights of order-k branches	(17),(18)	
Λ_k	Mean length of the longest stream to the divide from an order-k branch (height)	(95),(96),(97)(98)	5d
$M_{(i)}$	Magnitude of vertex i		9
$\overline{M}_k$	Empirical average of the magnitudes of order-k branches	(17),(18)	
M_k	Mean magnitude of an order-k branch	(34),(36),(40),(42),(59),(72),(73),(122)	5b,10b
$N_k[T]$	Number of branches of order k in a tree T	(8),(10),(78),(79),(80),(82)	5a,10a
$\mathcal{N}_k[K]$	Mean number of branches of order k in a tree of order K	(35),(36),(43),(44),(59),(72),(73),(122),(123)	
$T_{i,j}$	Tokunaga coefficients	(20)	
T_k	Tokunaga sequence	(23),(24)	
$\overline{S}_k$	Empirical average of the number of edges (vertices) in order-k branches	(17),(18)	
S_k	Mean number of edges (or vertices) in an order-k branch	(33),(50),(60),(118)	5b,10d

Exponent	Description	Eqns.	Figs.
d	Fractal dimension of a tree	(3),(4),(5),(85),(87)	11a,12a,14a
h	Hack's exponent	(4),(5),(88),(96),(97),(98)	11b,12b,14b,16
R_B	Horton exponent for mean branch numbers $\mathcal{N}_k[K]$	(7),(8),(10),(13),(27),(43),(44),(59),(70),(78),(79),(120),(100),(102)	5a,10a,13a
R_M	Horton exponent for mean branch magnitudes M_k	(7),(40),(42),(70)	5b,10b
R_A	Horton exponent for mean branch contributing areas A_k	(7),(12),(48),(70),(100),(102)	5c,10c
R_L	Horton exponent for mean branch lengths L_k	(7),(12),(51),(60),(70),(85)	5d
R_S	Horton exponent for mean combinatorial branch lengths S_k	(7),(13),(50),(70),(100),(102)	5b,10d
R_Λ	Horton exponent for mean lengths Λ_k of the longest stream to the divide (heights)	(7),(102)	5d,15
β_A	Exponent of the power law exceedance frequency of branch contributing areas $A_{(i)}$	(1),(2),(4),(5),(14),(99),(100),(102),(103)	16a
β_Λ	Exponent of the power law exceedance frequency of lengths of the longest stream to the divide (heights) $\Lambda_{(i)}$	(1),(2),(4),(5),(15),(101),(102),(104)	

means that the (random) structure of a river basin is determined by its order. For example, a basin with outlet of order three and a sub-basin of order three within a basin of order nine have, statistically, the same structure. This assumption is in the heart of analyses based on the Horton-Strahler orders; it has been imposed, explicitly or implicitly, in the mainstream studies of river networks (Horton, 1945; Rodriguez-Iturbe and Rinaldo, 2001; Shreve, 1966; De Vries et al., 1994; Peckham, 1995; Peckham and Gupta, 1999; Tarboton, 1996). A distribution that satisfies the coordination property is called coordinated. The *Horton pruning* $\mathcal{R}$ is a natural model for the change of resolution in a river network (Fig. 2). Indeed, better observations lead to detecting smaller streams, which increases the basin order. Pruning a basin by order is roughly equivalent to decreasing the resolution of stream detection. The Horton prune-invariance requires that the statistical structure of trees remains the same after zooming in or out (Sect. 4.3, Eq. (22)).

Self-similar distributions are abundant on spaces of rooted trees. Each self-similar distribution corresponds to a unique sequence of non-negative *Tokunaga coefficients* T_k, $k \geq 1$, equal to the mean number of tributaries of order $K-k$ within a stream of order K, for any K (Sects. 4.1, 4.4). At the same time, an arbitrary sequence of Tokunaga coefficients T_k corresponds to an infinite number of self-similar distributions (with the same mean numbers of side-tributaries). The well-established models such as Peano basin or Shreve's topologically random model are self-similar in the above sense.

A foundational result (Sect. 7, Thm. 1 and Cor. 1) states that self-similarity implies Horton's law for the mean branch numbers $\mathcal{N}_K$ with exponent R_B and for the mean branch magnitudes M_K with exponent $R_M = R_B$. Furthermore, a conventional hydrological assumption of equally distributed link lengths (Rodriguez-Iturbe and Rinaldo, 2001; Tarboton et al., 1989) yields Horton's laws for the mean branch contributing areas A_K, mean number S_K of links within a stream, and mean total stream length L_K. The corresponding Horton exponents are uniquely expressed via T_k. Section 7 and Appendix C examine the Horton's laws for mean branch attributes in the most general situation, with and without the equally distributed link length assumption. The Horton's laws for the mean attributes imply Horton's laws (with the same exponents) for the random attributes obtained by averaging over branches of a given order in a single tree (Sect. 3.7).

Horton's laws imply power-law frequencies of stream attributes and power-law relations between different attributes (Sects. 3.4, 3.5). This includes the power law frequencies of (1) for link contributing areas and the length of the longest channel from a link to the basin divide, as well as for the length of a random stream in a basin. This, in turn, leads to the conventional expressions (3) and (4) that involve the Hack's law and basin fractal dimension. Table 1 lists the attributes and exponents examined in this work, with references to the related equations and figures.

A self-similar Hierarchical Branching Process (HBP) introduced in Sect. 8 generates a particularly symmetric distribution of trees for a given Tokunaga sequence T_k. The HBP trees obey the strongest forms of Horton's laws for

Table 2 Selected scaling exponents (1st column) in critical Tokunaga model expressed via the model parameter $c \geq 1$ (2nd column), fractal dimension $\mathbf{d}$ (3rd column), and Hack's exponent $\mathbf{h}$ (4th column). Columns 5-7 show the values of the exponents for $c = 2.0, 2.3, 2.5$. Column 8 shows the values estimated in the OCN model. Columns 9 summarizes estimations in the observed river networks. The agreement of the exponents of the critical Tokunaga model with $c = 2.3$ with those observed from real basins is noted.

Exponent	Expressed via			Critical Tokunaga model			OCN[†]	Real basins[‡]
	c	$\mathbf{d}$	$\mathbf{h}$	$c = 2.0^*$	$c = 2.3$	$c = 2.5$		
$R_B = R_M = R_A$	$2c$	$2^{\mathbf{d}/(\mathbf{d}-1)}$	$2^{1/(1-\mathbf{h})}$	4	4.6	5.0	4	4.1 – 4.8
$R_S = R_L$	c	$2^{1/(\mathbf{d}-1)}$	$2^{\mathbf{h}/(1-\mathbf{h})}$	2	2.3	2.5	2	2.1 – 2.7
$\mathbf{d} = \dfrac{\log R_B}{\log R_L}$	$\log_c(2c)$	$\mathbf{d}$	$\mathbf{h}^{-1}$	2	1.832	1.756	2	1.7 – 2.0
$\mathbf{h} = \dfrac{\log R_L}{\log R_B}$	$\log_{2c} c$	$\mathbf{d}^{-1}$	$\mathbf{h}$	0.5	0.546	0.569	0.57	0.5 – 0.6
β_A	$\log_{2c} 2$	$1 - \mathbf{d}^{-1}$	$1 - \mathbf{h}$	0.5	0.454	0.431	0.43	0.4 – 0.5
β_Λ	$\log_c 2$	$\mathbf{d} - 1$	$\mathbf{h}^{-1} - 1$	1	0.832	0.756	0.8	0.65 – 0.9

* Equivalent to the critical binary Galton-Watson branching process with i.i.d. exponential edge lengths.

† Mean values estimated in simulated OCN basins. According to Rodriguez-Iturbe and Rinaldo (2001); Cieplak et al. (1998).

‡ According to Rodriguez-Iturbe and Rinaldo (2001); De Vries et al. (1994); Peckham (1995); Maritan et al. (1996); Rigon et al. (1996); Tarboton et al. (1988).

multiple stream attributes. A fast and simple recursive simulation algorithm allows one to generate networks of realistic size within seconds (Fig. 4). Multiple additional symmetries and a well-developed theoretical framework make the process an efficient modeling tool.

A special subfamily of HBPs, a one-parameter *critical Tokunaga model*, is specified by $T_k = (c-1)c^{k-1}$ for some $c \geq 1$ (Sect. 9). This model yields a simple relation among the Horton exponents:

$$2c = R_B = R_M = R_A > R_S = R_L = R_\Lambda = c, \tag{7}$$

where we list, left to right, the Horton exponents for the stream counts (R_B), magnitude (R_M), area (R_A), number of links in a stream (R_S), stream length (R_L), and lengths of the longest channel to the divide (R_Λ). The critical Tokunaga model provides a close fit to the data and scalings reported in river studies over the past decades. We illustrate this in the Beaver Creek basin of Fig. 2. Figure 5 shows seven Horton's laws and their respective fits by the critical Tokunaga model with $c = 2.3$. Specifically, we consider the following stream attributes parameterized by stream order $K = 1, \dots, 6$: the stream numbers N_K (panel a), the mean magnitude M_K and the mean number S_K of links in a stream (panel b), the mean contributing area A_K and the mean total channel length L_K^{tot} upstream (panel c), the mean length Λ_K of the longest channel to the divide and the mean stream length L_K (panel d). The fitting lines correspond to the theoretical model predictions (see figure caption) that we derive in Sects. 9, 11. We notice that the model predictions in panels (a) and (b) span the entire range of orders, while those in panels (c) and (d) only give asymptotic behavior at large orders. This explains some fairly large deviations between the data and fitting lines that one can notice at small orders in panel (d).

Table 2 summarizes the expressions for the Horton exponents and main scaling constants in the critical Tokunaga model and compare them with the respective quantities estimated in data and the OCN model.

It may seem remarkable that a model with a single parameter provides such a close fit to the variety of Horton's laws (and other attributes, as can be seen from the further discussion). This hints at deep symmetries in the structure of trees that describe river networks. The theory of random self-similar trees explains the mathematical origin of these symmetries and provides one with tools for future exploration.

2.3 Survey Organization

The rest of the paper is organized as follows. Section 3 presents main concepts and definitions. Tree representation of a stream network and graph-theoretic terminology (Fig. 6) used throughout the paper are introduced in Sect. 3.1. Sections 3.2, 3.3 define Horton-Strahler stream orders and related Horton's laws (in their simplest form) for stream attributes. The remainder of this section

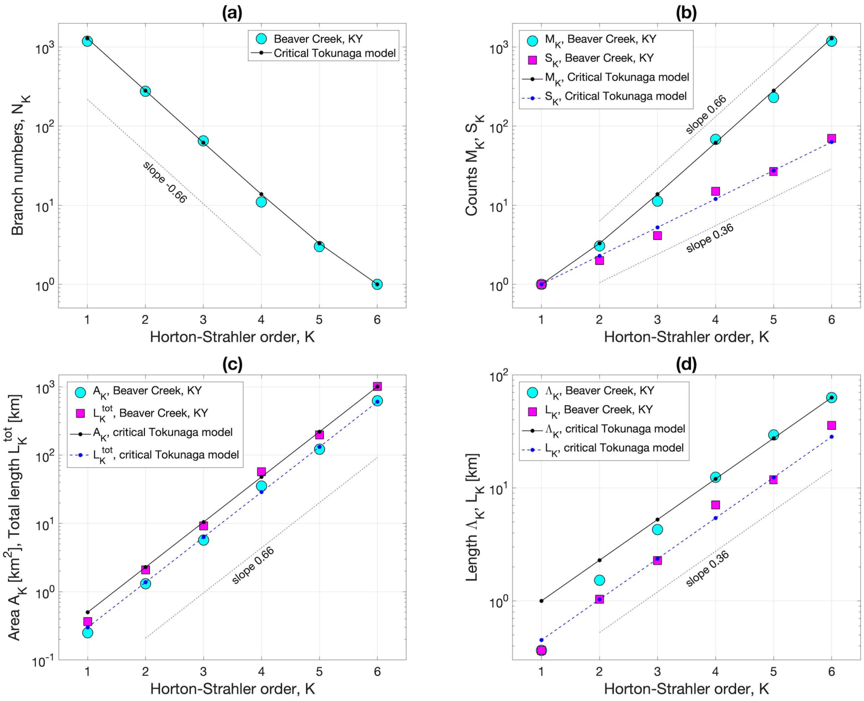

Fig. 5 Horton's laws in the stream network of Beaver creek, Floyd County, KY of Fig. 2. Symbols correspond to the observed attributes. Lines and dots show theoretical fit by a critical Tokunaga model (Sect. 9) with $c = 2.3$. (a) Stream numbers N_K. The model fit is given by (73); it has asymptotic slope $-\log_{10}(2c) \approx -0.66$, which is achieved here at small orders. (b) Stream magnitudes M_K (cyan circles) and number of links S_K in a stream (magenta squares). The fit for M_K is given by (73); it has asymptotic slope $\log_{10}(2c) \approx 0.66$, which is achieved here at large orders. For this model, $S_K = c^{K-1}$, which correspons to the slope $\log_{10} c \approx 0.36$. (c) Contributing areas A_K (cyan circles) and total upstream channel length $L_K^{\rm tot}$ (magenta squares). The fitting lines have theoretical slope $\log_{10}(2c) \approx 0.66$; see (76). The intercept is selected so that the fitting line coincides with an observed quantity at $K = 6$. (d) Lengths Λ_K of the longest stream to the divide (cyan circles) and lengths L_K of streams (magenta squares). The fitting lines have theoretical slope $\log_{10}(c) \approx 0.36$; see (60) and (95).

discusses essential heuristic implications of the Horton's laws for power-law frequencies of and relations among stream attributes.

Section 4 introduces the key technical tools of self-similarity analysis – Tokunaga coefficients, coordination of tree measures, Horton pruning, and Horton prune-invariance.

Section 5 recalls basic facts from the theory of generating functions and complex analysis that are used to establish our main results. Here we formalize the notion of Horton's laws for mean stream attributes by considering three consecutively stronger versions – (R), (Q), and (G) – of geometric variation. Later we refer to those as the *root* Horton's law, *quotient* Horton's law, and *geometric* Horton's law, respectively.

The key stream attributes and their relations are presented in Sect. 6. This includes the mean number S_K of edges in a branch of order K, the mean magnitude M_K of a branch of order K, the mean number $\mathcal{N}_k[K]$ of branches of order k in a tree of order K, the mean length L_K of a branch of order K, and the mean contributing area A_K of a branch of order K.

Section 7 presents the main results of this work – Horton's laws for mean branch attributes in a self-similar tree. We start in Sect. 7.1 with the geometric Horton's laws for the mean branch magnitudes M_k (Theorem 1) and mean branch numbers $\mathcal{N}_k[K]$ (Corollary 1). These laws hold in any self-similar tree under a mild constraint $\limsup_{k\to\infty} T_k^{1/k} < \infty$; they form a foundation for further development. Section 7.2 establishes quotient Horton's laws for S_k, L_k, and A_k under additional Assumption 1, which seems practically appealing to most applications. Section 7.3 discusses a special case of Horton's laws under the hydrologically relevant constraint of unit Horton exponent for edge lengths, which is a substantial generalization of the condition of equally distributed edge lengths that is well-documented in hydrologic observations. Section 8 introduces a self-similar Hierarchical Branching Process (HBP) – a computationally simple and analytically tractable model that generates trees with arbitrary Tokunaga sequences and obeys the (strongest) geometric Horton's laws for the mean branch counts, magnitudes, and lengths. An important one-parametric family of self-similar HBP – critical Tokunaga process with $T_k = (c-1)c^{k-1}$ for $c \geq 1$ – is examined in Sect. 9. We propose this as a conceptual model for river stream networks, as it provides a very close fit to the attributes and scalings reported in observations (see Table 2 and Fig. 5). A classical model with slightly relaxed constraints on the Tokunaga sequence, $T_k = ac^{k-1}$, is discussed in Appendix E.

Fractal dimension and Hack's law in self-similar HBP trees are examined in Sects. 10, 11, respectively. We show in particular that the fractal dimension **d** and Hack's exponent **h** are reciprocal to each other, which has been heuristically known for trees with exact Horton's laws since the 1980s (La Barbera and Rosso, 1989; De Vries et al., 1994; Peckham, 1995).

Section 12 illustrates the origin of power-law exceedance frequencies of edge and branch attributes in a tree that satisfies selected Horton's laws. Specifically, we consider the vertex contributing area $A_{(i)}$, the length $\Lambda_{(i)}$ of the longest stream from a vertex to the divide, and the length $L_{[i]}$ of a randomly selected branch. Section 13 provides concluding remarks. Proofs and most general results, which may not be of prime interest in applied analyses, are given in Appendices.

3 Horton's Laws and Their Implications: A Heuristic Approach

3.1 Tree Representation of River Networks

River studies commonly represent a stream network that drains a single basin (watershed, catchment) as a rooted binary tree with planar embedding. The basin *outlet* (point furthest downstream) corresponds to the tree root, *sources*

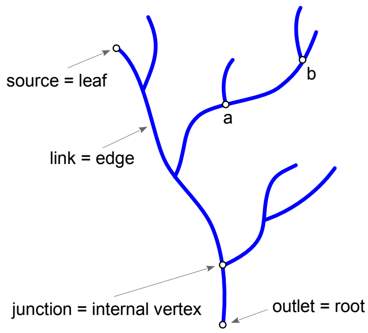

Fig. 6 Hydrologic vs. graph theoretic terminology. The figure shows, clockwise from the bottom, the basin outlet (tree root), a stream junction (internal vertex), link (edge), and a source (leaf). Vertex b is an offspring of a; and a is the parent for b.

(points furthest upstream) to leaves, *junctions* (points where two streams meet) to internal vertices, and *links* (stream segments between two successive nodes) to edges. A node j immediately upstream of a node i is called an *offspring* of i, and i is called the *parent* of j. Any node j upstream of i is called a *descendant* of i, and any node j downstream of i is called an *ancestor* of i. Figure 6 illustrates this correspondence.

In this work we use the graph-theoretic nomenclature, which provides a better link to the other systems examined using their tree representation (see Sect. 1). We assume that all examined trees belong to the space $\mathcal{T}$ of finite binary rooted trees, or to the space $\mathcal{L}$ of trees from $\mathcal{T}$ with positive edge lengths (Kovchegov and Zaliapin, 2020, Sect. 2.1). The models discussed in this work do not deal with planar embedding of trees, which is an important separate problem. We notice at the same time, that a suitable (non-physical) planar embedding of a given tree can be readily developed. One such embedding approach is used to illustrate synthetic HBP trees in Fig. 4.

3.2 Horton-Strahler Orders

The importance of vertices and their parental edges is measured by the Horton-Strahler order $K \geq 1$ (Horton, 1945; Strahler, 1957). We agree that each vertex and its parental edge (the unique edge that connects this vertex to its parent, or the immediate downstream edge) have the same order. The order assignment is done in a hierarchical fashion, from the leaves towards the root (that is, from the sources downstream). Specifically, each leaf (and its parental edge) is assigned order $K = 1$. When two edges of the same order K merge at a vertex, the vertex is assigned order $K + 1$. When two edges with different orders $K_1 > K_2$ merge at a vertex, the largest order prevails and the vertex is assigned order K_1. The connected sequence of vertices and their parental edges of the same order K is called a *branch* of order K. The Horton-Strahler orders are illustrated in Figs. 1, 2. We denote by $N_K = N_K[T]$ the number of branches of order K in a finite tree T.

The Horton-Strahler order of a tree is that of its root, or equivalently, the maximal order of its vertices (edges, branches). We show below that multiple fundamental regularities in the structure and dynamics of river networks are expressed in terms of the Horton-Strahler orders.

3.3 Horton's Laws

The observed stream counts N_K in a large basin are closely approximated by (Horton, 1945)

$$\frac{N_K}{N_{K+1}} = R_B \quad \Leftrightarrow \quad N_K \propto R_B^{-K} \tag{8}$$

for some *Horton exponent* $R_B \geq 2$. The lower bound on R_B follows immediately from the definition of Horton-Strahler orders, since it takes at least two branches of order K to create a single stream of order $K+1$. It has been noticed by (Strahler, 1957, p. 914) that the value of the empirical ratio R_B in river streams is between 3 and 5, and is usually close to 4. This has been strongly corroborated in numerous observational studies, e.g. Kirchner (1993), Shreve (1966), Leopold et al. (1992), Peckham (1995), Tarboton (1996), Turcotte (1997), Gupta and Waymire (1998), Zanardo et al. (2013), Rodriguez-Iturbe and Rinaldo (2001), and Mesa (2018).

In hydrogeomophology, a geometric scaling of any branch attribute with order, similar to that of Eq.(8), is called *Horton's law.* Horton's laws are documented for multiple physical and combinatorial quantities, including upstream area, magnitude (number of upstream sources), the total channel lengths, link slope, mean annual discharge, energy expenditure, etc. (Rodriguez-Iturbe and Rinaldo, 2001). These quantities often increase with order (unlike the branch counts N_K that decrease with order), which justifies a slightly different form of the respective Horton's laws. Specifically, consider the values Z_K obtained by averaging a selected attribute Z over branches of order K. Horton's law with exponent $R_Z \geq 1$ states that Z_K scale as

$$\frac{Z_{K+1}}{Z_K} = R_Z \quad \Leftrightarrow \quad Z_K \propto R_Z^K. \tag{9}$$

In both cases (8) and (9) the law is formulated in such a way that the Horton exponent is greater than unity. Informally, Z_K may represent a particular way to measure the branch "size", and the law (9) states that the order K of a branch is proportional to its logarithmic size $\ln(Z_K)$.

Horton's laws play an elemental role in statistical modeling of river basins, which rests upon empirical regularities that describe the frequencies of and relations among the key geometric and physical characteristics of individual streams. Remarkably, many such regularities heuristically follow from Horton's laws and are parameterized by the respective Horton exponents. Below we discuss several key power laws that are commonly observed in river networks.

3.4 Power-Law Relations Between Attributes

Suppose a stream attribute Z satisfies the Horton's law (9) with exponent R_Z, and the branch counts N_K satisfy the Horton's law (8) with exponent R_B. Then, using each of the laws to express K and equating these expressions, we find

$$Z_K \;\propto\; N_K^{-\alpha}, \quad \text{with} \quad \alpha = \frac{\ln R_Z}{\ln R_B}. \tag{10}$$

Similarly, suppose that the Horton's law (9) holds for selected river attributes Z and Y, with exponents R_Z and R_Y, respectively. Then, Z_K and Y_K are connected via a power-law relation

$$Z_K \;\propto\; Y_K^{\alpha}, \quad \alpha = \frac{\log R_Z}{\log R_Y}. \tag{11}$$

Equations (10), (11) are a punctuated (by discrete orders) version of a general power-law relation $Z \;\propto\; Y^{\alpha}$ that is abound among hydrologic quantities.

It is common to relate an attribute of interest to the basin area A. A well studied example is Hack's law that relates the length L of the longest stream in a basin to the basin area A via $L \;\propto\; A^{\mathbf{h}}$ with $\mathbf{h} \approx 0.6$ (Hack, 1957; Rigon et al., 1996; Rodriguez-Iturbe and Rinaldo, 2001). Assuming Horton's laws for the area and length of the longest stream, the parameter $\mathbf{h}$ is expressed via the respective Horton exponents as in Eq. (11):

$$\mathbf{h} = \frac{\log R_L}{\log R_A}. \tag{12}$$

3.5 Power-Law Frequencies of Link Attributes

Consider empirical frequencies of an attribute Z calculated at every edge (link) in a large tree (basin). We write $Z_{(i)}$ for the value of Z calculated at the i-th edge. Assume that Horton's law holds (i) for the examined attribute Z, with exponent R_Z and (for simplicity) proportionality constant equal to one; (ii) for the average number S_K of edges within a branch of order K, with exponent R_S; and (iii) for branch counts N_K with exponent R_B as in (8). The number of edges of order K in such a tree is given by $N_K S_K$. One can now heuristically approximate the expected frequencies of $Z_{(i)}$ by using the same value Z_K for any edge of order K and considering a limit of an infinitely large tree:

$$\#\{i : Z_{(i)} \geq R_Z^K\} \approx \sum_{j=K}^{\infty} N_j S_j \propto \sum_{j=K}^{\infty} \left(\frac{R_S}{R_B}\right)^j \propto \left(\frac{R_S}{R_B}\right)^K.$$

As before, this is a punctuated (by discrete order) version of a general power law relation

$$\#\{i : Z_{(i)} \geq z\} \;\propto\; z^{-\beta}, \quad \beta = \frac{\log R_B - \log R_S}{\log R_Z}. \tag{13}$$

Such power laws are reported for the upstream contributing area, stream lengths to the divide, water discharge, or energy expenditure. For example, analyses of Tarboton et al. (1989), Rodriguez-Iturbe et al. (1992), and Maritan et al. (1996) on river basins extracted from digital elevation models (DEM's) suggest

$$\#\{i : A_{(i)} \geq x\} \;\propto\; x^{-\beta_A} \quad \text{with} \quad \beta_A \approx 0.45, \tag{14}$$

$$\#\{i : \Lambda_{(i)} \geq x\} \;\propto\; x^{-\beta_\Lambda} \quad \text{with} \quad \beta_\Lambda \approx 0.8, \tag{15}$$

where $A_{(i)}$ is the area upstream of link i and $\Lambda_{(i)}$ is the distance from link i to the furthest source (or, equivalently, to the basin divide) measured along the channel network. Our analysis below shows that these power laws hold in self-similar trees and their exponents found via (13) fit the empirical exponents found in observations.

3.6 Modeling Physical Characteristics of a Stream

Classical hydrologic and geomorphologic studies of the mid 20th century revealed that the key physical characteristics of streams – such as stream width, depth, slope, and flow velocity, can be modeled as power functions of the stream magnitude (Leopold and Miller, 1956; Leopold et al., 1992; Dodov and Foufoula-Georgiou, 2004a,b, 2005),(Rodriguez-Iturbe and Rinaldo, 2001, Chapter 1). Specifically, data analysis suggests that physical characteristics of streams scale with the stream discharge Q defined as the volume of water flowing through a river stream. Such scalings are called *hydraulic-geometric relations.* For example, the velocity v through the stream can be approximated by $v \;\propto\; Q^a$, etc. The discharge Q, in turn, is a power-law function of the basin area: $Q \;\propto\; A^\beta$ (see Sect. 3.4). The value of exponent β depends on a precise definition of discharge (bankful, mean annual, etc.) Finally, the basin area A is closely approximated by the basin magnitude M, since it is natural to think of a stream network as a space-filling tree (see Sect. 7 for a formal treatment). Combining these observations, we find that the stream velocity can be modeled as a power-law function of the stream magnitude:

$$v \;\propto\; M^{a\beta}.$$

Hydraulic-geometric relations exist for other physical characteristics of a stream, including the average link slope s (e.g., Gupta and Waymire, 1989):

$$s \;\propto\; A^{-\theta} \;\propto\; M^{-\theta}, \quad \theta \approx 0.5.$$

In summary, rather unexpectedly, essential physical characteristics of a river network (e.g., stream velocity or link slope) can be estimated from purely combinatorial statistics of its tree representation (e.g., magnitude). Gupta (2017) asserts that *"Self-similarity in channel networks plays a foundational role in understanding the observed scaling, or power law relations, between peak flows and drainage areas"*. For example, the emergent scaling behavior opens up the opportinity to circumvent a large number of parameters governing production and transport of runoff along the stream channels and use basin's combinatorial characteristics for developing flood frequency relations and flood forecasting in ungauged basins (Gupta et al., 1994, 1996, 2010, 2007; Gupta, 2017). Accordingly, the results presented in this survey can inform modeling efforts aimed at physical quantities of the streams and the related processes and hazards. Gupta and Mesa (2014) discussed an alternative approach for establishing Horton's laws for river physical attributes (hydraulic-geometric variables) based on the the Buckingham π theorem and asymptotic self-similarity of first and second kinds (Barenblatt, 1996).

3.7 Beyond Heuristics

We observe that the above discussion in Sects. 3.3, 3.4, and 3.5 is heuristic, only maintaining a physical (but not mathematical) level of rigor. The very definition of Horton's law via Eqs.(8) and (9) is not instrumental for developing a useful theory. Indeed, since it is hard to expect that the exact equalities would hold in a range of practically interesting situations, one should accept an approximate nature of these statements and, hence, define what is meant by "approximate".

The approach adopted in this survey (and in most of the studies reviewed herewith) asserts that, for any fixed k, the branch number ratios N_k/N_{k+1} converge to the Horton's exponent R_B when the tree size increases. Similarly, the other branch attribute ratios Z_{k+1}/Z_k converge to the appropriate Horton exponents when both k and the tree size increase. These convergences involve random variables N_k and Z_k and hence should be understood in a proper probabilistic sense (Bhattacharya and Waymire, 2007). The difference between the treatment of the branch numbers N_k and other branch attributes Z_k is explained by the observation that N_k decreases with k, while all other branch attributes (examined here) increase with k.

As an intermediate step, we consider the *mean values* of the examined attributes with respect to the examined distribution of trees. Such a mean value should not be confused with the *average value* that is calculated over a collection of branches within a single random tree. For example, we consider below a random length $L_{[i]}$ of the ith branch. One can average the observed random lengths over the N_k branches of a given order k in a single tree T to obtain the empirical *average* of the branch lengths:

$$\overline{L}_k = \frac{1}{N_k} \sum_{i=1}^{N_k} L_{[i]}.$$

Importantly, $\overline{L}_k$ is a random variable that takes on a new value for each realization of a random tree. Finally, we consider the *mean branch length*

$$L_k = \mathsf{E}[L_{[i]}] \quad \text{for any } i \text{ because of coordination,}$$

which is a constant that only depends on the examined tree distribution.

Most of our results are formulated for the *mean* branch attributes. Importantly, the respective results for the random attributes readily follow from these mean results. To illustrate this implication, let $Z_{[i]}$ denote a random value of the examined attribute calculated for branch i in a random tree T, $\overline{Z}_k$ denote the empirical average of the attribute over the branches of order k in T, and Z_k denote the mean value of the attribute for a random branch of order k. Horton's law for the deterministic mean attribute is defined as a limit statement:

$$\lim_{k\to\infty} \frac{Z_{k+1}}{Z_k} = R_Z. \tag{16}$$

Section 5 discusses this and two other (weaker and stronger) forms of Horton's law that can hold in deterministic sequences of branch attributes. We have

$$\mathsf{E}\left[\overline{Z}_k\right] = \mathsf{E}\left[\mathsf{E}\left[\overline{Z}_k|N_k\right]\right] = \mathsf{E}\left[Z_{[i]}\right] = Z_k.$$

In a tree T of order K the branch numbers are bounded from below by $N_k \geq 2^{K-k}$ for any $k \leq K$, so $K \to \infty$ implies $N_k \to \infty$ for any fixed k with probability one. Hence, the Weak Law of Large Numbers (Bhattacharya and Waymire, 2007) asserts that

$$\overline{Z}_k \stackrel{p}{\to} Z_k \quad \text{for any } k \text{ as } K \to \infty, \tag{17}$$

where $\stackrel{p}{\to}$ denote convergence in probability. Accordingly,

$$\frac{\overline{Z}_{k+1}}{\overline{Z}_k} \stackrel{p}{\to} \frac{Z_{k+1}}{Z_k} \quad \text{for any } k \text{ as } K \to \infty, \tag{18}$$

where the (deterministic) fraction in the right-hand side converges to R_Z as in (16). Practically, statements (16) and (18) suggest that the empirical averages $\overline{Z}_k$ satisfy Horton's approximation of (9) in a sufficiently large tree (or a finite collection of such trees).

The above discussion applies to the branch magnitudes M_k, combinatorial and metric branch lengths S_k and L_k, branch contributing areas A_k, and the length of the longest stream to the divide (height) Λ_k. At the same time, the probabilistic limit results for the random branch counts N_k require more sophisticated techniques that are outside of the scope of this survey. Section 9.2 reviews limit laws for the random branch numbers in the critical Tokunaga model of Sect. 9.

4 Self-Similarity of River Networks

4.1 Tokunaga Coefficients

The Tokunaga coefficients complement the branch counts $N_K[T]$ in describing the structure of a tree T. The empirical Tokunaga coefficient $t_{i,j}[T]$ with $i < j$ is the average number of branches of order i that merge with a branch of order j in a finite tree T:

$$t_{i,j}[T] = \frac{N_{i,j}[T]}{N_j[T]}, \tag{19}$$

where $N_{i,j}[T]$ is the number of instances when an order-i branch merges with an order-j branch within T. The merging of branches of distinct orders is referred to as *side-branching*, and a branch that merges into a branch of a higher order is called a *side-branch*. Merging of two branches of the same order is called *principal branching*.

Assume that we fix a distribution μ on the space $\mathcal{T}$ of finite rooted binary trees. For example, one might consider a uniform distribution among trees with a given number of leaves, leading to the critical binary Galton-Watson random tree (Burd et al., 2000; Pitman, 2006). Then one can define the Tokunaga coefficient $T_{i,j}$ as the expected number of side-branches of order i per a randomly selected branch of order j (Dodds and Rothman, 1999; Tokunaga, 1966, 1978; Burd et al., 2000). This definition serves well the purpose of our study; we refer to Kovchegov and Zaliapin (2020) for a more general approach.

We can arrange the Tokunaga coefficients for trees of a given order K in an upper triangular matrix

$$\mathbb{T}_K = \begin{bmatrix} 0 & T_{1,2} & T_{1,3} & \dots & T_{1,K} \\ 0 & 0 & T_{2,3} & \dots & T_{2,K} \\ 0 & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 & T_{K-1,K} \\ 0 & 0 & \dots & 0 & 0 \end{bmatrix}. \tag{20}$$

For example, the Tokunaga coefficients calculated for the Beaver Creek basin illustrated in Figs. 2, 5 are given by

$$\mathbb{T}_6 = \begin{bmatrix} 0 & 1.06 & 2.40 & 8.91 & 15.33 & 44.00 \\ 0 & 0 & 0.92 & 3.64 & 8.67 & 20.00 \\ 0 & 0 & 0 & 2.00 & 4.00 & 9.00 \\ 0 & 0 & 0 & 0 & 0.33 & 4.00 \\ 0 & 0 & 0 & 0 & 0 & 1.00 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}. \tag{21}$$

4.2 Coordination

We assume that the structure of a river basin is determined by its order. This means, for example, that a basin with outlet of order three and a sub-basin of order three within a larger basin of order nine have, statistically, the same structure. This assumption is in the heart of the Horton-Strahler orders, and is imposed, explicitly or implicitly, in the mainstream studies of river networks (Shreve, 1966, 1969; Peckham, 1995; Rodriguez-Iturbe and Rinaldo, 2001). We refer to this assumption as *coordination*. Under the assumption of coordination the Tokunaga matrix $\mathbb{T}_K$ of Eq. (20) coincides with the upper-left $K \times K$ submatrix of the Tokunaga matrix $\mathbb{T}_M$ for any $M \geq K$, which explains the assumption name.

4.3 Tree Self-Similarity

Most generally, *self-similarity* is understood as statistical invariance of a river basin under rescaling Mandelbrot (1982); Turcotte (1997); Dodds and Rothman (2000); Rodriguez-Iturbe and Rinaldo (2001). A fundamental specific way to downscale a river basin of order K is to only consider its branches with highest orders between $K - k + 1$ and K for a given $k < K$. This results in a coarser basin, whose order (being computed according to the rules of Sect. 3.2) is k.

Formally, we consider the operation of *Horton pruning* $\mathcal{R} : \mathcal{L} \to \mathcal{L}$ that removes the leaves from a tree T together with their parental edges, followed by a *series reduction* that eliminates all degree two non-root vertices by merging the edges adjacent to them. It is readily seen that the Horton pruning reduces the tree order by 1. Moreover, the order of each branch is also reduced by 1 (with understanding that branches of order 1 are eliminated). We refer to Peckham (1995), Burd et al. (2000), and Kovchegov and Zaliapin (2016) for a comprehensive discussion.

A coordinated distribution μ on the space $\mathcal{T}$ of combinatorial trees is called *self-similar* if it is invariant with respect to the Horton pruning (Burd et al., 2000; Kovchegov and Zaliapin, 2016):

$$\mu\left(\mathcal{R}^{-1}(T)|T \neq \phi\right) = \mu(T) \quad \text{for any} \quad T \in \mathcal{T}. \tag{22}$$

Informally, consider a forest of trees, where each tree T occurs multiple times according to its probability $\mu(T)$. The forest is self-similar if after pruning each tree by $\mathcal{R}$ we obtain the same forest. This definition can be extended to trees with edge lengths from space $\mathcal{L}$. In this case, we allow the edge lengths to scale by a *scaling constant* $\zeta > 0$ after pruning. We refer to (Kovchegov and Zaliapin, 2020, Sect. 3) for a formal treatment.

We use a conventional abuse of terminology by saying that a tree T is self-similar; this means that T is a random tree drawn from a self-similar distribution μ.

4.4 Tokunaga Sequence

A coordinated self-similar measure necessarily satisfies the following Toeplitz property (Kovchegov and Zaliapin, 2016, 2020): there exists a *Tokunaga sequence* $\{T_k\}_{k=1,2,...}$ such that

$$T_{i,i+k} = T_k \qquad \text{for all} \qquad i, k > 0. \tag{23}$$

In this case, the Tokunaga matrix (20) is a Toeplitz matrix:

$$\mathbb{T}_K = \begin{bmatrix} 0 & T_1 & T_2 & \dots & T_{K-1} \\ 0 & 0 & T_1 & \dots & T_{K-2} \\ 0 & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 & T_1 \\ 0 & 0 & \dots & 0 & 0 \end{bmatrix}. \tag{24}$$

Sometimes we refer to T_k as *Tokunaga indices*, which creates no confusion with $T_{i,j}$ of Sect. 4.1. A comprehensive discussion of this approach can be found in the works of Kovchegov and Zaliapin (2016) and Burd et al. (2000). The Tokunaga indices T_k for the Beaver Creek of Figs. 2, 5 can be approximated by averaging the values along the diagonals of the initial Tokunaga matrix (21):

$$T_1 = 1.06, \quad T_2 = 3.51, \quad T_3 = 8.86, \quad T_4 = 17.665, \quad T_5 = 44.00. \tag{25}$$

We emphasize that self-similarity (Sect. 4.3) is a property of a distribution of trees on $\mathcal{T}$ or $\mathcal{L}$ and hence, formally, cannot be applied to a single tree. In applied analysis, however, one works with a single basin, or a finite sample of basins. The Tokunaga sequence T_k provides a fundamental connection between the properties of a distribution on an infinite collection of trees and easily computed attributes of a single tree.

We show below that the Tokunaga indices T_k provide enough information to find the mean values of all other branching attributes in random self-similar trees.

4.5 Tokunaga Two-Parameter Model

The first model for river networks that explicitly describes the network structure in terms of side-branch counts is due to Eiji Tokunaga (1978). It postulates

$$T_k = ac^{k-1}, \quad a, c > 0. \tag{26}$$

The observed river networks are closely approximated by the *Tokunaga model* (Tokunaga, 1978; Peckham, 1995; Kovchegov et al., 2021). The estimated parameters $a \approx 1.1$ and $c \approx 2.5$ of this model have shown to be independent of (or only weakly dependent on) river's geographic location (Peckham, 1995; Dodds and Rothman, 2000; Zanardo et al., 2013).

McConnell and Gupta (2008) have shown that the Tokunaga model obeys the quotient Horton's law for stream numbers (when tree size increases) with

$$R_B = \frac{(a+c+2)+\sqrt{(a+c+2)^2-8c}}{2}. \tag{27}$$

This result revealed, for the first time, the emergence of Horton's law from the tree side-branch structure. Our Theorem 1 below establishes the most general statement of this type, showing that (almost) *any* Tokunaga sequence implies the geometric Horton's laws for branch numbers and mean magnitudes. A detailed treatment of the two-parameter Tokunaga model (26) is given in Appendix E.

Burd et al. (2000) demonstrated that the Shreve's random topology model (Shreve, 1966, 1969), equivalent to the critical binary Galton-Watson tree with a fixed progeny, is a special case of the Tokunaga model with $(a, c) = (1, 2)$:

$$T_k = 2^{k-1} \quad \text{for} \quad k \geq 1.$$

Accordingly, it satisfies the geometric Horton's law for mean branch numbers with $R_B = 4$.

For a long time, the critical binary Galton-Watson tree has remained the only well-studied probability model for which self-similarity was rigorously established, and whose Horton-Strahler ordering has received attention in the literature (Shreve, 1966, 1969; Kemp, 1979; Tarboton et al., 1988; Wang and Waymire, 1991; Barndorff-Nielsen, 1993; Yekutieli and Mandelbrot, 1994; Peckham, 1995; Devroye and Kruszewski, 1994; Burd et al., 2000). Scott Peckham has explicitly noticed, by performing a high-precision extraction of river channels for Kentucky River, Kentucky and Powder River, Wyoming, that the Horton exponents and Tokunaga parameters for the observed rivers significantly deviate from those in the Galton-Watson model (Peckham, 1995). He reported values $R_B \approx 4.6$ and $(a, c) \approx (1.2, 2.5)$ and emphasized the importance of studying a broad range of Horton exponents and Tokunaga parameters.

The general interest to fractals and self-similar structures in natural sciences during the 1990s resulted in a quest, mainly inspired and led by Donald Turcotte, for Tokunaga self-similar trees of diverse origin (Gabrielov et al., 1999; Newman et al., 1997; Ossadnik, 1992; Pelletier and Turcotte, 2000; Turcotte, 1997; Turcotte et al., 1999, 1998; Yakovlev et al., 2005; Zanardo et al., 2013). As a result, the Tokunaga model and respective Horton's laws, with a broad range of parameters, have been empirically or rigorously found in numerous observed and modeled systems, well beyond river networks.

5 Generating Functions: A Tool for Establishing Horton's Laws for Mean Attributes

This section summarizes the basic facts about generating functions that are used below to derive asymptotic behavior and Horton's laws for the mean branch attributes in a self-similar tree.

The *generating function* $f(z)$ of a sequence $a_k \geq 0$, $k = 0, 1, \ldots$, of non-negative real numbers is defined as a *formal power series*

$$f(z) = \sum_{k=0}^{\infty} a_k \, z^k, \tag{28}$$

where z is a complex number, $z \in \mathbb{C}$. It is known (Wilf, 1992) that there exist such a real number $r \geq 0$ that the series in the right hand side (rhs) of (28) converges (to the function $f(z)$) for any $|z| < r$ and diverges for any $|z| > r$. The number r (which can be infinite) is called the *radius of convergence* of the sequence a_k. The value of r puts notable constraints on the asymptotic behavior of a_k. In general, the smaller the radius of convergence, the faster the growth of the coefficients. Roughly speaking, $0 < r < 1$ implies that the coefficients a_k increase geometrically, $r > 1$ that the coefficients decrease geometrically, and $r = 1$ that the coefficient vary at a rate slower than geometric (e.g., polynomially). The values $r = 0$ and $r = \infty$ imply a faster than geometric growth or decay, respectively.

The Cauchy-Hadamar theorem expresses the radius of convergence in terms of the series coefficients (Wilf, 1992):

$$\frac{1}{r} = \limsup_{k \to \infty} a_k^{1/k}. \tag{29}$$

We consider the following, consecutively stronger, forms of geometric growth ($r > 1$) or decay ($r < 1$) of the sequence a_k:

$$\frac{1}{r} = \lim_{k \to \infty} a_k^{1/k}, \tag{R}$$

$$\lim_{k \to \infty} \frac{a_k}{a_{k+1}} = r, \tag{Q}$$

$$\lim_{k \to \infty} a_k \, r^k = \alpha \tag{G}$$

The three limits are related as follows:

$$(\mathrm{R}) \Rightarrow (\mathrm{Q}) \Rightarrow (\mathrm{G}), \tag{30}$$

which means that the existence of (G) implies the existence of (Q), etc. In hydrogeomorphology, the familiar quotient limit (Q) is referred to as Horton's law for coefficients a_k with *Horton exponent* r. We also consider a weaker root limit (R) and a stronger geometric limit (G), and refer to them as the root and geometric Horton's laws, respectively. The limits (R), (Q) and (G) may or may not exist, but if either of them does, then it has the same Horton exponent r as in (29) and ensures the existence of the weaker limit(s), according to (30), with the same Horton exponent. As has been mentioned above, we use a convention that Horton exponents are greater than or equal to unity. Hence, if $r < 1$, we consider the reciprocal quotient in (Q), $\lim a_{k+1}/a_k = r^{-1}$, and change the other limits accordingly to make the Horton exponent equal to r^{-1}.

Often, the radius of convergence for a_k can be easily found from an explicit form of $f(z)$. Indeed, if $r > 0$, then the function $f(z)$ is analytic within the disk $|z| < r$ and has at least one *singularity* on the circle $|z| = r$, that is it has to diverge for at least one point on that circle (Wilf, 1992, Thm. 2.4.2). Thus, the radius of convergence equals to the modulus of a singularity closest to the origin. Furthermore, recalling that $a_k \geq 0$ we have

$$|f(z)| = \left|\sum_{k=0}^{\infty} a_k\, z^k\right| \leq \sum_{k=1}^{\infty} a_k\, |z|^k = f(|z|), \tag{31}$$

where the equality is only achieved for $z = |z|$. This means that the singularity closest to the origin lies on the real axis (although there might be other singularities with the same modulus.) This makes the search for such a singularity much easier: one can only consider the restriction of the function $f(z)$ to the real axis. In other words, despite the use of complex analysis in establishing some of our results, the applied examination of suitable generating functions can be done in the real domain.

One can examine the function $f(z)$ in (28) to obtain more precise information about the coefficients a_k. If $f(z)$ has no singularities inside the circle $|z| = \rho$, then (Ahlfors, 1953)

$$a_k = \frac{1}{2\pi i} \oint_{|z|=\rho} \frac{f(z)dz}{z^{k+1}}. \tag{32}$$

We have mentioned that in general neither of the limits (R), (Q) and (G) must exist. However, if the singularity of $f(z)$ nearest to the origin is simple enough, then these properties are satisfied.

Proposition 1 (Geometric Horton's Law for a Simple Pole Sequence) *Suppose $f(z) = \sum_{k=1}^{\infty} a_k z^k$ with $a_k \geq 0$ is analytic in the disk $|z| < \rho$ except for a single singularity that occurs at a positive real value $r < \rho$, which is a pole of multiplicity one (simple pole). Then the geometric Horton's law* (G)*, and hence the quotient Horton's law* (Q) *and the root Horton's law* (R)*, are satisfied for the coefficients a_k with Horton exponent r. Furthermore, if we define $g(z) = f(z)(z - r)$, then the coefficient in the geometric Horton's law is $\alpha = -g(r)/r$.*

Proof See Appendix A.

Another useful result states that if we can write $f(z) = g(z)h(z)$, and the radius of convergence of $g(z)$ is smaller than that of $h(z)$, then the coefficients of $f(z)$ satisfy the same Horton's laws as those of $g(z)$. A formal statement is given below.

Proposition 2 (Horton's Laws for Product Sequence) *Consider complex valued functions $f(z)$, $g(z)$, and $h(z) \not\equiv 0$ that are analytic around the origin with the following series expansions*

$$f(z) = \sum_{k=0}^{\infty} a_k z^k, \qquad g(z) = \sum_{k=0}^{\infty} b_k z^k, \qquad \text{and} \qquad h(z) = \sum_{k=0}^{\infty} c_k z^k,$$

where a_k, $b_k \in [0, \infty)$ and $c_k \in \mathbb{R}$ for all k. Denote by r_a, r_b and r_c the radii of convergence of the sequences a_k, b_k and c_k, respectively.

(i) Suppose that

$$f(z) = g(z)h(z) \quad \text{and} \quad r_c > r_b > 0.$$

Then, any of the Horton's laws (R), (Q), *and* (G) *for the sequence b_k implies the same law for the sequence a_k, with Horton exponent $R = 1/r_b = 1/r_a > 0$.*

(ii) Suppose, in addition, that $h(z) \neq 0$ for $|z| \leq r_b$. Then any of the Horton's laws (R), (Q), *and* (G) *for the sequence a_k implies the same law for the sequence b_k, with Horton exponent $R = 1/r_a = 1/r_b > 0$.*

Proof See Appendix B.

Below, we find the generating functions for the sequences of mean branching attributes relevant to our study. The respective radii of convergence provide information on the asymptotic behavior of the examined sequences. In particular, the generating function $M(z)$ for the mean branch magnitudes M_k has a simple pole closest to the origin, and hence it satisfies the geometric Horton's law by Proposition 1.

6 Self-Similar Trees: Main Attributes and Their Relations

We assume that a river basin is represented by a self-similar tree T with a Tokunaga sequence T_k. This means, in particular, that each branch of Horton-Strahler order j produces an mean of T_{j-i} side-branches of order i for each i such that $1 \leq i < j$. In this section we state the recursive relations for the essential branch attributes: the mean number $\mathcal{N}_k$ of branches of order k, the mean number S_k of edges in a branch of order k, the mean magnitude (number of descendant leaves) M_k of a branch of order k, the mean length L_k of a branch of order k, and the mean contributing area A_k of a branch of order k. These relations are mostly due to straightforward applications of the Wald's formula (Bhattacharya and Waymire, 2007); it asserts that for a random sum of N independent identically distributed (i.i.d.) random variables X_i we have

$$\mathsf{E}\left[X_1 + \cdots + X_N\right] = \mathsf{E}[X_1]\mathsf{E}[N].$$

The expectations are taken with respect to a self-similar measure μ on $\mathcal{L}$.

Let S_k denote the mean number of edges (or vertices) within a branch of order k. This attribute is also known as the mean number of links in a Strahler stream of order k (e.g., Peckham, 1995). It equals the mean number of side-branches that join this branch plus the branch starting vertex:

$$S_k = 1 + T_1 + \cdots + T_{k-1}, \quad k \geq 1. \tag{33}$$

The *mean magnitude* M_k is the mean number of leaves descendent to an order k branch. It can be represented as the sum of magnitudes of two order $k-1$ branches that created this branch (called *principal branches*), plus the magnitudes of all the side-branches. Hence $M_1 = 1$, and

$$M_k = 2\, M_{k-1} + \sum_{i=1}^{k-1} M_i\, T_{k-i}, \quad \text{for} \quad k > 1. \tag{34}$$

The mean number $\mathcal{N}_k[K]$ of branches of order k in a tree of order K, also known as the mean total number of Strahler streams (Peckham, 1995), equals twice the number of branches of order $k+1$ plus the number of instances when a branch of rank k joins a branch of a higher rank. Thus, for a tree of order K we have $\mathcal{N}_K[K] = 1$ and

$$\mathcal{N}_k[K] = 2\,\mathcal{N}_{k+1}[K] + \sum_{i=k+1}^{K} \mathcal{N}_i[K]\, T_{i-k}, \quad \text{for any } k < K. \tag{35}$$

Comparing equations (34) and (35) we find

$$\mathcal{N}_{K-k+1}[K] = M_k \quad \text{for all orders} \quad 1 \leq k \leq K. \tag{36}$$

Notice that we explicitly indicate the tree order K when working with the mean number of branches $\mathcal{N}_k[K]$, and do not do that for S_k and M_k. This is because the initial terms of the increasing sequences M_k and S_k coincide for different values of K, which is not the case for $\mathcal{N}_k[K]$ (e.g., $M_1[4] = M_1[5]$ but $\mathcal{N}_1[4] \neq \mathcal{N}_1[5]$).

Let ℓ_k denote the mean length of an edge of order k. Then the mean length L_k of an order k branch equals $L_k = S_k \ell_k$. By the *differential* contributing area (as is opposed to the *total* contributing area) of an edge we understand the area that drains directly to the edge (not via its upstream vertex). Assume that the mean differential contributing area of an edge of order k equals α_k. Then, for the total mean contributing area A_K of a tree of order $K \geq 1$, we have

$$A_K = \sum_{k=1}^{K} \alpha_k S_k \mathcal{N}_k[K], \tag{37}$$

where $S_k \mathcal{N}_k[K]$ is the number of edges of order k in a tree of order K.

Another important product that appears in (37) is $\delta_k = \alpha_k S_k$, which is the mean differential contributing area of a branch of order k. The total contributing area A_k can be expressed recursively by noticing that $A_1 = \alpha_1$ and

$$A_k = 2A_{k-1} + \alpha_k S_k + \sum_{i=1}^{k-1} A_i \, T_{k-i} \quad \text{for} \quad k \geq 2. \tag{38}$$

7 Horton's Laws in a Self-Similar Tree

This section establishes the main theoretical result of our work: geometric Horton's laws for the mean branch numbers $\mathcal{N}_k[K]$ and mean magnitudes M_k in a self-similar tree (Sect. 7.1). Next, we show (Sect. 7.2) how Horton's laws for other mean attributes follow from these ones under additional assumptions.

7.1 Geometric Horton's Law for Mean Branch Numbers and Magnitudes

Consider a Tokunaga sequence T_k and its generating function $T(z) = \sum\limits_{k=1}^{\infty} T_k \, z^k$. If we let $t_1 = T_1 + 2$, and $t_k = T_k$ for $k \geq 2$, then t_k takes into account the *side-branching* and *principal branching*. Let

$$\hat{t}(z) = -1 + \sum_{k=1}^{\infty} t_k \, z^k = -1 + 2z + T(z).$$

Observe that $\hat{t}(0) = -1$, and since $T_k \geq 0$ we have $\hat{t}(1/2) = T(1/2) \geq 0$. Furthermore, since

$$\frac{d}{dz}\hat{t}(z) = 2 + \sum_{k=1}^{\infty} kT_k \, z^{k-1} > 0$$

for all positive real values of z, the equation $\hat{t}(z) = 0$ has a unique real root w_0 of multiplicity one in the interval $(0, 1/2]$. Let r_T be the radius of convergence for $T(z)$ and define $R_T = r_T^{-1}$. We notice that $r_T > w_0$. The following result of Kovchegov and Zaliapin (2016) ensures that w_0 is the root of $\hat{t}(z)$ closest to the origin; this fact will be used below.

Lemma 1 *Suppose* $\limsup\limits_{k\to\infty} T_k^{1/k} < \infty$ *and let* w_0 *be the only real root of* $\hat{t}(z)$ *in the interval* $(0, 1/2]$*. Then, for any other root* $w \in \mathbb{C}$ *of* $\hat{t}(z)$*, we have* $|w| > w_0$.

The generating function for the magnitudes M_k is obtained by multiplying both sides in (34) by z^k and summing over all $k \geq 1$:

$$M(z) = \sum_{k=1}^{\infty} M_k \, z^k = z + 2zM(z) + M(z) \, T(z).$$

Thus,

$$M(z) = \frac{z}{1-2z-T(z)} = -\frac{z}{\hat{t}(z)}. \tag{39}$$

The function $M(z)$ is analytic with the exception of zeroes and singularities of $\hat{t}(z)$. Lemma 1 asserts that $w_0 \in (0, 1/2]$ is the closest to the origin root of $\hat{t}(z)$; recall that it has multipliciy one. Hence, w_0 is a simple pole of $M(z)$ and the only singularity of $M(z)$ within a disk $|z| < w_0 + \epsilon$ for a small enough $\epsilon > 0$. Consequently, the radius of convergence for $M(z)$ is $r_M = w_0$. We define $R_M = r_M^{-1}$. Proposition 1 implies that the geometric Horton's law holds for M_k. We formulate this result in the following theorem.

Theorem 1 (Geometric Horton's Law for Mean Branch Magnitudes) *Suppose that $r_T > 0$, that is $\limsup\limits_{k\to\infty} T_k^{1/k} < \infty$. Then, the geometric Horton's law for mean branch magnitudes M_k holds with Horton exponent $R_M = 1/w_0$, where w_0 is the only real root of the function $\hat{t}(z) = -1 + 2z + \sum\limits_{j=1}^{\infty} z^j T_j$ in the interval $\left(0, \frac{1}{2}\right]$. Specifically, the geometric Horton's law states that*

$$\lim_{k\to\infty} \left(M_k\, R_M^{-k}\right) = M \ < \infty, \tag{40}$$

where M is a positive real constant given by

$$M = -\frac{1}{w_0} \lim_{z\to w_0} \frac{z(z-w_0)}{\hat{t}(z)}. \tag{41}$$

The geometric Horton's law implies the quotient Horton's law

$$\lim_{k\to\infty} \frac{M_{k+1}}{M_k} = R_M. \tag{42}$$

Recalling (36) we notice that $\mathcal{N}_1[K] = M_K$ and hence obtain the asymptotic behavior for $\mathcal{N}_k[K]$.

Corollary 1 (Geometric Horton's Law for Mean Branch Numbers) *Under the assumption of Theorem 1 the geometric Horton's law holds for the mean branch numbers $\mathcal{N}_k[K]$ with Horton exponent $R_B = R_M = 1/w_0$. Specifically, the geometric Horton's law states that*

$$\lim_{K\to\infty} \left(\mathcal{N}_1[K]\, R_B^{-K}\right) = M \ < \infty, \tag{43}$$

where M is the same as in Thm. 1. This implies the quotient Horton's law: for each positive integer j we have

$$\lim_{K\to\infty} \frac{\mathcal{N}_j[K]}{\mathcal{N}_{j+1}[K]} = R_B \quad \text{or} \quad \lim_{K\to\infty} \frac{\mathcal{N}_j[K]}{\mathcal{N}_1[K]} = R_B^{1-j}. \tag{44}$$

Informally, Theorem 1 and Corollary 1 ensure that the geometric Horton's laws for mean branch magnitudes and mean branch counts hold with the same Horton exponent in "any" self-similar tree, that is in any coordinated tree with a well-defined Tokunaga sequence T_k. The assumption of non-zero radius of convergence in $T(z)$ eliminates obscure cases of super-exponential growth of T_k, such as $T_k = k!$ or $T_k = k^k$.

7.2 Horton's Laws for Other Mean Branch Attributes

Horton's laws for other mean branch attributes are obtained by examining the generating functions for the respective sequences and using the properties of the series T_k, ℓ_k, and α_k. The most general results that examine each type of the Horton's law (root, quotient, and geometric) under the assumption $\limsup\limits_{k\to\infty} T_k^{1/k} < \infty$ are formulated in Appendix C. This section illustrates a particular case of the quotient Horton's law (Q) for selected branch attributes under the following more stringent yet practically appealing assumption. Informally, it suggests that sequences T_k, ℓ_k and α_k behave "nicely".

Assumption 1 (Quotient Horton's law for T_k, ℓ_k, α_k) *Assume that the quotient Horton's law holds for T_k, ℓ_k and α_k:*

$$\lim_{k\to\infty} = \frac{T_k}{T_{k+1}} = c^{-1} > 0, \qquad \lim_{k\to\infty} = \frac{\ell_k}{\ell_{k+1}} = \lambda^{-1} > 0,$$
$$\textit{and} \qquad \lim_{k\to\infty} = \frac{\alpha_k}{\alpha_{k+1}} = \alpha^{-1} > 0. \tag{45}$$

Assumption 1 is satisfied for a multitude of natural choices for T_k, ℓ_k, and α_k, including inverse polynomial, constant, polynomial, and geometric series. The inequality $r_T > w_0$ implies that $c^{-1} > w_0$.

The asymptotic behavior for the mean total contributing areas A_k follows from that for M_k. First, we write the generating function for A_k via (38):

$$A(z) = \sum_{k=1}^{\infty} A_k z^k = 2zA(z) + \sum_{k=1}^{\infty} \alpha_k S_k z^k + A(z)T(z),$$

which yields, by (39),

$$A(z) = \frac{\sum\limits_{k=1}^{\infty} \alpha_k S_k z^k}{1 - 2z - T(z)} = M(z)\left(\sum_{k=1}^{\infty} \alpha_k S_k z^{k-1}\right) = -\frac{D(z)}{\hat{t}(z)}. \tag{46}$$

Here $D(z)$ is the generating function for the differential contributing areas $\delta_k = \alpha_k\, S_k$ of branches of order k. The radius of convergence of $D(z)$ and the asymptotic behavior of δ_k can be examined using Propositions 5,8 where ℓ_k need to be replaced with α_k.

Equation (46) implies the following convolution expression for A_k (Wilf, 1992):

$$A_k = \sum_{i=1}^{k} \alpha_{k+1-i} S_{k+1-i} M_i. \tag{47}$$

Observe that comparing equations (47) and (37) we arrive at $\mathcal{N}_{K-k+1}[K] = M_k$ that was first established in Eq. (36). We denote by r_A and r_D the radii of convergence of $A(z)$ and $D(z)$, respectively, and let $R_A = r_A^{-1}$, $R_D = r_D^{-1}$.

Proposition 3 [Quotient Horton's law for A_k] *Suppose that Assumption 1 holds. Then*

$$r_A = \min\{r_M,\, r_D\} = \min\left\{r_M,\, \alpha^{-1}\min\{1, r_T = c^{-1}\}\right\}.$$

The quotient Horton's law holds with the Horton exponent

$$R_A = \max\{R_M,\, R_D\} = \max\left\{w_0^{-1},\, \alpha\max\{1, c\}\right\},$$

that is

$$\lim_{k\to\infty}\frac{A_{k+1}}{A_k} = R_A. \tag{48}$$

We next examine the mean number S_k of edges within a branch of order k. The most straightforward practical way to obtain the asymptotic of S_k is via direct application of (33). The generating function approach clarifies the origin of the respective Horton's laws. Multiplying both sides in (33) by z^k and summing over $k = 1, 2, \ldots$, we obtain the generating function $S(z)$ of S_k:

$$S(z) = \sum_{k=1}^{\infty} S_k\, z^k = \sum_{k=1}^{\infty} z^k + \sum_{k=1}^{\infty}\left(\sum_{i=1}^{k-1} T_i\right) z^k = \frac{z\,(T(z)+1)}{1-z}. \tag{49}$$

The function $S(z)$ may become singular because of a singularity of $T(z)$ or the vanishing denominator $(1-z)$. The singularity of $S(z)$ closest to the origin is the smallest of $z = 1$ and the (only) singularity of $T(z)$. Let r_T, r_S denote the radiuses of convergence for the series $T(z)$ and $S(z)$, respectively, and define $R_T = r_T^{-1}$, $R_S = r_S^{-1}$.

Proposition 4 [Quotient Horton's law for S_k] *Suppose that Assumption 1 holds. Then $r_S = \min\{1, r_T = c^{-1}\}$ and the quotient Horton's law holds with the Horton exponent $R_S = \max\{1, c\} \geq 1$:*

$$\lim_{k\to\infty}\frac{S_{k+1}}{S_k} = R_S. \tag{50}$$

Next, consider the generating function of L_k denoted by $L(z)$:

$$L(z) = \sum_{k=1}^{\infty} L_k z^k = \sum_{k=1}^{\infty} S_k \ell_k z^k.$$

Let r_L denotes the radius of convergence for the series $L(z)$, and define $R_L = r_L^{-1}$.

Proposition 5 [Quotient Horton's law for L_k] *Suppose that Assumption 1 holds. Then $r_L = \lambda^{-1}\min\{1, r_T = c^{-1}\}$ and the quotient Horton's law holds with the Horton exponent $R_L = \lambda\max\{1, c\}$:*

$$\lim_{k\to\infty}\frac{L_{k+1}}{L_k} = R_L. \tag{51}$$

7.3 Quotient Horton's Laws in a River Basin

Here we use the results of Sect. 7.2 to formulate Horton's laws for branch attributes taking into account empirical constraints established for the observed river basins.

Observational studies suggest that the link lengths distribution in real rivers is independent of the position of the link within a basin (Tarboton et al., 1989; Rodriguez-Iturbe and Rinaldo, 2001). This corresponds to the assumption that the edge lengths are i.i.d. random variables with the same mean. We substantially relax this constraint in the following assumption.

Assumption 2 (Unit Quotient Horton's law for ℓ_k) *Assume that the quotient Horton's law with exponent of unity holds for the mean edge lengths ℓ_k:*

$$\lim_{k\to\infty} \frac{\ell_k}{\ell_{k+1}} = 1. \tag{52}$$

Assumption 2 is trivially satisfied in the case of i.i.d. edge lengths, where $\ell_k = \ell_{k+1}$ for all k. It also allows much wider variability of the edge lengths, including, for example, different length distribution for different orders and polynomial variation of the means $\ell_k \sim k^{\text{Const.}}$.

Assumption 2 strongly supports the existence of the unit quotient Horton's law for the mean differential contributing areas α_k of edges. Specifically, let ξ_k and ν_k be random variables that represent, respectively, the length and differential contributing area of a randomly selected edge of order k. We have $\mathsf{E}[\xi_k] = \ell_k$ and $\mathsf{E}[\nu_k] = \alpha_k$. Suppose there exist scalars $b > 0$ and $\sigma > 0$ such that

$$\nu_k = b\, \xi_k^\sigma \tag{53}$$

with σ and b being the same for all orders $k \geq 1$. Suppose, furthermore, that the random variables ξ_k scale with order, that is

$$\frac{\xi_k}{\ell_k} \stackrel{d}{=} \frac{\xi_1}{\ell_1}, \tag{54}$$

where $\stackrel{d}{=}$ denotes equality of distributions. This gives, in particular,

$$\xi_{k+1} \stackrel{d}{=} \frac{\ell_{k+1}}{\ell_k} \xi_k. \tag{55}$$

Then

$$\alpha_{k+1} = \mathsf{E}[\nu_{k+1}] = \mathsf{E}[b\xi_{k+1}^\sigma] = \mathsf{E}\left[b\left(\frac{\ell_{k+1}}{\ell_k}\right)^\sigma \xi_k^\sigma\right] = \left(\frac{\ell_{k+1}}{\ell_k}\right)^\sigma \alpha_k, \tag{56}$$

which implies

$$\lim_{k\to\infty} \frac{\alpha_k}{\alpha_{k+1}} = \lim_{k\to\infty} \left(\frac{\ell_k}{\ell_{k+1}}\right)^\sigma = 1. \tag{57}$$

The asymptotic of α_k without the scaling assumption (55) is examined in Appendix C.

Example 1 (Exponential edge lengths) *Suppose that the random edge length ξ_k is an exponential random variable with parameter $1/\ell_k$ so that $\mathsf{E}[\xi_k] = \ell_k$. The scaling assumption* (55) *is satisfied and the expected differential contributing area ν_k of an order k edge is given by*

$$\alpha_k = b\,\mathsf{E}[\xi_k^\sigma] = b\int\limits_0^\infty \frac{1}{\ell_k} x^\sigma e^{-x/\ell_k}\,dx = b\Gamma(\sigma+1)\ell_k^\sigma.$$

Accordingly, the relation (57) holds.

The relation (53) between edge length and differential contributing area and the scaling assumption (55) are sufficient but not necessary to obtain the quotient Horton's law for α_k under Assumption 2. To make our results applicable to a range of specific situations, where these assumptions may or may not hold, we make the following general assumption.

Assumption 3 (Unit Quotient Horton's law for ℓ_k, α_k) *Assume that the quotient Horton's law holds for T_k and the quotient Horton's law with exponent of* 1 *holds for the mean edge lengths ℓ_k and mean differential edge contributing areas α_k:*

$$\lim_{k\to\infty}\frac{T_k}{T_{k+1}} = c^{-1} > 0, \quad \lim_{k\to\infty}\frac{\ell_k}{\ell_{k+1}} = \lim_{k\to\infty}\frac{\alpha_k}{\alpha_{k+1}} = 1. \tag{58}$$

Proposition 6 [Quotient Horton's laws in a River Basin] *Suppose that Assumption 3 holds. Then the radii of convergence for the branch attributes are related as*

$$w_0 = r_M = r_A < r_L = r_S = c^{-1},$$

and, accordingly,

$$c = R_L = R_S < R_B = R_M = R_A = w_0^{-1}.$$

In particular, the following quotient Horton's laws hold

$$\lim_{K\to\infty}\frac{\mathcal{N}_j[K]}{\mathcal{N}_{j+1}[K]} = \lim_{k\to\infty}\frac{M_{k+1}}{M_k} = \lim_{k\to\infty}\frac{A_{k+1}}{A_k} = R_B = w_0^{-1} > 1, \tag{59}$$

$$\lim_{k\to\infty}\frac{L_{k+1}}{L_k} = \lim_{k\to\infty}\frac{S_{k+1}}{S_k} = R_L = c. \tag{60}$$

Proof Observe that Assumption 1 is satisfied with $\lambda = \alpha = 1$, recall that $w_0 < c^{-1}$, and apply Theorem 1, Corollary 1, and Propositions 3,4,5. In particular,

$$r_A = \min\{r_M,\, r_D\} = \min\big\{w_0,\, \min\{1, r_T = c^{-1}\}\big\} = w_0,$$
$$r_L = \min\{1, c^{-1}\} > w_0, \quad r_S = \min\{1, c^{-1}\} > w_0.$$

Finally, we observe that the total length $L_K^{\rm tot}$ of a tree of order K is treated similarly to the contributing area A_K, with α_k replaced by ℓ_k in definition (37). This means that under the hydrology-motivated Assumption 3 of this section, the sequence $L_K^{\rm tot}$ has the same Horton's law as A_K.

8 Self-Similar Hierarchical Branching Processes (HBP)

A flexible model that generates trees with arbitrary Tokunaga sequences T_k has been introduced and discussed by Kovchegov and Zaliapin (2018, 2020) and Kovchegov et al. (2021); it is called *Hierarchical Branching Processes* (HBP). Here we describe a self-similar version of the HBP.

Definition 1 (Self-similar Hierarchical Branching Process) *We say that $S(t)$ is a self-similar hierarchical branching process with a Tokunaga sequence $\{T_k\}$, and parameters $p \in (0,1)$, $\gamma > 0$ and $\zeta > 0$ if $S(t)$ is a multi-type branching process that develops in continuous time $t > 0$ according to the following rules:*

(i) The process $S(t)$ starts at $t = 0$ with a single progenitor (root branch) whose Horton-Strahler order (type) is $K \geq 1$ with probability $\pi_K = p(1-p)^{K-1}$.
(ii) Every branch of order $j \leq K$ produces offspring (side branches) of every order $i < j$ with rate $\gamma\,\zeta^{1-j}T_{j-i}$.
(iii) A branch of order j terminates with rate $\gamma\,\zeta^{1-j}$.
(iv) At its termination time, a branch of order $j \geq 2$ splits into two independent branches of order $j-1$.
(v) A branch of order $j = 1$ terminates without leaving offspring.
(vi) Generation of side branches and termination of distinct branches are independent.

8.1 Properties of Self-similar HBP Trees

The trajectories of the HBP are random trees from the space $\mathcal{L}$ of binary trees with edge lengths. A random tree generated by an HBP is called an HBP tree. Each process parameter completely specifies a particular attribute of a random HBP tree: the Tokunaga sequence T_k specifies the combinatorial structure of a tree of a given order; the probability p specifies the frequencies of trees of different orders; the constant ζ specifies the ratio of the mean lengths of branches of consecutive orders; and the rate γ specifies a unit of measurement for the tree edges. It has been shown by Kovchegov and Zaliapin (2018) that a random HBP tree T has the following properties:

Self-similarity. T is a self-similar tree with the Tokunaga sequence T_k and scaling constant ζ. This means that the distribution of T is invariant with respect to the Horton pruning as in (22), and the edge lengths of the random tree scale by ζ^{-1} after each Horton pruning.

Side-branching. The number $N[b] \geq 0$ of side-branches within a branch b of order K has geometric distribution

$$\mathsf{P}(N[b] = k) = q(1-q)^k, \quad \text{with} \quad q = S_K^{-1}, \tag{61}$$

where S_K is defined in (33). This implies, in particular, $\mathsf{E}[N[b]] = S_K - 1$.

Side-branch orders. Let $N_i[b] \geq 0$ be the number of side branches of order i within branch b. Conditioned on the total number $N[b]$ of side branches, the distribution of vector $(N_1[b], \ldots, N_{K-1}[b])$ is multinomial with $N[b]$ trials and success probabilities

$$\mathsf{P}(\text{side branch has order } i) = \frac{T_{K-i}}{S_K - 1}. \tag{62}$$

Branch and edge lengths. The length of an order K branch has exponential distribution with rate $\gamma\zeta^{1-K}$. The corresponding edge lengths ξ_K are i.i.d. exponential random variables with rate

$$\gamma\zeta^{1-K} S_K. \tag{63}$$

Accordingly,

$$\frac{\mathsf{E}[\xi_{k+1}]}{\mathsf{E}[\xi_k]} = \frac{\ell_{k+1}}{\ell_k} = \zeta \frac{S_k}{S_{k+1}}. \tag{64}$$

This means that the quotient Horton's law holds for S_k if and only if the quotient Horton's law holds for ℓ_k, and in this case $R_S = \zeta\lambda^{-1}$. The same equivalence holds for the root and geometric Horton's laws.

Geometric Horton's law for branch lengths. Definition 1(iii) implies the geometric Horton's law for the branch lengths L_k with $R_L = \zeta$. In fact, here we have a stronger statement that holds for any $k \geq 1$ (not only in the limit of large k):

$$L_k R_L^{-k} = \frac{\zeta^{k-1}}{\gamma}\zeta^{-k} = \frac{1}{\zeta\gamma}. \tag{65}$$

It follows, in particular, that

$$\frac{L_{k+1}}{L_k} = \zeta \quad \text{for any } k \geq 1.$$

Distribution of subtrees. Consider a random HBP tree T and fix K such that $K > 1$ and K is less than or equal to the order of T. Select a uniform random vertex $v \in T$ of order K, and consider two planted trees T_a and T_b descendant to v in T that have v as their root. Informally, we consider the pair of sibling trees at a random vertex of order K. It has been shown by (Kovchegov and Zaliapin, 2020, Lem. 16) that the joint distribution of the ordered statistics (K_1, K_2) of the orders (K_a, K_b) of these trees is given by

$$\mathsf{P}\left(K_1 = j, K_2 = m | K\right) = \begin{cases} S_K^{-1} & \text{if } j = m = K-1, \\ T_{K-j} S_K^{-1} & \text{if } j < m = K. \end{cases} \tag{66}$$

Moreover, the trees T_a and T_b are also HBP trees with the same parameters (γ, ζ) as T and orders given by (66). This result is essential for a fast recursive construction of T described in Sect. 8.2.

Independence of branches. Distinct branches have independent structure.

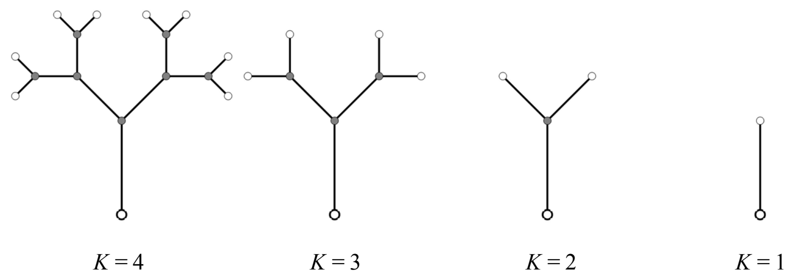

Fig. 7 Examples of perfect planted binary trees of orders $K = 1, ..., 4$.

8.2 Simulation of Self-similar HBP Trees

This section describes three algorithms for constructing HBP trees that do not involve time-dependent simulations. Each algorithm constructs a tree of a given order K. To construct a random HBP tree, one first generates a random order $K \geq 1$ according to π_K of Definition 1(i) and then constructs a tree of order K using either of the algorithms. Examples of HBP trees are shown in Fig. 4.

8.2.1 Algorithm A: Recursion by Tree Depth

This is the most straightforward algorithm that relays on the property (66) of subtree distributions described above in Sect. 8. The tree is constructed recursively, starting from the root and adding two principal subtrees at every recursion step.

Formally, a tree of order $K = 1$ consists of two vertices (root and leaf) connected by an edge of exponential length with rate γ. To construct a tree of a given order $K \geq 2$ we first use (66) to obtain the orders of its two principal subtrees T_a and T_b rooted at the stem vertex farthest from the root. The stem edge has exponential length with parameter $\gamma\zeta^{1-K}S_K$. To this stem we attach two trees, each of which is generated using the same approach via recursion.

Observe that the main branch of order K consists of a finite number of edges with probability one, meaning that after a finite number of recursion steps the two principal trees T_a and T_b will both have order $K - 1$. Extending this argument to consequently smaller orders we find that this recursive procedure does stop at a finite number of steps.

The algorithm is linear with respect to the tree size N (number of tree vertices) – it takes $O(N)$ time units to generate a tree of size N, or, equivalently, $O(R_B^K)$ time units to generate a tree of order K. Figure 8 illustrates the relations among the tree size N, order K, and generation timing t in seconds for the critical Tokunaga trees with $c = 2$ and orders $2 \leq K \leq 10$. This corresponds to sizes between $N = 4$ and $N = 1\,266\,454$. We generated 100 trees of each order. The time to create the largest trees using a 3.5 GHz desktop is about 100 sec.

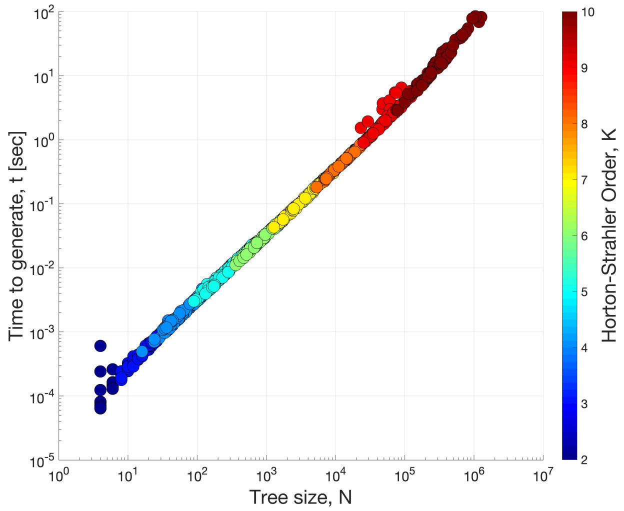

Fig. 8 Time t (in seconds) taken to generate HBP trees using the recursive Algorithm A of Sec. 8.2.1. The time t is shown as a function of tree size (number of vertices) N. Color represents the tree order K (see colorbar). This experiment corresponds to the critical Tokunaga process with $\gamma = 1$ and $c = 2$, which is equivalent to critical binary Galton-Watson tree with exponential edge lengths. We generated 100 trees of each order $2 \leq K \leq 10$. The computations were performed in Matlab on an Apple Desktop 3.5 GHz 6-Core Intel Xeon E5 with 32GB memory.

For example, a tree of order $K = 8$ has theoretical mean size $2\mathcal{N}_1[8] = 21,846$ according to (73). In our simulations the mean size of order-8 trees is $\bar{N} = 21\,445$, and the mean running time is $\bar{t} = 0.78$ sec.

It is clear from the algorithm description that the running time only depends on the Tokunaga sequence T_k and order K via the tree size, and is independent of the parameters γ and ζ.

8.2.2 Algorithm B: Recursion by Tree Order

This algorithm is given in Kovchegov and Zaliapin (2020). It uses a recursion by tree order – we start with a perfect binary tree of order K and add its side-branches of smaller orders.

Formally, a tree of order $K = 1$ consists of two vertices (root and leaf) connected by an edge of exponential length with rate γ. Assume now that we know how to construct a random tree of any order below $K \geq 2$. To construct a tree of order K, we start with a perfect (combinatorial) planted binary tree of depth K, which we call *skeleton*. The combinatorial shapes of such trees is illustrated in Fig. 7. All leaves in the skeleton have the same depth K, and all vertices at depth κ such that $1 \leq \kappa \leq K$ have the same Horton-Strahler order $K - \kappa + 1$. The root (at depth 0) has order K. Next, we assign lengths to the branches of the skeleton. Observe that each branch in a perfect tree consists of a single edge. To assign length to a branch b of order κ, with $1 \leq \kappa \leq K$, we generate a geometric number $N[b]$ according to (61) with parameter $q = S_\kappa^{-1}$ and then $N[b] + 1$ i.i.d. exponential lengths $\xi_{\kappa,i}$, $i = 1, \dots, N[b] + 1$, with the common rate $\gamma\zeta^{1-\kappa} S_\kappa$ according to (63). The total length of the branch b is $\xi_{\kappa,1} + \cdots + \xi_{\kappa,N[b]+1}$. The branch b has $N[b]$ side branches that are attached along b with spacings $\xi_{\kappa,i}$, starting from the branch point closest to the root. The order assignment for the side branches is done according to (62). We generate side branches (each has order below K) independently and attach them to the branch b. This completes the construction of a random tree of order K.

8.2.3 Algorithm C: Random Attachment Model

Here we construct a Markov tree process $\{\Upsilon_K\}_{K=1,2,\dots}$ corresponding to the HBP $S(t)$ following (Kovchegov and Zaliapin, 2020, Sect. 7.6). Each tree Υ_K is distributed as a tree generated by the self-similar HBP with Tokunaga sequence $\{T_k\}$ and parameters (γ, ζ), conditioned on its Horton-Strahler order being equal to K, and with its edge lengths scaled by ζ^{1-K}. This scaling is needed to ensure that $\Upsilon_K \subset \Upsilon_{K+1}$, when we consider each tree as a metric space of points connected by paths along the tree edges (Kovchegov and Zaliapin, 2020, Sect. 2.2). Accordingly, there exists the limit space, which informally can be considered an "infinite tree":

$$\Upsilon_\infty = \lim_{K\to\infty} \Upsilon_K = \bigcup_{K=1}^{\infty} \Upsilon_K.$$

Section 10 uses this construction to find the fractal dimension of Υ_∞.

Let Υ_1 be an I-shaped tree of Horton-Strahler order one, with the edge length distributed as an exponential random variable with parameter γ. Conditioned on Υ_K, the tree Υ_{K+1} is constructed according to the following transition rules. We attach new leaf edges to Υ_K at the points sampled by an inhomogeneous Poisson point process with the intensity $\rho_{j,K} = \gamma\zeta^{K-j} T_j$ along the edges of order $j \leq K$ in Υ_K. We also attach a pair of new leaf edges to each of the leaves in Υ_K. The lengths of all the newly attached leaf edges are i.i.d. exponential random variables with parameter $\gamma\zeta^K$ that are independent of the combinatorial shape and the edge lengths in Υ_K. Finally, we let the tree Υ_{K+1} consist of Υ_K and all the attached leaves and leaf edges.

By construction, a branch of order j in Υ_K becomes a branch of order $j+1$ in Υ_{K+1} after the attachment of new leave edges. The length of order j branch in Υ_K (and therefore, the length of order $j+1$ branch in Υ_{K+1}) is exponential random variable with parameter $\gamma\zeta^{K-j}$. Therefore, in a tree Υ_{K+1}, the number $n_{1,j+1}(K+1)$ of side-branches of order one in a branch of order $j+1$ has geometric distribution:

$$\begin{aligned}\mathsf{P}\big(n_{1,j+1}(K+1)=r\big) &= \frac{\gamma\zeta^{K-j}}{\gamma\zeta^{K-j}+\rho_{j,K}}\left(\frac{\rho_{j,K}}{\gamma\zeta^{K-j}+\rho_{j,K}}\right)^r \\ &= \frac{1}{1+T_j}\left(\frac{T_j}{1+T_j}\right)^r, \qquad r=0,1,2,\ldots \end{aligned} \tag{67}$$

with the mean value

$$\mathsf{E}\big[n_{1,j+1}(K+1)\big] = \frac{\rho_{j,K}}{\gamma\zeta^{K-j}} = T_j.$$

Therefore, after $i \geq 1$ rounds of attachments the mean number $n_{i,j+i}(M)$ of side-branches of order i in a branch of order $j+i$ in a tree Υ_M (where $M=K+i$ and $K \geq j$) is

$$\mathsf{E}\big[n_{i,j+i}(M)\big] = T_j.$$

8.2.4 Comparison of the Algorithms

The Algorithms A and B are best suited for numerical simulations of HBP trees, while the Random Attachment Model of Algorithm C has mainly a theoretical value. The Algorithm A (Sect. 8.2.1) is slightly simpler than the Algorithm B (Sect. 8.2.2), as it only involves generating a single edge length and merging two trees. However, the recursion by tree depth used in Algorithm A could make it computationally prohibitive. Heuristically, the expected value of the tree depth $\mathsf{depth}(T)$ can be approximated by the sum of combinatorial lengths (number of edges) of all orders:

$$\mathsf{E}[\mathsf{depth}(T)] \approx \sum_{k=1}^{K} S_k \sim \text{Const.} \times c^K \quad \text{as} \quad K \to \infty.$$

The expression $a_k \sim b_k$ as $k \to \infty$ means that $\lim\limits_{k\to\infty} a_k/b_k = 1$. This gives a coarse estimate on the recursion depth that is required to successfully use Algorithm A in generating large trees. For example, a critical binary Galton-Watson tree corresponds to $c=2$. Hence, the depth of a tree of order $K=10$ is about $2^{10} = 1\,024$. The mean size of such a tree is 349 525. All simulations in this work have been done using Algorithm A.

9 Critical Tokunaga Tree: A Model for River Networks

Recall that the analysis of the observed river networks suggests that the distribution of edge lengths is independent of their position within a tree (see Sect. 7.3). Formally, this corresponds to the assumption that the edge lengths ξ_k are i.i.d. random variables. For the HBP model, this assumption is satisfied only for a particular one-parametric class of trees, called *critical Tokunaga process*, that we describe in this section. The critical Tokunaga trees enjoy many additional symmetries as discussed by Kovchegov and Zaliapin (2018, 2019, 2020). The class is sufficiently broad and includes the critical binary Galton-Watson process with exponential edge lengths as a special case. Table 2 summarizes the main findings for the critical Tokunaga process and lists the values of its essential exponents fit to the observed river networks. Several examples of critical Tokunaga trees are shown in Fig. 4.

Definition 2 (Critical Tokunaga process) *For given $\gamma > 0$ and $c \geq 1$, we say that a self-similar hierarchical branching process $S(t)$ is a critical Tokunaga process with parameters (γ, c) if*

$$p = \frac{1}{2}, \quad \zeta = c, \quad \textit{and} \quad T_k = (c-1)\,c^{k-1}. \tag{68}$$

Similarly to the general HBP, we call a random tree generated by the critical Tokunaga process a critical Tokunaga tree. The parameter c completely determines the combinatorial structure of a random critical Tokunaga tree of a given order and scaling of the mean branch lengths. The rate γ specifies a unit of measurement.

It has been shown by (Kovchegov and Zaliapin, 2020, Thm. 14) that the critical Tokunaga process has the unit mean progeny, which explains the name critical, as in critical Galton-Watson process. These authors also have shown (Kovchegov and Zaliapin, 2020, Thm. 15) that the critical Tokunaga process with parameters (γ, $c = 2$) is equivalent to the critical binary Galton-Watson process with edge lengths distributed as independent exponential random variables with rate γ. One can also observe that $c = 1$ corresponds to $T_k = 0$, which results in a perfect binary tree (no side branching).

9.1 Horton's Laws for Mean Attributes

According to Def. 2, the distribution of orders in the critical Tokunaga process is $\pi_K = 2^{-K}$. The mean number of edges within a branch of order K is $S_k = 1+T_1+\cdots+T_{k-1} = c^{k-1}$. The edge lengths are i.i.d. exponential variables with common rate γ. Accordingly, $\ell_k = \gamma^{-1}$ and $L_k = \gamma^{-1}S_k = \gamma^{-1}c^{k-1}$. While the HBP does not have formally defined areas, the discussion in Sect. 7.3 and Example 1 suggest that one can set $\alpha_k = b\Gamma(\sigma+1)\gamma^{-\sigma}$ for some $b > 0$ and $\sigma > 0$.
The generating function $\hat{t}(z)$ is given by

$$\hat{t}(z) = -1 + 2z + \sum_{k=1}^{\infty}(c-1)c^{k-1}z^k = \frac{(1-2cz)(z-1)}{1-cz}. \tag{69}$$

The real root of $\hat{t}(z)$ nearest to the origin is $w_0 = (2c)^{-1}$. Assumption 3 trivially holds, and Proposition 6 implies the existence of the quotient Horton's laws for $\mathcal{N}_k[K]$, M_k, A_k, L_k, and S_k, with Horton exponents

$$c = R_L = R_S < R_M = R_N = R_A = 2c. \tag{70}$$

Stronger results are readily obtained by examining the generating functions. The generating function for the mean branch magnitudes M_k is

$$M(z) = -\frac{z}{\hat{t}(z)} = -\frac{z(1-cz)}{(1-2cz)(z-1)}, \tag{71}$$

and by Theorem 1

$$\mathcal{N}_{K-k+1}[K] = M_k = \frac{(2c)^k}{4c-2} + o\left((2c)^k\right). \tag{72}$$

An exact expression has been obtained using more powerful martingale techniques by (Kovchegov and Zaliapin, 2020, Eq. (133) of Cor. 4):

$$\mathcal{N}_{K-k+1}[K] = M_k = \frac{(2c)^k + 2c - 2}{4c-2}. \tag{73}$$

The generating function $D(z)$ for the mean differential contributing areas is

$$D(z) = \sum_{k=1}^{\infty} \alpha_k S_k z^k = d_0 \sum_{k=1}^{\infty} c^{k-1} z^k = \frac{d_0 z}{1-cz}, \tag{74}$$

with $d_0 = b\Gamma(\sigma+1)\gamma^{-\sigma}$. This leads to

$$A(z) = -\frac{D(z)}{\hat{t}(z)} = -\frac{d_0 z}{(1-2cz)(z-1)}. \tag{75}$$

Proposition 1 gives the asymptotic form of A_k:

$$A_k = \frac{d_0(2c)^k}{2c-1} + o\left((2c)^k\right). \tag{76}$$

According to this discussion, the geometric Horton's law (not only the quotient Horton's law) holds for $\mathcal{N}_k[K]$, M_k and A_k.

Comparing (72) and (76) we find

$$A_k \sim 2d_0 M_k \quad \text{as} \quad k \to \infty. \tag{77}$$

Figure 9(a) shows the relation among the contributing areas and magnitudes of vertices in a critical Tokunaga tree with $c = 2.5$ of order $K = 10$. This tree has 2 440 508 vertices. The asymptotic relation (77), which becomes in this case $A = 4M$, is closely followed for the high-order vertices, $K \geq 6$. At the same time, the low-order vertices may show pre-asymptotic behavior that results in substantial deviations from the asymptotic approximation $A = 4M$. Panel (b) shows the multiplicative error $A/(4M)$ as a function of the vertex order. One notices extreme errors, up to several orders of magnitude, for the orders below $K = 5$. This experiment shows that the conventional approximation $A \propto M$ has a substantial error at low orders and may be misleading.

Figure 10 illustrates Horton's laws for $N_k[T]$, M_k, A_k, and S_k in three critical Tokunaga trees or order $K = 9$ with $\gamma = 1$. We consider the perfect

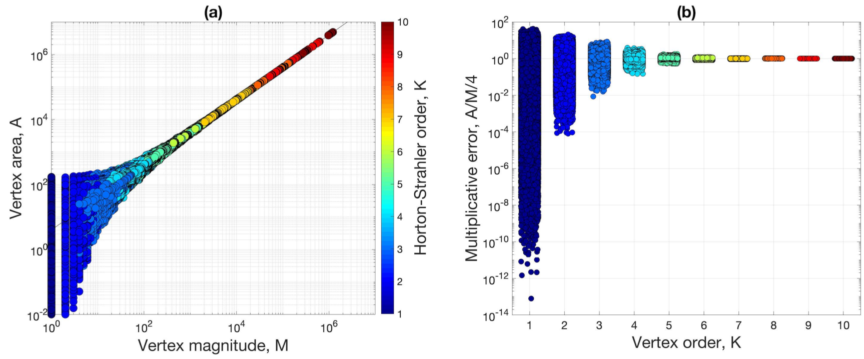

Fig. 9 Relation between vertex contributing area $A_{(i)}$ and vertex magnitude $M_{(i)}$ in a critical Tokunaga tree with $c = 2.5$ of order $K = 10$. (a) $A_{(i)}$ as a function of $M_{(i)}$. The theoretical asymptotic relation (77), $A = 4M$, is shown by black line. Colorcode corresponds to vertex order (see colorbar). The vertical axis is trimmed at 10^{-2}, although the minimal area is 3×10^{-13}. (b) Multiplicative error $A_{(i)}/(4M_{(i)})$ of the asymptotic approximation: unity corresponds to a perfect fit. The horizontal coordinates are given with a uniform jitter. Colorcode is the same as in panel (a), it duplicates the horizontal coordinate.

binary tree ($c = 1$, blue circles), a critical binary Galton-Watson tree ($c = 2$, red squares), and a tree with $c = 3$ (green triangles). Panel (a) illustrates Horton's law for the branch counts $N_k[T]$. Here the fitting lines start at $N_1[T]$ and have a constant slope of $-\log_{10}(2c)$ suggested by (78). We notice a very close fit for all examined orders.

Panel (b) refers to Horton's law for the average magnitudes M_k. The fitting lines are taken from (73) – the fit is ideal for the deterministic perfect binary tree (blue) and is very close for the two random trees (red and green). The only visible deviations from the theoretical quantities are observed for the high-order branches ($K > 6$) that correspond to small-sample averaging. For instance, recall from panel (a) that we only have an average of 3 branches of order $K = 8$ and 11 branches of order $K = 7$ in a critical binary Galton-Watson tree (red squares).

Panel (c) illustrates Horton's law for the average contributing areas A_k. We assume here that the differential contributing areas ν_k of edges are determined by the edge lengths ξ_k via $\nu_k = \xi_k^\sigma$ with $\sigma = 2$ and the contributing area of a branch is the sum of the differential contributing areas of all its descendant edges; see Sect. 7.3 for a discussion and examples of this approach. The fitting lines here correspond to the asymptotic expression (76) with $d_0 = \gamma^{-\sigma}\Gamma(\sigma+1) = \Gamma(3) = 2$. We expect them to fit the observed values for the intermediate range of orders – when the asymptotic approximation already works yet the sample size (number of branches of a given order) is still large enough. In random trees (red and green), the fitting lines provide almost perfect fit to the data for orders $2 \leq K \leq 6$ and show very small deviations at the higher orders. For the perfect binary tree (with random edge lengths and areas), where $c = 1$, the best fit is for the largest orders $K > 5$. The discrepancy at the low orders $K < 4$ is related to the fact that the asymptotic expression (76)

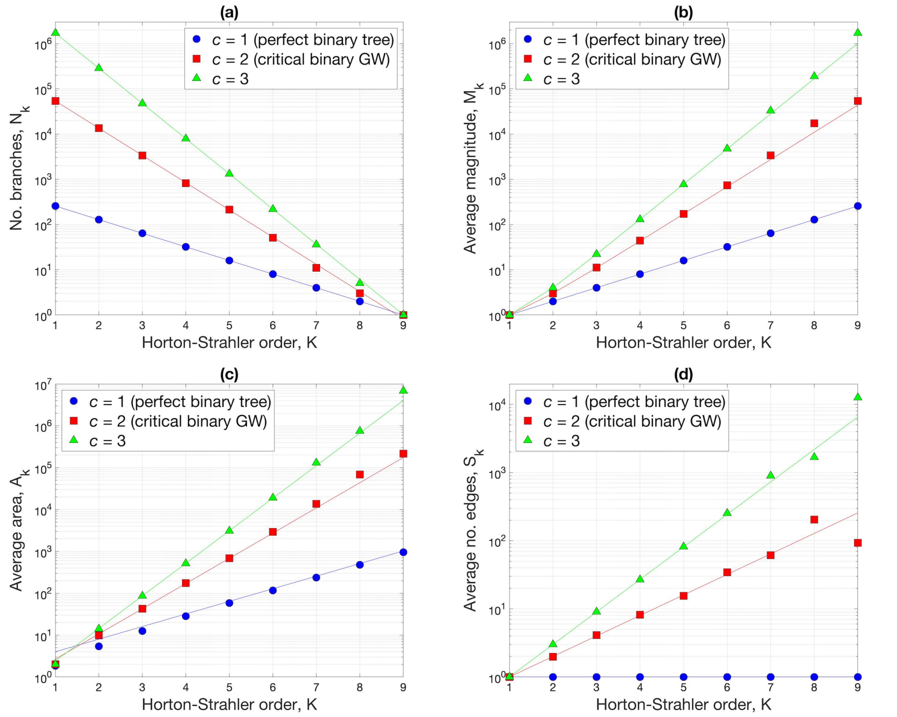

Fig. 10 Horton's laws in critical Tokunaga trees (Sec. 9). The figure illustrates three cases: a perfect binary tree, $c = 1$ (blue circles); a critical binary Galton-Watson tree, $c = 2$ (red squares); and $c = 3$ (green triangles). Symbols correspond to the average attributes estimated in a single realization of a tree. Lines correspond to theoretical predictions (see below). (a) Branch counts $N_k[T]$. Lines start at $N_1[T]$ and have the theoretical slopes $-\log_{10}(2c)$ of (72). (b) Average branch magnitudes M_k. Lines show the theoretical means of (73). (c) Average contributing areas A_k. Lines show the theoretical means of (76). (d) Average number of edges S_k. Lines show the theoretical means of (33).

suggests $A_1 \approx 2d_0c/(2c-1) = 4$, while the actual mean here (for any c) is $A_1 = \mathsf{E}[\xi^2] = 2$, with ξ being an exponential random variable with rate $\gamma = 1$. In general, observe that

$$A_1 \approx d_0 \frac{2c}{2c-1} \to d_0, \quad \text{as} \quad c \to \infty,$$

so the asymptotic expression (76) does provide a good fit to the data at low orders for large enough c, which we do observe for $c = 2, 3$ (red and green).

Panel(d) refers to Horton's law for the average number of edges S_k. The fitting lines show the theoretical values $S_k = 1 + T_1 + \cdots + T_{k-1}$. It follows from the properties of the HBP model that this is merely an exercise in sampling from the geometric distribution. The fit quality depends on the sample size and is very good for orders $K < 7$. The deviations at higher orders are due to small sample size. The perfect binary tree ($c = 1$, blue) has all branches consisting of a single edge.

The condition $T_{i,i+k} = T_k = a\,c^{k-1}$, which is slightly more general than that of (68), was first introduced in hydrology by Eiji Tokunaga (1978) in a study of river networks, hence the process name. In the present work, the constraint $a = c - 1$ is necessitated by the equality of the mean edge lengths, which requires the sequence λ_j to be geometric. The sequence of the Tokunaga coefficients then also has to be geometric, and satisfy $a = c - 1$. Interestingly, the constraint $a = c - 1$ appears in the *random self-similar network* (RSN) model introduced by Veitzer and Gupta (2000), which uses a purely topological algorithm of recursive local replacement of the network generators to construct random self-similar trees. Results of Chunikhina (2018a,b) imply that the critical Tokunaga model with $c = 2$ maximizes the entropy rate among the trees that satisfy the quotient Horton's law of stream numbers, and that the critical Tokunaga model with a fixed c maximizes the entropy rate among the trees that satisfy the quotient Horton's law for stream numbers with $R_B = 2c$.

9.2 Limit Laws for Random Branch Numbers

While this review mainly focuses on the mean branch numbers $\mathcal{N}_k[K]$, much stronger, distributional, results are available for the random branch numbers $N_k[T]$ in critical Tokunaga trees. This section assumes that we consider a critical Tokunaga tree with parameter c. The results refer to the combinatorial tree structure, and hence hold for an arbitrary measurement unit $\gamma > 0$.

Let Δ_K be a critical Tokunaga tree of order K, then the following Weak Law of Large Numbers holds (Kovchegov and Zaliapin, 2020, Sect. 7.6.3, Cor. 5): for any $k \geq 1$ we have

$$\frac{N_k[\Delta_K]}{N_{k+1}[\Delta_K]} \xrightarrow{p} R_B = 2c \quad \text{as} \quad K \to \infty, \tag{78}$$

where $\xrightarrow{p}$ denotes convergence in probability. Moreover, using the notations of Sect. 8.2.3 one can establish the following Strong Law of Large Numbers (Kovchegov and Zaliapin, 2020, Sect. 7.6.3, Thm. 16): for any $k \geq 1$ we have

$$\frac{N_k[\Upsilon_K]}{N_{k+1}[\Upsilon_K]} \xrightarrow{\text{a.s.}} R_B = 2c \quad \text{as} \quad K \to \infty, \tag{79}$$

where $\xrightarrow{\text{a.s.}}$ denotes convergence with probability one (almost sure convergence). The difference between these two laws is that (78) considers a sequence of independent trees Δ_K, while (79) refers to a sequence of trees Υ_K related such that $\mathcal{R}(\Upsilon_K) = \Upsilon_{K-1}$. The following distributional geometric Horton's law also holds (Kovchegov and Zaliapin, 2020, Sect. 7.6.3, Cor. 16):

$$R_B^{1-K} N_1[\Upsilon_K] \xrightarrow{\text{a.s.}} V_\infty(\Upsilon_\infty) \quad \text{as} \quad K \to \infty, \tag{80}$$

where $V_\infty(\Upsilon_\infty)$ is a finite and positive value that depends on a particular realization of the Markov process Υ_K. This random value allows different random trees to have very different sizes, while preserving the relative frequencies of branches of different orders. Specifically, the frequencies of branches of different orders in a tree of order K are approximated by a geometric distribution:

$$\frac{\#\{\text{branches of order } k\}}{\#\{\text{branches}\}} = (R_B - 1)R_B^{-k}(1 + o(1)) \tag{81}$$

for any $k \geq 1$ as $K \to \infty$.

Finally, one can prove the following Central Limit Theorem:

Theorem 2 (CLT for Branch Numbers in Critical Tokunaga Trees) *Let Δ_K be a critical Tokunaga tree of order $K > 0$ with parameter $c \geq 1$. Then, for every integer $k \geq 1$,*

$$\sqrt{N_1[\Delta_K]}\left(\frac{N_{k+1}[\Delta_K]}{N_k[\Delta_K]} - \frac{1}{2c}\right) \stackrel{d}{\to} \mathsf{N}\Big(0,\, (c-1)(2c)^{k-3}\Big) \quad \text{as} \quad K \to \infty, \tag{82}$$

where $\mathsf{N}(\mu, \sigma^2)$ denotes a Normal distribution with mean μ and variance σ^2, and $\stackrel{d}{\to}$ denotes convergence in distribution.

In a special case of the critical Tokunaga tree with $c = 2$, which is equivalent to the critical Galton-Watson process, similar limit results were established by Wang and Waymire (1991), Yamamoto (2017), and Kovchegov and Zaliapin (2020). These authors considered a random tree conditioned on the number N_1 of leaves and established the conditional limits of N_{k+1}/N_k as $N_1 \to \infty$. We notice that, trivially, $K \to \infty$ implies $N_1 \to \infty$ since a binary tree of order K has at least 2^{K-1} leaves. Hence, the limit results for an increasing order K follow from those for an increasing tree magnitude N_1. Vice versa, it can be shown that increasing N_1 implies an increasing order K. In summary, the above limit laws are equivalent under both forms of increasing tree size in critical Tokunaga trees. This equivalence issue will be treated formally elsewhere.

10 Fractal Dimension of Self-Similar HBP Trees

Consider a self-similar HBP $S(t)$ (Def. 1) with a Tokunaga sequence $\{T_k\}$ satisfying $\limsup_{k\to\infty} T_k^{1/k} < \infty$, and parameters $\gamma > 0$ and $\zeta > 1$. We use here the Random Attachment Model representation of the HBP process discussed in Sect. 8.2.3. The self-similarity of the HBP process (Sect. 8) suggests that the limit space Υ_∞ does not change its statistical properties after rescaling, which corresponds here to the Horton pruning. Let $\mathbf{d}$ denote its fractal dimension. That the limit space includes at least the root branch Υ_1 suggests $\mathbf{d} \geq 1$. Assume that $\mathbf{d} > \mathbf{1}$. Then, denoting the mean $\mathbf{d}$-dimensional volume of Υ_∞ by $\mathbf{vol}$, we have

$$\mathbf{vol} = \sum_{k=1}^{\infty} t_k \frac{\mathbf{vol}}{\zeta^{\mathbf{d}k}}. \tag{83}$$

This equation is obtained by splitting a tree Υ_∞ into the subtrees attached to its highest-order branch Υ_1. There is an average of $t_1 = T_1 + 2$ subtrees

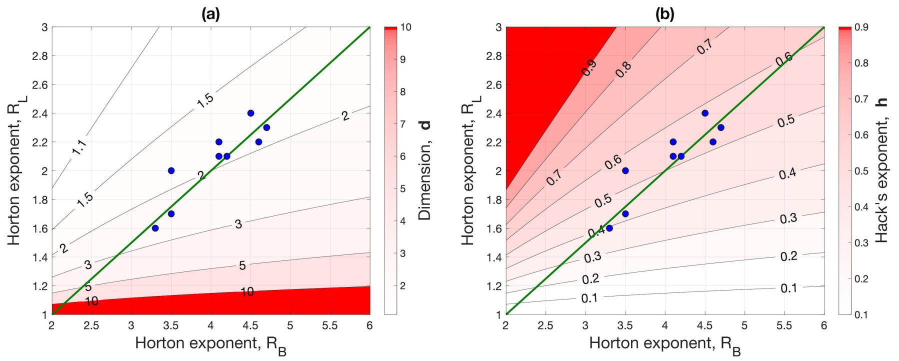

Fig. 11 Fractal dimension $\mathbf{d} = \max\{1, \frac{\log R_B}{\log R_L}\}$ (panel a) and Hack's exponent $\mathbf{h} = \mathbf{d}^{-1}$ (panel b) of self-similar HBP tree in the limit of infinite size as a function of the Horton exponents R_B and R_L. Selected levels of **d** and **h** are shown by marked black lines. Green line corresponds to the critical Tokunaga process of Sect. 9, for which $R_B = 2R_L$. Blue circles depict the pairs (R_B, R_L) estimated in nine real river basins by Tarboton et al. (1988), see also (Rodriguez-Iturbe and Rinaldo, 2001, Table 2.1).

distributed as Υ_∞ scaled by ζ^{-1}. In general, for each k, there will be an average of t_k subtrees distributed as Υ_∞ scaled by ζ^{-k}. Scaling the lengths by ζ^{-k} in the **d**-dimensional space results in scaling the volume by $\zeta^{-\mathbf{d}k}$. The **vol** term in (83) can be cancelled out, yielding

$$\hat{t}(\zeta^{-\mathbf{d}}) = 0, \tag{84}$$

and hence, $\zeta^{-\mathbf{d}} = w_0 = R_B^{-1}$. Finally, we find:

$$\mathbf{d} = \max\{1, \mathbf{d}_0\}, \quad \mathbf{d}_0 = -\frac{\log w_0}{\log \zeta} = \frac{\log R_B}{\log \zeta} = \frac{\log R_B}{\log R_L}. \tag{85}$$

This expression has been first obtained for river networks by La Barbera and Rosso (1989). Figure 11(a) illustrates the fractal dimension of self-similar HBP trees for $R_B \in [2, 6]$ and $R_L \in [1, 3]$, which are the ranges suitable for the studies of the observed river networks (Tarboton et al., 1988; Rodriguez-Iturbe and Rinaldo, 2001). The figure also shows the values (R_B, R_L) estimated for nine river basins by Tarboton et al. (1988). We notice the tendency of the estimated Horton exponents to cluster around the critical Tokunaga model for which $R_B = 2R_L$ (green line). Figure 12(a) illustrates the fractal dimension for the two-parameter Tokunaga model (Appendix E) with parameters (a, c). Here $\mathbf{d} = \frac{\log R_B}{\log c}$ and R_B is given by (120). To add to this plot the values from the observed river basins, for which the pairs (R_B, R_L) are known, we use $c = R_L$ and find a by solving (120):

$$a = R_B - c - 2 + \frac{2c}{R_B}. \tag{86}$$

Again, we see a tendency for the real basins to cluster around the critical Tokunaga model (green line) for which $a = c - 1$. Figure 13 shows the value

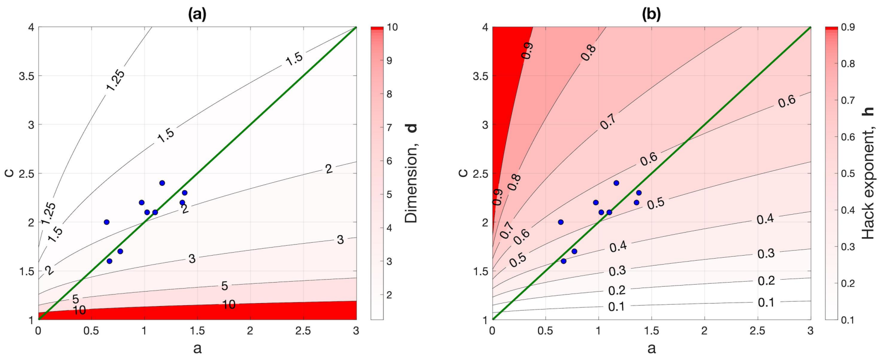

Fig. 12 Fractal dimension $\mathbf{d} = \max\{1, \frac{\log R_B}{\log c}\}$ (panel a) and Hack's exponent $\mathbf{h} = \mathbf{d}^{-1}$ (panel b) in the two-parameter Tokunaga model of Appendix E with parameters (a, c). Here $R_B = R_B(a, c)$ according to (120). Selected levels of $\mathbf{d}$ and $\mathbf{h}$ are shown by marked black lines. Green line corresponds to the critical Tokunaga process of Sect. 9, for which $a = c - 1$ and $R_B = 2c$. Blue circles depict the pairs (a, c) that correspond to the values (R_B, R_L) estimated in nine real river basins by Tarboton et al. (1988), see also (Rodriguez-Iturbe and Rinaldo, 2001, Table 2.1).

of the Horton exponent $R_B = R_B(a, c)$ and its reciprocal value $w_0 = w_0(a, c)$ according to (120).

Recalling that $R_B \geq 2$ and $\zeta > 1$ we find that the dimension $\mathbf{d}$ can take any value $\mathbf{d} \geq 1$. There exists an infinite collection of self-similar HBPs with a given value of $\mathbf{d}$, since there are infinitely many ways to select a Tokunaga sequence T_k with a given w_0. Recalling (70) we find the dimension of a critical Tokunaga tree with parameter $c > 1$ illustrated in Fig. 14(a):

$$\mathbf{d} = \frac{\log 2c}{\log c} = 1 + \log_c 2. \tag{87}$$

For the limit space Υ_∞ to be embedded into a plane, one need to ensure that $\mathbf{d} \leq 2$, which is equivalent to $w_0 \geq \zeta^{-2}$ or $R_B \leq \zeta^2$, and in the family of critical Tokunaga trees to $c \geq 2$. We also observe that the condition $w_0 = \zeta^{-2}$ corresponds to the space-filling tree with $\mathbf{d} = 2$ (see Newman et al. (1997)) and in the critical Tokunaga family this corresponds to $c = 2$ (see Kovchegov and Zaliapin (2018)), which is the critical binary Galton-Watson tree.

11 Hack's Law in Self-Similar HBP Trees

One of the fundamental scaling laws of hydrology is the Hack's law (Hack, 1957; Mesa and Gupta, 1987; Rigon et al., 1996; Rodriguez-Iturbe and Rinaldo, 2001) that relates the lengths L of the longest stream in a river basin to the basin contributing area A:

$$L \propto A^{\mathbf{h}}, \quad \mathbf{h} \approx 0.6. \tag{88}$$

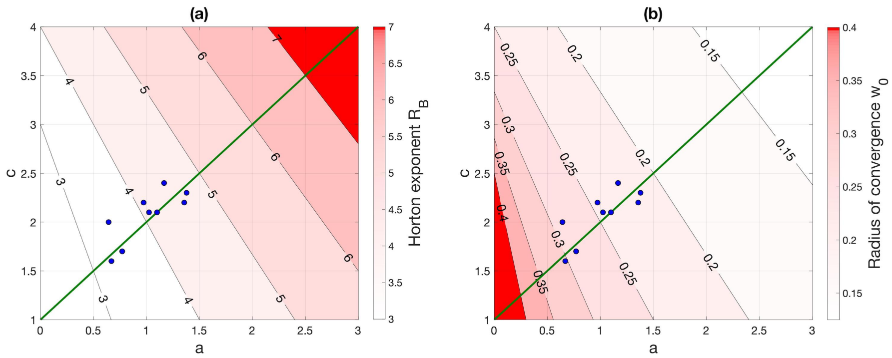

Fig. 13 Horton exponent R_B and the radius of convergence $w_0 = R_B^{-1}$ for $M(z)$ in the two-parameter Tokunaga model of Appendix E with parameters (a, c). Other notations are the same as in Fig. 12.

If T is the tree representing the stream network, then the length of the longest stream is the height of the tree T, denoted by $\textsc{height}(T)$ (Pitman, 2006; Kovchegov and Zaliapin, 2020). This section establishes the mean Hack's law in self-similar HBP trees.

Consider a tree T generated by a self-similar HBP with a Tokunaga sequence $\{T_k\}$ satisfying $\limsup_{k\to\infty} T_k^{1/k} < \infty$, and parameters $\gamma > 0$ and $\zeta > 1$. Let

$$\Lambda_k = \mathsf{E}\left[\textsc{height}(T) \,\Big|\, \mathsf{ord}(T) = k\right] \tag{89}$$

that represents the mean length of *the longest river stream* in a basin with the Horton-Strahler order k. Notice that, since $\textsc{height}\big(\mathcal{R}(T)\big) \le \textsc{height}(T)$,

$$\zeta\,\Lambda_{k-1} = \mathsf{E}\left[\textsc{height}\big(\mathcal{R}(T)\big) \,\Big|\, \mathsf{ord}(T) = k\right] \le \mathsf{E}\left[\textsc{height}(T) \,\Big|\, \mathsf{ord}(T) = k\right] = \Lambda_k. \tag{90}$$

Hence, since $\Lambda_1 = \gamma^{-1}$, we have $\Lambda_k \ge \gamma^{-1}\zeta^{k-1}$. Next, let

$$Y_1, Y_2, \ldots, Y_{N_1[T]}$$

denote the leaf lengths in the tree T. Then, since

$$\textsc{height}(T) \le \textsc{height}\big(\mathcal{R}(T)\big) + \max_{j=1,\ldots,N_1[T]} Y_j,$$

we have,

$$\begin{aligned}
\Lambda_k &\le \mathsf{E}\left[\textsc{height}\big(\mathcal{R}(T)\big) \,\Big|\, \mathsf{ord}(T) = k\right] + \mathsf{E}\left[\max_{j=1,\ldots,N_1[T]} Y_j \,\Big|\, \mathsf{ord}(T) = k\right] \\
&= \zeta\,\Lambda_{k-1} + \gamma^{-1}\mathsf{E}\left[\sum_{j=1}^{N_1[T]} \frac{1}{j} \,\Big|\, \mathsf{ord}(T) = k\right] \qquad (91) \\
&\le \zeta\,\Lambda_{k-1} + \gamma^{-1}\mathsf{E}\left[1 + \log\big(N_1[T]\big) \,\Big|\, \mathsf{ord}(T) = k\right] \\
&\le \zeta\,\Lambda_{k-1} + \gamma^{-1} + \gamma^{-1}\log\big(\mathsf{E}\big[N_1[T] \,\big|\mathsf{ord}(T) = k\big]\big) \qquad (92)
\end{aligned}$$

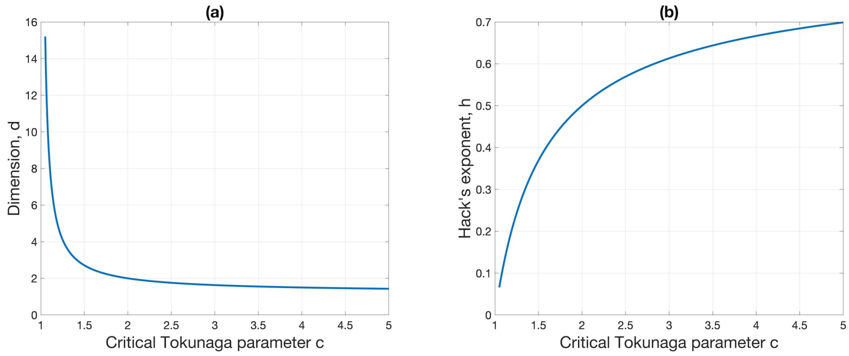

Fig. 14 Fractal dimension $\mathbf{d} = \frac{\log 2c}{\log c}$ (panel a) and Hack's exponent $\mathbf{h} = \mathbf{d}^{-1}$ (panel b) in the critical Tokunaga model with parameter $c > 1$.

by Wald's equation, the Coupon Collector Problem, and finally, the Jensen's inequality (Bhattacharya and Waymire, 2007). Recall the geometric Horton's law (43) for the leaf count in a self-similar process

$$\mathcal{N}_1[k] = M_k = M\,R_B^k + o\left(R_B^k\right).$$

Hence, equations (90) and (91) imply

$$0 \;\le\; \Lambda_k - \zeta\,\Lambda_{k-1} \;\le\; \gamma^{-1} k \log R_B + \beta$$

for some constant β, and

$$0 \;\le\; \frac{\Lambda_k}{\Lambda_{k-1}} - \zeta \;\le\; \gamma^{-1}\frac{k\log R_B + \beta}{\Lambda_{k-1}} \le \frac{k\log R_B + \beta}{\zeta^{k-2}} \to 0 \;\text{ as } k \to \infty. \quad (93)$$

Accordingly,

$$\log \Lambda_k = \sum_{j=2}^{k} \log\left(\frac{\Lambda_k}{\Lambda_{k-1}}\right) + \log \Lambda_1 = (k-1)\log\zeta + \sum_{j=2}^{k}\log(1+\mathcal{E}_j) - \log\gamma, \quad (94)$$

where $0 \le \mathcal{E}_j \le (k\log R_B + \beta)\zeta^{1-k}$, and therefore, $\sum\limits_{j=2}^{\infty}\log(1+\mathcal{E}_j)$ converges to a constant. We therefore conclude that the geometric Horton's law holds for Λ_k with Horton exponent $R_\Lambda = R_L = \zeta$:

$$\Lambda_k \sim \text{Const.} \times \zeta^k. \quad (95)$$

This and the geometric Horton's law for the mean branch magnitudes M_k implies the *Hack's law* for self-similar HBP:

$$\Lambda_k \sim \text{Const.} \times \left(M_k\right)^{\mathbf{h}}, \qquad \text{where} \quad \mathbf{h} = \frac{\log\zeta}{\log R_B}. \quad (96)$$

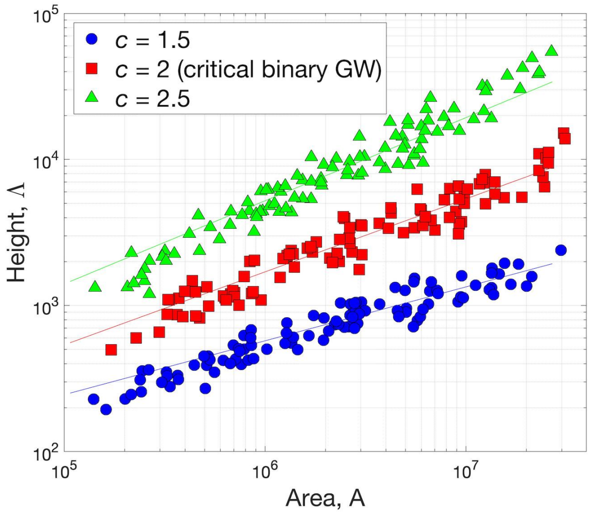

Fig. 15 Hack's law in critical Tokunaga trees with parameters $c = 1.5, 2, 2.5$ and $\gamma = 1$. Symbols correspond to individual simulated trees. Lines show theoretical slopes $\mathbf{h} = \log(c)/\log(2c)$. The differential contributing area ν of an edge is calculated via the edge length ξ as $\nu = \xi^2$. For each (c, γ) we simulated 100 independent trees of different orders: $11 \le K \le 14$ for $c = 1.5$, $9 \le K \le 11$ for $c = 2$, and $8 \le K \le 10$ for $c = 2.5$. In each simulated tree, the area $A_{(i)}$ and height $\Lambda_{(i)}$ are reported for a random vertex from the stream of the highest order.

An asymptotic equivalence $A_k \sim \text{Const.} \times M_k$ would imply the mean Hack's law in its classical form, relating the longest river channel to the basin's area:

$$\Lambda_k \sim \text{Const.} \times \left(A_k\right)^{\mathbf{h}}, \quad \text{where} \quad \mathbf{h} = \frac{\log \zeta}{\log R_B}. \tag{97}$$

Our analysis in Sect. 7.2 asserts that such equivalence holds as soon as $r_D > r_M = w_0$. This is so, for example, in the critical Tokunaga process of Sect. 9, or more generally under the hydrologic Assumption 3. The reciprocity of the Hack's law exponent $\mathbf{h}$ and the fractal dimension $\mathbf{d}$ has been heuristically established by Peckham (1995), La Barbera and Rosso (1989). Figures 11(b) and 12(b) show the exponent $\mathbf{h}$ in the self-similar HBP as a function of (R_B, R_L) and in the two-parameter Tokunaga model as a function of (a, c), respectively. The nine observed river basins from Tarboton et al. (1988) (blue dots) have the Hack's exponent within the range $0.39 < \mathbf{h} < 0.6$.

The Hack's law for the critical Tokunaga processes with parameter $c > 1$ takes the form

$$\Lambda_k \sim \text{Const.} \times \left(A_k\right)^{\mathbf{h}}, \quad \text{where} \quad \mathbf{h} = \frac{\log c}{\log(2c)} = \frac{1}{1 + \log_c 2}. \tag{98}$$

The value of $\mathbf{h} = \mathbf{h}(c)$ is illustrated in Fig. 14(b) for $1 < c < 5$. Here, for $c = 2.5$, $\mathbf{h} \approx 0.57\ldots$. Figure 15 illustrates the Hack's law in simulated critical

Tokunaga trees with $c = 1.5, 2, 2.5$ by showing the scatter between the tree contributing area A and height Λ.

12 Scaling Laws in Self-similar Trees

We have discussed in Introduction that Horton's laws imply a variety of power laws for the frequencies of edge attributes in a large tree and power-law relations between different attributes. This section illustrates these general observations with specific selected examples. We make here the hydrologic assumption of equality of the mean edge length ℓ_k for all k and equality of the differential edge contributing areas α_k for all k. We also assume a geometric Horton's law for T_k with $r_T < 1$. These assumptions are conventionally accepted in the hydrologic literature and are justified by field observations (Rodriguez-Iturbe and Rinaldo, 2001). The assumptions can be relaxed (with more technical work) if needed.

We observe that by Proposition 9 the areas A_k in this case satisfy geometric Horton's law with Horton exponent R_A such that $R_A = R_B = R_M$. Moreover, by Propositions 7,8 the combinatorial branch lengths S_k and metric branch lengths L_k also satisfy geometric Horton's laws with Horton exponents R_S and R_L such that $R_S = R_L < R_B = R_M$.

12.1 Power Laws for Edge Attributes

Recall (Sect. 3.5) that the geometric Horton's laws for branch counts $N_k[T]$, edge counts S_k, and an arbitrary branch attribute Z_k imply a power-law frequency distribution for the edge attribute $Z_{(i)}$ with power index given by (13). Section 7 establishes a variety of Horton's laws for a self-similar tree, including those for branch counts $N_k[T]$ and edge counts S_k. This means that any attribute that satisfies the geometric Horton's law is expected to have a power-law frequency distribution when examined on individual edges. The well-studied hydrological examples are power law frequency statistics for link contributing areas (14) and distance to the divide (15); see Tarboton et al. (1989), Rodriguez-Iturbe et al. (1992), and Maritan et al. (1996).

Consider a critical Tokunaga HBP with parameters $\gamma = 1$ and $c > 1$ (Sect. 9). According to (70) we have $R_B = R_A = 2c$ and $R_S = c$, hence establishing

$$\#\{i : A_{(i)} \geq x\} \;\propto\; x^{-\beta_A} \tag{99}$$

with

$$\beta_A = \frac{\log R_B - \log R_S}{\log R_A} = \frac{\log(2c) - \log c}{\log(2c)} = 1 - \frac{\log c}{\log(2c)} = 1 - \mathbf{h}. \tag{100}$$

Here, $c = 2.3$ corresponds to $\beta_A \approx 0.45$ reported in analyses of river networks (Rodriguez-Iturbe et al., 1992; Rodriguez-Iturbe and Rinaldo, 2001). The last equality, $\beta_A = 1 - \mathbf{h}$, holds in a general self-similar HBP tree under the

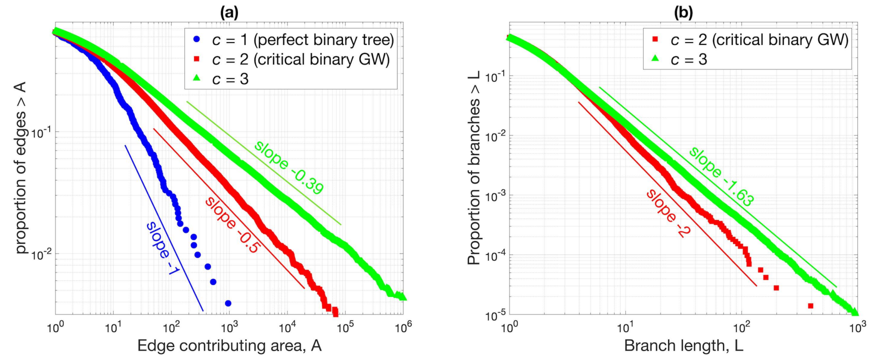

Fig. 16 Power laws in critical Tokunaga trees with parameters $c = 1, 2, 3$ and $\gamma = 1$. Symbols refer to the empirical counts. Lines show theoretical slopes. (a) Power law for the exceedance frequencies of edge contributing areas $A_{(i)}$. The lines show the theoretical slope $\beta_A = -(1 - \mathbf{h}(c))$, with convention $\mathbf{h}(1) = 0$. (b) Power law for the exceedance frequencies of branch lengths $\Lambda_{[i]}$. The lines show the theoretical slope $-\mathbf{d}(c)$.

assumptions listed in the beginning of Sect. 12. This relation is well-known in the analysis of river basins (Rodriguez-Iturbe and Rinaldo, 2001, Eq. (2.215)).

Figure 16(a) shows empirical exceedance frequencies for edge areas $A_{(i)}$ calculated in three critical Tokunaga trees with $c = 1, 2, 3$ and $\gamma = 1$. The distributions have power-law tails (seen as linear segments in the double logarithmic plot) with power indices given by $\beta_A(c) = 1 - \mathbf{h}(c)$, according to (100).

Similarly, the geometric Horton's law (95) for tree heights Λ_k implies a power-law distribution of the distances $\Lambda_{(i)}$ from link i to the most distant source along the tree (i.e., heights of the edges using the graph-theoretic terminology):

$$\#\{i : \Lambda_{(i)} \geq x\} \;\propto\; x^{-\beta_\Lambda} \tag{101}$$

with

$$\beta_\Lambda = \frac{\log R_B - \log R_S}{\log R_\Lambda} = \mathbf{d} - 1 = \frac{\beta_A}{\mathbf{h}}. \tag{102}$$

This distribution can be alternatively derived by writing the Hack's law (97) on individual edges:

$$\#\{i : \Lambda_{(i)} \geq x\} = \#\{i : A^{\mathbf{h}}_{(i)} \geq \text{Const.} \times x\} \propto x^{-\beta_A/\mathbf{h}}. \tag{103}$$

The relation (102) is well-documented in the analysis of natural river basins as discussed by (Rodriguez-Iturbe and Rinaldo, 2001, Sect. 2.9.3, Eq. (2.185)). In a critical Tokunaga tree with parameter c we have

$$\beta_\Lambda = \frac{\log(2c) - \log c}{\log c} = \log_c 2. \tag{104}$$

Here for $c = 2.3$ we have $\beta_\Lambda \approx 0.83$.

12.2 Power Laws for Branch Attributes

Tree self-similarity, and associated Horton's laws, also imply power laws for attributes calculated for random branches. The argument is very similar to that used to establish power laws for the edge attributes. We consider here the branch lengths $L_{[i]}$, where the lower bracketed index $[i]$ indicates that we calculate the length of a uniformly randomly selected branch (and distinguishes this attribute from the mean length L_K of an order-K branch). In a limit of a large tree,

$$\#\{i : L_{[i]} \geq R_L^K\} \;\propto\; \sum_{j=K}^{\infty} N_K = \sum_{j=K}^{\infty} R_B^{-j} \propto R_B^{-K}.$$

This is a punctuated (by discrete order) version of a general power law relation

$$\#\{i : L_{[i]} \geq z\} \;\propto\; z^{-\frac{\log R_B}{\log R_L}} = z^{-\mathbf{d}}. \tag{105}$$

Appendix H shows a rigorous derivation of the relative frequencies of branch lengths $L_{[i]}$ in a critical Tokunaga tree, which leads to essentially the same result. Figure 16(b) shows the exceedance frequency for branch lengths $L_{[i]}$ calculated in two critical Tokunaga trees with $c = 2, 3$ and $\gamma = 1$. The distributions have power-law tails (expressed as linear segments in the double logarithmic plot) with power indices given by $-\mathbf{d}(c)$.

13 Discussion

A solid body of observational, modeling, and theoretical studies ascribe Horton's laws, power-law distributions of tree attributes, and power-law relations between attributes to the self-similar structure of a tree that represents the examined system (Gupta and Waymire, 1989; Peckham, 1995; Gupta et al., 1996; Tarboton, 1996; Gupta and Waymire, 1998; Cieplak et al., 1998; Peckham and Gupta, 1999; Turcotte, 1997; Dodds and Rothman, 2000; Pelletier and Turcotte, 2000; Veitzer and Gupta, 2000; Gupta et al., 2007; Rodriguez-Iturbe and Rinaldo, 2001; Mesa, 2018). Here we review a recently formulated theory of random self-similar trees (Kovchegov and Zaliapin, 2020; Kovchegov et al., 2021) that suggests a rigorous treatment of the emergence of Horton's laws and related scalings in river networks and other dendritic systems.

Self-similarity is defined here (Sect. 4) as invariance of a coordinated tree distribution with respect to the operation of Horton pruning (cutting the source streams); this definition is justified by the empirical and modeling evidence of the past decades. Horton's laws are rigorously defined as limit statements about random or mean values of the examined branch attributes (Sects. 3.7,5,9.2). We show that self-similarity guarantees the (strongest) geometric Horton's laws for mean branch numbers and mean magnitudes (Theorem 1 and Corollary 1). Horton's laws of different strengths for multiple other mean attributes follow under additional natural assumptions (Sect. 7). Each Horton's law for a mean attribute (e.g., mean branch number $\mathcal{N}_1[K]$) implies the respective Horton's law for its random counterpart (e.g., random branch number $N_1[T]$). We have examined several commonly studied branch attributes (Table 1) whose scaling laws are well documented in the literature and have shown that the proposed self-similar model closely reproduces the scalings and exponents reported in observational studies (Table 2).

Our definition of tree self-similarity unifies several alternative definitions that have been introduced in studies of dendritic systems. Burd et al. (2000) define self-similarity in Galton-Watson trees as the Horton prune-invariance. This is a special case of our definition since the Galton-Watson trees are coordinated (Kovchegov and Zaliapin, 2020). Peckham (1995) and Newman et al. (1997) define self-similarity as Toeplitz property for the Tokunaga coefficients. This property, which only considers the Tokunaga coefficients and not the entire tree distribution, follows from and is weaker than our definition (Sect. 4.4). Moreover, Kovchegov and Zaliapin (2020) showed that the Toeplitz property alone, without coordination (Sect. 4.2), allows for a multitude of obscure measures that are hardly useful in practice. Gupta and Waymire (1989) and Peckham and Gupta (1999) suggested a concept of *statistical self-similarity* that requires a random stream attribute Z to have distribution that scales with order. It can be shown (Kovchegov and Zaliapin, 2020, Sect. 7) that (i) statistical self-similarity for some attributes (e.g., for any discrete attribute) may only hold asymptotically, and (ii) multiple attributes, including stream length, magnitude, and total basin length, are statistically self-similar in the limit of an infinitely large basin that is self-similar according to our definition.

The results reviewed herein contribute to a long-standing debate on the "inevitability" of Horton's laws in river networks (Shreve, 1966, 1969; Kirchner, 1993; McConnell and Gupta, 2008), and suggest that Horton prune-invariance is a useful paradigm for systems that exhibit such laws. The family of self-similar distributions is extremely rich and flexible. It includes the famous random topology model of Shreve (1966), which is equivalent to the critical binary Galton-Watson process with given progeny (Burd et al., 2000; Pitman, 2006), and closely fits the multitude of existing hydrologic observations summarized by Maritan et al. (1996), Turcotte (1997), Dodds and Rothman (1999), Dodds and Rothman (2000), Rodriguez-Iturbe and Rinaldo (2001), and Gupta et al. (2007); see Table 2. The self-similar family extends way beyond the hydrological constraints, allowing one to study self-similar trees with edge lengths that depend on the position within the hierarchy, having arbitrary fractal dimension $\mathbf{d} \in (1, \infty)$, and Horton branch exponent $R_B \in (2, \infty)$. For instance, the HBP (Sect. 8) might be a suitable model for phylogenetic trees (Aldous, 2001; Blum and François, 2006) or dendritic structures generated by Diffusion Limited Aggregation (DLA) (Vicsek, 1984; Ossadnik, 1992; Newman et al., 1997). While this survey focuses on binary trees, the self-similarity definition and main results are readily extended to trees with multiple branching. The proposed approach emphasizes the importance of Tokunaga coefficients T_k that have been well-known in the literature (Tokunaga, 1978; Ossadnik, 1992; Peckham, 1995; Newman et al., 1997; Gabrielov et al., 1999; Turcotte, 1999;

Pelletier and Turcotte, 2000; Holliday et al., 2008; Yoder et al., 2013) and which also play a distinct role in the presented theory. Namely, each self-similar measure is characterized by an infinite sequence $\{T_1, T_2, T_3, \dots\}$ of Tokunaga coefficients, and each such sequence corresponds to an infinite number of self-similar measures, which gives an idea of the richness of the self-similar family.

Notwithstanding this richness, the essential scalings established in the hydrological literature are closely fit by a one-parameter family of critical Tokunaga trees with $T_k = (c-1)c^{k-1}$ (Table 2). This empirical constraint has been known for long time (Tokunaga, 1978; Peckham, 1995; Veitzer and Gupta, 2000), with the special case $c = 2$ corresponding to the Shreve's random topology model (Burd et al., 2000). However, only very recently a rigorous understanding has been gained of its theoretical importance. For example, this sequence is necessary to generate tree distributions that are time-invariant, critical, and having i.i.d. edge lengths (Kovchegov and Zaliapin, 2018, 2019, 2020; Kovchegov et al., 2021).

The critical Tokunaga model presents an ultimately symmetric class of trees characterized by coordination, Horton prune-invariance, criticality, time-invariance, and identically distributed link lengths (and hence local contributing areas). Despite these multiple constraints, this class is surprisingly rich, extending from perfect binary trees ($c = 1$) to the famous Shreve's random topology model ($c = 2$) to the structures reminiscent of the observed river networks ($c \approx 2.3$) and beyond. While offering a convenient theoretical and modeling paradigm, the critical Tokunaga model is merely a subclass of a much broader family of self-similar trees that might better accommodate for various problem-specific data features. An applied study can use the self-similar theory to either focus on the symmetries of the critical Tokunaga family, or explore deviations from this stiff parameterization, both of which may have physical underpinnings.

This survey focuses on the results that concern the static structure of examined systems (e.g., river networks). Self-similarity of this structure might provide tangible constraints on the additional geometric attributes (e.g., channel slopes, junction angles, etc.) (Stark et al., 2009; Devauchelle et al., 2012), the dynamical processes that evolve along its static fabric (Mesa and Mifflin, 1986; Gupta et al., 1994, 1996; Menabde et al., 2001; Mantilla et al., 2006; Lashermes and Foufoula-Georgiou, 2007; Zaliapin et al., 2010; Gupta et al., 2010; Ramirez, 2012; Czuba and Foufoula-Georgiou, 2014) or control its formation and evolution (Seybold et al., 2007; Singh et al., 2015; Ranjbar et al., 2018).

The goal of this work was to review the recent results concerning tree self-similarity and present a simple model that explains a variety of scaling laws that are central for the studies of dendritic systems of diverse origin. Our prime illustration of the power of the proposed approach (Table 2, Fig. 5) uses empirical scaling laws of hydrogeomorphology that have been established and independently verified by multiple researchers since the 1940s. Accordingly, we intentionally avoided new data analyses and took all empirical constraints from the existing literature. An original data analysis performed through the prism of the proposed modeling approach is a topic of future research.

The proposed approach to modeling dendritic systems based on random self-similar trees is subject to further testing and verification using data from diverse fields (see Sect. 1). There are several avenues for approaching such testing. One can test the foundational principle of the theory – tree self-similarity – that combines the tree coordination and Horton prune-invariance. The coordination property can be either rigorously verified (as in Galton-Watson trees) or heuristically accepted (as in river networks). The self-similarity is then tested either by checking the Horton prune-invariance property, which might be more suitable for theoretical models, or by verifying the Toeplitz property of the Tokunaga coefficients, which can be readily done for the empirical Tokunaga coeficients $t_{i,j}$ of (19) via the ANOVA framework (Scheffe, 1999). Independently, one can check the theory's predictions. This includes (i) Horton's laws (e.g. Horton's law for branch numbers); (ii) Power law distribution of attributes calculated at individual vertices (e.g. vertex magnitudes); (iii) Power law relation between distinct attributes (e.g., Hack's law); and (iv) System's fractal dimension. Specifically, one would check if the above laws hold and if their scaling exponents are related in the way predicted by the theory (e.g., in critical Tokunaga trees the key exponents take only two distinct values). Finally, one can directly test if an observed tree (or forest of trees) can be approximated by a self-similar model. A likelihood approach to such direct statistical testing is being developed by the authors and will be presented elsewhere.

A Proof of Proposition 1

We have, for any $\Delta \in (0, r)$ (Ahlfors, 1953):

$$a_k = \frac{1}{2\pi i} \oint_{|z|=\Delta} \frac{f(z)dz}{z^{k+1}}. \tag{106}$$

By the Residue Theorem (Ahlfors, 1953), we obtain, for any $\gamma \in (r, \rho)$

$$\frac{1}{2\pi i} \oint_{|z|=\gamma} \frac{f(z)dz}{z^{k+1}} = \mathsf{Res}\left(\frac{f(z)}{z^{k+1}}; 0\right) + \mathsf{Res}\left(\frac{f(z)}{z^{k+1}}; r\right) \tag{107}$$

$$= a_k + \mathsf{Res}\left(\frac{f(z)}{z^{k+1}}; r\right). \tag{108}$$

Therefore,

$$a_k = \frac{1}{2\pi i} \oint_{|z|=\gamma} \frac{f(z)dz}{z^{k+1}} - \mathsf{Res}\left(\frac{f(z)}{z^{k+1}}; r\right), \tag{109}$$

where

$$\left| \oint_{|z|=\gamma} \frac{f(z)dz}{z^{k+1}} \right| \le \frac{\max_{|z|=\gamma} f(z)}{\gamma^k} = o\left(r^{-k}\right). \tag{110}$$

Consider $g(z) = (z - r)f(z)$. It is known that (Ahlfors, 1953)

$$\mathsf{Res}\,(f(z); r) = g(r), \tag{111}$$

and hence,

$$\mathsf{Res}\left(\frac{f(z)}{z^{k+1}};r\right)=\frac{g(r)}{r^{k+1}}=\frac{g(r)}{r}r^{-k}. \tag{112}$$

Accordingly, we obtain

$$a_k=-\frac{g(r)}{r}r^{-k}+o(r^{-k}), \tag{113}$$

which completes the proof.

B Proof of Proposition 2

First, we prove statement (i).

We begin with the root Horton's law (R). Suppose $\lim\limits_{k\to\infty}\big(b_k\big)^{1/k}=R>0$, then for a given $\epsilon\in(0,1/r_c)\subset(0,R)$, there exist $C_0,C_1>0$ such that

$$C_0(R-\epsilon)^k\le b_k\le C_1(R+\epsilon)^k \qquad \text{for all } k=0,1,\ldots.$$

Then,

$$R-\epsilon\le\liminf_{k\to\infty}\left(\sum_{j=0}^{k}b_{k-j}\,c_j\right)^{1/k}=\liminf_{k\to\infty}\big(a_k\big)^{1/k}$$

and

$$\limsup_{k\to\infty}\big(a_k\big)^{1/k}=\limsup_{k\to\infty}\left(\sum_{j=0}^{k}b_{k-j}\,c_j\right)^{1/k}\le R+\epsilon$$

since there exists $D>0$ such that $c_k\le D(1/r_c+\epsilon)^k$ for all k. Hence, since ϵ can be taken arbitrarily small, $\lim\limits_{k\to\infty}\big(a_k\big)^{1/k}=R$.

Next, we consider the quotient Horton's law (Q). Suppose $\lim\limits_{k\to\infty}\frac{b_{k+1}}{b_k}=R$. Then, by the Dominated Convergence Theorem,

$$\lim_{k\to\infty}\frac{a_k}{b_k}=\lim_{k\to\infty}\sum_{j=0}^{k}\frac{b_{k-j}}{b_k}\,c_j=\sum_{j=0}^{\infty}R^{-j}\,c_j<\infty$$

as $R^{-1}=r_b<r_c$. Hence, $\lim\limits_{k\to\infty}\frac{a_{k+1}}{a_k}=R$.

Finally, we consider the geometric Horton's law (G). Suppose $\lim\limits_{k\to\infty}b_kR^{-k}=\beta$ for some $\beta>0$. Then, by the Dominated Convergence Theorem,

$$\lim_{k\to\infty}a_kR^{-k}=\lim_{k\to\infty}\sum_{j=0}^{k}R^{j-k}b_{k-j}\,R^{-j}c_j=\beta\sum_{j=0}^{\infty}R^{-j}\,c_j=\alpha.$$

Statement (ii) follows form (i) if we write $g(z)=f(z)\tilde{h}(z)$ with $\tilde{h}(z)=\frac{1}{h(z)}$ analytic for $|z|\le r_b+\epsilon$ for some $\epsilon>0$.

C Horton's Laws for Mean Branch Attributes S_k, L_k, and A_k

Recall that the mean branch length S_k is expressed via the Tokunaga coefficients as

$$S_k=1+\sum_{i=1}^{k-1}T_i,\quad k\ge1, \tag{114}$$

and the generating function for this sequence is given by (49). The following statement is proven using Proposition 2.

Proposition 7 **[Asymptotic behavior of S_k]**
(a) If $r_T > 1$ then $r_S = 1$ and $S_k \sim 1 + T(1) = \hat{t}(1) = 1 + \sum_{k=1}^{\infty} T_k$. Accordingly, the geometric Horton's law (G) *holds.*
(b) If $r_T < 1$ then $r_S = r_T$ and S_k have the same asymptotic as T_k. Namely, the same Horton's laws hold for T_k and S_k, with the same Horton exponent. In particular, if $\lim\limits_{k\to\infty}(T_{k+1}/T_k) = R_T > 1$, then $\lim\limits_{k\to\infty}(S_{k+1}/S_k) = R_T$.
(c) If $r_T = 1$ then $r_S = 1$. In this case the sequence S_k increases to infinity at a sub-exponential rate. The geometric Horton's law (G) *does not hold. The quotient Horton's law* (Q) *and root Horton's law* (R) *may or may not hold depending on the form of T_k. See examples below.*

Example 2 **(Relation between T_k and S_k)**

(a) Finite Tokunaga sequence. *Suppose $T_1 > 0$ and $T_k = 0$ for any $k > 1$. Then $T(z) = T_1\, z$ and $r_T = \infty$ as in Prop. 7(a). According to* (49), *$S(z) = z(T_1\, z+1)(1-z)^{-1}$ and $r_S = 1$. In this case $S_1 = 1$ and $S_k = T_1 + 1$ for any $k > 1$.*

(b) Harmonic Tokunaga sequence. *Suppose $T_k = 1/k$ for any k. Then*

$$T(z) = \sum_{k=1}^{\infty} k^{-1} z^k = -\ln(1-z)$$

and $r_T = 1$ as in Prop. 7(c). According to (49), *$S(z) = z(1 - \ln(1-z))(1-z)^{-1}$ and $r_S = 1$. In this case $S_k = 1 + \sum_{i=1}^{k-1} k^{-1} \sim \ln(k)$ as $k \to \infty$. The quotient Horton's law* (Q) *holds with $R_S = 1$.*

(c) Constant Tokunaga sequence. *Suppose $T_k = 1$ for any k. Then $T(z) = \sum_{k=1}^{\infty} z^k = z(1-z)^{-1}$ and $r_T = 1$ as in Prop. 7(c). According to* (49), *$S(z) = z(1-z)^{-2}$ and $r_S = 1$. In this case $S_k = k$. The quotient Horton's law* (Q) *holds with $R_S = 1$.*

(d) Linear Tokunaga sequence. *Suppose $T_k = k$ for any k. Then $T(z) = \sum_{k=1}^{\infty} k\, z^k = z(1-z)^{-2}$ and $r_T = 1$ as in Prop. 7(c). According to* (49), *$S(z) = z(z^2 - z + 1)(1-z)^{-3}$ and $r_S = 1$. In this case $S_k = \frac{k(k-1)}{2}$. The quotient Horton's law* (Q) *holds with $R_S = 1$.*

(e) Tokunaga sequence that does not satisfy Horton's law. *Let $T_k = 2^{-k}$ if k is not a full square, and $T_k = 2^{\sqrt{k}}$ otherwise. Here $r_T = 1$ and even the root Horton's law* (R) *does not hold for T_k. We have $S_{j^2} \sim 2^j$ and so $T_{j^2}/S_{j^2} \to 1$. At the same time, $T_{j^2-1}/S_{j^2-1} \to 0$. This means that the limit of T_k/S_k does not exist, and so the limit of $S_{k+1}/S_k = 1 + T_k/S_k$, which is equivalent to quotient Horton's law* (Q), *does not exist. The root Horton's law* (R) *also does not hold since $\left(2^{\sqrt{k}}\right)^{1/k} \to 1$, while $\left(2^{-k}\right)^{1/k} = 2^{-1}$.*

(f) Geometric Tokunaga sequence. *Suppose $T_k = (c-1)c^{k-1}$ for any k with some $c > 1$. Then $T(z) = (c-1)\sum_{k=1}^{\infty} c^{k-1}\, z^k = (c-1)z(1-cz)^{-1}$ and $r_T = c^{-1}$ as in Prop. 7(b). According to* (49), *$S(z) = z(1-cz)^{-1}$ and $r_S = c^{-1}$. In this case $S_k = c^{k-1}$.*

The generating function of the mean branch lengths L_k is given by

$$L(z) = \sum_{k=1}^{\infty} L_k z^k = \sum_{k=1}^{\infty} S_k \ell_k z^k.$$

If we assume root Horton's law (R) for ℓ_k with Horton exponent λ^{-1}, then as $L_k = S_k \ell_k$, we have $r_L = \lambda^{-1} r_S$ whenever $\lambda > 0$, and $r_L = \infty$ if $\lambda = 0$.

Proposition 8 **[Asymptotic behavior of L_k]**
Suppose the root Horton's law holds for ℓ_k with Horton exponent λ^{-1}, $\lambda \in (0, \infty)$.
(a) If $r_T > 1$ then $r_L = \lambda^{-1}$. The same Horton's laws hold for ℓ_k and L_k.
(b) If $r_T < 1$ then $r_L = \lambda^{-1} r_T$. If some Horton's laws hold for T_k and ℓ_k, then the weakest of those holds for L_k.
(c) If $r_T = 1$ then $r_L = \lambda^{-1}$. The geometric Horton's law (G), *the quotient Horton's law* (Q) *and the root Horton's law* (R) *may or may not hold for L_k.*

Example 3 (Relation between T_k and L_k)
If ℓ_k scales geometrically with k, i.e., there is a scalar $\lambda > 0$ such that $\ell_k = \ell_1\lambda^{k-1}$, then

$$L(z) = \ell_1\lambda^{-1}\sum_{k=1}^{\infty} S_k\lambda^k z^k = \ell_1\lambda^{-1}\, S\big(\lambda z\big).$$

Here, if $r_T > 1$, and therefore $T(z)$ converges to a finite value $T(1)$ at $z = 1$, we have

$$L_k = S_k\ell_k \;\sim\; \ell_1\lambda^{k-1}\big(1 + T(1)\big) \;= \ell_1\lambda^{k-1}\hat{t}(1).$$

In this case, the geometric Horton's law (G) *holds with Horton exponent λ^{-1}. If $T_k = (c-1)c^{k-1}$ for some $c > 1$, then $r_T = c^{-1} < 1$ and $S_k = c^{k-1}$. Hence, $L_k = \ell_1(\lambda\, c)^{k-1}$ and the geometric Horton's law* (G) *holds for L_k with Horton exponent $(\lambda\, c)^{-1}$.*

The generating function $A(z)$ for the contributing areas is given by

$$A(z) = \frac{\sum\limits_{k=1}^{\infty}\alpha_k S_k z^k}{1 - 2z - T(z)} = M(z)\left(\sum_{k=1}^{\infty}\alpha_k S_k z^{k-1}\right) = -\frac{D(z)}{\hat{t}(z)}. \tag{115}$$

Proposition 9 [Asymptotic behavior of A_k]
Suppose the root Horton's law holds for α_k with Horton exponent α^{-1}, $\alpha \in (0,\infty)$. Recall that $\delta_k = \alpha_k S_k$.
(a) If $r_T > 1$ then $r_D = \alpha^{-1}$. The same Horton's laws hold for α_k and δ_k. If $\alpha^{-1} < w_0$ then $r_A = \alpha^{-1}$ and $A_k \sim$ Const. $\times\, \alpha_k$. If $\alpha^{-1} > w_0$ then $r_A = w_0$ and $A_k \sim$ Const. $\times\, M_k$.
(b) If $r_T < 1$ then $r_D = \alpha^{-1}r_T$. If $r_D < w_0$ then $r_A = r_D$ and $A_k \sim$ Const. $\times\, \delta_k$. If $r_D < w_0$ and some of the Horton's laws hold for T_k and α_k, the weakest of those holds for δ_k and A_k. If $r_D > w_0$ then $r_A = w_0$ and $A_k \sim$ Const. $\times\, M_k$.
(c) If $r_T = 1$ then $r_D = \alpha^{-1}$. The geometric Horton's law (G)*, the quotient Horton's law* (Q) *and the root Horton's law* (R) *may or may not hold for δ_k. If $\alpha^{-1} < w_0$ then $r_A = \alpha^{-1}$ and $A_k \sim$ Const. $\times\, \delta_k$. If $\alpha^{-1} > w_0$ then $r_A = w_0$ and $A_k \sim$ Const. $\times\, M_k$.*

D Relation Between Edge Lengths and Differential Contributing Areas

Consider a random variable ξ_k representing the length of a randomly selected edge of order $k \geq 1$. Suppose there exists $C \geq 0$ such that for all $k \geq 1$, the expectation $\mathsf{E}[\xi_k] = \ell_k$ and the standard deviation $\mathsf{SD}(\xi_k)$ satisfy

$$\frac{\mathsf{SD}(\xi_k)}{\ell_k} \leq C.$$

Let ν_k be a random variable representing the differential contributing area of the edge, and suppose there exist scalars $b > 0$ and $\sigma > 1$ such that $\delta_k = b\,\xi_k^\sigma$ with σ and b being the same for all orders $k \geq 1$. Then, by Jensen's inequality and the size biasing method (Bhattacharya and Waymire, 2007), we have

$$\begin{aligned} \ell_k^\sigma &\leq \mathsf{E}[\xi_k^\sigma] = \ell_k\mathsf{E}[\zeta_k^{\sigma-1}] \leq \ell_k\mathsf{E}[\zeta_k]^{\sigma-1} \\ &= \ell_k\left(\frac{\mathsf{E}[\xi_k^2]}{\ell_k}\right)^{\sigma-1} = \ell_k^\sigma\left(1 + \frac{\mathsf{Var}(\xi_k)}{\ell_k^2}\right)^{\sigma-1} \leq (1 + C^2)^{\sigma-1}\ell_k^\sigma, \end{aligned} \tag{116}$$

where ζ_k is a random variable distributed as $P(\zeta_k \in A) = \frac{1}{\ell_k}\mathsf{E}\big[\xi_k\,\mathbf{1}_A(\xi_k)\big]$.

Hence, the expected contributing area $\alpha_k = b\,\mathsf{E}[\xi_k^\sigma]$ of an order k link satisfies

$$b\ell_k^\sigma \leq \alpha_k \leq b(1 + C^2)^{\sigma-1}\ell_k^\sigma.$$

Accordingly, the root Horton's law holds for the lengths implies that for the areas:

$$\lim_{k\to\infty} \ell_k^{1/k} = \lambda^{-1} \quad \Rightarrow \quad \lim_{k\to\infty} \alpha_k^{1/k} = \lambda^{-\sigma}.$$

Our analysis does not establish the quotient Horton's law, since

$$(1+C^2)^{1-\sigma}\left(\frac{\ell_{k+1}}{\ell_k}\right)^{\sigma} \le \frac{\alpha_{k+1}}{\alpha_k} \le (1+C^2)^{\sigma-1}\left(\frac{\ell_{k+1}}{\ell_k}\right)^{\sigma}.$$

At the same time, the gap between the low and upper bounds above can be small, hence implying the quotient Horton's law with a practical level of accuracy. For instance, if $\sigma = 2$ and $C = 1$, then that gap is $[1/2, 2]$. One also can add to that the observation that increasing/decreasing sequence of lengths corresponds to the increasing/decreasing sequence of areas, which bounds the Horton exponent by 1 from below of above.

E Tokunaga Two-parameter Model

This section discusses a slight relaxation of the critical Tokunaga process constraint $T_k = (c-1)c^{k-1}$. Specifically, we consider the sequence of Tokunaga coefficients that has been introduced in river studies by Tokunaga (1966, 1978, 1984):

$$\frac{T_{k+1}}{T_k} = c \qquad \text{or} \qquad T_k = a\,c^{k-1} \qquad \text{for} \quad a, c > 0. \tag{117}$$

The trees satisfying (117) are usually called *Tokunaga trees*. The Tokunaga trees have been shown to closely approximate multiple observed branching structures beyond river networks (Peckham, 1995; McConnell and Gupta, 2008; Zanardo et al., 2013; Dodds and Rothman, 2000; Gabrielov et al., 1999; Kovchegov and Zaliapin, 2016; Newman et al., 1997; Ossadnik, 1992; Pelletier and Turcotte, 2000; Turcotte et al., 1998). A perfect binary tree is a Tokunaga tree with $a = 0$ and arbitrary c. The critical binary Galton-Watson tree corresponds to $(a, c) = (1, 2)$ (Burd et al., 2000). The critical Tokunaga process of Sect. 4.1 corresponds to a special case $a = c - 1$. The geometric behavior of the Tokunaga's indices allows one to find an explicit form of the generating function $T(z)$ and makes the branching analysis particularly straightforward.

If $c = 1$, then $S_k = 1 + \sum_{j=1}^{k-1} a = 1 + a(k-1)$. If $c \neq 1$, then

$$\begin{aligned} S_k &= 1 + \sum_{j=1}^{k-1} T_j = 1 + \sum_{j=1}^{k-1} a c^{j-1} = 1 + a\frac{c^{k-1}-1}{c-1} \\ &= \begin{cases} a(c-1)^{-1}c^{k-1} + O(1), & c > 1 \\ 1 + a(1-c)^{-1} + O(c^k), & c < 1 \end{cases} \end{aligned} \tag{118}$$

Next, we have

$$\hat{t}(z) = -1 + 2z + \frac{az}{1-cz} = \frac{-1 + (a+c+2)z - 2cz^2}{1-cz}, \tag{119}$$

whose two real roots are

$$w_{1,0} = \frac{(a+c+2) \pm \sqrt{(a+c+2)^2 - 8c}}{4c},$$

with $w_1 > w_0$. The smallest root w_0 has been reported in multiple works (e.g., Peckham, 1995; McConnell and Gupta, 2008). Here we have

$$R_B = w_0^{-1} = 2\,c\,w_1 = \frac{(a+c+2) + \sqrt{(a+c+2)^2 - 8c}}{2}. \tag{120}$$

Accordingly,

$$M(z) = -\frac{z}{\hat{t}(z)} = \frac{z(1-cz)}{2c(z-w_0)(z-w_1)}, \tag{121}$$

and, by Theorem 1,

$$\begin{aligned}\mathcal{N}_{K-k+1}[K] = M_k &= \frac{1}{w_0^k}\frac{(1-cw_0)}{2c(w_1-w_0)} + o\left(\frac{1}{w_0^k}\right)\\ &= \frac{1}{w_0^k}\frac{(2-c-a)+\sqrt{(2+c+a)^2-8c}}{4\sqrt{(2+c+a)^2-8c}} + o\left(\frac{1}{w_0^k}\right),\end{aligned} \tag{122}$$

The exact expression is derived in Appendix F:

$$\mathcal{N}_{K-k+1}[K] = M_k = w_0^{-k+1}\left[1+(1-c\,w_1)\sum_{i=1}^{k-1}\left(\frac{w_0}{w_1}\right)^i\right]. \tag{123}$$

To examine the branch contributing areas A_k one need to make additional assumptions about the process. We adopt here the hydrological constraint of Assumption 2 of Sect. 7.3:

$$\lim_{k\to\infty}\frac{\ell_{k+1}}{\ell_k} = 1. \tag{124}$$

Combining this with (64) gives

$$\lim_{k\to\infty}\frac{\ell_{k+1}}{\ell_k} = \zeta\lim_{k\to\infty}\frac{S_k}{S_{k+1}} = 1. \tag{125}$$

The geometric form (117) of the Tokunaga coefficients implies that the quotient Horton's law holds for T_k. Proposition 4 then ensures that the quotient Horton's law also holds for S_k with Horton exponent $R_S = \max\{1,c\} \geq 1$. Hence, $R_S = \zeta$ and $\zeta \geq 1$, which implies that the mean branch length is non-decreasing with order. The field observations (Rodriguez-Iturbe and Rinaldo, 2001) strongly suggest that the stream length increases geometrically with order. This implies $\zeta > 1$ and hence $\zeta = c > 1$, which also means geometric growth of T_k.

The edge lengths in the HBP have exponential distribution. We use Example 1 to find

$$\lim_{k\to\infty}\frac{\alpha_{k+1}}{\alpha_k} = \lim_{k\to\infty}\left(\frac{\ell_{k+1}}{\ell_k}\right)^\sigma = 1. \tag{126}$$

Accordingly, Assumption 3 is satisfied here with $c > 1$, and Proposition 6 guarantees the existence of the quotient Horton's law for A_k.

One can obtain a stronger result by recalling

$$A(z) = -\frac{D(z)}{\hat{t}(z)}. \tag{127}$$

Proposition 8 applied to α_k states that the radius of convergence for $D(z)$ is $r_D = r_T = c^{-1} > w_0$. This means that the asymptotic behavior of A_k is determined by the simple pole of $\hat{t}(z)^{-1}$. In other words, $A_k \sim AM_k$, where the proportionality constant is given by

$$A = \frac{D(w_0)(1-cw_0)}{2w_0c(w_1-w_0)}. \tag{128}$$

Accordingly, the geometric Horton's law holds for $\mathcal{N}_k[K]$, M_k, and A_k with the Horton exponent $R_B = R_M = R_A = w_0^{-1}$.

F Exact Form of Mean Magnitudes M_k in Tokunaga Two-parameter Model

If $T_k = a\,c^{k-1}$, equation (34) implies

$$M_k = 2\,M_{k-1} + a\sum_{i=1}^{k-1} c^{k-i-1} M_i \qquad (k \geq 2) \tag{129}$$

and therefore,

$$M_{k+1} = (2+a)\,M_k + a\sum_{i=1}^{k-1} c^{k-i} M_i \qquad (k \geq 2). \tag{130}$$

Equations (129) and (130) yield the following recursion

$$M_{k+1} - c\,M_k = (2+a)\,M_k - 2c\,M_{k-1}$$

which simplifies to

$$M_{k+1} - (2+a+c)\,M_k + 2c\,M_{k-1} = 0. \tag{131}$$

The recurrence relation (131) is solved by finding the roots of its characteristic equation

$$x^2 - (2+a+c)x + 2c = 0. \tag{132}$$

The roots of (132) equal R_B and $2+a+c-R_B$. See (120). Therefore,

$$M_k = c_1 R_B^{k-1} + c_2(2+a+c-R_B)^{k-1},$$

where the initial conditions $M_1 = 1$ and $M_2 = 2 + T_1 = 2 + a$ yield

$$c_1 = \frac{R_B - c}{2R_B - 2 - a - c} \quad \text{and} \quad c_2 = \frac{R_B - 2 - a}{2R_B - 2 - a - c}.$$

Hence,

$$\mathcal{N}_1[K] = M_K = \frac{R_B - c}{2R_B - 2 - a - c} R_B^{K-1} + \frac{R_B - 2 - a}{2R_B - 2 - a - c}(2+a+c-R_B)^{K-1}. \tag{133}$$

Notice that, by (120),

$$R_B > 2 + a + c - R_B > 0.$$

Recall that $2 + a + c = 2c(w_0 + w_1)$ and $R_B = w_0^{-1} = 2cw_1$. Thus, $2 + a + c - R_B = w_1^{-1}$, and (133) can be rewritten as follows

$$M_K = \frac{2w_1 - 1}{2(w_1 - w_0)} w_0^{1-K} + \frac{1 - 2w_0}{2(w_1 - w_0)} w_1^{1-K}.$$

Finally, for $1 \leq j \leq K$, (133) yields

$$\mathcal{N}_j[K] = \mathcal{N}_1[K-j+1] = \frac{R_B - c}{2R_B - 2 - a - c} R_B^{K-j} + \frac{R_B - 2 - a}{2R_B - 2 - a - c}(2+a+c-R_B)^{K-j}. \tag{134}$$

G Exact Quotient Horton's Law for Mean Branch Counts, Magnitudes

Assume that the quotient Horton's law for the mean branch counts $\mathcal{N}_1[K]$, and hence for the mean magnitudes M_K, holds exactly, that is (using the fact that $M_1 = 1$):

$$M_K = R_M^{K-1}. \tag{135}$$

Then,

$$M(z) = \frac{z}{1 - R_M z}$$

which leads to

$$t(z) = -\frac{z}{M(z)} = -1 + R_M z \quad \text{and} \quad T(z) = (R_M - 2)z.$$

This implies that the only self-similar model with exact quotient Horton's law for the mean branch counts and magnitudes corresponds to the Tokunaga sequence

$$T_1 = R_M - 2, \quad T_k = 0 \quad \text{for} \quad k > 1.$$

H Power Law Frequency for Branch Lengths in Critical Tokunaga Trees

Here we examine the frequencies of branch lengths $L_{[i]}$ in critical Tokunaga trees. First, we prove the following technical lemma that establishes a power law decay of a function that later will be interpreted as the survival function of the branch lengths.

Lemma 2 *Define*

$$\phi(x) = \sum_{k=0}^{\infty} \alpha^{-k} \exp\left(-\beta^{-k} x\right), \quad x > 0$$

for some constants $\alpha > 1$ and $\beta > 1$. Then,

$$\phi(x) = b(x) x^{-\frac{\log \alpha}{\log \beta}}, \tag{136}$$

where $b(x)$ is a function bounded from above and from below by positive constants.

Proof Consider

$$\phi(x) = \sum_{k=0}^{\infty} \alpha^{-k} e^{-\beta^{-k} x}, \quad x > 0.$$

Let

$$\sigma(x) = \sum_{k=-\infty}^{\infty} \alpha^{-k} e^{-\beta^{-k} x}, \quad x > 0.$$

Since $\lim\limits_{x \to \infty} \frac{\phi(x)}{\sigma(x)} = 1$, we have

$$\phi(x) = \sigma(x)\big(1 + o(1)\big) \quad \text{as} \quad x \to \infty.$$

Next, observe that

$$\sigma(x) = \alpha\, \sigma(\beta x),$$

and therefore

$$\frac{\Big(\sigma(\beta x)\Big)^{-\frac{\log \beta}{\log \alpha}}}{\beta x} = \frac{\Big(\sigma(x)\Big)^{-\frac{\log \beta}{\log \alpha}}}{x}.$$

Hence,

$$\pi(y) = e^{-y}\Big(\sigma(e^y)\Big)^{-\frac{\log\beta}{\log\alpha}}$$

is a positive continuous periodic function with period $\log\beta$ that is bounded away from $+\infty$ and from 0. Observe that

$$\sigma(x) = \Big(x\pi(\log x)\Big)^{-\frac{\log\alpha}{\log\beta}}.$$

Accordingly, $\phi(x)$ can be expressed as

$$\phi(x) = \sigma(x)\big(1+o(1)\big) = b(x)x^{-\frac{\log\alpha}{\log\beta}},$$

where

$$b(x) = \Big(\pi(\log x)\Big)^{-\frac{\log\alpha}{\log\beta}}\big(1+o(1)\big)$$

is a positive function, bounded from above and from below, by positive quantities. This completes the proof.

Recall from Sect. 9.2, Eq. (81) that the empirical frequencies of branch orders in a critical Tokunaga tree of order K are approximately geometric:

$$\{\text{proportion of branches of order } k\} = (R_B-1)R_B^{-k}(1+o(1)) \tag{137}$$

for any $k \geq 1$ as $K \to \infty$. The lengths of branches of order k are i.i.d. exponential random variables with rate $\gamma\zeta^{1-k}$. Accordingly, the relative proportion of the lengths of order-k branches that exceed a given value $x > 0$ is $\exp(-\gamma\zeta^{1-k}x)$. Taking into account the relative proportions of branches of different orders, we find the relative proportion of branches with length exceeding x:

$$\begin{aligned}
\frac{\#\{i : L_{[i]} \geq x\}}{\text{total no. branches}} &= \sum_{k=1}^{\infty}(R_B-1)R_B^{-k}\exp(-\gamma\zeta^{1-k}x) \\
&= (1-R_B^{-1})\sum_{k=0}^{\infty}R_B^{-k}\exp(-\gamma\zeta^{-k}x) \\
&= (1-R_B^{-1})b(\gamma x)(\gamma x)^{-\frac{\log R_B}{\log\zeta}} \\
&\propto b(\gamma x)x^{-\frac{\log R_B}{\log\zeta}},
\end{aligned}$$

where $b(x)$ is a positive function bounded from zero and infinity. Here the next to the last step uses Lemma 2 with $\alpha = R_B$ and $\beta = \zeta$.

Acknowledgements We are grateful to Ed Waymire, Alejandro Tejedor, and Victor Peñaranda for discussions related to this survey. We thank three anonymous reviewers and Associate Editor, Alik Ismail-Zadeh, for their editorial considerations and comments that helped us to improve the work and expand its scope.

Conflict of Interest

The authors declare that they have no conflict of interest.

References

Ahlfors LV (1953) Complex analysis: An introduction to the theory of analytic functions of one complex variable. McGraw-Hill, New York

Aldous DJ (2001) Stochastic models and descriptive statistics for phylogenetic trees, from Yule to today. Statistical Science 16(1):23–34, DOI 10.1214/ss/998929474

Allegre CJ, Mouel JLL, Provost A (1982) Scaling rules in rock fracture and possible implications for earthquake prediction. Nature 297(5861):47–49, DOI https://doi.org/10.1038/297047a0

Badii R, Politi A (1999) Complexity: Hierarchical structures and scaling in physics, vol 6. Cambridge University Press, 318 pp

Baiesi M, Paczuski M (2004) Scale-free networks of earthquakes and aftershocks. Physical Review E 69(6):066106, DOI https://doi.org/10.1103/PhysRevE.69.066106

Balister P, Balogh J, Bertuzzo E, Bollobás B, Caldarelli G, Maritan A, Mastrandrea R, Morris R, Rinaldo A (2018) River landscapes and optimal channel networks. Proceedings of the National Academy of Sciences 115(26):6548–6553, DOI https://doi.org/10.1073/pnas.1804484115

Barenblatt GI (1996) Scaling, self-similarity and intermediate asymptotics. Cambridge University Press

Barndorff-Nielsen OE (1993) River networks: a brief guide to the literature for statisticians and probabilists. In Chaos and Networks: Statistical and Probabilistic Aspects, Chapman & Hall, pp 276–287

Barnes EA, Power ME, Foufoula-Georgiou E, Hondzo M, Dietrich WE (2007) Upscaling river biomass using dimensional analysis and hydrogeomorphic scaling. Geophysical Research Letters 34(24)

Baró J (2020) Topological properties of epidemic aftershock processes. Journal of Geophysical Research: Solid Earth 125:5, DOI https://doi.org/10.1029/2019JB018530

Ben-Zion Y (2008) Collective behavior of earthquakes and faults: Continuum-discrete transitions, progressive evolutionary changes, and different dynamic regimes. Reviews of Geophysics 46:4, DOI https://doi.org/10.1029/2008RG000260

Bhattacharya RN, Waymire EC (2007) A basic course in probability theory, vol 69. Springer, New York

Blum MG, François O (2006) Which random processes describe the tree of life? A large-scale study of phylogenetic tree imbalance. Systematic Biology 55(4):685–691, DOI https://doi.org/10.1080/10635150600889625

Burd G, Waymire EC, Winn RD (2000) A self-similar invariance of critical binary Galton-Watson trees. Bernoulli 6(1):1–21, DOI https://doi.org/10.2307/3318630

Burridge R, Knopoff L (1967) Model and theoretical seismicity. Bulletin of the Seismological Society of America 57(3):341–371, DOI https://doi.org/10.1785/BSSA0570030341

Cassot F, Lauwers F, Fouard C, Prohaska S, Lauwers-Cances V (2006) A novel three-dimensional computer-assisted method for a quantitative study of microvascular networks of the human cerebral cortex. Microcirculation 13(1):1–18, DOI 10.1080/10739680500383407

Chunikhina EV (2018a) Entropy rates for horton self-similar trees. Chaos 28(8):081104, DOI https://doi.org/10.1063/1.5048965

Chunikhina EV (2018b) Information theoretical analysis of self-similar trees. PhD thesis, Oregon State University

Cieplak M, Giacometti A, Maritan A, Rinaldo A, Rodriguez-Iturbe I, Banavar JR (1998) Models of fractal river basins. Journal of Statistical Physics 91(1-2):1–15, DOI https://doi.org/10.1023/A:1023069201470

Czuba JA, Foufoula-Georgiou E (2014) A network-based framework for identifying potential synchronizations and amplifications of sediment delivery in river basins. Water Resources Research 50(5):3826–3851, DOI 10.1002/2013WR014227

Davidsen J, Stanchits S, Dresen G (2007) Scaling and universality in rock fracture. Physical Review Letters 98(12):125502, DOI https://doi.org/10.1103/PhysRevLett.98.125502

De Vries H, Becker T, Eckhardt B (1994) Power law distribution of discharge in ideal networks. Water Resources Research 30(12):3541–3543, DOI https://doi.org/10.1029/94WR02178

Devauchelle O, Petroff AP, Seybold HF, Rothman DH (2012) Ramification of stream networks. Proceedings of the National Academy of Sciences 109(51):20832–20836, DOI https://doi.org/10.1073/pnas.1215218109

Devroye L, Kruszewski P (1994) A note on the Horton-Strahler number for random trees. Inform Processing Lett 56:95–99

Dodds P, Rothman D (1999) Unified view of scaling laws for river networks. Physical Review E 59(5):4865–4877, DOI https://doi.org/10.1103/PhysRevE.59.4865

Dodds P, Rothman D (2000) Scaling, universality, and geomorphology. Ann Rev Earth and Planet Sci 28:571–610, DOI https://doi.org/10.1146/annurev.earth.28.1.571

Dodov B, Foufoula-Georgiou E (2004a) Generalized hydraulic geometry: Derivation based on a multiscaling formalism. Water resources research 40(6)

Dodov B, Foufoula-Georgiou E (2004b) Generalized hydraulic geometry: Insights based on fluvial instability analysis and a physical model. Water Resources Research 40(12)

Dodov B, Foufoula-Georgiou E (2005) Fluvial processes and streamflow variability: Interplay in the scale-frequency continuum and implications for scaling. Water Resources Research 41(5)

Drmota M, Prodinger H (2006) The register function for t-ary trees. ACM Transactions on Algorithms 2(3):318–334, DOI https://doi.org/10.1145/1159892.1159894

Flajolet P, Raoult JC, Vuillemin J (1979) The number of registers required for evaluating arithmetic expressions. Theoretical Computer Science 9(1):99–125, DOI https://doi.org/10.1016/0304-3975(79)90009-4

Gabrielov AM, Levshina TA, Rotwain IM (1990) Block model of earthquake sequence. Physics of the Earth and Planetary Interiors 61(1-2):18–28, DOI https://doi.org/10.1016/0031-9201(90)90091-B

Gabrielov AM, Newman WI, Turcotte DL (1999) Exactly soluble hierarchical clustering model: Inverse cascades, self-similarity, and scaling. Physical Review E 60(5):5293, DOI https://doi.org/10.1103/PhysRevE.60.5293

Gangodagamage C, Barnes E, Foufoula-Georgiou E (2007) Scaling in river corridor widths depicts organization in valley morphology. Geomorphology 91(3-4):198–215

Grant EHC, Lowe WH, Fagan WF (2007) Living in the branches: Population dynamics and ecological processes in dendritic networks. Ecology Letters 10(2):165–175, DOI https://doi.org/10.1111/j.1461-0248.2006.01007.x

Gupta VK (2017) Scaling Theory of Floods for Developing a Physical Basis of Statistical Flood Frequency Relations. Research Encyclopedia of Natural Hazard Science, Oxford, URL `https://oxfordre.com/naturalhazardscience/view/10.1093/acrefore/9780199389407.001.0001/acrefore-9780199389407-e-301`

Gupta VK, Mesa OJ (2014) Horton laws for hydraulic-geometric variables and their scaling exponents in self-similar Tokunaga river networks. Nonlinear Processes in Geophysics 21(5):1007–1025, DOI https://doi.org/10.5194/npg-21-1007-2014

Gupta VK, Waymire EC (1989) Statistical self-similarity in river networks parameterized by elevation. Water Resources Research 25(3):463–476, DOI https://doi.org/10.1029/WR025i003p00463

Gupta VK, Waymire EC (1998) Some mathematical aspects of rainfall, landforms and floods. In: Barndorff-Nielsen OE, Gupta VK, Perez-Abreu V, Waymire EC (eds) Stochastic Methods in Hydrology: Rainfall, Landforms and Floods, World Scientific, Singapore, Singapore, pp 129–172

Gupta VK, Mesa OJ, Dawdy DR (1994) Multiscaling theory of flood peaks: Regional quantile analysis. Water Resources Research 30(12):3405–3421, DOI https://doi.org/10.1029/94WR01791

Gupta VK, Castro SL, Over TM (1996) On scaling exponents of spatial peak flows from rainfall and river network geometry. Journal of Hydrology 187(1-2):81–104, DOI https://doi.org/10.1016/S0022-1694(96)03088-0

Gupta VK, Troutman BM, Dawdy DR (2007) Towards a nonlinear geophysical theory of floods in river networks: An overview of 20 years of progress. Nonlinear Dynamics in Geosciences pp 121–151

Gupta VK, Mantilla R, Troutman BM, Dawdy D, Krajewski WF (2010) Generalizing a nonlinear geophysical flood theory to medium-sized river networks. Geophysical Research Letters 37:11, DOI https://doi.org/10.1029/2009GL041540

Gutenberg B, Richter CF (1954) Seismicity of the Earth. Princeton University Press, Princeton, NJ

Hack JT (1957) Studies of longitudinal stream profiles in Virginia and Maryland, vol 294. US Government Printing Office

Herrmann HJ, Roux S (2014) Statistical models for the fracture of disordered media. Elsevier

Holliday JR, Turcotte DL, Rundle JB (2008) Self-similar branching of aftershock sequences. Physica A: Statistical Mechanics and its Applications 387(4):933–943, DOI https://doi.org/10.1016/j.physa.2007.09.045

Horton RE (1945) Erosional development of streams and their drainage basins; hydrophysical approach to quantitative morphology. Bulletin of Geophysical Society of America 56:275–370, DOI https://doi.org/10.1130/0016-7606(1945)56[275:EDOSAT]2.0.CO;2

Kagan YY (1973) A probabilistic description of the seismic regime. Izv Acad Sci USSR, Phys Solid Earth 213:219

Kagan YY, Knopoff L (1976) Statistical search for non-random features of the seismicity of strong earthquakes. Physics of the Earth and Planetary Interiors 12(4):291–318, DOI https://doi.org/10.1016/0031-9201(76)90026-1

Kagan YY, Knopoff L (1981) Stochastic synthesis of earthquake catalogs. Journal of Geophysical Research: Solid Earth 86(B4):2853–2862, DOI https://doi.org/10.1029/JB086iB04p02853

Kassab GS (2000) The coronary vasculature and its reconstruction. Annals of Biomedical Engineering 28(8):903–915, DOI https://doi.org/10.1114/1.1308494

Keilis-Borok V (2002) Earthquake prediction: State-of-the-art and emerging possibilities. Annual Review of Earth and Planetary Sciences 30(1):1–33, DOI 10.1146/annurev.earth.30.100301.083856

Kemp R (1979) The average number of registers needed to evaluate a binary tree optimally. Acta Informatica 11:363–372, DOI https://doi.org/10.1007/BF00289094

Kirchner JW (1993) Statistical inevitability of Horton's laws and the apparent randomness of stream channel networks. Geology 21(7):591–594, DOI https://doi.org/10.1130/0091-7613(1993)021⟨0591:SIOHSL⟩2.3.CO;2

Kovchegov Y, Zaliapin I (2016) Horton's law in self-similar trees. Fractals 24:1650017, DOI https://doi.org/10.1142/S0218348X16500171

Kovchegov Y, Zaliapin I (2018) Tokunaga self-similarity arises naturally from time invariance. Chaos 28(4):041102, DOI https://doi.org/10.1063/1.5029937

Kovchegov Y, Zaliapin I (2019) Random self-similar trees and a hierarchical branching process. Stochastic Processes and Their Applications 129(7):2528–2560, DOI https://doi.org/10.1016/j.spa.2018.07.015

Kovchegov Y, Zaliapin I (2020) Random self-similar trees: A mathematical theory of Horton's laws. Probability Surveys 17:1–213, DOI https://doi.org/10.1214/19-PS331

Kovchegov Y, Zaliapin I, Foufoula-Georgiou E (2021) Critical Tokunaga model for river networks, in review

La Barbera P, Rosso R (1989) On the fractal dimension of stream networks. Water Resources Research 25(4):735–741, DOI https://doi.org/10.1029/WR025i004p00735

Lashermes B, Foufoula-Georgiou E (2007) Area and width functions of river networks: New results on multifractal properties. Water Resour Res 43:W09405, DOI 10.1029/2006WR005329

Lashermes B, Foufoula-Georgiou E, Dietrich WE (2007) Channel network extraction from high resolution topography using wavelets. Geophys Res Lett 34:L23S04, DOI 10.1029/2007GL031140

Leopold LB, Miller JP (1956) Ephemeral streams: Hydraulic factors and their relation to the drainage net, vol 282-A. US Government Printing Office, 37 pp

Leopold LB, Wolman MG, Miller J (1992) Fluvial processes in geomorphology. Dover Publications, Inc., New York, USA

Malamud BD, Morein G, Turcotte DL (1998) Forest fires: An example of self-organized critical behavior. Science 281(5384):1840–1842, DOI 10.1126/science.281.5384.1840

Malamud BD, Turcotte DL, Guzzetti F, Reichenbach P (2004a) Landslide inventories and their statistical properties. Earth Surface Processes and Landforms 29(6):687–711, DOI https://doi.org/10.1002/esp.1064

Malamud BD, Turcotte DL, Guzzetti F, Reichenbach P (2004b) Landslides, earthquakes, and erosion. Earth and Planetary Science Letters 229(1-2):45–59, DOI https://doi.org/10.1016/j.epsl.2004.10.018

Mandelbrot BB (1982) The Fractal Geometry of Nature, vol 1. W. H. Freeman and Company, New York

Mantilla R, Gupta VK, Mesa OJ (2006) Role of coupled flow dynamics and real network structures on Hortonian scaling of peak flows. Journal of Hydrology 322(1-4):155–167, DOI https://doi.org/10.1016/j.jhydrol.2005.03.022

Maritan A, Rinaldo A, Rigon R, Giacometti A, Rodriguez-Iturbe I (1996) Scaling laws for river networks. Physical Review E 53(2):1510, DOI https://doi.org/10.1103/PhysRevE.53.1510

McConnell M, Gupta V (2008) A proof of the Horton's law of stream numbers for the Tokunaga model of river networks. Fractals 16:227–233, DOI https://doi.org/10.1142/S0218348X08003958

Menabde M, Veitzer S, Gupta V, Sivapalan M (2001) Tests of peak flow scaling in simulated self-similar river networks. Advances in Water Resources 24(9-10):991–999, DOI https://doi.org/10.1016/S0309-1708(01)00043-4

Mesa OJ (2018) Cuatro modelos de redes de drenaje, revista de la academia colombiana de ciencias exactas. Físicas y Naturales 42(165):379–391, DOI https://doi.org/10.18257/raccefyn.641

Mesa OJ, Gupta VK (1987) On the main channel length-area relationship for channel networks. Water Resources Research 23(11):2119–2122, DOI https://doi.org/10.1029/WR023i011p02119

Mesa OJ, Mifflin ER (1986) On the relative role of hillslope and network geometry in hydrologic response. In: Scale problems in hydrology, Springer, Dordrecht, pp 1–17, DOI https://doi.org/10.1007/978-94-009-4678-1_1

Molz FJ, Liu HH, Szulga J (1997) Fractional brownian motion and fractional Gaussian noise in subsurface hydrology: A review, presentation of fundamental properties, and extensions. Water Resources Research 33(10):2273–2286

Neuman SP (1990) Universal scaling of hydraulic conductivities and dispersivities in geologic media. Water Resources Research 26(8):1749–1758, DOI https://doi.org/10.1029/WR026i008p01749

Newman WI, Turcotte DL, Gabrielov AM (1997) Fractal trees with side-branching. Fractals 5:603–614, DOI https://doi.org/10.1142/S0218348X97000486

Ogata Y (1998) Space-time point-process models for earthquake occurrences. Annals of the Institute of Statistical Mathematics 50(2):379–402, DOI https://doi.org/10.1023/A:1003403601725

Omori F (1894) On the after-shocks of earthquakes. J Coll Sci, Imp Univ, Japan 7:111–200

Ossadnik P (1992) Branch order and ramification analysis of large diffusion-limited-aggregation clusters. Physical Review A 45(2):1058, DOI https://doi.org/10.1103/PhysRevA.45.1058

Passalacqua P, Trung TD, Foufoula-Georgiou E, Sapiro G, Dietrich WE (2010) A geometric framework for channel network extraction from LiDAR: Nonlinear diffusion and geodesic paths. J Geophys Res 115:F01002, DOI 10.1029/2009JF001254

Peckham SD (1995) New results for self-similar trees with applications to river networks. Water Resources Research 31(1):1023–1029, DOI https://doi.org/10.1029/94WR03155

Peckham SD, Gupta VK (1999) A reformulation of Horton's laws for large river networks in terms of statistical self-similarity. Water Resources Research 35(9):2763–2777, DOI https://doi.org/10.1029/1999WR900154

Pelletier JD, Turcotte DL (2000) Shapes of river networks and leaves: are they statistically similar? Philosophical Transactions of the Royal Society of London B: Biological Sciences 355(1394):307–311, DOI https://doi.org/10.1098/rstb.2000.0566

Pitman J (2006) Combinatorial Stochastic Processes. Ecole d'été de probabilités de Saint-Flour XXXII-2002, Lectures on Probability Theory and Statistics, Springer

Ramirez J (2012) Population persistence under advection-diffusion in river networks. Journal of Mathematical Biology 65(5):919–942, DOI https://doi.org/10.1007/s00285-011-0485-6

Ranjbar S, Hooshyar M, Singh A, Wang D (2018) Quantifying climatic controls on river network branching structure across scales. Water Resources Research 54(10):7347–7360, DOI https://doi.org/10.1029/2018WR022853

Rigon R, Rinaldo A, Rodriguez-Iturbe I, Bras RL, Ijjasz-Vasquez E (1993) Optimal channel networks: A framework for the study of river basin morphology. Water Resources Research 29(6):1635–1646, DOI https://doi.org/10.1029/92WR02985

Rigon R, Rodriguez-Iturbe I, Maritan A, Giacometti A, Tarboton D, Rinaldo A (1996) On Hack's law. Water Resources Research 32(11):3367–3374, DOI https://doi.org/10.1029/96WR02397

Rinaldo A, Rodriguez-Iturbe I, Rigon R, Bras RL, Ijjasz-Vasquez E, Marani A (1992) Minimum energy and fractal structures of drainage networks. Water Resources Research 28(9):2183–2195, DOI https://doi.org/10.1029/92WR00801

Rodriguez-Iturbe I, Rinaldo A (2001) Fractal river basins: chance and self-organization. Cambridge University Press

Rodriguez-Iturbe I, Ijjasz-Vasquez EJ, Bras RL, Tarboton DG (1992) Power law distributions of discharge mass and energy in river basins. Water Resources Research 28(4):1089–1093, DOI https://doi.org/10.1029/91WR03033

Roering JJ, Kirchner JM, Dietrich WE (1999) Evidence for nonlinear, diffusive sediment transport on hillslopes and implications for landscape morphology. Water Resources Research 35(3):853–870, DOI https://doi.org/10.1029/1998WR900090

Saichev A, Helmstetter A, Sornette D (2005) Power-law distributions of offspring and generation numbers in branching models of earthquake triggering. Pure and Applied Geophysics 162(6):1113–1134, DOI https://doi.org/10.1007/s00024-004-2663-6

Sapozhnikov V, Foufoula-Georgiou E (1996) Self-affinity in braided rivers. Water Resources Research 32(5):1429–1439, DOI https://doi.org/10.1029/96WR00490

Scheffe H (1999) The analysis of variance, vol 72. Wiley, John & Sons

Scheidegger AE (1967) A stochastic model for drainage patterns into an intramontane treinch. Hydrological Sciences Journal 12(1):15–20, DOI https://doi.org/10.1080/02626666709493507

Seybold H, Andrade JS, Herrmann HJ (2007) Modeling river delta formation. Proceedings of the National Academy of Sciences 104(43):16804–16809, DOI https://doi.org/10.1073/pnas.0705265104

Shreve RL (1966) Statistical law of stream numbers. J Geol 74(1):17–37, DOI https://doi.org/10.1086/627137

Shreve RL (1969) Stream lengths and basin area in topologically random channel networks. J Geol 77:397–414, DOI https://doi.org/10.1086/628366

Singh A, Reinhardt L, Foufoula-Georgiou E (2015) Landscape reorganization under changing climatic forcing: Results from an experimental landscape. Water Resour Res 51(6):4320–4337, DOI 10.1002/2015WR017161

Soloviev A, Ismail-Zadeh A (2003) Models of dynamics of block-and-fault systems. In: Nonlinear Dynamics of the Lithosphere and Earthquake Prediction, Springer, Berlin, Heidelberg, pp 71–139

Stark CP, Foufoula-Georgiou E, Ganti V (2009) A nonlocal theory of sediment buffering and bedrock channel evolution. J Geophys Res 114:F01029, DOI 10.1029/2008JF000981

Strahler AN (1957) Quantitative analysis of watershed geomorphology. Trans Am Geophys Un 38:913–920, URL `https://doi.org/10.1029/TR038i006p00913`

Takayasu H, Nishikawa I, Tasaki H (1988) Power-law mass distribution of aggregation systems with injection. Physical Review A 37(8):3110, DOI https://doi.org/10.1103/PhysRevA.37.3110

Tarboton DG (1996) Fractal river networks, Horton's laws and Tokunaga cyclicity. Journal of Hydrology 187(1):105–117, DOI https://doi.org/10.1016/S0022-1694(96)03089-2

Tarboton DG, Bras RL, Rodriguez-Iturbe I (1988) The fractal nature of river networks. Water Resources Research 24:1317–1322, DOI https://doi.org/10.1029/WR024i008p01317

Tarboton DG, Bras RL, Rodriguez-Iturbe I (1989) Scaling and elevation in river networks. Water Resources Research 25(9):2037–2051, DOI https://doi.org/10.1029/WR025i009p02037

Tejedor A, Longjas A, Zaliapin I, Foufoula-Georgiou E (2015a) Delta channel networks: 1. A graph-theoretic approach for studying connectivity and steady state transport on deltaic surfaces. Water Resources Research 51(6):3998–4018

Tejedor A, Longjas A, Zaliapin I, Foufoula-Georgiou E (2015b) Delta channel networks: 2. Metrics of topologic and dynamic complexity for delta comparison, physical inference, and vulnerability assessment. Water Resources Research 51(6):4019–4045

Tejedor A, Longjas A, Edmonds DA, Zaliapin I, Georgiou TT, Rinaldo A, Foufoula-Georgiou E (2017) Entropy and optimality in river deltas. Proceedings of the National Academy of Sciences 114(44):11651–11656

Tessler ZD, Vorosmarty CJ, Grossberg M, Gladkova I, Aizenman H, Syvitski J, Foufoula-Georgiou E (2015) Profiling risk and sustainability in coastal deltas of the world. Science 349(6248):638–643, DOI 10.1126/science.aab3574

Tokunaga E (1966) The composition of drainage networks in Toyohira river basin and valuation of Horton's first law. Geophysical Bulletin, Hokkaido University 15:1–19

Tokunaga E (1978) Consideration on the composition of drainage networks and their evolution. Geographical Reports of Tokyo Metropolitan University 13:1–27

Tokunaga E (1984) Ordering of divide segments and law of divide segment numbers. Transaction Japanese Geomorphological Union 5:71–77

Turcotte DL (1997) Fractals and chaos in geology and geophysics. Cambridge University Press

Turcotte DL (1999) Self-organized criticality. Reports on Progress in Physics 62(10):1377, DOI https://doi.org/10.1088/0034-4885/62/10/201

Turcotte DL, Pelletier JD, Newman WI (1998) Networks with side-branching in biology. Journal of Theoretical Biology 193(4):577–592, DOI https://doi.org/10.1006/jtbi.1998.0723

Turcotte DL, Malamud BD, Morein G, Newman WI (1999) An inverse cascade model for self-organized critical behavior. Physica A 268:629–643, DOI https://doi.org/10.1016/S0378-4371(99)00092-8

Turcotte DL, Malamud BD, Guzzetti F, Reichenbach P (2002) Self-organization, the cascade model, and natural hazards. Proceedings of the National Academy of Sciences 99(suppl 1):2530–2537, DOI https://doi.org/10.1073/pnas.012582199

Utsu T, Ogata Y, Matsuura R (1995) The centenary of the Omori formula for a decay law of aftershock activity. Journal of Physics of the Earth 43(1):1–33, DOI https://doi.org/10.4294/jpe1952.43.1

Veitzer SA, Gupta VK (2000) Random self-similar river networks and derivations of generalized Horton's laws in terms of statistical simple scaling. Water Resour Res 36(4):1033–1048, DOI https://doi.org/10.1029/1999WR900327
Vere-Jones D (1976) A branching model for crack propagation. Pure and Applied Geophysics 114(4):711–725, DOI https://doi.org/10.1007/BF00875663
Vicsek T (1984) Pattern formation in diffusion-limited aggregation. Physical Review Letters 53(24):2281, DOI https://doi.org/10.1103/PhysRevLett.53.2281
Wang SX, Waymire EC (1991) A large deviation rate and central limit theorem for Horton ratios. SIAM Journal on Discrete Mathematics 4(4):575–588, DOI https://doi.org/10.1137/0404050
West GB, Brown JH, Enquist BJ (1997) A general model for the origin of allometric scaling laws in biology. Science 276(5309):122–126, DOI 10.1126/science.276.5309.122
Wilf HS (1992) Generatingfunctionology. Philadelphia, PA, USA, freely available from `http://www.math.upenn.edu/~wilf/DownldGF.html`
Yakovlev G, Newman WI, Turcotte DL, Gabrielov A (2005) An inverse cascade model for self-organized complexity and natural hazards. Geophysical Journal International 163(2):433–442, DOI https://doi.org/10.1111/j.1365-246X.2005.02717.x
Yamamoto K (2017) Central limit theorem for the Horton-Strahler bifurcation ratio of general branch order. Journal of Applied Probability 54(4):1111–1124, DOI https://doi.org/10.1017/jpr.2017.55
Yekutieli I, Mandelbrot BB (1994) Horton-Strahler ordering of random binary trees. J Phys A: Math Gen 27(2):285–293
Yoder MR, Aalsburg JV, Turcotte DL, Abaimov SG, Rundle JB (2013) Statistical variability and Tokunaga branching of aftershock sequences utilizing BASS model simulations. Pure and Applied Geophysics 170(1-2):155–171, DOI https://doi.org/10.1007/s00024-011-0411-2
Zaliapin I, Ben-Zion Y (2013) Earthquake clusters in southern California I: Identification and stability. Journal of Geophysical Research: Solid Earth 118(6):2847–2864, DOI https://doi.org/10.1002/jgrb.50179
Zaliapin I, Keilis-Borok V, Ghil M (2003) A Boolean delay equation model of colliding cascades. Part II: Prediction of critical transitions. Journal of Statistical Physics 111(3):839–861, DOI https://doi.org/10.1023/A:1022802432590
Zaliapin I, Gabrielov A, Keilis-Borok V, Wong H (2008) Clustering analysis of seismicity and aftershock identification. Physical Review Letters 101(1):018501, DOI https://doi.org/10.1103/PhysRevLett.101.018501
Zaliapin I, Foufoula-Georgiou E, Ghil M (2010) Transport on river networks: A dynamic tree approach. Journal of Geophysical Research: Earth Surface 115(F2):F00A15, DOI https://doi.org/10.1029/2009JF001281
Zanardo S, Zaliapin I, Foufoula-Georgiou E (2013) Are American rivers Tokunaga self-similar? New results on fluvial network topology and its climatic dependence. Journal of Geophysical Research: Earth Surface 118:1–18, DOI https://doi.org/10.1029/2012JF002392